



DAGSTUHL REPORTS

Volume 13, Issue 2, February 2023

Scheduling (Dagstuhl Seminar 23061) <i>Nicole Megow, Benjamin J. Moseley, David Shmoys, Ola Svensson, Sergei Vassilvitskii, and Jens Schlöter</i>	1
Programming Language Processing (Dagstuhl Seminar 23062) <i>Michael Pradel, Baishakhi Ray, Charles Sutton, and Eran Yahav</i>	20
From Big Data Theory to Big Data Practice (Dagstuhl Seminar 23071) <i>Martin Farach-Colton, Fabian Daniel Kuhn, Ronitt Rubinfeld, and Przemysław Uznański</i>	33
Challenges and Perspectives in Deep Generative Modeling (Dagstuhl Seminar 23072) <i>Vincent Fortuin, Yingzhen Li, Kevin Murphy, Stephan Mandt, and Laura Manduchi</i>	47
Agents on the Web (Dagstuhl Seminar 23081) <i>Olivier Boissier, Andrei Ciortea, Andreas Harth, Alessandro Ricci, and Danai Vachtsevanou</i>	71
Resilient Software Configuration and Infrastructure Code Analysis (Dagstuhl Seminar 23082) <i>Jürgen Cito, Ruzica Piskac, Mark Santolucito, Andy Zaidman, and Daniel Sokolowski</i>	163
Algorithmic Foundations of Programmable Matter (Dagstuhl Seminar 23091) <i>Aaron Becker, Sándor Fekete, Irina Kostitsyna, Matthew J. Patitz, Damien Woods, and Ioannis Chatzigiannakis</i>	183
A Human-Computer Interaction Perspective to Drive Change towards Sustainable Future (Dagstuhl Perspectives Workshop 23092) <i>Susanne Boll, Nic Bidwell, Kaisa Väänänen, and Marc Hassenzahl</i>	199

ISSN 2192-5283

Published online and open access by

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <https://www.dagstuhl.de/dagpub/2192-5283>

Publication date

September, 2023

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <https://dnb.d-nb.de>.

License

This work is licensed under a Creative Commons Attribution 4.0 International license (CC BY 4.0).



In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.

The copyright is retained by the corresponding authors.

Aims and Scope

The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.

In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:

- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e. g. summaries from panel discussions or open problem sessions.

Editorial Board

- Elisabeth André
- Franz Baader
- Daniel Cremers
- Goetz Graefe
- Reiner Hähnle
- Barbara Hammer
- Lynda Hardman
- Oliver Kohlbacher
- Steve Kremer
- Rupak Majumdar
- Heiko Mantel
- Albrecht Schmidt
- Wolfgang Schröder-Preikschat
- Raimund Seidel (*Editor-in-Chief*)
- Heike Wehrheim
- Verena Wolf
- Martina Zitterbart

Editorial Office

Michael Wagner (*Managing Editor*)
Michael Didas (*Managing Editor*)
Jutka Gasiorowski (*Editorial Assistance*)
Dagmar Glaser (*Editorial Assistance*)
Thomas Schillo (*Technical Assistance*)

Contact

Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de
<https://www.dagstuhl.de/dagrep>

Digital Object Identifier: 10.4230/DagRep.13.2.i

Scheduling

Nicole Megow^{*1}, Benjamin J. Moseley^{*2}, David Shmoys^{*3},
Ola Svensson^{*4}, Sergei Vassilvitskii^{*5}, and Jens Schlöter^{†6}

- 1 Universität Bremen, DE. nicole.megow@uni-bremen.de
- 2 Carnegie Mellon University – Pittsburgh, US. moseleyb@andrew.cmu.edu
- 3 Cornell University – Ithaca, US. david.shmoys@cornell.edu
- 4 EPFL – Lausanne, CH. ola.svensson@epfl.ch
- 5 Google – New York, US. sergeiv@gmail.com
- 6 Universität Bremen, DE. jschloet@uni-bremen.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23061 “Scheduling”. The seminar focused on the emerging models for beyond-worst case algorithm design, in particular, recent approaches that incorporate learning. This includes models for the integration of learning into algorithm design that have been proposed recently and that have already demonstrated advances in the state-of-art for various scheduling applications: (i) scheduling with error-prone learned predictions, (ii) data-driven algorithm design, and (iii) stochastic and Bayesian learning in scheduling.

Seminar February 5–10, 2023 – <https://www.dagstuhl.de/23061>

2012 ACM Subject Classification Theory of computation → Scheduling algorithms; Mathematics of computing → combinatorial optimization; Theory of computation → Approximation algorithms analysis

Keywords and phrases scheduling, mathematical optimization, approximation algorithms, learning methods, uncertainty

Digital Object Identifier 10.4230/DagRep.13.2.1

1 Executive Summary


Nicole Megow (Universität Bremen, DE)

Benjamin J. Moseley (Carnegie Mellon University – Pittsburgh, US)

David Shmoys (Cornell University – Ithaca, US)

Ola Svensson (EPFL – Lausanne, CH)

Sergei Vassilvitskii (Google – New York, US)

License  Creative Commons BY 4.0 International license
© Nicole Megow, Benjamin J. Moseley, David Shmoys, Ola Svensson, and Sergei Vassilvitskii

This Dagstuhl Seminar was the seventh in a series of Dagstuhl “Scheduling” seminars (since 2008). Scheduling is a major research field that is studied from a practical and theoretical perspective in computer science, mathematical optimization, and operations research. Applications range from traditional production scheduling and project planning to the newly arising resource management tasks in the advent of internet technology and shared resources. Despite the remarkable progress on algorithmic theory for fundamental scheduling problems, new questions gain greater prominence due to the rise of new applications and new methodologies.

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Scheduling, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 1–19

Editors: Nicole Megow, Benjamin J. Moseley, David Shmoys, Ola Svensson, Sergei Vassilvitskii, and Jens Schlöter



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

At this meeting, we focused on the emerging models for **beyond-worst case analysis**, especially for models that incorporate learning. Decades of research have produced algorithmic discoveries along with impossibility results for scheduling. The primary directions have been focused on achieving the best performance in worst-case scenarios, in terms of run time, memory usage, and proximity to the optimum. Unfortunately, it has often been observed that such algorithms do not necessarily perform the best in practice. Closing the gap between theory and practice is often impossible using the standard approach due to lower bound barriers on worst-case instances.

To bridge the gap between theory and practice, there has been an effort to understand algorithms through a new lens, called beyond-worst-case algorithm analysis. Initial work in this area has produced exciting algorithmic insights and has given convincing explanations for why some algorithms work well in practice. Moreover, the area has led to interesting new algorithmic insights and theoretical directions.

Several of the emerging beyond-worst-case analysis models incorporate learning. Developing the understanding of the models have the promise of discovering new algorithmic insights into scheduling problems and improved performance. This seminar focused on the following three themes.

- *Scheduling with Error-Prone Learned Predictions.* This theme focuses on the settings where the algorithm has access to (e.g. machine-learned) predictions. In this model, an algorithm is given access to a prediction about the problem instance. The performance of the algorithm is parameterized by the quality of this prediction. Typically, an algorithm is given access to an error-prone prediction as machine learning is often imperfect. This prediction can then be leveraged to make algorithmic decisions. Ideally, (1) the predictions result in better performance than the best worst-case bound; (2) the algorithm never performs asymptotically worse than the best worst-case algorithm even if the prediction error is large; and (3) in-between, there is a graceful degradation in performance as the error in the prediction becomes worse. For example, the competitive ratio, approximation ratio or running time can be parameterized by the prediction error.
- *Data-Driven Algorithm Design.* Practitioners often do not use algorithms with the best worst-case performance guarantees directly. Rather, they investigate several algorithms and optimize over them to find the one that works the best for the given application. That is, input data is used to determine the best algorithm (empirically) to use for a given application. Motivated by this, recent research seeks to give a theoretical foundation for this kind of data-driven algorithm selection. In this setting, a family or class of algorithms are defined for a problem. This family is constructed by the algorithm designer. The designer has access to prior instances of the problem that are drawn from a fixed but unknown probability distribution over instances. The goal is to choose an algorithm from the family that will have the best expected performance. That is, to learn from prior instances and decide on the best algorithm for future instances. This model is closely related to practice. Indeed, it is often the case that a scheduling system will have access to job instances from prior days when making scheduling decisions on the current day.
- *Stochastic and Bayesian Learning in Scheduling.* Similar to the learning-augmented setting, we wish to understand how to best incorporate uncertain knowledge about the future in algorithm design. The main difference lies in that, while the learning-augmented framework makes no assumptions on the quality of predictions, the stochastic setting assumes that the data is drawn from a probability distribution. The distribution can either be known upfront or can be learned over time. Questions of interest include:

- If we know the distribution of data, how can we best use it in the design of algorithms?
- If the distribution of data is learned over time, how can we best make decisions while being adaptive to new knowledge of the data?

The scheduling community has recently made striking contributions to both these questions. When the data distribution is known up front – distribution of processing times in this case – there has been major progress on algorithms for scheduling jobs on multiple machines. This is an interesting practical and theoretical question as the processing time of one job may affect the completion of another. An interesting extension to the above question is when the algorithm is allowed some *adaptivity* in a second stage: after computing a schedule based on the distributional knowledge as above, the real processing times are revealed and the algorithm can adapt by making small changes to the schedule. It is known that some settings that near optimal schedule algorithms exist if jobs arrive from underlying distributions, but in the worst-case there are strong lower bounds showing algorithms must be far from optimal. The goal is to better understand the when and how algorithms should optimize over probability distributions.

Organization of the Seminar. The seminar brought together 43 researchers from theoretical computer science, mathematical optimization, and operations research. The participants consisted of both senior and junior researchers, including a number of postdocs and advanced PhD students.

During the five days of the seminar, 29 talks of different lengths took place. Six keynote speakers gave an overview of the state-of-the art of the respective area resp. presented recent highlight results in 60 minutes:

- Siddhartha Banerjee: We Need to Talk About How we Talk About Online Decision-Making
- Sami Davies: Scheduling with Communication Delays
- Silvio Lattanzi: Clustering with Advice
- Claire Mathieu: Stable matching in Practice
- Debmalya Panigrahi: Santa Claus with Predictions
- Sergei Vassilvitskii: Five(ish) Years of Algorithms with Predictions.

The remaining slots were filled with shorter talks of 30 minutes on various topics related to scheduling, resource allocation, and related problems, from the perspective of coping with uncertainty, learning, and applications in practice.

Further, in the beginning of the week, open problem sessions were held. One of the open problem sessions was devoted to problems that would have been of interest to Gerhard Woeginger, pivotal member of the community that passed away in 2022. Throughout the week, a few sessions with spotlight talks of 8 minutes gave participants the chance to announce recent results and invite for discussions. The schedule left ample free time that was actively used for fruitful discussions and joint research.

Outcome. Organizers and participants regard the seminar as a great success. The seminar achieved the goal to bring together the related communities, share the state-of-the art research and discuss the current major challenges. The talks were excellent and stimulating; participants actively met in working groups in the afternoon and evenings. It was remarked positively that a significant number of younger researchers (postdocs and PhD students) participated and integrated well.

Acknowledgements. The organizers wish to express their gratitude towards the Scientific Directorate and the administration of the Dagstuhl Center for their great support for this seminar.

2 Table of Contents

Executive Summary

<i>Nicole Megow, Benjamin J. Moseley, David Shmoys, Ola Svensson, and Sergei Vassilvitskii</i>	1
--	---

Overview of Talks

Paging with Succinct Predictions <i>Antonios Antoniadis</i>	6
Flow Time Scheduling with Uncertain Processing Time <i>Yossi Azar</i>	6
Better Trees for Santa Claus <i>Étienne Bamas</i>	7
We Need to Talk About How we Talk About Online Decision-Making <i>Siddhartha Banerjee</i>	7
Scheduling with communication delays <i>Sami Davies</i>	8
Stochastic Configuration Balancing <i>Franziska Eberle</i>	8
Decentralized Multi-agent Learning in Bipartite Queuing Systems <i>Daniel Freund</i>	9
Online Matching with Reusable Capacities <i>Vineet Goyal</i>	9
Online Load Balancing Beyond ℓ_p Norms <i>Thomas Kesselheim, Marco Molinaro, and Sahil Singla</i>	10
Online Algorithms with Multiple Predictions <i>Amit Kumar</i>	10
Clustering with Advice <i>Silvio Lattanzi</i>	11
Algorithms with Prediction Portfolios <i>Thomas Lavastida</i>	11
Predictions for Online Min-Sum Scheduling Problems <i>Alex Lindermayr</i>	12
DAG Scheduling Problems <i>Alberto Marchetti-Spaccamela</i>	12
Stable matching in practice <i>Claire Mathieu</i>	12
Online Rounding of Bipartite Matchings <i>Seffi Naor</i>	13
Learning-augmented Assignment: Santa Claus does Load Balancing <i>Debmalya Panigrahi</i>	13
Evaluating Stochastic Score Functions <i>Kevin Schewior</i>	14

Learning-Augmented Query Policies for Minimum Spanning Tree with Uncertainty <i>Jens Schlöter</i>	14
New bounds for approximation algorithms for graph burning <i>Jiri Sgall</i>	15
Hiking through the complexity landscape with Gerhard – a personal reminiscence <i>Martin Skutella</i>	15
Stochastic Minimum-Norm Combinatorial Optimization <i>Chaitanya Swamy</i>	15
Fast Graph Algorithms with Learned Duals <i>Ali Vakilian</i>	16
Five(ish) Years of Algorithms with Predictions <i>Sergei Vassilvitskii</i>	17
Search infected nodes in uncertain graphs <i>Jose Verschae</i>	17
Bayesian Scheduling: Analysis of simple policies <i>Tjark Vredeveld</i>	17
Minimizing Completion Times for Stochastic Jobs via Batched Free Times <i>Rudy Zhou</i>	18
Participants	19

3 Overview of Talks

3.1 Paging with Succinct Predictions

Antonios Antoniadis (University of Twente, NL)

License © Creative Commons BY 4.0 International license
© Antonios Antoniadis

Joint work of Antonios Antoniadis, Joan Boyar, Marek Eliás, Lene M. Favrholdt, Ruben Hoeksma, Kim S. Larsen, Adam Polak, Bertrand Simon

Main reference Antonios Antoniadis, Joan Boyar, Marek Eliás, Lene M. Favrholdt, Ruben Hoeksma, Kim S. Larsen, Adam Polak, Bertrand Simon: “Paging with Succinct Predictions”, CoRR, Vol. abs/2210.02775, 2022.

URL <https://doi.org/10.48550/arXiv.2210.02775>

We study learning-augmented paging from the new perspective of requiring the least possible amount of predicted information. More specifically, the predictions obtained alongside each page request are limited to one bit only. We consider two natural such setups: (i) discard predictions, in which the predicted bit denotes whether or not it is “safe” to evict this page, and (ii) phase predictions, where the bit denotes whether the current page will be requested in the next phase (for an appropriate partitioning of the input into phases). We develop learning-augmented algorithms for each of the two setups and establish that our algorithms are essentially best possible. We believe that succinct predictions are of interest for other problems beyond paging.

3.2 Flow Time Scheduling with Uncertain Processing Time

Yossi Azar (Tel Aviv University, IL)

License © Creative Commons BY 4.0 International license
© Yossi Azar

Joint work of Yossi Azar, Stefano Leonardi, Noam Touitou

Main reference Yossi Azar, Stefano Leonardi, Noam Touitou: “Flow Time Scheduling with Uncertain Processing Time”, CoRR, Vol. abs/2103.05604, 2021.

URL <https://arxiv.org/abs/2103.05604>

We consider the problem of online scheduling on a single machine to minimize unweighted and weighted flow time. The existing algorithms for these problems require exact knowledge of the processing time of each job. This assumption is crucial, as even a slight perturbation of the processing time would lead to polynomial competitive ratio. However, this assumption very rarely holds in real-life scenarios. We present a competitive algorithm (the competitive ratio is a function of the distortion) for unweighted flow time that does not require knowledge of the distortion in advance. For the weighted flow time we present competitive algorithms but, in this case, we need to know (an upper bound on) the distortion in advance. This is joint work with Stefano Leonardi and Noam Touitou based on papers that appeared in STOC 21 and SODA 2022.

3.3 Better Trees for Santa Claus

Étienne Bamas (EPFL – Lausanne, CH)

License © Creative Commons BY 4.0 International license

© Étienne Bamas

Joint work of Étienne Bamas, Lars Rohwedder

Main reference Étienne Bamas, Lars Rohwedder: “Better Trees for Santa Claus”, CoRR, Vol. abs/2211.14259, 2022.

URL <https://doi.org/10.48550/arXiv.2211.14259>

A notorious open problem in approximation algorithms is whether there exists a constant factor approximation for MaxMin Fair Allocation of indivisible items (also known as the Santa Claus problem). Bateni, Charikar, and Guruswami [STOC’09] introduced the MaxMin Arborescence problem as an important special case: Given a directed graph with sources and sinks we have to find vertex disjoint arborescences rooted in the sources such that at each non-sink vertex of an arborescence the out-degree is at least k , where k is to be maximized. This problem is of particular interest, since it appears to capture much of the difficulty of the general Santa Claus problem. Indeed, the progress made by Bateni et al. was quickly generalized by Chakrabarty, Chuzhoy, and Khanna [FOCS’09] to the general case. These two results remain the state-of-the-art for both problems, and they yield a polylogarithmic approximation in quasi-polynomial time.

In this talk, I will present the main ideas behind an exponential improvement to this, a $\text{poly}(\log \log n)$ -approximation in quasi-polynomial time for the MaxMin Arborescence problem.

3.4 We Need to Talk About How we Talk About Online Decision-Making

Siddhartha Banerjee (Cornell University – Ithaca, US)

License © Creative Commons BY 4.0 International license

© Siddhartha Banerjee

Scheduling, and other problems involving online allocation of resources, are topics of great interest across many academic communities. However, the huge diversity in underlying models and methodologies means that existing assumptions/algorithms/guarantees are difficult to understand and compare (and often not very useful...).

I will try to present the stochastic control viewpoint on these problems, and discuss how there seems to be a fundamental divide between the view of online algorithms in CS and in controls. I will then present a sample-path coupling technique, which provides a simple way of reasoning about online algorithms, and regret guarantees against any chosen benchmark. I will describe how this framework gives new algorithms and insights for a variety of problems, including (time permitting):

1. *Constant regret* algorithms (i.e., having additive loss compared to the hindsight optimal solution which is *independent of the horizon and budget*) for several widely-studied settings including online packing, load balancing, dynamic pricing, assortment optimization, and online bin packing.
2. Incorporating side information and historical data in these settings (and achieve constant regret with as little as a single data trace).
3. Fundamental tradeoffs in multi-objective settings (in particular, for fairness in online allocation).

3.5 Scheduling with communication delays

Sami Davies (Northwestern University – Evanston, US)

License  Creative Commons BY 4.0 International license
 Sami Davies

Joint work of Sami Davies, Janardhan Kulkarni, Thomas Rothvoss, Sai Sandeep, Jakub Tarnawski, Yiaho Zhang

I'll discuss progress on scheduling with communication delays. In this setting, if two dependent jobs are scheduled on different machines, a delay must pass between their execution times. The question of whether constant factor approximation algorithms exist in this setting was one of the biggest open problems in scheduling theory. We effectively answered this question by (1) finding polylog approximation algorithms when the delay is uniform between dependent jobs and (2) proving super-constant hardness when the delay is non-uniform between dependent jobs.

3.6 Stochastic Configuration Balancing

Franziska Eberle (London School of Economics, GB)

License  Creative Commons BY 4.0 International license
 Franziska Eberle

Joint work of Franziska Eberle, Anupam Gupta, Nicole Megow, Benjamin Moseley, Rudy Zhou

Main reference Franziska Eberle, Anupam Gupta, Nicole Megow, Benjamin Moseley, Rudy Zhou: “Configuration Balancing for Stochastic Requests”, CoRR, Vol. abs/2208.13702, 2022.

URL <https://doi.org/10.48550/arXiv.2208.13702>

The configuration balancing problem with stochastic requests generalizes many well-studied resource allocation problems such as load balancing and virtual circuit routing. In it, we have m resources and n requests. Each request has multiple possible configurations, each of which increases the load of each resource by some amount. The goal is to select one configuration for each request to minimize the makespan: the load of the most-loaded resource. In our work, we focus on a stochastic setting, where we only know the distribution for how each configuration increases the resource loads, learning the realized value only after a configuration is chosen. We develop both offline and online algorithms for configuration balancing with stochastic requests. When the requests are known offline, we give a non-adaptive policy for configuration balancing with stochastic requests that $O(\frac{\log m}{\log \log m})$ -approximates the optimal adaptive policy. In particular, this closes the adaptivity gap for this problem as there is an asymptotically matching lower bound even for the very special case of load balancing on identical machines. When requests arrive online in a list, we give a non-adaptive policy that is $O(\log m)$ competitive. Again, this result is asymptotically tight due to information-theoretic lower bounds for very special cases (e.g., for load balancing on unrelated machines). Finally, we show how to leverage adaptivity in the special case of load balancing on related machines to obtain a constant-factor approximation offline and an $O(\log \log m)$ -approximation online. A crucial technical ingredient in all of our results is a new structural characterization of the optimal adaptive policy that allows us to limit the correlations between its decisions.

3.7 Decentralized Multi-agent Learning in Bipartite Queuing Systems

Daniel Freund (MIT – Cambridge, US)

License © Creative Commons BY 4.0 International license
© Daniel Freund

Joint work of Daniel Freund, Thodoris Lykouris, Wentao Weng

Main reference Daniel Freund, Thodoris Lykouris, Wentao Weng: “Efficient decentralized multi-agent learning in asymmetric queuing systems”, CoRR, Vol. abs/2206.03324, 2022.

URL <https://doi.org/10.48550/arXiv.2206.03324>

Learning in multi-agent systems often poses significant challenges due to interference between agents. In particular, unlike classical stochastic systems, the performance of an agent’s action is not drawn i.i.d. from some distribution but is directly affected by the (unobserved) actions of the other agents. This is the reason why most collaborative multi-agent learning approaches aim to globally coordinate all agents’ actions to evade this interference.

In this talk, we focus on agents in a decentralized bipartite queuing system, where N agents request service from K servers. Prior decentralized approaches aim to globally identify a coordinated schedule or do not take advantage of the bipartite structure, which leads to significant shortcomings: performance that degrades exponentially in the number of servers, requirement of shared randomness and unique identifiers, and computationally demanding algorithms. In contrast, we provide a low-complexity algorithm that is run decentrally by each agent, avoids the shortcomings of “global coordination” and leads the queuing system to have efficient performance in asymmetric bipartite queuing systems while also having additional robustness properties.

3.8 Online Matching with Reusable Capacities

Vineet Goyal (Columbia University – New York, US)

License © Creative Commons BY 4.0 International license
© Vineet Goyal

Joint work of Vineet Goyal, Garud Iyengar, Rajan Udhwani

Main reference Vineet Goyal, Garud Iyengar, Rajan Udhwani: “Asymptotically Optimal Competitive Ratio for Online Allocation of Reusable Resources”. CoRR aVol. bs/2002.02430 (2022)

URL <https://doi.org/10.48550/arXiv.2002.02430>

Online matching problems are at the heart of resource allocation given the inherent demand uncertainty. For instance, in a typical setting users arrive sequentially to a platform and the platform needs to make irrevocable matching decisions in an online manner. We consider fundamental generalizations of the classical variants in order to incorporate some of the natural stochasticity in resource usage that arises in many applications.

This talk mainly focuses on understanding the impact of reusability of resource capacities – a key aspect of resource allocation in applications such as cloud computing and sharing economies. Here allocated resources are required by users for some a priori unknown (stochastic) durations. Resources are returned after use and are available for re-allocation. We propose a new policy that achieves the best possible guarantee of $(1-1/e - o(1))$ under reasonable assumptions. Further, in the process of analyzing this policy we develop a novel framework of analysis that is useful more broadly in other settings with stochastic resource consumption.

3.9 Online Load Balancing Beyond ℓ_p Norms

Thomas Kesselheim (Universität Bonn, DE), Marco Molinaro, and Sahil Singla

License © Creative Commons BY 4.0 International license
© Thomas Kesselheim, Marco Molinaro, and Sahil Singla

Joint work of Thomas Kesselheim, Marco Molinaro, Sahil Singla

Main reference Thomas Kesselheim, Marco Molinaro, Sahil Singla: “Online and Bandit Algorithms Beyond p Norms”, in Proc. of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023, pp. 1566–1593, SIAM, 2023.

URL <https://doi.org/10.1137/1.9781611977554.ch58>

The classic problem of online makespan minimization can be understood as minimizing the ℓ_∞ norm of the vector of machine loads. It was extended to ℓ_p norms already more than 25 years ago. We study the problem beyond ℓ_p norms.

We show that general norms admit good algorithms as long as the norm can be approximated by a function that is “gradient-stable”, a notion that we introduce. Roughly it says that the gradient of the function should not drastically decrease in any component as we increase the input vector.

In particular, we give the first $O(\log^2 m)$ -competitive algorithm for online load balancing with respect to an arbitrary monotone symmetric norm. Our techniques extend to applications beyond symmetric norms as well, e.g., to Online Vector Scheduling and to Online Generalized Assignment with Convex Costs.

The set of techniques can also be applied in stochastic settings, e.g., in which the sequence of jobs arrives in random order. Given that they can be understood as based on duality, it would be interesting to see if they also can be used with error-prone predictions.

3.10 Online Algorithms with Multiple Predictions

Amit Kumar (Indian Institute of Technology – New Dehli, IN)

License © Creative Commons BY 4.0 International license
© Amit Kumar

Joint work of Keerti Anand, Rong Ge, Amit Kumar, Debmalya Panigrahi

Main reference Keerti Anand, Rong Ge, Amit Kumar, Debmalya Panigrahi: “Online Algorithms with Multiple Predictions”, CoRR, Vol. abs/2205.03921, 2022.

URL <https://doi.org/10.48550/arXiv.2205.03921>

This paper studies online algorithms augmented with multiple machine-learned predictions. While online algorithms augmented with a single prediction have been extensively studied in recent years, the literature for the multiple predictions setting is sparse. In this paper, we give a generic algorithmic framework for online covering problems with multiple predictions that obtains an online solution that is competitive against the performance of the best predictor. Our algorithm incorporates the use of predictions in the classic potential-based analysis of online algorithms. We apply our algorithmic framework to solve classical problems such as online set cover, (weighted) caching, and online facility location in the multiple predictions setting. Our algorithm can also be robustified, i.e., the algorithm can be simultaneously made competitive against the best prediction and the performance of the best online algorithm (without prediction).

3.11 Clustering with Advice

Silvio Lattanzi (Google – Barcelona, ES)

License © Creative Commons BY 4.0 International license
© Silvio Lattanzi

In this talk, starting from practical questions, we motivate different models of machine learning advice and present new algorithms that leverage the additional information to obtain stronger guarantees. In particular, we start by describing the semi-supervised active clustering framework and how one can recover convex and non-convex clusters in this setting. Then we describe how partial knowledge about input instances can be leveraged to obtain better guarantees in online correlation clustering and the classic k-means problem.

3.12 Algorithms with Prediction Portfolios

Thomas Lavastida (University of Texas – Dallas, US)

License © Creative Commons BY 4.0 International license
© Thomas Lavastida

Joint work of Michael Dinitz, Sungjin Im, Thomas Lavastida, Benjamin Moseley, Sergei Vassilvitskii

Main reference Michael Dinitz, Sungjin Im, Thomas Lavastida, Benjamin Moseley, Sergei Vassilvitskii: “Algorithms with Prediction Portfolios”, in Proc. of the Advances in Neural Information Processing Systems, 2022.

URL https://openreview.net/forum?id=wwWCZ7sER_C

The research area of algorithms with predictions has seen recent success showing how to incorporate machine learning into algorithm design to improve performance when the predictions are correct, while retaining worst-case guarantees when they are not. Most previous work has assumed that the algorithm has access to a single predictor. However, in practice, there are many machine learning methods available, often with incomparable generalization guarantees, making it hard to pick a best method a priori. In this work we consider scenarios where multiple predictors are available to the algorithm and the question is how to best utilize them.

Ideally, we would like the algorithm’s performance to depend on the quality of the best predictor. However, utilizing more predictions comes with a cost, since we now have to identify which prediction is the best. We study the use of multiple predictors for a number of fundamental problems, including matching, load balancing, and non-clairvoyant scheduling, which have been well-studied in the single predictor setting. For each of these problems we introduce new algorithms that take advantage of multiple predictors, and prove bounds on the resulting performance.

3.13 Predictions for Online Min-Sum Scheduling Problems

Alex Lindermayr (Universität Bremen, DE)

License  Creative Commons BY 4.0 International license
 Alex Lindermayr

- Joint work of** Alexandra Lassota, Nicole Megow, Martin Rapp, Jens Schlöter
- Main reference** Alexandra Lassota, Alexander Lindermayr, Nicole Megow, Jens Schlöter: “Minimalistic Predictions to Schedule Jobs with Online Precedence Constraints”, CoRR, Vol. abs/2301.12863, 2023.
- URL** <https://doi.org/10.48550/arXiv.2301.12863>
- Main reference** Alexander Lindermayr, Nicole Megow: “Permutation Predictions for Non-Clairvoyant Scheduling”, in Proc. of the SPAA '22: 34th ACM Symposium on Parallelism in Algorithms and Architectures, Philadelphia, PA, USA, July 11 – 14, 2022, pp. 357–368, ACM, 2022.
- URL** <https://doi.org/10.1145/3490148.3538579>
- Main reference** Alexander Lindermayr, Nicole Megow, Martin Rapp: “Speed-Oblivious Online Scheduling: Knowing (Precise) Speeds is not Necessary”, CoRR, Vol. abs/2302.00985, 2023.
- URL** <https://doi.org/10.48550/arXiv.2302.00985>

We consider the problem of scheduling jobs to minimize the total weighted completion time, and present models for various uncertainties. For uncertain processing requirements, also called non-clairvoyant scheduling, we present and recall predictions and results to improve the lower bound of 2 on the competitive ratio on a single machine. For online precedence constraints, we present bounds for different prediction models which try to overcome a simple $\Omega(n)$ lower bound in the worst-case setting. For uncertain processing speeds of unrelated machines, lower bounds rule out competitive ratios better than $\Omega(m)$ for m machines. Here we present two models and algorithms, which overcome this lower bound: speed predictions give predictions on the uncertain speeds, and speed-ordered machines ensure a single machine order which all job-dependent speeds respect, i.e., for any job j , the speed of j on machine i no less than on machine i' if and only if i comes before i' in the order.

3.14 DAG Scheduling Problems

Alberto Marchetti-Spaccamela (Sapienza University of Rome, IT)

License  Creative Commons BY 4.0 International license
 Alberto Marchetti-Spaccamela

- Joint work of** Sanjoy Baruah, Vincenzo Bonifaci, Alberto Marchetti-Spaccamela, Nicole Megow, Jens Schlöter, Martin Skutella, Sebastian Stiller, Leen Stougie, Andreas Wiese

The Directed Acyclic Graph (DAGs) is a popular representation to describe the structure of parallel applications and to model the execution of multi-threaded programs that is widely used in cloud computing and in real-time systems. In this talk I review recent results on DAG scheduling considering different models and focusing on complexity and approximation.

3.15 Stable matching in practice

Claire Mathieu (CNRS – Paris, FR)

License  Creative Commons BY 4.0 International license
 Claire Mathieu

Stable matching methods, based on the algorithm designed by Gale and Shapley, are used around the world in many applications. the college assigned to the applicant in their preference list; robustness; running time; etc.

After a brief theoretical review, we present issues arising in practice, (1) in the context of college admissions in France since 2018, and (2) in the context of upcoming medical studies specialization in France starting in 2024.

3.16 Online Rounding of Bipartite Matchings

Seffi Naor (Technion – Haifa, IL)

License © Creative Commons BY 4.0 International license
© Seffi Naor

Joint work of Niv Buchbinder, Seffi Naor, Aravind Srinivasan, David Wajc

Main reference Niv Buchbinder, Joseph Naor, David Wajc: “A Randomness Threshold for Online Bipartite Matching, via Lossless Online Rounding”, CoRR, Vol. abs/2106.04863, 2021.

URL <https://doi.org/10.48550/arXiv.2106.04863>

Two complementary facets of the online bipartite matching problem are discussed. (1) For numerous online bipartite matching problems, such as edge-weighted matching and matching under two-sided vertex arrivals, state-of-the-art fractional algorithms outperform their randomized integral counterparts. Thus, a natural question is whether we can achieve lossless online rounding of fractional solutions in this setting. Even though lossless online rounding is impossible in general, randomized algorithms do induce fractional algorithms of the same competitive ratio, which by definition are losslessly roundable online. This motivates the addition of constraints that decrease the “online integrality gap”, thus allowing for lossless online rounding. We characterize a set of non-convex constraints which allow for such lossless online rounding and allow for better competitive ratios than yielded by deterministic algorithms. (2) In a different vein, we study the problem of rounding fractional bipartite matchings in online settings. We assume that a fractional solution is already generated for us online by a black box (via a fractional algorithm, or some machine-learned advice) and provided as part of the input, which we then wish to round. We provide improved bounds on the rounding ratio and discuss several applications.

3.17 Learning-augmented Assignment: Santa Claus does Load Balancing

Debmalya Panigrahi (Duke University – Durham, US)

License © Creative Commons BY 4.0 International license
© Debmalya Panigrahi

Assignment problems are among the most well-studied in online algorithms. In these problems, a sequence of items arriving online must be assigned among a set of agents so as to optimize a given objective. This encompasses scheduling problems for minimizing makespan, p-norms, and other objectives, as well as fair division problems such as the Santa Claus problem and Nash welfare maximization. One common feature is that many of these problems are characterized by strong worst-case lower bounds in the online setting. To circumvent these impossibility results, recent research has focused on using additional (learned) information about the problem instance and this has led to dramatic improvements in the competitive ratio over the worst case. In this talk, I will first survey some of this literature (Lattanzi et al., SODA 20; Li and Xian, ICML 21; Banerjee et al., SODA 22; Barman et al., AAAI 22) that addresses specific problems in this domain. I will then proceed to describe recent work

with Ilan Cohen that brings these problems under one umbrella: we give a single algorithmic framework for learning-augmented online assignment for a large class of maximization and minimization objectives.

3.18 Evaluating Stochastic Score Functions

Kevin Schewior (University of Southern Denmark – Odense, DK)

License  Creative Commons BY 4.0 International license
© Kevin Schewior

Joint work of Kevin Schewior, Benedikt M. Plank

Main reference Benedikt M. Plank, Kevin Schewior: “Simple Algorithms for Stochastic Score Classification with Small Approximation Ratios”, CoRR, Vol. abs/2211.14082, 2022.

URL <https://doi.org/10.48550/arXiv.2211.14082>

We revisit the Stochastic Score Classification (SSC) problem introduced by Gkenosis et al. (ESA 2018): There are n tests. Each test j can be conducted at cost c_j , and it succeeds independently with probability p_j . Further, a partition of the (integer) interval $\{0, \dots, n\}$ into a number of smaller intervals is known. The goal is to conduct tests so as to determine that interval from the partition in which the number of successful tests lies while minimizing the expected cost. Ghuge et al. (IPCO 2022) independently showed that a polynomial-time constant-factor approximation algorithm exists. We present a simple polynomial-time $(3 + 2 * \sqrt{2})$ -approximation algorithm and highlight (potential) connections to scheduling problems. This is joint work with Benedikt Plank.

3.19 Learning-Augmented Query Policies for Minimum Spanning Tree with Uncertainty

Jens Schlöter (Universität Bremen, DE)

License  Creative Commons BY 4.0 International license
© Jens Schlöter

Joint work of Thomas Erlebach, Murilo Santos de Lima, Nicole Megow, Jens Schlöter

Main reference Thomas Erlebach, Murilo Santos de Lima, Nicole Megow, Jens Schlöter: “Learning-Augmented Query Policies for Minimum Spanning Tree with Uncertainty”, in Proc. of the 30th Annual European Symposium on Algorithms, ESA 2022, September 5-9, 2022, Berlin/Potsdam, Germany, LIPIcs, Vol. 244, pp. 49:1–49:18, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.

URL <https://doi.org/10.4230/LIPIcs.ESA.2022.49>

We study how to utilize (possibly erroneous) predictions in a model for computing under uncertainty in which an algorithm can query unknown data. Our aim is to minimize the number of queries needed to solve the minimum spanning tree problem, a fundamental combinatorial optimization problem that has been central also to the research area of explorable uncertainty. For all integral $\gamma \geq 2$, we present algorithms that are γ -robust and $(1 + \frac{1}{\gamma})$ -consistent, meaning that they use at most γOPT queries if the predictions are arbitrarily wrong and at most $(1 + \frac{1}{\gamma})OPT$ queries if the predictions are correct, where OPT is the optimal number of queries for the given instance. Moreover, we show that this trade-off is best possible. Furthermore, we argue that a suitably defined hop distance is a useful measure for the amount of prediction error and design algorithms with performance guarantees that degrade smoothly with the hop distance. Our results demonstrate that untrusted predictions can circumvent the known lower bound of 2, without any degradation of the worst-case ratio. To obtain our results, we provide new structural insights for the

minimum spanning tree problem that might be useful in the context of query-based algorithms regardless of predictions. In particular, we generalize the concept of witness sets – the key to lower-bounding the optimum – by proposing novel global witness set structures and completely new ways of adaptively using those.

3.20 New bounds for approximation algorithms for graph burning

Jiri Sgall (Charles University – Prague, CZ)

License © Creative Commons BY 4.0 International license
© Jiri Sgall

Joint work of Jiri Sgall, Matej Lieskovsky, Andreas E. Feldmann

Graph burning is the following process: We start with a graph with no node burned. At time $t=1$ we choose a node and burn it. At each time step $t > 1$, all neighbours of already burned nodes are also burned, and we choose one arbitrary additional node to burn. The process stops when all nodes of the graph have been burned. The burning number is the minimal number of steps needed for all nodes of a graph to be burned.

We sketch a randomized 2.32-approximation algorithm and a lower bound of $4/3$ for burning number of arbitrary graphs. This improves the previous upper bound of 3, for lower bound only NP-hardness was known.

3.21 Hiking through the complexity landscape with Gerhard – a personal reminiscence

Martin Skutella (TU Berlin, DE)

License © Creative Commons BY 4.0 International license
© Martin Skutella

This talk is dedicated to the memory of my friend and esteemed colleague Gerhard Woeginger, in thankful admiration for his outstanding scientific contributions, his fine taste of problems, his unfailingly inspiring lectures, his valuable advice over the course of 25 years, and, last but not least, for his great sense of humor.

3.22 Stochastic Minimum-Norm Combinatorial Optimization

Chaitanya Swamy (University of Waterloo, CA)

License © Creative Commons BY 4.0 International license
© Chaitanya Swamy

Joint work of Sharat Irahimpur, Chaitanya Swamy

Main reference Sharat Irahimpur, Chaitanya Swamy: “Approximation Algorithms for Stochastic Minimum-Norm Combinatorial Optimization”, in Proc. of the 61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020, pp. 966–977, IEEE, 2020.

URL <https://doi.org/10.1109/FOCS46700.2020.00094>

We develop a framework for designing approximation algorithms for a wide class of (1-stage) stochastic-optimization problems with norm-based objective functions. We introduce the model of stochastic minimum-norm combinatorial optimization, wherein the costs involved are random variables with given distributions, and we are given a monotone, symmetric

norm f . Each feasible solution induces a random multidimensional cost vector whose entries are independent random variables, and the goal is to find a solution that minimizes the expected f -norm of the induced cost vector. This is a very rich class of objectives, containing all ℓ_p norms, as also Top- l norms (sum of l largest coordinates in absolute value), which enjoys various closure properties.

Our chief contribution is a framework for designing approximation algorithms for stochastic minimum-norm optimization, which has two key components: (i) A reduction showing that one can control the expected f -norm by simultaneously controlling a (small) collection of expected Top- l norms; and (ii) Showing how to tackle the minimization of a single expected Top- l -norm by leveraging techniques used to deal with minimizing the expected maximum, circumventing the difficulties posed by the non-separable nature of Top- l norms.

We apply our framework to obtain strong approximation guarantees for two concrete problem settings: (1) stochastic load balancing, wherein jobs have random processing times and the induced cost vector is the machine-load vector; and (2) stochastic spanning tree, where edges have stochastic weights and the cost vector consists of the edge-weight variables of edges in the spanning tree returned.

This is joint work with Sharat Irahimpur.

3.23 Fast Graph Algorithms with Learned Duals

Ali Vakilian (TTIC – Chicago, US)

License © Creative Commons BY 4.0 International license
© Ali Vakilian

Joint work of Justin Y. Chen, Sandeep Silwal, Ali Vakilian, Fred Zhang

Main reference Justin Y. Chen, Sandeep Silwal, Ali Vakilian, Fred Zhang: “Faster Fundamental Graph Algorithms via Learned Predictions”, CoRR, Vol. abs/2204.12055, 2022.

URL <https://doi.org/10.48550/arXiv.2204.12055>


We consider the question of speeding up classic graph algorithms with machine-learned predictions. In this model, algorithms are furnished with extra advice learned from past or similar instances. Given the additional information, we aim to improve upon the traditional worst-case run-time guarantees. Our contributions are the following:

- (i) We give a faster algorithm for minimum-weight bipartite matching via learned duals, improving the recent result by Dinitz, Im, Lavastida, Moseley and Vassilvitskii (NeurIPS, 2021);
- (ii) We extend the learned dual approach to the single-source shortest path problem (with negative edge lengths), achieving an almost linear runtime given sufficiently accurate predictions which improves upon the classic fastest algorithm due to Goldberg (SIAM J. Comput., 1995);
- (iii) We provide a general reduction-based framework for learning-based graph algorithms, leading to new algorithms for degree-constrained subgraph and minimum-cost 0-1 flow, based on reductions to bipartite matching and the shortest path problem.

Finally, we give a set of general learnability theorems, showing that the predictions required by our algorithms can be efficiently learned in a PAC fashion.

3.24 Five(ish) Years of Algorithms with Predictions

Sergei Vassilvitskii (Google – New York, US)

License  Creative Commons BY 4.0 International license
© Sergei Vassilvitskii

Worst-case analysis has proven invaluable for understanding aspects of both the complexity and practicality of algorithms. In some cases, however, we do not face worst-case scenarios, and the question arises of how we can tune our algorithms to work even better on the kinds of instances we are likely to see, while keeping a rigorous formal framework similar to what we have developed through worst-case analysis.

We give an overview of a recent trend that develops algorithms parameterized by additional parameters which capture “the kinds of instances we are likely to see,” and obtains a finer grained analysis of algorithms’ performance. We will give examples of re-analyzing classical algorithms through this lens, as well as developing new algorithms that expose new structural insights about the problems.

3.25 Search infected nodes in uncertain graphs

Jose Verschae (PUC – Santiago de Chile, CL)

License  Creative Commons BY 4.0 International license
© Jose Verschae

Joint work of Jose Verschae, I. S. Beaudry, José Baboun, M. Castro, A. Jara, Benjamín Rubio

In the context of COVID-19, methods for traceability and search of active cases played a vital role. One such method uses PCR samples in the sewage network, aiming to locate the appearance of new infections quickly.

Given a representation of the sewer network as an acyclic directed graph, we seek to design a search strategy that finds new infected nodes while minimizing the number of samples needed in the worst case. This problem has been extensively studied in the context of perfect information, that is, if the graph is known. However, in reality, we find much information missing from the network. In this talk, we will propose a model to address the problem when the network is only partially known and present some partial results. In particular, we give tight bounds on the number of samples needed when the network is a planar graph of bounded degree.

3.26 Bayesian Scheduling: Analysis of simple policies

Tjark Vredeveld (Maastricht Univ. School of Business & Economics, NL)

License  Creative Commons BY 4.0 International license
© Tjark Vredeveld


Joint work of Tjark Vredeveld, Sebastian Marban, Cyriel Rutten

Bayesian scheduling is an extension of stochastic scheduling in which there is uncertainty about the system parameters. By processing jobs, we can learn about the true values of these parameters. We consider the basic scheduling problem of non-preemptively scheduling jobs on a single machine so as to minimize the expected total completion time. We consider a setting in which there are m classes of (a finite number of) jobs and the processing

times of jobs of the same class are independent and identically exponentially distributed with an unknown parameter. The initial beliefs on this parameter are modelled as a prior distribution and after processing a job of the class, the posterior distribution models the updated beliefs on the parameter. For this setting, optimal policies based on Gittins-indices exist (see, e.g., Hamada and Glazebrook, 1993). However, computing these indices may be computationally challenging. In this talk, we consider two simple policies, based on SEPT (Shortest Expected Processing Time), the optimal policy for the variant in which the parameters of the exponential distribution are known. We consider one version of SEPT, where the expected processing time are based on the initial belief and one version that updates the expected value each time a job has been processed. We can show that the first version of SEPT is at most a factor of m away from the optimal policy and that this is tight. Moreover, for the second variant we conjecture that the bound is also tight.

3.27 Minimizing Completion Times for Stochastic Jobs via Batched Free Times

Rudy Zhou (Carnegie Mellon University – Pittsburgh, US)

License  Creative Commons BY 4.0 International license
© Rudy Zhou

Joint work of Anupam Gupta, Benjamin Moseley, Rudy Zhou

Main reference Anupam Gupta, Benjamin Moseley, Rudy Zhou: “Minimizing Completion Times for Stochastic Jobs via Batched Free Times”, CoRR, Vol. abs/2208.13696, 2022.

URL <https://doi.org/10.48550/arXiv.2208.13696>

In this talk, we consider the classic problem of minimizing the expected total completion time of jobs on m identical machines in the setting where the sizes of the jobs are stochastic. Specifically, the size of each job is a random variable whose distribution is known to the algorithm, but whose realization is revealed only after the job is scheduled. We give a $O(m^{1/2}poly(\log n))$ -approximation for stochastic jobs which have Bernoulli processing times. This is the first approximation for this problem that is both independent of the variance in the job sizes, and is sublinear in the number of machines m . Our algorithm is based on a novel reduction from minimizing the total completion time to a natural makespan-like objective, which we call the weighted free time. We hope this free time objective will be useful in further improvements to this problem, as well as other stochastic scheduling problems.

Participants

- Antonios Antoniadis
University of Twente –
Enschede, NL
- Yossi Azar
Tel Aviv University, IL
- Etienne Bamas
EPFL – Lausanne, CH
- Siddhartha Banerjee
Cornell University – Ithaca, US
- Sanjoy Baruah
Washington University –
St. Louis, US
- Sami Davies
Northwestern University –
Evanston, US
- Christoph Dürr
Sorbonne University– Paris, FR
- Franziska Eberle
London School of Economics, GB
- Daniel Freund
MIT – Cambridge, US
- Naveen Garg
Indian Institute of Technology –
New Delhi, IN
- Vineet Goyal
Columbia University –
New York, US
- Thomas Kesselheim
Universität Bonn, DE
- Samir Khuller
Northwestern University –
Evanston, US
- Amit Kumar
Indian Institute of Technology –
New Delhi, IN
- Silvio Lattanzi
Google – Barcelona, ES
- Thomas Lavastida
University of Texas – Dallas, US
- Alexander Lindermayr
Universität Bremen, DE
- Alberto Marchetti-Spaccamela
Sapienza University of Rome, IT
- Claire Mathieu
CNRS – Paris, FR
- Nicole Megow
Universität Bremen, DE
- Benjamin J. Moseley
Carnegie Mellon University –
Pittsburgh, US
- Seffi Naor
Technion – Haifa, IL
- Debmalya Panigrahi
Duke University – Durham, US
- Kirk Pruhs
University of Pittsburgh, US
- Lars Rohwedder
Maastricht University, NL
- Thomas Rothvoss
University of Washington –
Seattle, US
- Kevin Schewior
University of Southern Denmark –
Odense, DK
- Jens Schlöter
Universität Bremen, DE
- Jiri Sgall
Charles University – Prague, CZ
- David Shmoys
Cornell University – Ithaca, US
- Martin Skutella
TU Berlin, DE
- Frits C. R. Spieksma
TU Eindhoven, NL
- Clifford Stein
Columbia University –
New York, US
- Leen Stougie
CWI – Amsterdam, NL
- Ola Svensson
EPFL – Lausanne, CH
- Chaitanya Swamy
University of Waterloo, CA
- Marc Uetz
University of Twente –
Enschede, NL
- Ali Vakilian
TTIC – Chicago, US
- Sergei Vassilvitskii
Google – New York, US
- Jose Verschae
PUC – Santiago de Chile, CL
- Tjark Vredeveld
Maastricht Univ. School of
Business & Economics, NL
- Andreas Wiese
TU München, DE
- Rudy Zhou
Carnegie Mellon University –
Pittsburgh, US



Programming Language Processing

Michael Pradel^{*1}, Baishakhi Ray^{*2}, Charles Sutton^{*3}, and Eran Yahav^{*4}

- 1 Universität Stuttgart, DE. michael@binaervervarianz.de
- 2 Columbia University – New York, US. rayb@cs.columbia.edu
- 3 Google – Mountain View, US. charlessutton@google.com
- 4 Technion – Haifa, IL. yahave@gmail.com

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23062 “Programming Language Processing” (PLP). The seminar brought together researchers and practitioners from three communities—software engineering, programming languages, and natural language processing—providing a unique opportunity for cross-fertilization and inter-disciplinary progress. We discussed machine learning models of code, integrating learning-based and traditional program analysis, and learning from natural language information associated with software. The seminar led to a better understanding of the commonalities and differences between natural and programming languages, and an understanding of the challenges and opportunities in industry adoption of PLP.

Seminar February 5–10, 2023 – <https://www.dagstuhl.de/23062>

2012 ACM Subject Classification Computing methodologies → Artificial intelligence; Computing methodologies → Machine learning; Software and its engineering

Keywords and phrases ML4PL, ML4SE, Neural Software Analysis

Digital Object Identifier 10.4230/DagRep.13.2.20


1 Executive Summary

Michael Pradel

Baishakhi Ray

Charles Sutton

Eran Yahav

License  Creative Commons BY 4.0 International license
© Michael Pradel, Baishakhi Ray, Charles Sutton, and Eran Yahav

Our 5-day Dagstuhl Seminar on “Programming Language Processing” (PLP) brought together researchers and practitioners from the software engineering, programming languages, and natural language processing communities. The seminar focused on activities prepared ahead by the participants, such as talks, demos of tools and challenges, and tutorials, as well as informal discussions anchored around the prepared activities. We provided each participant who wanted to present their work an opportunity for doing so. In addition, we asked specific people to present specific topics, e.g., experts of a particularly relevant subfield to prepare a tutorial or creators of a particularly relevant tool to give a tool demo.

In addition to talks and informal discussions, there were several break-out sessions during which participants discussed specific topics in smaller groups and eventually reported back to the other participants. In particular, we had break-out sessions on the following topics:

* Editor / Organizer



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Programming Language Processing, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 20–32

Editors: Michael Pradel, Baishakhi Ray, Charles Sutton, and Eran Yahav



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

- How (if at all) do AI programming assistants change programming?
- Interpreting neural models of code.
- Do we still need per-task models, or do large language models solve it all?
- What software engineering tasks are not yet explored (sufficiently) by neural models?
- How should and will computer science education change in response to ML-based coding tools?
- What kinds of guarantees can we expect, and do we want, from ML-based developer tools? What human factors in interacting with ML systems are relevant?
- How can learned models use existing tools, e.g., compilers and interpreters, to improve their predictions?

As a result of the seminar, several participants plan to launch various follow-up activities, such as joint publications and transferring promising ideas from academia to industry.

2 Table of Contents

Executive Summary

Michael Pradel, Baishakhi Ray, Charles Sutton, and Eran Yahav 20

Overview of Talks

Mining Idioms

Rui Abreu 23

Crafting Code Suggestions Using Large Language Models

Edward E. Aftandilian and Albert Ziegler 23

Counterfactual Explanations for Models of Code

Jürgen Cito 24

Synthesizing Correctness Properties

Elizabeth Dinella 24

DiCoder: Code Completion via Dialog

Aryaz Eghbali 25

Automated Repair for Static Warnings: PLMs (Codex) vs Template-based Meta-programming

Khashayar Etemadi Someoliayi 25

Controlling Large Language Models to Generate Secure and Vulnerable Code

Jingruan He 26

Explaining Code Intelligence to Bridge the Gap Between Testing and Debugging

Reyhaneh Jabbarvand 26

LExecutor: Learning-Guided Execution

Michael Pradel 27

How to incorporate semantics in LLM pretraining

Baishakhi Ray 27

On the role of data in Neural Bug Detection and Repair

Cedric Richter 28

Two Benchmarks for ML4Code: GLUE Code and RunBugRun

Romain Robbes 29

Combining compiler IRs with machine learning

Baptiste Rozière 29

Large language models and program invariants

Charles Sutton 30

Customized Models or Generic Code Language Models?

Lin Tan 30

An Empirical Study of Deep Learning Models for Vulnerability Detection

Wei Le 31

Participants 32

3 Overview of Talks

3.1 Mining Idioms

Rui Abreu (Meta Platforms – Bellevue, US)

License © Creative Commons BY 4.0 International license
© Rui Abreu

Joint work of Aishwarya Sivaraman, Rui Abreu, Andrew Scott, Tobi Akomolede, Satish Chandra

Main reference Aishwarya Sivaraman, Rui Abreu, Andrew Scott, Tobi Akomolede, Satish Chandra: “Mining Idioms in the Wild”, in Proc. of the 44th IEEE/ACM International Conference on Software Engineering: Software Engineering in Practice, ICSE (SEIP) 2022, Pittsburgh, PA, USA, May 22-24, 2022, pp. 187–196, IEEE, 2022.

URL <https://doi.org/10.1109/ICSE-SEIP55303.2022.9794062>

Existing code repositories contain numerous instances of code patterns that are idiomatic ways of accomplishing a particular programming task. Sometimes, the programming language in use supports specific operators or APIs that can express the same idiomatic imperative code much more succinctly. However, those code patterns linger in repositories because the developers may be unaware of the new APIs or have not gotten around to them. Detection of idiomatic code can also point to the need for new APIs.

We share our experiences in mine idiomatic patterns from the Hack repo at Facebook. We found that existing techniques either cannot identify meaningful patterns from syntax trees or require test-suite-based dynamic analysis to incorporate semantic properties to mine useful patterns. The key insight of the approach proposed in this talk – Jezero – is those semantic idioms from a large codebase can be learned from canonicalized dataflow trees. We propose a scalable, lightweight static analysis-based approach to construct a tree that is well suited to mine semantic idioms using nonparametric Bayesian methods.

Our experiments with Jezero on Hack code show a clear advantage of adding canonicalized dataflow information to ASTs: Jezero was significantly more effective than a baseline that did not have the dataflow augmentation in being able to effectively find refactoring opportunities from unannotated legacy code.

3.2 Crafting Code Suggestions Using Large Language Models

Edward E. Aftandilian (GitHub – San Francisco, US) and Albert Ziegler (GitHub – San Francisco, US)

License © Creative Commons BY 4.0 International license
© Edward E. Aftandilian and Albert Ziegler

Large Language Models are great at putting out code, but if you want the right code, you have to ask the right question. Concretely, you have to turn your application task (e.g., I want a test for this function) into a completion task (i.e., a string whose most likely completion will give you such a test). Building it is the object of the emerging discipline of promptcrafting.

We state some general principles for promptcrafting, distinguishing between communicating the task (write a test now) and the background information for this task (the function, its dependencies, the test style in the project, etc.). We discuss typical constraints to promptcrafting (context window size, latency, generality), and desiderata of the produced prompt, in particular the “Little Red Rule” of normality.

Then we go into detail of how we construct prompts in GitHub Copilot by breaking up the different types of background information into discrete “wishes” organized in a wish

list that can be quickly optimized over. We describe the strategies we use to construct our “wishes”, and how these prompt crafting strategies improve the performance of the underlying OpenAI Codex model for our application task of generating code editing suggestions.

3.3 Counterfactual Explanations for Models of Code

Jürgen Cito (TU Wien, AT)

License © Creative Commons BY 4.0 International license
© Jürgen Cito

Joint work of Jürgen Cito, Isil Dillig, Vijayaraghavan Murali, Satish Chandra

Main reference Jürgen Cito, Isil Dillig, Vijayaraghavan Murali, Satish Chandra: “Counterfactual Explanations for Models of Code”, in Proc. of the 44th IEEE/ACM International Conference on Software Engineering: Software Engineering in Practice, ICSE (SEIP) 2022, Pittsburgh, PA, USA, May 22-24, 2022, pp. 125–134, IEEE, 2022.

URL <https://doi.org/10.1109/ICSE-SEIP55303.2022.9794112>

Machine learning (ML) models play an increasingly prevalent role in many software engineering tasks. However, because most models are now powered by opaque deep neural networks, it can be difficult for developers to understand why the model came to a certain conclusion and how to act upon the model’s prediction. Motivated by this problem, this paper explores counterfactual explanations for models of source code. Such counterfactual explanations constitute minimal changes to the source code under which the model “changes its mind”. We integrate counterfactual explanation generation into models of source code in a real-world setting at Meta. We describe considerations that impact both the ability to find realistic and plausible counterfactual explanations, as well as the usefulness of such explanations to the developers that use the model. In a series of experiments, we investigate the efficacy of our approach on three different models, each based on a BERT-like architecture operating over source code, including models to detect performance regressions, test plan quality, and taint propagation.

3.4 Synthesizing Correctness Properties

Elizabeth Dinella (University of Pennsylvania – Philadelphia, US)

License © Creative Commons BY 4.0 International license
© Elizabeth Dinella

Joint work of Elizabeth Dinella, Gabriel Ryan, Aaditya Naik, Todd Mytkowicz, Shuvendu Lahiri, Mayur Naik

Main reference Elizabeth Dinella, Gabriel Ryan, Todd Mytkowicz, Shuvendu K. Lahiri: “TOGA: A Neural Method for Test Oracle Generation”, in Proc. of the 44th IEEE/ACM 44th International Conference on Software Engineering, ICSE 2022, Pittsburgh, PA, USA, May 25-27, 2022, pp. 2130–2141, ACM, 2022.

URL <https://doi.org/10.1145/3510003.3510141>

Automatically checking for software correctness is a well studied and important problem in software engineering. Despite many successful works, there are still significant barriers to full automation. Namely, manually expressing a precise notion of correctness for which the tools can check for. A program analysis tool cannot check for a property which solely exists in the developer’s mind. In this talk I will present works to automatically synthesize such properties. I explore synthesis of both pre and post conditions and evaluate in the automated testing domain. I will present learning techniques for synthesis as well as datasets for future research in this area.

3.5 DiCoder: Code Completion via Dialog

Aryaz Eghbali (Universität Stuttgart, DE)

License © Creative Commons BY 4.0 International license
© Aryaz Eghbali

Joint work of Aryaz Eghbali, Michael Pradel

Large Language Models (LLMs) pretrained on code are assisting developers by suggesting code completions. Since the quality of the prediction can greatly be influenced by the prompt, and because of prompt size limitations, providing appropriate context for these models has been challenging. Current methods rely on heuristic approaches to select context from various locations in code, which fail in cases where there exist relevant information present in the same project, but in non-trivial locations. Although some predictions are not using the proper API or are using them in a wrong way, the output of LLMs are in many cases close enough to the desired prediction that can lead to correct completions if provided with hints about the APIs. We propose a novel approach, called DiCoder, for using the output of the model to guide the process of gathering relevant context from the same project. DiCoder iteratively prompts the model with the most relevant context that it has gathered until the most recent iteration, and then uses tokens in the completion to retrieve further relevant context.

3.6 Automated Repair for Static Warnings: PLMs (Codex) vs Template-based Metaprogramming

Khashayar Etemadi Someoliayi (KTH Royal Institute of Technology – Stockholm, SE)

License © Creative Commons BY 4.0 International license
© Khashayar Etemadi Someoliayi

Main reference Khashayar Etemadi Someoliayi, Nicolas Yves Maurice Harrant, Simon Larsén, Haris Adzemovic, Henry Luong Phu, Ashutosh Verma, Fernanda Madeiral, Douglas Wikstrom, Martin Monperrus: “Sorald: Automatic Patch Suggestions for SonarQube Static Analysis Violations”, IEEE Transactions on Dependable and Secure Computing, pp. 1–1, 2022.

URL <https://doi.org/10.1109/TDSC.2022.3167316>

Static analyzers create warnings regarding parts of code that can potentially cause bugs. As addressing the warnings can be very time-consuming, developers sometimes ignore these warnings. This causes significant quality degradation to the program in the long term. Automated program repair (APR) tools are proposed to deal with static warnings and modify the program to preserve their semantics and make them warning free. Existing APR techniques modify programs using fixed metaprogramming templates that address certain types of static warnings. In this work, we assess if using PLMs can replace traditional APR tools for static warnings. We create a prototype PLM-based APR tool that fixes SonarQube static warnings. Our prototype tool fixes warnings with 7 different types in a dataset of five real-world projects, while the state-of-the-art template-based tool fixes warnings of 4 types on the same dataset. This proves the flexibility and usefulness PLM-based APR tools in this area.

3.7 Controlling Large Language Models to Generate Secure and Vulnerable Code

Jingxuan He (ETH Zürich, CH)

License © Creative Commons BY 4.0 International license
© Jingxuan He

Joint work of Jingxuan He, Martin T. Vechev

Main reference Jingxuan He, Martin T. Vechev: “Controlling Large Language Models to Generate Secure and Vulnerable Code”, CoRR, Vol. abs/2302.05319, 2023.

URL <https://doi.org/10.48550/arXiv.2302.05319>

Large language models (LLMs) are increasingly pretrained on large codebases and used for code generation tasks. A fundamental limitation of LLMs is that they can generate secure and vulnerable code, but the users cannot control this. In this work, we formulate a new problem called controlled code generation, which allows users to input a boolean property into LLMs to control if the output code is secure or vulnerable. We propose svGen, which learns soft prompts for solving controlled code generation. Our extensive evaluation on a wide range of vulnerabilities shows the effectiveness of svGen.

3.8 Explaining Code Intelligence to Bridge the Gap Between Testing and Debugging

Reyhaneh Jabbarvand (University of Illinois – Urbana-Champaign, US)

License © Creative Commons BY 4.0 International license
© Reyhaneh Jabbarvand

Testing and Debugging are an inseparable part of the software engineering and development pipeline. While performed separately, they are connected in such a way that debugging relies on the result of test execution to localize and ultimately fix the bugs. Without accurate, automated, and explainable oracles, detecting and localizing the bugs could be cumbersome. We proposed SEER, an automated technique that leverages deep learning and attention analysis to produce accurate and explainable oracles. To build the ground truth, SEER jointly embeds unit tests and code into a unified vector space, in such a way that the neural representation of tests are similar to that of code they pass on them, but dissimilar to the code they fail on them. The classifier built on top of this vector representation serves as the oracle to generate “fail” labels, when test inputs detect a bug in code or “pass” labels, otherwise. The extensive experiments on applying SEER to more than 5K unit tests from a diverse set of open-source Java projects show that the produced oracle is (1) effective in predicting the fail or pass labels, achieving an overall accuracy, precision, recall, and F1 measure of 93%, 86%, 94%, and 90%, (2) generalizable, predicting the labels for the unit test of projects that were not in training or validation set with negligible performance drop, and (3) efficient, detecting the existence of bugs in only 6.5 milliseconds on average.

3.9 LExecutor: Learning-Guided Execution

Michael Pradel (Universität Stuttgart, DE)

License © Creative Commons BY 4.0 International license
© Michael Pradel

Joint work of Michael Pradel, Beatriz Souza

Main reference Beatriz Souza, Michael Pradel: “LExecutor: Learning-Guided Execution”, CoRR, Vol. abs/2302.02343, 2023.

URL <https://doi.org/10.48550/arXiv.2302.02343>

Executing code is essential for various program analysis tasks, e.g., to detect bugs that manifest through exceptions or to obtain execution traces for further dynamic analysis. However, executing an arbitrary piece of code is often difficult in practice, e.g., because of missing variable definitions, missing user inputs, and missing third-party dependencies. This talk presents LExecutor, a learning-guided approach for executing arbitrary code snippets in an underconstrained way. The key idea is to let a neural model predict missing values that otherwise would cause the program to get stuck, and to inject these values into the execution. For example, LExecutor injects likely values for otherwise undefined variables and likely return values of calls to otherwise missing functions. We evaluate the approach on Python code from popular open-source projects and on code snippets extracted from Stack Overflow. The neural model predicts realistic values with an accuracy between 80.1% and 94.2%, allowing LExecutor to closely mimic real executions. As a result, the approach successfully executes significantly more code than any available technique, such as simply executing the code as-is. For example, executing the open-source code snippets as-is covers only 4.1% of all lines, because the code crashes early on, whereas LExecutor achieves a coverage of 50.1%.

3.10 How to incorporate semantics in LLM pretraining

Baishakhi Ray (Columbia University – New York, US)

License © Creative Commons BY 4.0 International license
© Baishakhi Ray

Understanding code semantics is significant for code modeling tasks such as software vulnerability detection, code clone detection, security analysis and code generation. Here we talk about ways how model can learn code semantics while pertaining.

- We present a novel self-supervised model focusing on identifying (dis)similar functionalities of source code. Different from existing works, our approach does not require a huge amount of randomly collected datasets. Rather, we design structure-guided code transformation algorithms to generate synthetic code clones and inject real-world security bugs, augmenting the collected datasets in a targeted way. We propose to pre-train the Transformer model with such automatically generated program contrasts to better identify similar code in the wild and differentiate vulnerable programs from benign ones. To better capture the structural features of source code, we propose a new cloze objective to encode the local tree-based context (e.g., parents or sibling nodes). We pre-train our model with a much smaller dataset, the size of which is only 5% of the state-of-the-art models’ training datasets, to illustrate the effectiveness of our data augmentation and the pre-training approach. The evaluation shows that, even with much less data, DISCO can still outperform the state-of-the-art models in vulnerability and code clone detection tasks.

- For a generative setting, we leverage such syntax guided transformation in a de-noising encoder-decoder setting. We inject semantically similar transformation as “noise” and the decoder learns to denoise and retrieve the original code. Learning to generate equivalent, but more natural code, at scale, over large corpora of open-source code, without explicit manual supervision, helps the model learn to both ingest and generate code. We fine-tune our model in three generative Software Engineering tasks: code generation, code translation, and code refinement with limited human-curated labeled data and achieve state-of-the-art performance rivaling CodeT5. We show that our pre-trained model is especially competitive at zero-shot and few-shot learning, and better at learning code properties (e.g., syntax, data flow).
- Incorporating Dynamic Code Feature: We discuss how building machine learning (ML) models toward learning program execution semantics so they can remain robust against transformations in program syntax and generalize to various program analysis tasks and security applications. The corresponding research tools, such as Trex, StateFormer, and NeuDep, have outperformed commercial tools and prior arts by up to 117x in speed and by 35% in precision and have helped identify security vulnerabilities in real-world firmware that run on billions of devices.

3.11 On the role of data in Neural Bug Detection and Repair

Cedric Richter (Universität Oldenburg, DE)

License  Creative Commons BY 4.0 International license
© Cedric Richter

Joint work of Cedric Richter, Heike Wehrheim

Main reference Cedric Richter, Heike Wehrheim: “Can we learn from developer mistakes? Learning to localize and repair real bugs from real bug fixes”, CoRR, Vol. abs/2207.00301, 2022.

URL <https://doi.org/10.48550/arXiv.2207.00301>

Real bug fixes seem to be the perfect source for training neural bug detection and repair models. Yet, existing real bug fix datasets are often too small to effectively train data-hungry neural approaches. For this reason, existing approaches often rely on artificial bugs that can be easily produced via a mutation operator at large scales. However, neural bug detection and repair approaches trained purely on mutants usually underperform when confronted with real bugs.

To address this shortcoming, we introduce a novel dataset of 31k real bug fixes for training. We utilize the dataset to evaluate the impact of artificially generated mutants and real bug fixes on the training of neural bug detection and repair approaches. We find that training on real bug fixes can significantly improve the ability of our model to detect and repair real bugs, while training on mutants is still necessary to achieve high performing models.

3.12 Two Benchmarks for ML4Code: GLUE Code and RunBugRun

Romain Robbes (*CNRS – Bordeaux, FR & University of Bordeaux, FR*)

License © Creative Commons BY 4.0 International license
© Romain Robbes

Joint work of Romain Robbes, Anjan Karmakar, Julian Prenner, Miltiadis Allamanis

Main reference Anjan Karmakar, Miltiadis Allamanis, Romain Robbes: “JEMMA: An Extensible Java Dataset for ML4Code Applications”, CoRR, Vol. abs/2212.09132, 2022.

URL <https://doi.org/10.48550/arXiv.2212.09132>

We present two benchmarks and datasets that are designed to help the community progress on goals that we think are important.

- GLUE Code (Global and Local Understanding Evaluation of Code) is geared towards the development of models that use a global context beyond a code snippet. Indeed, one of our studies shows that 60% of method calls are project-specific, and 40% come from a distant context. GLUE Code is based on the JEMMA dataset of source code projects, a dataset of 50,000 projects (derived from 50K-C) that include significant post-processing to add source code representations, call graphs, and static analysis tool data. GLUE Code includes tasks that require a model to go beyond the current code snippet and include larger context (file, package, callers/callees). GLUE Code users can use JEMMA to assemble the context they need to solve the GLUE Code tasks. In this way, GLUE Code and JEMMA allow users to experiment with a variety of source code contexts. Find more details on JEMMA at: <https://arxiv.org/abs/2212.09132>.
- RunBugRun is a large-scale, executable, and multi-lingual dataset to incentivize Automated Program Repair models to leverage runtime information in their design. RunBugRun is derived from CodeNet; it includes 450,000 (carefully curated) executable bug/fix pairs that can be validated via running tests. Generated patches can be compiled and executed. RunBugRun includes bug/fix pairs in C, C++, Python, Java, JavaScript, Go, Ruby, and PHP. The bug/fix pairs are also labeled with respect to the kind of changes they include. Initial results on two baselines show both that there is room for future improvement, and the potential of transfer learning from common to uncommon languages. Find more details on RunBugRun at: https://github.com/giganticode/run_bug_run.

3.13 Combining compiler IRs with machine learning

Baptiste Rozière (*Meta AI – Paris, FR*)

License © Creative Commons BY 4.0 International license
© Baptiste Rozière

Joint work of Baptiste Rozière, Marc Szafraniec, Gabriel Synnaeve, Ruba Mutasim, David Pichardie, Patrick Labatut, Hugh Leather, François Charton

Main reference Marc Szafraniec, Baptiste Rozière, Hugh Leather, François Charton, Patrick Labatut, Gabriel Synnaeve: “Code Translation with Compiler Representations”, CoRR, Vol. abs/2207.03578, 2022.

URL <https://doi.org/10.48550/arXiv.2207.03578>

We leverage low-level compiler intermediate representations (IR) to improve code translation. Traditional transpilers rely on syntactic information and handcrafted rules, which limits their applicability and produces unnatural-looking code. Applying neural machine translation (NMT) approaches to code has successfully broadened the set of programs on which one can get a natural-looking translation. However, they treat the code as sequences of text tokens, and still do not differentiate well enough between similar pieces of code which have different semantics in different languages. The consequence is low quality translation, reducing the

practicality of NMT, and stressing the need for an approach significantly increasing its accuracy. Here we propose to augment code translation with IRs, specifically LLVM IR, with results on the C++, Java, Rust, and Go languages. Our method improves upon the state of the art for unsupervised code translation, increasing the number of correct translations by 11% on average, and up to 79% for the Java → Rust pair with greedy decoding. With beam search, it increases the number of correct translations by 5.5% in average.

Additionally, we train models with high performance on the problem of IR decompilation, generating programming source code from IR, and study using IRs as intermediary pivot for translation. We also show that IR decompilation can be used to simplify source code, or for code repair. As the retrieved source code can be compiled to IR again and compared to the input IR, we can sometimes guarantee the correctness of the output assuming that the compiler is correct.

3.14 Large language models and program invariants

Charles Sutton (Google – Mountain View, US)

License  Creative Commons BY 4.0 International license
© Charles Sutton

Joint work of Kexin Pei, David Bieber, Kensen Shi, Charles Sutton, Pengcheng Yin

Identifying invariants is an important program analysis task with applications towards program understanding, vulnerability analysis, and formal verification. Existing tools for identifying invariants rely on dynamic analysis, requiring traces collected from multiple executions in order to produce reliable invariants. We study the application of large language models to invariant prediction, finding that models trained on source code and fine-tuned for invariant generation can perform invariant prediction as static rather than dynamic analysis. Using a scratchpad approach where invariants are predicted sequentially through a program gives the best performance, finding invariants statically of quality comparable to those obtained by a dynamic analysis tool with access to five program traces.

3.15 Customized Models or Generic Code Language Models?

Lin Tan (Purdue University – West Lafayette, US)

License  Creative Commons BY 4.0 International license
© Lin Tan

Joint work of Nan Jiang, Thibaud Lutellier, Yiling Lou, Dan Goldwasser, Xiangyu Zhang, and Kevin Liu

Main reference Nan Jiang, Kevin Liu, Thibaud Lutellier, Lin Tan: “Impact of Code Language Models on Automated Program Repair”, CoRR, Vol. abs/2302.05020, 2023.

URL <https://doi.org/10.48550/arXiv.2302.05020>

Main reference Nan Jiang, Thibaud Lutellier, Yiling Lou, Lin Tan, Dan Goldwasser, Xiangyu Zhang: “KNOD: Domain Knowledge Distilled Tree Decoder for Automated Program Repair”, CoRR, Vol. abs/2302.01857, 2023.

URL <https://doi.org/10.48550/arXiv.2302.01857>

This talk presents a novel customized neural-network model KNOD for fixing software bugs automatically. KNOD contains (1) a novel three-stage tree decoder, which directly generates Abstract Syntax Trees of patched code according to the inherent tree structure, and (2) a novel domain-rule distillation, which leverages syntactic and semantic rules and teacher-student distributions to explicitly inject the domain knowledge into the decoding procedure during both the training and inference phases. KNOD outperforms existing program-repair techniques on three widely-used benchmarks.

It then discusses the pros and cons of building customized models such as KNOD for a task versus using and fine-tuning generic code language models for the same task. These discussions are based on two recent ICSE 2023 papers “KNOD: Domain Knowledge Distilled Tree Decoder for Automated Program Repair” and “Impact of Code Language Models on Automated Program Repair”. Some surprising results include that the best code language model as is, fixes 72% more bugs than the state-of-the-art deep-learning-based program repair techniques.

3.16 An Empirical Study of Deep Learning Models for Vulnerability Detection

Wei Le (Iowa State University – Ames, US)

License © Creative Commons BY 4.0 International license
© Wei Le

Joint work of Wei Le, Benjamin Steenhoek, Mahbubur Rahman, Richard Jiles

Main reference Benjamin Steenhoek, Md Mahbubur Rahman, Richard Jiles, Wei Le: “An Empirical Study of Deep Learning Models for Vulnerability Detection”, CoRR, Vol. abs/2212.08109, 2022.

URL <https://doi.org/10.48550/arXiv.2212.08109>

Deep learning (DL) models of code have recently reported great progress for vulnerability detection. In some cases, DL-based models have outperformed static analysis tools. Although many great models have been proposed, we do not yet have a good understanding of these models. This limits the further advancement of model robustness, debugging, and deployment for the vulnerability detection. In this paper, we surveyed and reproduced 9 state-of-the-art (SOTA) deep learning models on 2 widely used vulnerability detection datasets: Devign and MSR. We investigated 6 research questions in three areas, namely model capabilities, training data, and model interpretation. We experimentally demonstrated the variability between different runs of a model and the low agreement among different models’ outputs. We investigated models trained for specific types of vulnerabilities compared to a model that is trained on all the vulnerabilities at once. We explored the types of programs DL may consider “hard” to handle. We investigated the relations of training data sizes and training data composition with model performance. Finally, we studied model interpretations and analyzed important features that the models used to make predictions. We believe that our findings can help better understand model results, provide guidance on preparing training data, and improve the robustness of the models. All of our datasets, code, and results are available at <https://doi.org/10.6084/m9.figshare.20791240>.

Participants

- Rui Abreu
Meta Platforms – Bellevue, US
- Edward E. Aftandilian
GitHub – San Francisco, US
- Jürgen Cito
TU Wien, AT
- Premkumar T. Devanbu
University of California –
Davis, US
- Elizabeth Dinella
University of Pennsylvania –
Philadelphia, US
- Aryaz Eghbali
Universität Stuttgart, DE
- Khashayar Etemadi Someoliayi
KTH Royal Institute of
Technology – Stockholm, SE
- Yoav Goldberg
Bar-Ilan University –
Ramat Gan, IL
- Lars Grunske
HU Berlin, DE
- Jingxuan He
ETH Zürich, CH
- Maliheh Izadi
TU Delft, NL
- Reyhaneh Jabbarvand
University of Illinois –
Urbana-Champaign, US
- Wei Le
Iowa State University –
Ames, US
- Martin Monperrus
KTH Royal Institute of
Technology – Stockholm, SE
- Alex Polozov
Google – Mountain View, US
- Michael Pradel
Universität Stuttgart, DE
- Baishakhi Ray
Columbia University –
New York, US
- Cedric Richter
Universität Oldenburg, DE
- Romain Robbes
CNRS – Bordeaux, FR &
University of Bordeaux, FR
- Baptiste Rozière
Meta AI – Paris, FR
- Jan Arne Sparka
HU Berlin, DE
- Charles Sutton
Google – Mountain View, US
- Lin Tan
Purdue University – West
Lafayette, US
- Eran Yahav
Technion – Haifa, IL
- Albert Ziegler
GitHub – San Francisco, US



From Big Data Theory to Big Data Practice

Martin Farach-Colton^{*1}, Fabian Daniel Kuhn^{*2}, Ronitt Rubinfeld^{*3},
and Przemysław Uznański^{*4}

1 Rutgers University – Piscataway, US. martin@farach-colton.com

2 Universität Freiburg, DE. kuhn@cs.uni-freiburg.de

3 MIT – Cambridge, US. ronitt@csail.mit.edu

4 Pathway – Wrocław, PL. przemek@pathway.com

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23071 “From Big Data Theory to Big Data Practice”. Some recent advances in the theory of algorithms for big data – sublinear/local algorithms, streaming algorithms and external memory algorithms – have translated into impressive improvements in practice, whereas others have remained stubbornly resistant to useful implementations. This seminar aimed to glean lessons for those aspect of these algorithms that have led to practical implementation to see if the lessons learned can both improve the implementations of other theoretical ideas and to help guide the next generation of theoretical advances.

Seminar February 12–17, 2023 – <https://www.dagstuhl.de/23071>

2012 ACM Subject Classification Information systems → Data structures; Theory of computation → Data structures design and analysis; Theory of computation → Distributed algorithms; Theory of computation → Streaming, sublinear and near linear time algorithms

Keywords and phrases external memory, local algorithms, sublinear algorithms

Digital Object Identifier 10.4230/DagRep.13.2.33

1 Executive Summary

Martin Farach-Colton

Fabian Daniel Kuhn

Ronitt Rubinfeld

Przemysław Uznański

License  Creative Commons BY 4.0 International license

© Martin Farach-Colton, Fabian Daniel Kuhn, Ronitt Rubinfeld, and Przemysław Uznański

As data has grown faster than RAM, the theory of algorithms has expanded to provide approaches for tackling such problems. These fall into three broad categories:

- Streaming and semi-streaming algorithms
- Sublinear or local algorithms
- External memory algorithms

Each of these areas has a vibrant literature, and many of the results from the theory literature have made their way into practice. Other results are not suitable for implementation and deployment.

This Dagstuhl Seminar’s aim was to address several questions by bringing together algorithmicists from these subcommunities, as well as algorithms engineers. Specifically, the aim was to address the following questions:

* Editor / Organizer



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

From Big Data Theory to Big Data Practice, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 33–46

Editors: Martin Farach-Colton, Fabian Daniel Kuhn, Ronitt Rubinfeld, and Przemysław Uznański



DAGSTUHL
REPORTS

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

- What themes emerge from considering practical algorithms from the theory literature?
- Can we use these insights to create new models or to capture interesting new optimization criteria?

By bringing together researchers in these disparate areas and by including researchers in algorithms engineering, we hope to have brought to light these deep connections. The goals were to:

- Extract shared lessons to help guide theoretical research towards practical solutions;
- Create a feedback loop where commonalities of practical solutions can help guide future theoretical research;
- Help cross-pollinate these research areas.

Organization of the Seminar

The seminar brought together 38 researchers from theoretical computer science and systems research, both from academia and from industry. The participants consisted of both senior and junior researchers, including a number of postdocs and advanced PhD students.

During the four¹ days of the seminar, 20 talks of different lengths took place. Speakers were given no time constraints, but all the talks took between 30 to 60 minutes. We value the freedom given to the speakers, and the audience enjoyed the relaxed schedule as well. Each day afternoons were spent on fruitful discussions and open-ended open problems sessions.

Outcome

The seminar was seen as a success by organizers and participants. It brought together relevant communities, shared state-of-the-art research, and discussed challenges. The talks were of high quality and stimulating, encouraging participants to actively engage in working groups during the afternoon and evenings. It was particularly appreciated that a significant number of younger researchers (postdocs and PhD students) participated and integrated well. The organizers would like to express their gratitude to the Scientific Directorate and the administration of the Dagstuhl Center for their strong support during the seminar.

¹ The seminar had to be cut short by one day due to unforeseen logistic issues.

2 Table of Contents

Executive Summary

<i>Martin Farach-Colton, Fabian Daniel Kuhn, Ronitt Rubinfeld, and Przemysław Uznański</i>	33
--	----

Overview of Talks

Paging on Shared Caches <i>Kunal Agrawal</i>	37
Contention resolution without collision detection <i>Michael A. Bender</i>	37
Data Structures on the Ultra-Wide Word RAM <i>Philip Bille</i>	37
Mosaic Pages <i>Alexander Conway</i>	38
Sampling Edges in Sublinear Time <i>Talya Eden</i>	38
Tiny Pointers <i>Martin Farach-Colton</i>	39
Atomic Power in Forks: A Super-Logarithmic Lower Bound for Implementing Butterfly Networks in the Nonatomic Binary Fork-Join Model <i>Riko Jacob</i>	39
Online List Labeling: Breaking the $\log^2(n)$ Barrier <i>Hanna Koplós</i>	40
Linear Probing Revisited: The Demise of Clustering <i>William Kuszmaul</i>	40
Cardinality Estimation Using Gumbel Distribution <i>Aleksander Łukasiewicz</i>	41
Big data algorithms for distance computation <i>Yasamin Nazari</i>	41
Locality and Volume in the Context of Algorithms for Dynamic Data Sets <i>Krzysztof Nowicki</i>	42
Adaptive Adversaries vs. Randomized Algorithms: A Few Challenges <i>Krzysztof Onak</i>	42
Massively-Parallel Computing in a Heterogenous Regime <i>Rotem Oshman</i>	42
IcebergHT: High Performance PMEM Hash Tables Through Stability and Low Associativity Authors: Prashant Pandey <i>Prashant Pandey</i>	43
Can sublinear graph algorithms impact practice? <i>Comandur Seshadhri</i>	44
Locality in online, dynamic, sequential, and distributed graph algorithms <i>Jukka Suomela</i>	44

Dynamic Graph Sketching: To Infinity And Beyond <i>David Tench</i>	44
Reactive incremental algorithms with Pathway. <i>Przemysław Uznański</i>	45
Participants	46

3 Overview of Talks

3.1 Paging on Shared Caches

Kunal Agrawal (Washington University – St. Louis, US)

License © Creative Commons BY 4.0 International license
© Kunal Agrawal

The talk presents results on using shared caches where p disjoint request sequences are trying to access a cache concurrently.

3.2 Contention resolution without collision detection

Michael A. Bender (Stony Brook University, US)

License © Creative Commons BY 4.0 International license
© Michael A. Bender

Joint work of Michael A. Bender, Tsvi Kopelowitz, William Kuszmaul, Seth Pettie

Main reference Michael A. Bender, Tsvi Kopelowitz, William Kuszmaul, Seth Pettie: “Contention resolution without collision detection”, in Proc. of the Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020, pp. 105–118, ACM, 2020.

URL <https://doi.org/10.1145/3357713.3384305>

In this talk we revisit the classical problem of randomized exponential backoff on a multiple-access channel. We show how to achieve constant throughput even when the players cannot detect collisions, i.e, simultaneous transmissions, on the channel.

3.3 Data Structures on the Ultra-Wide Word RAM

Philip Bille (Technical University of Denmark – Lyngby, DK)

License © Creative Commons BY 4.0 International license
© Philip Bille

Joint work of Philip Bille, Inge Li Gørtz, Tord Stordalen

Main reference Philip Bille, Inge Li Gørtz, Tord Stordalen: “Predecessor on the Ultra-Wide Word RAM”, in Proc. of the 18th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2022, June 27-29, 2022, Tórshavn, Faroe Islands, LIPIcs, Vol. 227, pp. 18:1–18:15, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.

URL <https://doi.org/10.4230/LIPIcs.SWAT.2022.18>

We consider the predecessor problem on the ultra-wide word RAM model of computation, which extends the word RAM model with ultrawords consisting of w^2 bits [TAMC, 2015]. The model supports arithmetic and boolean operations on ultrawords, in addition to scattered memory operations that access or modify w (potentially non-contiguous) memory addresses simultaneously. The ultra-wide word RAM model captures (and idealizes) modern vector processor architectures. Our main result is a simple, linear space data structure that supports predecessor in constant time and updates in amortized, expected constant time. This improves the space of the previous constant time solution that uses space in the order of the size of the universe. Our result holds even in a weaker model where ultrawords consist of $w^{(1+\varepsilon)}$ bits for any $\varepsilon > 0$. It is based on a new implementation of the classic x-fast trie data structure of Willard [Inform. Process. Lett. 17(2), 1983] combined with a new dictionary data structure that supports fast parallel lookups.

3.4 Mosaic Pages

Alexander Conway (VMware Research – Palo Alto, US)

License © Creative Commons BY 4.0 International license
© Alexander Conway

Joint work of Krishnan Gosakan, Jaehyun Han, William Kuszmaul, Ibrahim N. Mubarek, Nirjhar Mukherjee, Karthik Sriram, Guido Tagliavini, Evan West, Michael A. Bender, Abhishek Bhattacharjee, Alex Conway, Martin Farach-Colton, Jayneel Gandhi, Rob Johnson, Sudarsun Kannan, Donald E. Porter

Main reference Krishnan Gosakan, Jaehyun Han, William Kuszmaul, Ibrahim N. Mubarek, Nirjhar Mukherjee, Karthik Sriram, Guido Tagliavini, Evan West, Michael A. Bender, Abhishek Bhattacharjee, Alex Conway, Martin Farach-Colton, Jayneel Gandhi, Rob Johnson, Sudarsun Kannan, Donald E. Porter: “Mosaic Pages: Big TLB Reach with Small Pages”, in Proc. of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 3, ASPLOS 2023, Vancouver, BC, Canada, March 25-29, 2023, pp. 433–448, ACM, 2023.

URL <https://doi.org/10.1145/3582016.3582021>

In this talk, I present an algorithmic approach to co-designing TLB hardware and the paging mechanism to increase TLB reach without the fragmentation issues incurred by huge pages. Along the way, I’ll introduce a new hash-table design that overcomes existing tradeoffs, and achieves better performance than state-of-the-art hash tables both in theory and in practice. Key to these results are “tiny pointers,” an algorithmic technique for compressing pointers.

3.5 Sampling Edges in Sublinear Time

Talya Eden (Bar-Ilan University – Ramat Gan, IL)

License © Creative Commons BY 4.0 International license
© Talya Eden

Joint work of Talya Eden, Shyam Narayanan, Jakub Tetek

Main reference Talya Eden, Shyam Narayanan, Jakub Tetek: “Sampling an Edge in Sublinear Time Exactly and Optimally”, in Proc. of the 2023 Symposium on Simplicity in Algorithms, SOSA 2023, Florence, Italy, January 23-25, 2023, pp. 253–260, SIAM, 2023.

URL <https://doi.org/10.1137/1.9781611977585.ch23>

Sampling edges from a graph in sublinear time is a fundamental problem and a powerful subroutine for designing sublinear-time algorithms. Suppose we have access to the vertices of the graph and know a constant-factor approximation to the number of edges. An algorithm for pointwise ε -approximate edge sampling with complexity $O(n/\sqrt{\varepsilon m})$ has been given by Eden and Rosenbaum [SOSA 2018]. This has been later improved by Tětek and Thorup [STOC 2022] to $O(n \log(\varepsilon^{-1})/\sqrt{m})$. At the same time, $\Omega(n/\sqrt{m})$ time is necessary. We close the problem, by giving an algorithm with complexity $O(n/\sqrt{m})$ for the task of sampling an edge exactly uniformly.

3.6 Tiny Pointers

Martin Farach-Colton (Rutgers University – Piscataway, US)

License © Creative Commons BY 4.0 International license
© Martin Farach-Colton

Joint work of Michael A. Bender, Alex Conway, Martin Farach-Colton, William Kuszmaul, Guido Tagliavini

Main reference Michael A. Bender, Alex Conway, Martin Farach-Colton, William Kuszmaul, Guido Tagliavini: “Tiny Pointers”, in Proc. of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023, pp. 477–508, SIAM, 2023.

URL <https://doi.org/10.1137/1.9781611977554.ch21>

We introduce a new data-structural object that we call the tiny pointer. In many applications, traditional $\log n$ -bit pointers can be replaced with $o(\log n)$ -bit tiny pointers at the cost of only a constant-factor time overhead and a small probability of failure. We develop a comprehensive theory of tiny pointers, and give optimal constructions for both fixed-size tiny pointers (i.e., settings in which all of the tiny pointers must be the same size) and variable-size tiny pointers (i.e., settings in which the average tiny-pointer size must be small, but some tiny pointers can be larger). If a tiny pointer references an element in an array filled to load factor $1 - \delta$, then the optimal tiny-pointer size is $\Theta(\log \log \log n + \log \delta^{-1})$ bits in the fixed-size case, and $\Theta(\log \delta^{-1})$ expected bits in the variable-size case. Using tiny pointers, it is possible to improve succinctness bounds on a variety of classic data-structure problems.

3.7 Atomic Power in Forks: A Super-Logarithmic Lower Bound for Implementing Butterfly Networks in the Nonatomic Binary Fork-Join Model

Riko Jacob (IT University of Copenhagen, DK)

License © Creative Commons BY 4.0 International license
© Riko Jacob

Joint work of Michael T. Goodrich, Riko Jacob, Nodari Sitchinava

Main reference Michael T. Goodrich, Riko Jacob, Nodari Sitchinava: “Atomic Power in Forks: A Super-Logarithmic Lower Bound for Implementing Butterfly Networks in the Nonatomic Binary Fork-Join Model”, in Proc. of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 – 13, 2021, pp. 2141–2153, SIAM, 2021.

URL <https://doi.org/10.1137/1.9781611976465.128>

We prove an $\Omega(\log n \log \log n)$ lower bound for the span of implementing the n input, $\log n$ -depth FFT circuit (also known as butterfly network) in the nonatomic binary fork-join model. In this model, memory-access synchronizations occur only through fork operations, which spawn two child threads, and join operations, which resume a parent thread when its child threads terminate. Our bound is asymptotically tight for the nonatomic binary fork-join model, which has been of interest of late, due to its conceptual elegance and ability to capture asynchrony. Our bound implies super-logarithmic lower bound in the nonatomic binary fork-join model for implementing the butterfly merging networks used, e.g., in Batcher’s bitonic and odd-even mergesort networks. This lower bound also implies an asymptotic separation result for the atomic and nonatomic versions of the fork-join model, since, as we point out, FFT circuits can be implemented in the atomic binary fork-join model with span equal to their circuit depth.

3.8 Online List Labeling: Breaking the $\log^2(n)$ Barrier

Hanna Komlós (Rutgers University – New Brunswick, US)

License  Creative Commons BY 4.0 International license
 Hanna Komlós

Joint work of Michael A. Bender, Alex Conway, Martin Farach-Colton, Hanna Komlós, William Kuszmaul, Nicole Wein

Main reference Michael A. Bender, Alex Conway, Martin Farach-Colton, Hanna Komlós, William Kuszmaul, Nicole Wein: “Online List Labeling: Breaking the $\log^2 n$ Barrier”, in Proc. of the 63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 – November 3, 2022, pp. 980–990, IEEE, 2022.

URL <https://doi.org/10.1109/FOCS54457.2022.00096>

The online list-labeling problem is a classical combinatorial problem with a large literature of upper bounds, lower bounds, and applications. The goal is to store a dynamically-changing set of n items in an array of m slots, while maintaining the invariant that the items appear in sorted order, and while minimizing the relabeling cost, defined to be the number of items that are moved per insertion/deletion. There has long existed a gap between the lower and upper bounds in the most algorithmically interesting part of the problem’s parameter space. We present our recent results, which narrow this gap for the first time in nearly 4 decades.

3.9 Linear Probing Revisited: The Demise of Clustering

William Kuszmaul (MIT – Cambridge, US)

License  Creative Commons BY 4.0 International license
 William Kuszmaul

Joint work of Michael A. Bender, Bradley C. Kuszmaul, William Kuszmaul

Main reference Michael A. Bender, Bradley C. Kuszmaul, William Kuszmaul: “Linear Probing Revisited: Tombstones Mark the Demise of Primary Clustering”, in Proc. of the 62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022, pp. 1171–1182, IEEE, 2021.

URL <https://doi.org/10.1109/FOCS52979.2021.00115>

The linear-probing hash table is one of the oldest and most widely used data structures in computer science. However, linear probing also famously comes with a major drawback: as soon as the hash table reaches a high memory utilization, elements within the hash table begin to cluster together, causing insertions to become slow. This clustering phenomenon, which was first discovered by Donald Knuth in 1962, increases the expected time per insertion to $\Theta(x^2)$ (rather than the more desirable $\Theta(x)$) in a hash table that is a $1 - 1/x$ fraction full.

A natural question is whether one can somehow reduce clustering. In this paper, we establish an even stronger statement: the classical linear-probing hash table (even as it was first implemented in the 1950s) already has less clustering than the classical results would seem to suggest. As insertions and deletions are performed over time, the tombstones left behind by deletions cause the combinatorial structure of the hash table to stabilize in a way that eliminates clustering. This means that, for some versions of linear probing, the amortized expected time per operation is actually $\tilde{O}(x)$. We also present a new version of linear probing that avoids clustering entirely, achieving $O(x)$ expected time per operation.

3.10 Cardinality Estimation Using Gumbel Distribution

Aleksander Łukasiewicz (University of Wrocław, PL)

License © Creative Commons BY 4.0 International license
© Aleksander Łukasiewicz

Joint work of Aleksander Łukasiewicz, Przemysław Uznański

Main reference Aleksander Łukasiewicz, Przemysław Uznanski: “Cardinality Estimation Using Gumbel Distribution”, in Proc. of the 30th Annual European Symposium on Algorithms, ESA 2022, September 5-9, 2022, Berlin/Potsdam, Germany, LIPIcs, Vol. 244, pp. 76:1–76:13, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.

URL <https://doi.org/10.4230/LIPIcs.ESA.2022.76>

Cardinality estimation is the task of approximating the number of distinct elements in a large dataset with possibly repeating elements. LogLog and HyperLogLog (c.f. Durand and Flajolet [ESA 2003], Flajolet et al. [Discrete Math Theor. 2007]) are small space sketching schemes for cardinality estimation, which have both strong theoretical guarantees of performance and are highly effective in practice. This makes them a highly popular solution with many implementations in big-data systems (e.g. Algebird, Apache DataSketches, BigQuery, Presto and Redis). However, despite having simple and elegant formulation, both the analysis of LogLog and HyperLogLog are extremely involved – spanning over tens of pages of analytic combinatorics and complex function analysis. We propose a modification to both LogLog and HyperLogLog that replaces discrete geometric distribution with the continuous Gumbel distribution. This leads to a very short, simple and elementary analysis of estimation guarantees, and smoother behavior of the estimator.

3.11 Big data algorithms for distance computation

Yasamin Nazari (Universität Salzburg, AT)

License © Creative Commons BY 4.0 International license
© Yasamin Nazari

Joint work of Greg Bodwin, Michael Dinitz, Jakub Łącki, Yasamin Nazari

Main reference Yasamin Nazari: “Distributed, Parallel and Dynamic Distance Structures”, Dissertation, Baltimore, 2021.

URL <https://jscholarship.library.jhu.edu/bitstream/handle/1774.2/64360/NAZARI-DISSERTATION-2021.pdf>

In recent years, there has been a growing interest in designing algorithms in models that capture different aspects of modern computational systems and big data processing. This talk focuses on graph algorithms in several such models, namely: distributed models, dynamic models, fault tolerant settings and the massively parallel computation model (MPC). Specifically, we focus on graph distance computation in these models. We study several graph theoretic structures that preserve approximate distances, but tradeoff this approximation factor with size, query time, or the number of hops on the approximate shortest paths. We then show how these combinatorial structures can be utilized for faster distance computation in distributed or dynamic settings. We conclude with applications in other related problems such as clustering and routing.

3.12 Locality and Volume in the Context of Algorithms for Dynamic Data Sets

Krzysztof Nowicki (Pathway – Paris, FR)

License © Creative Commons BY 4.0 International license
© Krzysztof Nowicki

Joint work of Aleksander B. G. Christiansen, Krzysztof D. Nowicki, Eva Rotenberg

Main reference Aleksander B. G. Christiansen, Krzysztof D. Nowicki, Eva Rotenberg: “Improved Dynamic Colouring of Sparse Graphs”, CoRR, Vol. abs/2211.06858, 2022.

URL <https://doi.org/10.48550/arXiv.2211.06858>

In this talk I’ll explain how to make algorithms for dynamic data sets out of distributed algorithms. The talk covers recent improvements on theory of algorithms for big data sets [Improved Dynamic Colouring of Sparse Graphs Aleksander B. G. Christiansen, Krzysztof D. Nowicki, Eva Rotenberg] and addresses some possible practical applications of presented approach [at Pathway.com].

3.13 Adaptive Adversaries vs. Randomized Algorithms: A Few Challenges

Krzysztof Onak (Boston University, US)

License © Creative Commons BY 4.0 International license
© Krzysztof Onak

Joint work of Krzysztof Onak, Omri Ben-Eliezer, Talya Eden, Rajesh Jayaram, Jakub Łącki, Slobodan Mitrović, Piotr Sankowski

Streaming algorithms and data structures often leverage randomness to lower their utilization of resources such as memory and computation. I will talk about the possible danger coming from the fact that estimates or replies provided by a randomized algorithm can leak information about its internal randomness. This information could be exploited via adaptive updates to break the guarantees of the algorithm that hold for any fixed sequences of updates with high probability.

Is there a way to make these types of randomized algorithms perform well even in the presence of adaptive updates? In this talk I will share both positive and negative news, ranging from simple techniques for moment estimation in the streaming model to a generic framework that shows that explicitly maintaining a collection of samples from nearly uniform distribution cannot be achieved with a small number of edits. This is a relatively new direction of research and many fundamental questions remain open.

3.14 Massively-Parallel Computing in a Heterogeneous Regime

Rotem Oshman (Tel Aviv University, IL)

License © Creative Commons BY 4.0 International license
© Rotem Oshman

Joint work of Orr Fischer, Adi Horowitz, Rotem Oshman

Main reference Orr Fischer, Adi Horowitz, Rotem Oshman: “Massively Parallel Computation in a Heterogeneous Regime”, CoRR, Vol. abs/2302.14692, 2023.

URL <https://doi.org/10.48550/arXiv.2302.14692>

Massively-parallel graph algorithms have received extensive attention over the past decade, with research focusing on three memory regimes: the superlinear regime, the near-linear regime, and the sublinear regime. The sublinear regime is the most desirable in practice,

but conditional hardness results point towards its limitations. In this work we study a *heterogeneous* model, where the memory of the machines varies in size. We focus mostly on the heterogeneous setting created by adding a single near-linear machine to the sublinear MPC regime, and show that even a single large machine suffices to circumvent most of the conditional hardness results for the sublinear regime: for graphs with n vertices and m edges, we give (a) an MST algorithm that runs in $O(\log \log(m/n))$ rounds; (b) an algorithm that constructs an $O(k)$ -spanner of size $O(n^{1+1/k})$ in $O(1)$ rounds; and (c) a maximal-matching algorithm that runs in $O(\sqrt{\log(m/n)} \log \log(m/n))$ rounds. We also observe that the best known near-linear MPC algorithms for several other graph problems which are conjectured to be hard in the sublinear regime (minimum cut, maximal independent set, and vertex coloring) can easily be transformed to work in the heterogeneous MPC model with a single near-linear machine, while retaining their original round complexity in the near-linear regime. If the large machine is allowed to have *superlinear* memory, all of the problems above can be solved in $O(1)$ rounds.

3.15 IcebergHT: High Performance PMEM Hash Tables Through Stability and Low Associativity Authors: Prashant Pandey

Prashant Pandey (University of Utah – Salt Lake City, US)

License © Creative Commons BY 4.0 International license
© Prashant Pandey

Joint work of Prashant Pandey, Michael A. Bender, Alex Conway, Martin Farach-Colton, William Kuszmaul, Guido Tagliavini, Rob Johnson


Main reference Prashant Pandey, Michael A. Bender, Alex Conway, Martin Farach-Colton, William Kuszmaul, Guido Tagliavini, Rob Johnson: “IcebergHT: High Performance PMEM Hash Tables Through Stability and Low Associativity”, CoRR, Vol. abs/2210.04068, 2022.

URL <https://doi.org/10.48550/arXiv.2210.04068>

Modern hash table designs strive to minimize space while maximizing speed. The most important factor in speed is the number of cache lines accessed during updates and queries. This is especially important on PMEM, which is slower than DRAM and in which writes are more expensive than reads. This paper proposes two stronger design objectives: stability and low-associativity. A stable hash table doesn’t move items around, and a hash table has low associativity if there are only a few locations where an item can be stored. Low associativity ensures that queries need to examine only a few memory locations, and stability ensures that insertions write to very few cache lines. Stability also simplifies scaling and crash safety. We present IcebergHT, a fast, crash-safe, concurrent, and space-efficient hash table for PMEM based on the design principles of stability and low associativity. IcebergHT combines in-memory metadata with a new hashing technique, iceberg hashing, that is (1) space efficient, (2) stable, and (3) supports low associativity. In contrast, existing hash-tables either modify numerous cache lines during insertions (e.g. cuckoo hashing), access numerous cache lines during queries (e.g. linear probing), or waste space (e.g. chaining). Moreover, the combination of (1)-(3) yields several emergent benefits: IcebergHT scales better than other hash tables, supports crash-safety, and has excellent performance on PMEM (where writes are particularly expensive). In our benchmarks, IcebergHT inserts are 50% to $3\times$ faster than state-of-the-art PMEM hash tables Dash and CLHT and queries are 20% to $2\times$ faster. IcebergHT space overhead is 17%, whereas Dash and CLHT have space overheads of $2\times$ and $3\times$, respectively. IcebergHT also exhibits linear scaling and is crash safe. In DRAM, IcebergHT outperforms state-of-the-art hash tables libcuckoo and CLHT by almost $2\times$ on insertions while offering good query throughput and much better space efficiency.

3.16 Can sublinear graph algorithms impact practice?

Comandur Seshadhri (University of California – Santa Cruz, US)

License  Creative Commons BY 4.0 International license
© Comandur Seshadhri

This talk is about sublinear graph algorithms applied to the analysis of real-world graphs. We discuss a provable sublinear algorithms for estimating the degree distribution in sublinear time. To analyze this algorithm, we define fatness indices of the degree distribution. The running time of the algorithm is parameterized by these indices. The algorithm is also practical.

We discuss the overall challenge of designing practical sublinear graph algorithms.

3.17 Locality in online, dynamic, sequential, and distributed graph algorithms

Jukka Suomela (Aalto University, FI)

License  Creative Commons BY 4.0 International license
© Jukka Suomela

Joint work of Amirreza Akbari, Henrik Lievonen, Darya Melnyk, Joona Sarkijarvi, Jukka Suomela
Main reference Amirreza Akbari, Henrik Lievonen, Darya Melnyk, Joona Sarkijarvi, Jukka Suomela: “Online Algorithms with Lookaround”, CoRR, Vol. abs/2109.06593, 2021.
URL <https://arxiv.org/abs/2109.06593>

I will discuss the notion of locality in the context of graph problems, from four different perspectives: online graph algorithms, dynamic graph algorithms, sequential distributed algorithms, and parallel distributed algorithms. I will use the graph coloring problem as a running example, and I will explore settings like this:

- Online graph algorithms: The adversary reveals the graph one node at a time. How far do you need to see around a node until it is safe to pick its color?
- Dynamic graph algorithms: The adversary constructs the graph one edge at a time, and you need to maintain a valid coloring after each addition. How far around the new edge do you need to modify the solution?
- Distributed graph algorithms: Each node has to choose its own color simultaneously in parallel based on its local neighborhood. How far does it need to see?

While these are different questions in general, we can show that there are families of graph problems for which all these notions are equal to each other.

3.18 Dynamic Graph Sketching: To Infinity And Beyond

David Tench (Rutgers University – Piscataway, US)

License  Creative Commons BY 4.0 International license
© David Tench

Existing graph stream processing systems must store the graph explicitly in RAM which limits the scale of graphs they can process. The graph semi-streaming literature offers algorithms which avoid this limitation via linear sketching data structures that use small (sublinear) space, but these algorithms have not seen use in practice to date. In this talk I

will explore what is needed to make graph sketching algorithms practically useful, and as a case study present a sketching algorithm for connected components and a corresponding high-performance implementation. Finally, I will give an overview of the many open problems in this area, focusing on improving query performance of graph sketching algorithms.

3.19 Reactive incremental algorithms with Pathway.

Przemysław Uznański (Pathway – Wrocław, PL)

License © Creative Commons BY 4.0 International license
© Przemysław Uznański

In this talk I discuss the ideas of incremental data processing, reactive algorithms, and their connections to dynamic algorithms, distributed algorithms, and big data processing.

Participants

- Kunal Agrawal
Washington University –
St. Louis, US
- Alkida Balliu
Gran Sasso Science Institute –
L'Aquila, IT
- Michael A. Bender
Stony Brook University, US
- Philip Bille
Technical University of Denmark
– Lyngby, DK
- Sebastian Brandt
CISPA – Saarbrücken, DE
- Vincent Cohen-Addad
Google – Paris, FR
- Alexander Conway
VMware Research –
Palo Alto, US
- Martin Dietzfelbinger
TU Ilmenau, DE
- Anne Driemel
Universität Bonn, DE
- Talya Eden
Bar-Ilan University –
Ramat Gan, IL
- Martin Farach-Colton
Rutgers University –
Piscataway, US
- Inge Li Gørtz
Technical University of Denmark
– Lyngby, DK
- Riko Jacob
IT University of
Copenhagen, DK
- Rob Johnson
VMware – Palo Alto, US
- Hanna Komlós
Rutgers University – New
Brunswick, US
- Adrian Kosowski
Pathway – Paris, FR
- Fabian Daniel Kuhn
Universität Freiburg, DE
- William Kuszmaul
MIT – Cambridge, US
- Aleksander Lukaszewicz
University of Wroclaw, PL
- Petra Mutzel
Universität Bonn, DE
- Yasamin Nazari
Universität Salzburg, AT
- Krzysztof Nowicki
Pathway – Paris, FR
- Dennis Olivetti
Gran Sasso Science Institute –
L'Aquila, IT
- Krzysztof Onak
Boston University, US
- Rotem Oshman
Tel Aviv University, IL
- Prashant Pandey
University of Utah –
Salt Lake City, US
- Ely Porat
Bar-Ilan University –
Ramat Gan, IL
- Simon J. Puglisi
University of Helsinki, FI
- Sofya Raskhodnikova
Boston University, US
- Eva Rotenberg
Technical University of Denmark
– Lyngby, DK
- Ronitt Rubinfeld
MIT – Cambridge, US
- Christian Scheideler
Universität Paderborn, DE
- Comandur Seshadhri
University of California –
Santa Cruz, US
- Christian Sohler
Universität Köln, DE
- Tatiana Starikovskaya
ENS – Paris, FR
- Jukka Suomela
Aalto University, FI
- David Tench
Rutgers University –
Piscataway, US
- Przemyslaw Uznanski
University of Wroclaw, PL



Challenges and Perspectives in Deep Generative Modeling

Vincent Fortuin^{*1}, Yingzhen Li^{*2}, Kevin Murphy^{*3},
Stephan Mandt^{*4}, and Laura Manduchi^{†5}

- 1 University of Cambridge, GB. vbf21@cam.ac.uk
- 2 Imperial College London, GB. yingzhen.li@imperial.ac.uk
- 3 Google Research – Mountain View, US. kpmurphy@google.com
- 4 University of California – Irvine, US. mandt@uci.edu
- 5 ETH Zürich, CH. laura.manduchi@inf.ethz.ch

Abstract

Deep generative models, such as variational autoencoders, generative adversarial networks, normalizing flows, and diffusion probabilistic models, have attracted a lot of recent interest. However, we believe that several *challenges* hinder their more widespread adoption: (C1) the difficulty of objectively evaluating the generated data; (C2) challenges in designing scalable architectures for fast likelihood evaluation or sampling; and (C3) challenges related to finding reproducible, interpretable, and semantically meaningful latent representations. In this Dagstuhl Seminar, we have discussed these open problems in the context of real-world *applications* of deep generative models, including (A1) generative modeling of scientific data, (A2) neural data compression, and (A3) out-of-distribution detection. By discussing challenges C1–C3 in concrete contexts A1–A3, we have worked towards identifying commonly occurring problems and ways towards overcoming them. We thus foresee many future research collaborations to arise from this seminar and for the discussed ideas to form the foundation for fruitful avenues of future research. We proceed in this report by summarizing the main results of the seminar and then giving an overview of the different contributed talks and working group discussions.

Seminar February 12–17, 2023 – <https://www.dagstuhl.de/23072>

2012 ACM Subject Classification Computing methodologies → Unsupervised learning; Computing methodologies → Kernel methods; Computing methodologies → Learning in probabilistic graphical models

Keywords and phrases deep generative models, representation learning, generative modeling, neural data compression, out-of-distribution detection

Digital Object Identifier 10.4230/DagRep.13.2.47

1 Executive Summary

Vincent Fortuin (University of Cambridge, UK)

Yingzhen Li (Imperial College London, UK)

Kevin Murphy (Google Research – Mountain View, US)

Stephan Mandt (University of California – Irvine, US)

License © Creative Commons BY 4.0 International license
© Vincent Fortuin, Yingzhen Li, Kevin Murphy, and Stephan Mandt

Premise

Since the inception of variational autoencoders, generative adversarial networks, normalizing flows, and diffusion models, the field of deep generative modeling has grown rapidly and consistently over the years. Especially in recent years, this has led to great advances in

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Challenges and Perspectives in Deep Generative Modeling, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 47–70

Editors: Vincent Fortuin, Yingzhen Li, Kevin Murphy, Stephan Mandt, and Laura Manduchi



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

generating images, speech and text, as well as great promises in generating structured data such as 3D objects, videos, and molecules. However, we believe that current research has not sufficiently addressed several fundamental challenges related to evaluating and scaling these models, as well as interpreting their latent structure. These challenges have different manifestations in different applications. For example, while a variational autoencoder’s sensitivity to changing data distributions can induce long code lengths and poor image reconstructions in neural compression, the same feature can be a positive attribute in detecting anomalies.

We believe that it is most beneficial to understand the challenges of deep generative models in their practical contexts. For this reason, we have invited a combination of researchers working on foundations of generative models and researchers working on specialized applications to this Dagstuhl Seminar. Thus, by integrating different communities, we have made a step towards identifying generalizable solutions across domains that spur innovation and new research.

As the main challenges of current deep generative modeling approaches we have identified the evaluation of generative models, performing scalable inference in such models, and improving the interpretability and robustness of the models’ learned latent representations.

As example applications, we have considered three *application areas* that draw on generative modeling and that show various manifestations of the aforementioned *challenges*. Concretely, we consider applications in modeling scientific data, neural data compression, and out-of-distribution detection.

Structure of the seminar

We have created an open and inclusive atmosphere where participants from different communities could mingle and exchange ideas, leaving enough room for serendipitous encounters and ad-hoc discussions. We have catalyzed this process by inviting the participants to give short talks on either models (for the researchers) or problems (for the practitioners) as a basis for subsequent discussions. We then had panel discussions and round-tables regarding different topics that the participants could self-assign to, in order to match their common interests.

To promote interactions among researchers especially between those who may have not known each other, we have randomly paired researchers and practitioners into pairs and small groups and assigned them small tasks, such as coming up with a short abstract that would combine their interests. These types of activities have ultimately planted the seeds for different future collaborations and fostered a sense of togetherness among the participants.

Main observations from the talks

The content of the talks is covered in more detail in the next section, but we want to take the opportunity here to highlight recurring patterns and topics that emerged.

One main observation was that while large generative models, such as diffusion models or large language models, yield impressive performance and can solve many tasks that we would naïvely not have expected them to solve well (e.g., diffusion models sorting lists or solving sudokus and large language models performing logical reasoning), we lack a proper

theoretical understanding of these models and can thus not guarantee their safety or reliability. This makes it particularly dangerous to use these models in critical applications, such as healthcare.

Moreover, many domains have specific requirements that are well-known to practitioners, but often ignored by machine learning researchers, e.g., non-iid data, safety constraints, prior knowledge, interpretability, or causal assumptions. While there are sub-fields of machine learning research studying these problems, most off-the-shelf methods do not readily provide solutions.

Finally, generative modeling holds great promise for areas such as neural compression or anomaly/out-of-distribution detection, but the practical improvements achieved by generative approaches in these domains remain limited. We will need more targeted collaborations between experts in generative modeling and these problem settings to make tangible real-world progress, some of which will have hopefully been sparked by this seminar.

Main takeaways from the working groups

Our working group sessions self-assembled spontaneously around key topics of interest that had emerged from the talks and informal discussions during the breaks. They focused on prior knowledge, continual learning, and anomaly detection.

Firstly, when it comes to domain knowledge, one working group tried to develop a categorization of different types and came up with physical constraints, symmetries, logic, ontologies, and factual knowledge. All of these require different approaches to incorporate them into generative models, so the developers of the model should be cognizant of the type of domain knowledge the practitioners might have. Moreover, eliciting the prior knowledge from the experts can be hard and cumbersome, and an elicitation strategy should be designed together with the model itself.

Secondly, continual learning is well-studied in the supervised setting, but less so in the unsupervised one. However, in the age of large generative models that are very expensive to train, continually expanding their generative abilities without having to retrain them from scratch becomes paramount. Since no explicit supervised objective function is available to measure the learning progress or potential forgetting, new solutions need to be developed to efficiently learn continually without catastrophic forgetting in the generative context.

Lastly, anomaly detection is a hard problem that has been studied in the statistical literature for decades, but novel powerful generative models harbor the promise of estimating quantities such as the compressibility or Kolmogorov complexity of data points, which might be used to more effectively detect outliers, out-of-distribution examples, and anomalous inputs.

2 Table of Contents

Executive Summary

<i>Vincent Fortuin, Yingzhen Li, Kevin Murphy, and Stephan Mandt</i>	47
--	----

Overview of Talks



Large Language Models vs. Large AI Models <i>Gerard de Melo</i>	52
Image membership in generative models <i>Sina Däubener</i>	53
Disentangling Style and Content for Neural Topic Models <i>Sophie Fellenz</i>	53
Modeling mixed-tailed distributions with Normalizing Flow and convergence of the ELBO of VAEs to a sum of three entropies <i>Asja Fischer</i>	54
Gaussian Process Variational Autoencoders <i>Vincent Fortuin</i>	55
Active search in structured spaces with domain-specific similarities <i>Thomas Gärtner</i>	55
Deep Generative Models in Healthcare <i>Julia Vogt</i>	56
Do we care about non-iid data for generative models? <i>Matthias Kirchler</i>	56
CMSSG: Heterogeneous image data integration with Causal Multi-Source StyleGAN <i>Christoph Lippert</i>	56
Self-Supervised Learning beyond Vision and Language <i>Maja Rudolph</i>	58
Towards Runtime-Efficient Neural Compression <i>Stephan Mandt</i>	58
Informed Representation Learning with Deep Generative Models <i>Laura Manduchi</i>	59
Towards Anytime Computation in Deep Architectures <i>Eric Nalisnick</i>	61
The Future (R)evolution of Generative AI <i>Björn Ommer</i>	61
Where to Diffuse, how to diffuse, and how to get back? <i>Rajesh Ranganath</i>	62
Universal Critics <i>Lucas Theis</i>	62
Getting the most **out** of your representations <i>Karen Ullrich</i>	63
Interventional causal representation learning with deep generative models <i>Yixin Wang</i>	63

Assaying Out-Of-Distribution Generalization in Transfer Learning <i>Florian Wenzel</i>	64
Trading Information between Latents in Hierarchical Variational Autoencoders <i>Robert Bamler</i>	65
Challenges in Generative Language Modeling <i>Alexander Rush</i>	66
Fun with Foundation Models and Amortized Inference <i>Frank Wood</i>	66
Languages for the Next 700 Application Domains in AI <i>Jan-Willem van de Meent</i>	67
Working groups	
Continual learning of deep generative models <i>Sophie Fellenz, Sina Däubener, Gerard de Melo, Florian Wenzel, and Frank Wood</i>	67
Priors in deep generative modeling <i>Vincent Fortuin, Thomas Gärtner, Matthias Kirchler, and Eric Nalisnick</i>	68
The role of domain knowledge in deep generative models <i>Vincent Fortuin, Thomas Gärtner, Matthias Kirchler, Christoph Lippert, Laura Manduchi, Guy Van den Broeck, Julia Vogt, and Florian Wenzel</i>	69
Anomaly detection using Kolmogorov complexities <i>Marius Kloft, Asja Fischer, and Lucas Theis</i>	69
Participants	70

3 Overview of Talks

3.1 Large Language Models vs. Large AI Models

Gerard de Melo (*Hasso-Plattner-Institut, Universität Potsdam, DE*)

License  Creative Commons BY 4.0 International license
 Gerard de Melo

Language models and other generative models of symbolic sequences have a long history that can be traced back to early studies on prediction probabilities for written language such as those by Shannon [6]. Later on, in the 1980s, their statistics started becoming crucial components of systems for machine translation, speech recognition, and optical character recognition. In the 2000s, the widespread availability of Web-scale word n-gram statistics opened up many new opportunities for language model-driven applications [3].

To generate outputs that resemble human-written text remarkably well at a superficial level, it suffices to sample from very simple n-gram language models. With the advent of neural network-driven language models [1], the generalization abilities improved further. Finally, the series of milestone successes of large language models, most notably the powerful GPT family of neural models [4], has led to generative models of language with unprecedented abilities to generalize to new tasks simply by following instructions.

Along with these improved prediction capabilities, I argue that another notable paradigm shift has emerged. Early studies had already shown that language models are not just mere models of linguistic well-formedness, but rather valuable sources of knowledge [7, 8]. With the powerful capabilities of the GPT models, people started to expect them to serve as universal engines for question answering and broader AI tasks.

While language models are normally supposed to produce *plausible* text, current models are increasingly expected to satisfy a number of *additional desiderata*. For instance, there is a need for models that provide only statements that are deemed factually accurate and trustworthy [2]. There are widespread calls for such models to avoid toxicity and bias [5]. With their deployment in commercial search engines and other mainstream applications, current models are expected to avoid outputs that may lead to harmful effects, for instance by refraining from responding in ways that could pose a risk for the mental health of human interlocutors and by refusing to carry out tasks related to illegal activities.

Thus, large language models are no longer just models of language but general-purpose AI models, leading to an urgent need for us to develop improved generative modeling techniques with substantially better constraint satisfaction and uncertainty estimation.

References

- 1 Yoshua Bengio, Réjean Ducharme, Pascal Vincent, Christian Jauvin. A Neural Probabilistic Language Model. *Journal of Machine Learning Research* vol. 3 (2003), pp. 11-1155. 2003.
- 2 Rajarshi Bhowmik, Gerard de Melo. Be Concise and Precise: Synthesizing Open-Domain Entity Descriptions from Facts. *Proceedings of The Web Conference 2019*. ACM, 2019.
- 3 Thorsten Brants, Alex Franz. *Web 1T 5-gram Version 1*. LDC2006T13. Linguistic Data Consortium, Philadelphia, 2006.
- 4 OpenAI. GPT-4 Technical Report. *arXiv* 2303.08774. 2023.
- 5 Lena Schwertmann, Manoj Prabhakar Kannan Ravi, Gerard de Melo. Model-Agnostic Bias Measurement in Link Prediction. *Findings of the Association for Computational Linguistics: EAACL 2023*. Association for Computational Linguistics, 2023.
- 6 Claude E. Shannon. Prediction and entropy of printed English. *The Bell System Technical Journal* vol. 30, no. 1, pp. 50–64. IEEE, 1951.

- 7 Niket Tandon, Gerard de Melo. Information Extraction from Web-Scale N-Gram Data. In: *Proceedings of the Web N-gram Workshop at SIGIR 2010*. ACM, 2010.
- 8 Niket Tandon, Gerard de Melo, Gerhard Weikum. Deriving a Web-Scale Common Sense Fact Database. In: *Proceedings of AAAI 2011*. AAAI, 2011.

3.2 Image membership in generative models

Sina Däubener (Ruhr-Universität Bochum, DE)

License © Creative Commons BY 4.0 International license
© Sina Däubener

Joint work of Sina Däubener, Asja Fischer, Mike Laszkiewicz, Denis Lukovnikov, Jonas Ricker, Simon Damm, Joel Frank, Thorsten Eisenhofer, Lea Schönherr, Dorothea Kolossa, Thorsten Holz

Images from current state-of-the-art generative models have unarguably led to impressive results. However, this comes with potential threats such as the spread of misinformation through generated images or unauthorized style imitations of certain artists.

To tackle the first problem, I'll first present works from our group and collaborators, which do successful deep fake detection in the frequency domain of images. In the second part, I'll talk about current work in progress where we try to (invisibly) watermark images, such that a model fine tuned on these images picks up on the watermark. This would allow us to certify that works from a certain artist have in fact been used.

3.3 Disentangling Style and Content for Neural Topic Models

Sophie Fellenz (RPTU – Kaiserslautern, DE)

License © Creative Commons BY 4.0 International license
© Sophie Fellenz


Joint work of Sophie Fellenz, Mayank Kumar Nagda

Neural topic models are used to quickly get an overview of the central themes in large text corpora. They are typically trained only on content words, whereas other words related to syntax or style of the text are removed in preprocessing. Rather than relying on manually curated and static stop word lists, we propose a data-driven way of dynamically identifying the style component in text data. The idea is to differentiate between long-range and short-range dependencies in text data. Following previous work in linguistics and cognitive science we posit that content words tend to have long-range dependencies whereas style or syntax words have short-range dependencies. Topic models are good at learning topics by disregarding sequential information, exclusively focusing on long-range dependencies. Language models however process text sequentially to predict the next word given the immediate short-range context. In this talk I present a method for combining both types of models to automatically distinguish syntactic and semantic words. Instead of simply removing the syntactic words, we can also group them and use them for the text analysis alongside the semantic words. Results show that this data-driven way of separating the words leads to higher topic quality and better feature selection on a range of datasets. This may lead to better ways of disentangling content and style in text data in the future, aiding controlled text generation for longer texts and overcoming the current bottleneck in text generation which is the size of input prompt. It was discussed how the topics in neural topic models can be seen as experts and the topic distribution of each document as a product of experts. This view point could help

to develop topic models with a flexible number of topics where topics could be added as new experts. An interesting direction might be to look at hierarchical topic models as a collection of experts where more general experts are located at the top of the hierarchy and combinations of experts or more specific experts would be at the bottom of the hierarchy. Furthermore, topic models could be integrated with foundation models for language. To do this, the tokenization and input format for topic models need to be unified and a general mechanism for conditioning on and extracting topics during text generation needs to be developed.

3.4 Modeling mixed-tailed distributions with Normalizing Flow and convergence of the ELBO of VAEs to a sum of three entropies

Asja Fischer (Ruhr-Universität Bochum, DE)

License  Creative Commons BY 4.0 International license
© Asja Fischer

Main reference Jörg Lücke, Dennis Forster, Zhenwen Dai: “The Evidence Lower Bound of Variational Autoencoders Converges to a Sum of Three Entropies”, CoRR, Vol. abs/2010.14860, 2020.

URL <https://arxiv.org/abs/2010.14860>

Main reference Mike Laszkiewicz, Johannes Lederer, Asja Fischer: “Marginal Tail-Adaptive Normalizing Flows”, in Proc. of the International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA, Proceedings of Machine Learning Research, Vol. 162, pp. 12020–12048, PMLR, 2022.

URL <https://proceedings.mlr.press/v162/laszlkiewicz22a.html>

This talk highlights two recent theoretical results about deep generative models.

The first part is based on the work of Laszkiewicz et al. (2022) and introduces an approach for normalizing flows that allows to model distributions with heavy- as well as light-tailed marginals. We prove that the marginal tailedness of an autoregressive flow can be controlled via the tailedness of the marginals of its base distribution (i.e. the distribution of the hidden variables). This theoretical insight leads us to a novel type of flows that are based on a three-step procedure: first, estimating the marginal tail indices, second, accordingly defining a set of heavy-tailed and a set of light-tailed base distribution, and third, training a normalizing flow with data-driven linear layers.

The second part is based on recent work Damm et al. (2023). Here we show that for standard (i.e., Gaussian) VAEs the ELBO converges to a value given by the sum of three entropies: the (negative) entropy of the prior distribution, the expected (negative) entropy of the observable distribution, and the average entropy of the variational distributions (the latter is already part of the ELBO). The result implies that the ELBO can for standard VAEs often be computed in closed-form at stationary points while the original ELBO requires numerical approximations of integrals.

Both works serve as illustrative examples of the importance of improving our theoretical understanding of deep generative models to gain robust, exact, and efficient generative models.

3.5 Gaussian Process Variational Autoencoders

Vincent Fortuin (University of Cambridge, GB)

License © Creative Commons BY 4.0 International license
© Vincent Fortuin

Joint work of Vincent Fortuin, Dmitry Baranchuk, Gunnar Raetsch, Stephan Mandt
Main reference Vincent Fortuin, Dmitry Baranchuk, Gunnar Raetsch, Stephan Mandt: “GP-VAE: Deep Probabilistic Time Series Imputation”, in Proc. of the Twenty Third International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, Vol. 108, pp. 1651–1661, PMLR, 2020.

URL <https://proceedings.mlr.press/v108/fortuin20a.html>

Variational autoencoders (VAEs) are performant deep generative models, based on principled Bayesian inference techniques. However, in practice, people often use them with isotropic Gaussian priors over the latents, which makes all the different data points independent from each other. This often does not match our true prior beliefs, especially when working with time series data. In this talk, I gave a brief overview of some recent approaches using Gaussian processes (GPs) as priors in VAEs, highlighting their tradeoffs and historical development. During the discussion, we discussed some newer follow-up work that uses Kalman filters for the GP inference in the latent space. We also discussed the choice of kernel and what the computational tradeoffs are between a full variational GP and a variational Gauss-Markov process.

3.6 Active search in structured spaces with domain-specific similarities

Thomas Gärtner (Technische Universität Wien, AT)

License © Creative Commons BY 4.0 International license
© Thomas Gärtner

Generating structured data has many important real-world applications such as generating molecular graphs for (de novo) drug discovery or genome sequences of bacteriophages to combat antibiotic resistant bacteria. The design space of potentially interesting structures in such applications is typically huge, it has, for instance, been estimated that there are more than 10^{30} different phages and even more small, drug-like molecules. This is in stark contrast to the amount of structures known to have desired properties in these domains which is often less than one hundred; in a lead discovery task for an antagonist of a particular integrin thought to play a key role in idiopathic pulmonary fibrosis 24 compounds were known to bind well from previous biological assays. For such applications, we propose an efficient active learning algorithm for generative models with domain-specific similarity measures. Similarity measures defined by domain experts are often not positive semi-definite and thus cannot be utilised by Hilbert-space kernel methods. They instead require more general Krein-space kernel methods which admit efficient learning by adapting Nyström approximations. Our active learning algorithm adapts the distribution of generated structures using a learned conditional exponential family model and leads to diverse sets of novel structures.

3.7 Deep Generative Models in Healthcare

Julia Vogt (ETH Zürich, CH)

License  Creative Commons BY 4.0 International license
© Julia Vogt

In recent years, enormous progress has been made to gather as much information as possible about an individual patient. The continuous adoption and integration of electronic medical records, linkage of data sources, and the advent of new diagnostic and digital monitoring technologies have led to an overwhelming amount of heterogeneous and multimodal clinical data. The ultimate aim is to utilize all this vast information for a medical treatment tailored to an individual patient's needs. To achieve this goal, we develop new generative machine learning techniques capable of dealing with the challenges arising in the medical application domain. The methods we develop cover for example multimodal data integration, transparent model development, or probabilistic clustering models.

3.8 Do we care about non-iid data for generative models?

Matthias Kirchler (Hasso-Plattner-Institut, Universität Potsdam, DE)

License  Creative Commons BY 4.0 International license
© Matthias Kirchler

Joint work of Matthias Kirchler, Christoph Lippert, Marius Kloft
Main reference Matthias Kirchler, Christoph Lippert, Marius Kloft: “Training Normalizing Flows from Dependent Data”, CoRR, Vol. abs/2209.14933, 2022.
URL <https://doi.org/10.48550/arXiv.2209.14933>

Current learning algorithms for generative models generally assume that data points are sampled independently, an assumption that is frequently violated in practice. We propose a likelihood objective of normalizing flows incorporating dependencies between the data points, for which we derive a flexible and efficient learning algorithm suitable for different dependency structures. Respecting dependencies between observations can improve empirical results on synthetic and real-world data, leading to higher statistical power in a downstream application to genome-wide association studies. So far, we have focused on normalizing flows due to their explicit likelihood modeling, but we can extend similar modeling approaches to other generative models. We have also assumed that the dependency structure is at least partially known in advance – we work on relaxing that assumption and learning low-rank approximations of dependency structure from the data.

3.9 CMSSG: Heterogeneous image data integration with Causal Multi-Source StyleGAN

Christoph Lippert (Hasso-Plattner-Institut, Universität Potsdam, DE)

License  Creative Commons BY 4.0 International license
© Christoph Lippert

Joint work of Wei-Cheng Lai, Matthias Kirchler, Hadya Yassin, Jana Fehr, Alexander Rakowski, Hampus Olsson, Ludger Starke, Jason M. Millward, Sonia Waiczies, Christoph Lippert

Introduction: Generative Adversarial Networks (GANs) have emerged as powerful tools for generating realistic images, with conditional and causal models enabling fine-grained control over latent factors. In the medical domain, data scarcity and the need to integrate information

from diverse sources present challenges for existing generative models, often resulting in low-quality images ill-suited for medical applications. To address this issue, we propose the Causal Multi-Source StyleGAN (CMSSG), an algorithm that leverages prior knowledge over the data distribution in the form of causal graphs over image covariates and conditioning to integrate heterogeneous data sources with differing underlying distributions. CMSSG learns from multiple data sources with divergent causal structures in parallel, effectively synthesizing the learned distributions for data generation. We present a proof-of-concept experiment demonstrating CMSSG’s ability to generate hand-written digit images with varying morphological features. Additionally, we apply CMSSG to generate brain MR images with heterogeneous characteristics from the UK Biobank and the Alzheimer’s Disease Neuroimaging Initiative (ADNI) datasets, illustrating its capacity to capture brain anatomical variations. Our proposed algorithm offers a promising direction for unbiased data generation from disparate sources.

Methodology: Our approach consists of two components: the causal component and the multi-source component. The causal component learns causal relationships between clinical and demographic characteristics and brain MRI features, facilitating the synthesis of images with specific attributes. The multi-source component enables the generation of synthetic data from multiple datasets with distinct causal models, fostering the creation of more diverse data with various characteristics from different distributions.

CMSSG learns from multiple data sources with divergent causal structures in parallel, effectively synthesizing the learned distributions for data generation. To the best of our knowledge, our work is the first to address multi-source heterogeneity in GANs within a principled causal framework.

Experiments and Results: We validate our CMSSG method by generating hand-written digits with distinct morphological features. We then apply CMSSG to generate brain MRIs with specific clinical and demographic characteristics. We train our model to learn causal relationships from two datasets in parallel: the UK Biobank dataset, focusing on the relationships between demographic characteristics and MRIs, and the Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohort, examining the relationships between clinical dementia features and MRIs. Using CMSSG, we generate synthetic brain MRIs with controlled age, sex, brain volumes, and cognitive function (normal or impaired).

Our results demonstrate that CMSSG can synthesize high-resolution brain MRIs while realistically manipulating causal structures within the images. Although the Frechet Inception Distance (FID) score from CMSSG does not outperform CausalStyleGANs with a single causal model, it provides a new opportunity to manipulate multi-source causal covariates.

Conclusion: In conclusion, the Causal Multi-Source StyleGAN (CMSSG) represents a novel approach to address the challenges of data scarcity and biased datasets in medical image generation. By leveraging causal graphs and conditioning, CMSSG integrates heterogeneous data sources with differing underlying distributions to generate high-quality, diverse medical images. Our experiments demonstrate the efficacy of CMSSG in generating hand-written digit images and brain MRIs with specific clinical and demographic characteristics.

Future work will involve collaboration with medical experts for comprehensive quality assessment and exploration of potential applications of synthetic medical images. Additionally, further experiments should be conducted to improve the quality of synthetic images from joint causal covariates and design appropriate metrics for evaluating the causality of GANs in respect to anatomical factors. This will ensure that the model learns the correct anatomical pattern with the aging process.

3.10 Self-Supervised Learning beyond Vision and Language

Maja Rudolph (Cornell University – Ithaca, US)

License  Creative Commons BY 4.0 International license
© Maja Rudolph

Self-supervised learning has emerged as a powerful paradigm for machine learning, especially for drawing insights from unlabeled data. The key idea is to introduce auxiliary prediction tasks and to train a deep model to solve these auxiliary tasks. If the tasks are designed well, the trained model will be useful for a number of purposes, such as anomaly detection, feature extraction, and forecasting. Unfortunately, most successful approaches for SSL rely on domain-specific inductive biases and are, therefore, limited to individual use cases. In this talk, I present advanced self-supervised learning losses that facilitate domain-general self-supervised learning beyond images and text. Exponential family embeddings, for example, generalize word embeddings to provide insight into a wide range of applications. They are a useful tool for studying zebrafish brains in neuroscience, studying shopping behavior in economics, or studying language evolution in computational social science. Similarly, neural transformation learning (NTL) is a new general-purpose tool for self-supervised anomaly detection. While related methods in computer vision typically require image transformations such as rotations, blurring, or flipping, NTL automatically learns the best transformations from the data and generalizes self-supervised AD to almost any data type.

3.11 Towards Runtime-Efficient Neural Compression

Stephan Mandt (University of California – Irvine, US)

License  Creative Commons BY 4.0 International license
© Stephan Mandt

Joint work of Stephan Mandt, Yibo Yang, Robert Bamler

Main reference Yibo Yang, Robert Bamler, Stephan Mandt: “Improving Inference for Neural Image Compression”, in Proc. of the Advances in Neural Information Processing Systems, Vol. 33, pp. 573–584, Curran Associates, Inc., 2020.

URL https://proceedings.neurips.cc/paper_files/paper/2020/file/066f182b787111ed4cb65ed437f0855b-Paper.pdf

Neural image and video compression models have proven superior performance in rate-distortion and rate-perception tradeoffs compared to their classical counterparts. However, while most research still focuses on improving rate-distortion tradeoffs, neural compression models are currently much too resource-inefficient to deploy in real-world environments. This talk seeks to review strategies to maintain the strong performance of neural compression methods while aiming to reduce their runtime efficiency by 1-2 orders of magnitude. To this end, we propose three modeling and algorithmic improvements: (1) introducing lightweight decoders, (2) improving encoding at training time using semi-amortized variational inference, and (3) establishing probabilistic circuits as new models for efficient entropy coding in lossy compression.

Background: The internet and the world’s IT systems could not exist in their current form without data compression. With video streaming dominating consumer internet traffic, every percent of gained performance improvement will have a large economic impact. Neural codecs are potentially also better suited for new data formats, such as light fields for AR/VR applications, lidar data, or multi-view video. These technologies’ fast evolution will demand rapid prototyping, making learnable codecs appealing. Neural codecs also lack the common

block-coding visual artifacts and can be “supervised” to allocate more bits to certain features of interest. Neural codecs are more flexible than traditional codecs and can be optimized for superior perceptual quality or other custom metrics at much lower bitrates.

While the focus of the neural compression community has been largely on improving the tradeoff between bitrate and distortion (or perceptual quality), neural compression methods are currently 1-2 orders of magnitude slower than their classical counterparts. This makes their real-time deployment highly impractical, e.g., downloading and unpacking a large machine learning data set such as ImageNet or decoding video in real-time. By drawing on resource-efficient architectures, iterative inference, and new entropy coding schemes based on parsimonious models, we argue that the community should seek to maintain the strong performance of neural compression methods while increasing their runtime efficiency ideally by 1-2 orders of magnitude, removing one of the major obstacles from widespread deployment of neural codecs in the real world.

Many compression applications, such as video streaming on Youtube, impose strict runtime limitations upon decoding while allowing a much larger computational budget upon encoding. In “Improving Inference for Neural Image Compression” [Yang, Bamler, Mandt, NeurIPS 2020], we exploited this asymmetry by improving the encoding process using a larger computational budget while leaving the decoder untouched. To this end, we searched for an improved discrete latent representation at test time using an annealing scheme. This way, we obtained 15-20% rate savings without modifying the decoder. Our result suggests that iterative inference may achieve state-of-the-art compression performance with more lightweight decoders.

In contrast to most existing work focusing on architectural improvements in non-linear transform coding, we stress the importance of algorithmic advances that have broad applicability to existing methods, e.g., by exploiting the asymmetrical resource budgets for encoding and decoding via iterative inference and/or by developing new paradigms for entropy coding. Ways to improve the resource efficiency of neural codecs include drawing on lightweight decoders, iterative encoding at training time, and advances in parsimonious generative models. Our goals are to accelerate transform coding while also proposing new architectures for efficient entropy coding based on recent work on lossless compression with probabilistic circuits.

3.12 Informed Representation Learning with Deep Generative Models

Laura Manduchi (ETH Zürich, CH)

License © Creative Commons BY 4.0 International license

© Laura Manduchi

Joint work of Laura Manduchi, Thomas M. Sutter, Alain Ryser, Julia E. Vogt, Ricards Marcinkevics, Fabian Laumer, Joachim M. Buhmann, Kieran Chin-Cheong

One of the most popular frameworks to extract meaningful information from a vast amount of unlabeled datasets is representation learning. The latter should encode valuable information for downstream tasks, such as classification, regression, and visualization. However, there are many circumstances where purely data-driven approaches lead to unsatisfactory results that are inconsistent with the domain knowledge. On the other hand, deep generative models can encode physical laws and constraints into their generative process to obtain preferred representations of data, enabling exploratory analysis of complex data types. In this talk, I explored several approaches to incorporate domain knowledge, in the form of constraints,

probabilistic relations, and prior distributions, in VAEs for static and temporal data with a focus on clustering. First, I introduced two different gaussian mixture prior distributions used in VAEs to enforce a clustering structure in the latent space. The first one [1] employs a categorical prior distribution on the clusters, while the second one [2] uses a differentiable hypergeometric distribution to overcome the i.i.d. assumption of the input data. I then introduced the inclusion of domain knowledge in the form of probabilistic relations and survival information to obtain representations with a clear clustering structure for time series and survival data, using a mixture of Weibull distributions [3]. I then focused on instance-level constraints to guide the learning process toward a preferred configuration using high-dimensional data, using a Conditional Gaussian Mixture Model [4]. Last but not least, I showed how the proposed techniques can be applied to real-world medical applications with a focus on cardiology. In cardiovascular medicine, to correctly quantify cardiac function and diagnose dysfunction, expensive and time-consuming medical imaging methods are often required. This might lead to a lack of available diagnostic modalities and inadequate patient care. The use of machine learning models based on affordable and minimally invasive diagnostics, such as echocardiography, might serve as a valuable assistant tool to enhance health care for people with cardiovascular diseases. However, a lack of large labeled datasets in the medical field prevents the use of supervised deep learning techniques. Therefore, there is a need for informed representation learning algorithms to leverage prior information on unlabeled datasets of cardiac ultrasound videos. The learned representation can then be used to solve further downstream tasks, such as diagnosis, anomaly detection, and denoising [5], [6].

References

- 1 Jiang, Z., Zheng, Y., Tan, H., Tang, B., Zhou, H. (2016). Variational Deep Embedding: An Unsupervised and Generative Approach to Clustering. International Joint Conference on Artificial Intelligence.
- 2 T. M Sutter, L. Manduchi, A. Ryser J. E. Vogt (2023). Learning Group Importance using the Differentiable Hypergeometric Distribution, International Conference on Learning Representations.
- 3 Manduchi, L., Marcinkevics, R., Massi, M.C., Gotta, V., Müller, T., Vasella, F., Neidert, M.C., Pfister, M., Vogt, J.E. (2022). A Deep Variational Approach to Clustering Survival Data. International Conference on Learning Representations.
- 4 Manduchi, L., Chin-Cheong, K., Michel, H., Wellmann, S., Vogt, J.E. (2021). Deep Conditional Gaussian Mixture Model for Constrained Clustering. Neural Information Processing Systems.
- 5 Ryser, A., Manduchi, L., Laumer, F., Michel, H., Wellmann, S., Vogt, J.E.. (2022). Anomaly Detection in Echocardiograms with Dynamic Variational Trajectory Models. Proceedings of the 7th Machine Learning for Healthcare Conference, in Proceedings of Machine Learning Research.
- 6 Laumer, F., Amrani, M., Manduchi, L., Beuret, A., Rubi, L., Dubatovka, A., Matter, C.M., Buhmann, J.M. (2022). Weakly supervised inference of personalized heart meshes based on echocardiography videos. Medical image analysis.

3.13 Towards Anytime Computation in Deep Architectures

Eric Nalisnick (University of Amsterdam, NL)

License © Creative Commons BY 4.0 International license
© Eric Nalisnick

Joint work of Eric Nalisnick, Allingham, James

Main reference James U. Allingham, Eric Nalisnick: “A Product of Experts Approach to Early-Exit Ensembles”.
Workshop on Dynamic Neural Networks at ICML, 2022.

URL https://dynn-icml2022.github.io/papers/paper_10.pdf

Predictive models often need to be evaluated in dynamic, uncertain computational conditions. For example, a model deployed to a mobile device should be useful, despite a lack of computational power. Moreover, the same model ideally should provide better performance on devices with richer computational resources. Hence, there is a need for “early-exit” architectures that allow computation to be halted early, before the model has run to completion. Current architectures provide little-to-no strong guarantees about how these intermediate predictions relate to the final prediction produced by the full model.

In this talk, I describe a construction that provide one such guarantee. Specifically, we guarantee that the probability of the mode under the full model monotonically increases in the intermediate solutions as more computation is done. This is achieved by a product-of-experts approach, as its predictive distribution takes the form of an intersection of the experts. We demonstrate that this architecture can be realized for both real-valued regression and multi-class classification.

3.14 The Future (R)evolution of Generative AI

Björn Ommer (LMU München, DE)

License © Creative Commons BY 4.0 International license
© Björn Ommer

Main reference Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer: “High-Resolution Image Synthesis with Latent Diffusion Models”, in Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022, pp. 10674–10685, IEEE, 2022.

URL <https://doi.org/10.1109/CVPR52688.2022.01042>

Recently, deep generative modeling has become the most prominent paradigm for learning powerful representations of our (visual) world and for generating novel samples thereof. Consequently, this has already become the main building block for numerous algorithms and practical applications. This talk will contrast the most commonly used generative models to date with a particular focus on denoising diffusion probabilistic models, the core of the currently leading approaches to visual synthesis. Despite their enormous potential, these models come with their own specific limitations. We will then discuss a solution, latent diffusion models, a.k.a. “Stable Diffusion”, that significantly improves the efficiency of diffusion models. Now billions of training samples can be summarized in compact representations of just a few gigabytes so that the approach runs on consumer GPUs. We will then discuss recent extensions that cast an interesting perspective on future generative models. In particular, retrieval augmentation during inference promises to significantly reduce model sizes by having powerful likelihood models focus on the composition of a scene rather than learning the training data. We will then highlight key aspects in which generative modeling will change in the future.

3.15 Where to Diffuse, how to diffuse, and how to get back?

Rajesh Ranganath (NYU Courant Institute of Mathematical Science, US)

License © Creative Commons BY 4.0 International license
© Rajesh Ranganath

Main reference Raghav Singhal, Mark Goldstein, Rajesh Ranganath: “Where to Diffuse, How to Diffuse, and How to Get Back: Automated Learning for Multivariate Diffusions”, CoRR, Vol. abs/2302.07261, 2023.

URL <https://doi.org/10.48550/arXiv.2302.07261>

Generative models have been making large leaps in both quantitative fidelity and qualitative appeal. One class of models that has been a driving force behind these leaps is diffusion-based generative models. Diffusion-based generative models, or diffusions models for short, work by first corrupting data towards a known, fixed stationary distribution and training a model to undo those corruptions, and thus providing a means to generate data by uncorrupting a sample from the stationary distribution. The choice of corruption or inference process affects both likelihoods and sample quality. For example, it has been shown that extending the inference diffusion with auxiliary variables, making them multivariate, leads to improved sample quality. However, deriving training algorithms for each new inference diffusion is onerous requiring manually deriving stationary distributions and transition kernels. To simplify the training, we provide a recipe for likelihood training of multivariate diffusion models. In the first step, we derive a lower bound on the likelihood. Next, we show how the terms in the lower bound can be automatically computed and show how to parametrize inference diffusions using results from Markov chain Monte Carlo to target a specific stationary distribution. We study several different inference diffusions and demonstrate how to learn and the value of learning inference diffusions.

3.16 Universal Critics

Lucas Theis (Google – London, GB)

License © Creative Commons BY 4.0 International license
© Lucas Theis

What distinguishes a realistic image from an unrealistic image? Despite tremendous progress in our ability to generate realistic images, we still lack functions that can reliably detect artifacts in images. Such a function would be of great interest in a variety of applications such as outlier detection, perceptual quality evaluation, neural compression, or neural rendering. The field of algorithmic probability provides many insights on a closely related question; namely, when is data a plausible sample from a distribution P ? However, these results are not widely known in the machine learning community. In this short presentation, I will discuss how Kolmogorov complexity can inspire “universal critics” – functions that are able to detect unrealistic images without being trained on corrupted data.

3.17 Getting the most ****out**** of your representations

Karen Ullrich (Meta – New York, US)

License © Creative Commons BY 4.0 International license
© Karen Ullrich

Joint work of Karen Ullrich, Yangjun Ruan, Daniel Severo, James Townsend, Ashish Khisti, Arnaud Doucet, Alireza Makhzani, Yann Dubois, Benjamin Bloem-Reddy, Chris J Maddison

Main reference Yangjun Ruan, Karen Ullrich, Daniel Severo, James Townsend, Ashish Khisti, Arnaud Doucet, Alireza Makhzani, Chris J. Maddison: “Improving Lossless Compression Rates via Monte Carlo Bits-Back Coding”, in Proc. of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event, Proceedings of Machine Learning Research, Vol. 139, pp. 9136–9147, PMLR, 2021.

URL <http://proceedings.mlr.press/v139/ruan21a.html>

Main reference Yann Dubois, Benjamin Bloem-Reddy, Karen Ullrich, Chris J Maddison: “Lossy Compression for Lossless Prediction”, in Proc. of the Advances in Neural Information Processing Systems, Vol. 34, pp. 14014–14028, Curran Associates, Inc., 2021.

URL https://proceedings.neurips.cc/paper_files/paper/2021/file/7535bbb91c8fde347ad861f293126633-Paper.pdf

The goal of source compression is to map any outcome of a discrete random variable $x \sim p_d(x)$ in a finite symbol space $x \in S$ to its shortest possible binary representation. Given a tractable model probability mass function (PMF) $p(x)$ that approximates $p_d(x)$, entropy coders provide such an optimal mapping. As a result, the task of source compression is simplified to identifying a good model PMF for the data at hand.

Even though the setup as described is the most commonly used one, there are restrictions to it. Entropy coders can only process one dimensional variables and process them sequentially. Hence the structure of the entropy coder implies a sequential structure of the data. This is a problem when compressing sets instead of sequences. In the first part of the talk, I present an optimal codec for sets [1]. The problem we encounter for sets can be generalized for many other structural priors in data. In the second part of the talk I thus investigate the problem. We generalize rate distortion theory for structural data priors and develop a strategy to learn codecs for this data [2].

References

- 1 D. Severo, J. Townsend, A. Khisti, A. Makhzani and K. Ullrich. “Compressing Multisets with Large Alphabets.” IEEE Journal on Selected Areas in Information Theory 2023.
- 2 Dubois, Y., Bloem-Reddy, B., Maddison, C.J. “Lossy Compression for Lossless Prediction.” Neural Information Processing Systems 2021.

3.18 Interventional causal representation learning with deep generative models

Yixin Wang (University of Michigan – Ann Arbor, US)

License © Creative Commons BY 4.0 International license
© Yixin Wang

Joint work of Yixin Wang, Kartik Ahuja, Divyat Mahajan, Yoshua Bengio, Michael I Jordan

Main reference Yixin Wang, Michael I. Jordan: “Desiderata for Representation Learning: A Causal Perspective”, CoRR, Vol. abs/2109.03795, 2021.

URL <https://arxiv.org/abs/2109.03795>

Main reference Kartik Ahuja, Yixin Wang, Divyat Mahajan, Yoshua Bengio: “Interventional Causal Representation Learning”, CoRR, Vol. abs/2209.11924, 2022.


URL <https://doi.org/10.48550/arXiv.2209.11924>

Causal representation learning seeks to extract high-level latent factors from low-level sensory data. Most existing methods rely on fitting deep generative model to observational data, leveraging structural assumptions (e.g., conditional independence) to identify the latent

factors. However, interventional data is prevalent across applications. Can interventional data facilitate causal representation learning? We explore this question in this talk. The key observation is that interventional data often carries geometric signatures of the latent factors’ support (i.e. what values each latent can possibly take). For example, when the latent factors are causally connected, interventions can break the dependency between the intervened latents’ support and their ancestors’. Leveraging this fact, we prove that the latent causal factors can be identified up to permutation and scaling given data from perfect do interventions. Moreover, we can achieve block affine identification, namely the estimated latent factors are only entangled with a few other latents if we have access to data from imperfect interventions. These results highlight the unique power of interventional data in causal representation learning; they can enable provable identification of latent factors without any assumptions about their distributions or dependency structure.

3.19 Assaying Out-Of-Distribution Generalization in Transfer Learning

Florian Wenzel (Amazon Web Services – Tübingen, DE)

License  Creative Commons BY 4.0 International license
© Florian Wenzel

Joint work of Florian Wenzel, Andrea Dittadi, Peter Vincent Gehler, Carl-Johann Simon-Gabriel, Max Horn, Dominik Zietlow, David Kernert, Chris Russell, Thomas Brox, Bernt Schiele, Bernhard Schölkopf, Francesco Locatello


Main reference Florian Wenzel, Andrea Dittadi, Peter Vincent Gehler, Carl-Johann Simon-Gabriel, Max Horn, Dominik Zietlow, David Kernert, Chris Russell, Thomas Brox, Bernt Schiele, Bernhard Schölkopf, Francesco Locatello: “Assaying Out-Of-Distribution Generalization in Transfer Learning”, CoRR, Vol. abs/2207.09239, 2022.

URL <https://doi.org/10.48550/arXiv.2207.09239>

Since out-of-distribution generalization is a generally ill-posed problem, various proxy targets (e.g., calibration, adversarial robustness, algorithmic corruptions, invariance across shifts) were studied across different research programs resulting in different recommendations. While sharing the same aspirational goal, these approaches have never been tested under the same experimental conditions on real data. In this paper, we take a unified view of previous work, highlighting message discrepancies that we address empirically, and providing recommendations on how to measure the robustness of a model and how to improve it. To this end, we collect 172 publicly available dataset pairs for training and out-of-distribution evaluation of accuracy, calibration error, adversarial attacks, environment invariance, and synthetic corruptions. We fine-tune over 31k networks, from nine different architectures in the many- and few-shot setting. Our findings confirm that in- and out-of-distribution accuracies tend to increase jointly, but show that their relation is largely dataset-dependent, and in general more nuanced and more complex than posited by previous, smaller scale studies.

3.20 Trading Information between Latents in Hierarchical Variational Autoencoders

Robert Bamler (Universität Tübingen, DE)

License  Creative Commons BY 4.0 International license

© Robert Bamler

Joint work of Robert Bamler, Tim Xiao

Motivated by the recent success stories of generative modeling with diffusion models, we reconsider the training objective of hierarchical variational autoencoders (VAEs). By keeping the so-called forward process (from data to latent representations) fixed and recurrent, diffusion models are able to efficiently scale up training to very deep probabilistic models with many layers of latents, leading to impressive generative modeling performances. However, it is not so clear how to use diffusion models for applications that make use of the forward process, such as representation learning or data reconstruction tasks, as the fixed forward diffusion process progressively reduces mutual information with the original data up to the point of almost no correlation. For these types of applications, VAEs with their learned inference models are often a more natural choice of model architecture. Motivated by the empirical observation that deep hierarchies of layers of latent variables are crucial to the performance of diffusion models, we reconsider the trade-off between reconstruction error (“distortion”) and information content of the latents (“rate”) in hierarchical VAEs, i.e., VAEs with more than one layer of latents.

We observe first that the general rate/distortion trade-off of beta-VAEs [Alemi et al., ICML 2018] can be refined by splitting up the rate term into a sum of contributions from each layer of latents. Importantly, however, this separation is only possible if the inference model proceeds in the same direction as the generative model, i.e., opposite to the direction of the otherwise analogous forward process in diffusion models. Splitting the rate into layer-wise contributions allows practitioners to control the information content of each layer individually by introducing individual layer-wise Lagrange multipliers (“beta hyperparameters”). We argue that this increased control is useful in practice by grouping application domains of (hierarchical) VAEs into three categories depending on whether they use (i) only the generative model (generative tasks), (ii) only the inference model (representation learning tasks), or (iii) both (reconstruction tasks). We show both by deriving theoretical performance bounds, as well as by large-scale empirical evaluations that the optimal distribution of rate between layers of latents is different for the three categories of applications, and we provide practical guidance for choosing reasonable values for the beta hyperparameters for each category.

The main open question that we will consider in the future is if the proposed hierarchical rate/distortion theory can be used to train VAEs where the size of the latent representation does not necessarily match between the inference and the generative process. Here, the highest-level latents could represent the length of the next lower-level latent representation. Such a variable-length information bottleneck would allow training VAEs with transformer architectures for text, e.g., building on the work by Henderson and Fehr (arXiv:2207.13529). Treating the length as a (higher-level) latent variable would allow training scenarios where the length of the reconstructed text does not necessarily match the length of the original text. The hope is that this would allow a VAE to more freely rephrase text, and that the individual beta-hyperparameters would allow controlling the length and the diversity of reconstructed text separately.

3.21 Challenges in Generative Language Modeling

Alexander Rush (Cornell University – Ithaca, US)

License  Creative Commons BY 4.0 International license
© Alexander Rush

My talk will argue that Generative models are the defining element of modern NLP. I will describe some of the recent usage of generative models for NLP. Specifically the now popular knowledge that they are extremely impressive tools that are central to the field. Understanding that this is different in spirit than the main focus of the seminar, I will try to start a discussion as to why NLP generative models are somehow less powerful than other generative approaches. Given this context, I will describe some of the remaining modeling challenges in using these systems. Specifically, language models have demonstrated the ability to generate highly fluent text; however, they still require additional scaffolding to maintain coherent high-level structure (e.g., story progression). Using the model criticism in latent space we can evaluate the high-level structure by comparing distributions between real and generated data in a latent space obtained according to an assumptive generative process. Different generative processes identify specific failure modes of the underlying model. We perform experiments on three representative aspects of high-level discourse – coherence, coreference, and topicality – and find that transformer-based language models are able to capture topical structures but have a harder time maintaining structural coherence or modeling coreference structures. Based on these conclusions, I pose questions about how we might update our models for language and ask whether richer generative processes might better capture some aspects of language current systems are missing.

3.22 Fun with Foundation Models and Amortized Inference


Frank Wood (University of British Columbia – Vancouver, CA)

License  Creative Commons BY 4.0 International license
© Frank Wood

In this talk, I will discuss the work of my UBC PLAI group and spin-out Inverted AI on foundation models of behavior, images, and video. In particular, I will talk about ways to get such models to “do what you want them to do” via amortized inference after they have been trained. I will spend most of my time introducing and talking about ITRA, a model I think can become the GPT of behavior, and how we use inference, including a novel algorithm called “critic SMC,” alongside reinforcement learning as inference techniques, to “tune” ITRA to stay on the road in new places, not collide with other agents, and produce trajectories that are achievable in the dynamics sense by specific vehicle classes. I will then discuss very related techniques for amortized conditional inference in image and video generative models, work that has led to state-of-the-art inpainting and conditional image generation results requiring no fine-tuning of a pre-trained VAE image model and stunning recent results on very long duration photorealistic video generation arising from meta-learning a flexible conditioning DDPM-based video generative model architecture.

3.23 Languages for the Next 700 Application Domains in AI

Jan-Willem van de Meent (University of Amsterdam, NL)


License  Creative Commons BY 4.0 International license
© Jan-Willem van de Meent

This will be a talk about where AI has arrived today, where we could be going in the next few years, and the role that probabilistic approaches to AI have to play in these developments. I will discuss where I see opportunities in applications of AI to computational design in the physical sciences, and I will discuss how programming language design can help realize these opportunities, with particular attention to our recent work on languages for inference programming, which opens up opportunities for new model and inference designs, both in the context of simulation-based inference and in the context of deep generative models.

4 Working groups

4.1 Continual learning of deep generative models

Sophie Fellenz (RPTU – Kaiserslautern, DE), Sina Däubener (Ruhr-Universität Bochum, DE), Gerard de Melo (Hasso-Plattner-Institut, Universität Potsdam, DE), Florian Wenzel (Amazon Web Services – Tübingen, DE), and Frank Wood (University of British Columbia – Vancouver, CA)

License  Creative Commons BY 4.0 International license
© Sophie Fellenz, Sina Däubener, Gerard de Melo, Florian Wenzel, and Frank Wood

Given that foundation models are increasing in size, the training time for these models is also increasing, which makes frequent retraining impractical. A more desirable alternative would be to continually update existing models with new data. This way we could build agents that learn continuously (life-long learning) and make learning more efficient. The main challenge we identified here is “catastrophic forgetting”. How can we make sure that the model incorporates new information without forgetting what it already knows? Many techniques have been proposed in the supervised setting where new labels can be added over time, but we posit that in order to solve continual learning in the supervised setting, it first needs to be solved in the unsupervised setting. We discussed techniques such as fine-tuning, functional regularization, context augmentation, storing part of the data (core sets) or storing a part of learned parameters. None of these are satisfactory solutions as they cannot guarantee to prevent catastrophic forgetting and are hard to optimize or control in practice. Sequential Monte Carlo or streaming variational Bayes are theoretically possible but do not scale in practice. We also discussed hybrid approaches that learn representations using neural networks and apply a Kalman Filter or similar on the condensed representations. As interesting directions we furthermore identified active learning on prompts and adversarial interactions.

4.2 Priors in deep generative modeling

Vincent Fortuin (University of Cambridge, GB), Thomas Gärtner (Technische Universität Wien, AT), Matthias Kirchler (Hasso-Plattner-Institut, Universität Potsdam, DE), and Eric Nalisnick (University of Amsterdam, NL)

License  Creative Commons BY 4.0 International license
© Vincent Fortuin, Thomas Gärtner, Matthias Kirchler, and Eric Nalisnick

This group discussed the problem of how to elicit priors for deep generative models from domain experts and how to encode them in the model. The running example was the problem of learning a generative model for antibiotic drugs from a small set of existing molecules (on the order of a few hundred). Standard deep generative modeling would probably not be data-efficient enough to learn a sufficiently expressive model from such a small dataset, but the hope would be to use prior knowledge from domain experts, such as chemists. The problem is that this knowledge is often rather vague in the chemist’s mind, for instance, some rough intuition for what kind of functional groups typical antibiotics should have, or how large or aromatic they are. One promising idea to elicit this prior knowledge from the expert is to show them molecules and then query their beliefs about them, either by asking them whether this looks like an antibiotic (binary feedback) or how much it looks like an antibiotic (continuous feedback). The molecules to show them can be samples from the model before training, so essentially from the prior predictive, which would then allow directly tuning the prior of the model based on the feedback. Alternatively, we can use a large dataset of unlabeled chemicals to collect feedback on and then use that to distill a prior. Both of these approaches have the disadvantage that most of the molecules we would show the expert would probably not look like antibiotics, so we would waste a lot of their time for not a lot of information. One would probably need to use some acquisition function, similar to active learning, to try and only ask about the most informative compounds in each iteration. We also discussed in this vein that the process could be made easier by treating the expert knowledge as a likelihood instead of a prior, that is, first training the generative model based on the small dataset and then finetuning its generations with the human feedback, which would hopefully create more interesting structures than the prior predictive. Moreover, if we ask the expert to make changes to the generated molecules to make them more drug-like, we could use a distribution over these changes as a prior for the score function in score-based models such as diffusion models. We also discussed that we could try to use weak supervision signals such as human-defined similarity measures or hand-designed features to build a classifier from the human feedback and then train the generative model with classifier-guidance instead of a proper prior. As a mechanism to incorporate the prior knowledge into the model, except for the aforementioned proper prior distributions or classifier-guidance, we also discussed rejection sampling, importance sampling, and probabilistic circuits. Overall, we concluded that directly incorporating vague human knowledge into a proper prior distribution seems hard and that approaches based on iterative human feedback are probably more promising.

4.3 The role of domain knowledge in deep generative models

Vincent Fortuin (University of Cambridge, GB), Thomas Gärtner (Technische Universität Wien, AT), Matthias Kirchler (Hasso-Plattner-Institut, Universität Potsdam, DE), Christoph Lippert (Hasso-Plattner-Institut, Universität Potsdam, DE), Laura Manduchi (ETH Zürich, CH), Guy Van den Broeck (UCLA, US), Julia Vogt (ETH Zürich, CH), and Florian Wenzel (Amazon Web Services – Tübingen, DE)

License © Creative Commons BY 4.0 International license
 © Vincent Fortuin, Thomas Gärtner, Matthias Kirchler, Christoph Lippert, Laura Manduchi, Guy Van den Broeck, Julia Vogt, and Florian Wenzel

In this working group we discussed different types of domain knowledge and how it can be incorporated into deep generative models. Firstly, we discussed why it can be useful to incorporate domain knowledge and agreed that it might improve data-efficiency, enable extrapolation beyond the training dataset, and increase trustworthiness of the model. Secondly, we discussed different kinds of domain knowledge and how they could respectively be implemented in deep generative models. Knowledge about causality can be incorporated using causal graphs or structural equation models in some latent space, which then implies disentanglement of the latent factors corresponding to causal factors. More generally, known probabilistic relationships can be represented in the form of prior distributions, either in the latent space or the data space directly. Physical constraints, invariances, and symmetries can often be incorporated directly through the choice of model architecture, for instance, CNNs in the case of images. Facts about the world can be incorporated via knowledge graphs or database retrieval mechanisms, while logical statements can be incorporated through fuzzy or probabilistic logic. Finally, ontologies can be incorporated via hierarchical modeling. Overall, we concluded that one of the main challenges is still to design a latent space in which representations carry semantic meaning, since many of these types of domain knowledge would need to be incorporated into such a latent space. This is loosely related to the problem of symbol grounding from continuous distributed representations.

4.4 Anomaly detection using Kolmogorov complexities

Marius Kloft (RPTU – Kaiserslautern, DE), Asja Fischer (Ruhr-Universität Bochum, DE), and Lucas Theis (Google – London, GB)

License © Creative Commons BY 4.0 International license
 © Marius Kloft, Asja Fischer, and Lucas Theis

Ideally, one would threshold the log-(pseudo)likelihood ratio $s = \log(p/q)$ of the distribution of the normal data p and the distribution of the anomalous data q for provably optimal anomaly detection. In practice, q is unknown, and one resorts to thresholding $\log(p)$. Steinwart (2005) showed that this could be equivalent to thresholding $\log(p/q)$ where q is a uniform distribution of anomalies. In practice, however, it has been observed that anomalies typically are “simpler” (easier to compress) – a phenomenon known as “Occam’s razor”. We propose to replace $\log(q)$ in s by an approximation of k , the Kolmogorov complexity, which – intuitively speaking – measures the likeliness of an instance occurring in nature. Furthermore, as an extension, one could integrate the Kolmogorov complexity with an estimate of q based on some observed anomalies.

Participants

- Robert Bamler
Universität Tübingen, DE
- Ryan Cotterell
ETH Zürich, CH
- Sina Däubener
Ruhr-Universität Bochum, DE
- Gerard de Melo
Hasso-Plattner-Institut,
Universität Potsdam, DE
- Sophie Fellenz
RPTU – Kaiserslautern, DE
- Asja Fischer
Ruhr-Universität Bochum, DE
- Vincent Fortuin
University of Cambridge, GB
- Thomas Gärtner
Technische Universität Wien, AT
- Matthias Kirchler
Hasso-Plattner-Institut,
Universität Potsdam, DE
- Marius Kloft
RPTU – Kaiserslautern, DE
- Yingzhen Li
Imperial College London, GB
- Christoph Lippert
Hasso-Plattner-Institut,
Universität Potsdam, DE
- Stephan Mandt
University of California –
Irvine, US
- Laura Manduchi
ETH Zürich, CH
- Eric Nalisnick
University of Amsterdam, NL
- Björn Ommer
LMU München, DE
- Rajesh Ranganath
NYU Courant Institute of
Mathematical Science, US
- Maja Rudolph
Bosch Center for AI –
Pittsburgh, US
- Alexander Rush
Cornell University – Ithaca, US
- Lucas Theis
Google – London, GB
- Karen Ullrich
Meta – New York, US
- Jan-Willem van de Meent
University of Amsterdam, NL
- Guy Van den Broeck
UCLA, US
- Julia Vogt
ETH Zürich, CH
- Yixin Wang
University of Michigan –
Ann Arbor, US
- Florian Wenzel
Amazon Web Services –
Tübingen, DE
- Frank Wood
University of British Columbia –
Vancouver, CA



Agents on the Web

Olivier Boissier^{*1}, Andrei Ciortea^{*2}, Andreas Harth^{*3},
Alessandro Ricci^{*4}, and Danai Vachtsevanou^{†5}

- 1 Ecole des Mines – St. Étienne, FR. olivier.boissier@emse.fr
- 2 Universität St. Gallen, CH. andrei.ciortea@unisg.ch
- 3 Fraunhofer IIS – Nürnberg, DE. andreas@harth.org
- 4 Università di Bologna, IT. a.ricci@unibo.it
- 5 Universität St. Gallen, CH. danai.vachtsevanou@unisg.ch

Abstract

Recent standardization on the Web of Things and (Social) Linked Data unlocks new practical use cases and new opportunities for research on Web-based multi-agent systems. While existing research on multi-agent systems can contribute to engineering adaptive and flexible Web-based systems, increased deployment of systems following the recent standards can bring new insight into engineering large-scale and open multi-agent systems. These developments motivate the need for a broader perspective that can only be achieved through a concerted effort of the research communities on *Web Architecture and Web of Things*, *Semantic Web and Linked Data*, and *Autonomous Agents and Multi-Agent Systems*. Thus, the main objective of the *Dagstuhl Seminar 23081 on Agents on the Web* was to investigate these new research opportunities, to support the transfer of knowledge and results across the different communities, and to create a network of leading scholars and practitioners around these topics. This report documents the seminar's program and outcomes. To continue the joint work after the seminar, the seminar participants created the *W3C Autonomous Agents on the Web (WebAgents) Community Group*.

Seminar February 19–24, 2023 – <https://www.dagstuhl.de/23081>

2012 ACM Subject Classification Information systems → World Wide Web; Computing methodologies → Multi-agent systems; Human-centered computing → Ubiquitous computing

Keywords and phrases Web Architecture, Web of Things, Semantic Web, Linked Data, Multi-Agent Systems

Digital Object Identifier 10.4230/DagRep.13.2.71


1 Executive Summary

Olivier Boissier (Ecole des Mines – St. Étienne, FR)

Andrei Ciortea (Universität St. Gallen, CH)

Andreas Harth (Fraunhofer IIS – Nürnberg, DE)

Alessandro Ricci (Università di Bologna, IT)

License  Creative Commons BY 4.0 International license
© Olivier Boissier, Andrei Ciortea, Andreas Harth, and Alessandro Ricci

The recent evolution towards hypermedia-driven services, Linked Data, and the Web of Things is turning the Web into a homogeneous hypermedia fabric that interconnects everything – devices, physical objects, documents, or abstract concepts. The latest standards on the Web of Things and (Social) Linked Data allow automated software clients to *navigate*, *query*, *observe*, and *act on* this uniform hypermedia fabric. Use cases that have long been envisioned

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Agents on the Web, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 71–162

Editors: Olivier Boissier, Andrei Ciortea, Andreas Harth, Alessandro Ricci, and Danai Vachtsevanou



DAGSTUHL
REPORTS

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

for artificial agents on the Web – as published in the original Semantic Web vision in 2001 and going back to the early days of the Internet – are now closer to their practical implementation and deployment. Nevertheless, the engineering of such modern Web-based systems poses research challenges that have yet to be addressed.

Today's Web-based systems are inherently complex, heterogeneous, and increasingly dynamic (e.g., especially in the context of the Web of Things). Moreover, the World Wide Web was designed to be a decentralized system that spans not only geographical boundaries but also organizational boundaries. In such settings, traditional paradigms for engineering Web-based systems become impractical. Many of the research questions underlying these challenges – such as how to design software agents able to cope with complex and dynamic environments, or how to design and govern interactions in decentralized systems – have been investigated in research on autonomous agents and multi-agent systems.¹ To design a new generation of Web-based systems, we thus need a broader perspective that can only be achieved through a concerted effort of several research communities – which, for the purpose of this seminar, we identified as the communities around *Web Architecture and Web of Things*, *Semantic Web and Linked Data*, and *Autonomous Agents and Multi-Agent Systems*.

The *Dagstuhl Seminar 23081 on Agents on the Web* continued to investigate the research opportunities identified in the *Dagstuhl Seminar 21072 on Autonomous Agents on the Web* in order to consolidate the discussions and to continue the transfer of knowledge and results across the involved research communities. Concretely, the seminar pursued the following objectives:

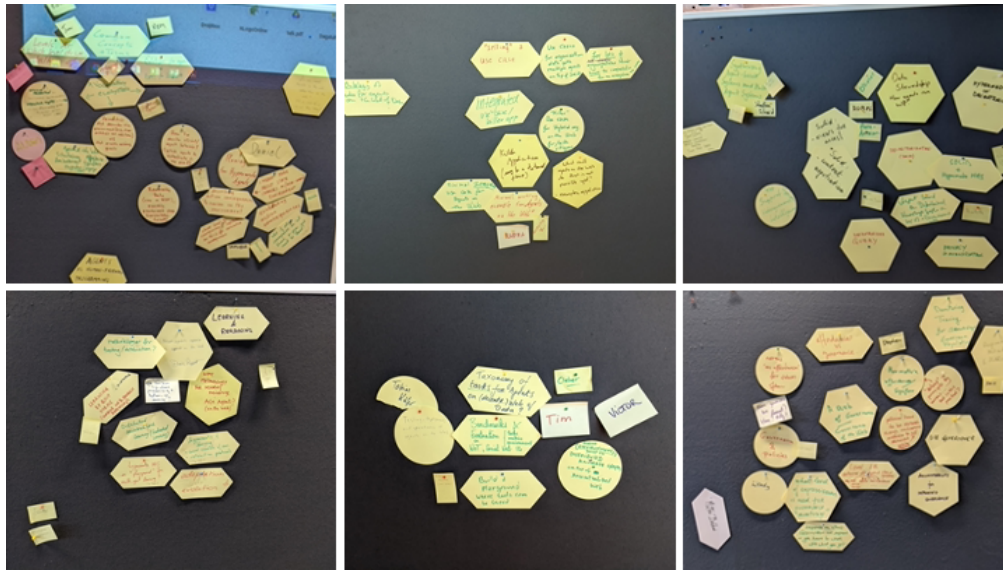
- to identify and align the various research threads in the targeted communities that are relevant for advancing the research on Web-based multi-agent systems;
- to work towards a shared conceptualization and shared theoretical underpinnings for Web-based multi-agent systems;
- to identify representative use cases in different domains that would help demonstrate the potential impact of this joint research effort on society and the economy;
- to evaluate the state of technologies available for prototyping and deploying Web-based multi-agent systems.

The main motivation for this seminar was to consolidate a network of senior and young researchers around the topics. To continue the joint work after the seminar, participants created the *W3C Autonomous Agents on the Web (WebAgents) Community Group*.²

The Seminar Format

The seminar brought together 39 participants across the above-mentioned research communities. The 5-day seminar was a blend of presentations, live demonstrators, and group work structured around three types of sessions:

- *presentations of position statements*: sessions organized on the first day based on position statements submitted by the seminar participants;
- *Demos & Tech sessions*: sessions organized on the afternoon of the second day for presenting demonstrators and technologies relevant to the seminar;
- *working group sessions*: sessions organized throughout the week for focused group work on specific topics identified during the seminar.



■ **Figure 1** Concept boards used to organize the seminar topics into working groups.

The seminar started with presentations of position statements submitted by participants to help bootstrap the discussions. Participants were invited to read all submitted position statements before the seminar and to prepare a 1-minute presentation of their position statements. To help create a narrative throughout the day, a selection of topics was invited for 5-minute presentations. The first day ended with a brainstorming session, in which participants used several concept boards to organize the seminar topics and to form working groups for the rest of the week.

The first half of the second day started with working group sessions, and the afternoon was reserved for the Demos & Tech sessions. The objective of these sessions was to ground the conceptual discussions from the working groups and to paint a picture of what can already be achieved with existing technologies. In total, these sessions attracted 16 demonstrators, out of which 7 submitted abstracts for this report (see Section 4).

The rest of the week continued with working group sessions. Beginning of each day, participants were invited to pitch new ideas for working groups. In addition, a synchronization session was organized in the middle of the week to review the progress, reinforce bridges across working groups, and reorganize the working groups if needed. In total, seven working groups were created and their consolidated discussion summaries are presented in Section 5.

Overview of the Report

This report is organized into four main parts. Section 3 includes the position statements submitted by seminar participants that were presented during the first day. Section 4 includes a list of abstracts for several of the demonstrators presented during the seminar. Section 5 includes the reports submitted by the working groups created during the seminar.

¹ For a broad overview of this area, we refer the interested reader to: Gerhard Weiss (ed.), *Multiagent Systems (Second Edition)*, MIT Press, ISBN 978-0262018890, 2013.

² <https://www.w3.org/community/webagents/>

2 Table of Contents

Executive Summary

Olivier Boissier, Andrei Ciortea, Andreas Harth, and Alessandro Ricci 71

Overview of Position Statements

How to Ensure Agents on the Web are Behaving as Expected?

Cleber Jorge Amaral 77

Swarm Intelligence for Agents in Smart Edge Environments

Jean-Paul Calbimonte 78

On Information and Interactions on the Web

Victor Charpenay 80

Towards Decentralized Applications based on Protocols, Norms, and Accountability

Amit K. Chopra 81

Principled Design of Web-based Multi-Agent Systems

Andrei Ciortea 82

Towards a Low Entry Barrier for Agents on the Web

Stephen Craneheld 84

Society = Autonomy + Adaptation

Jérôme Euzenat 86

Multi-Agent Systems on the Web as a Special Kind of Knowledge Graphs Ecosystems

Catherine Faron 86

Governing Communities of Autonomous Agents and People on the Web Using Social Norms

Nicoletta Fornara 87

Finding the Critical MAS(S): Resources and Representations Needed for Weaving a Web that Hosts Linked Multi-Agent Systems

Fabien Gandon 88

About the Place of Agents in the Web

Jomi Fred Hübner 90

Towards an Analysis-oriented Perspective on Agent-oriented Abstractions for the Web

Timotheus Kampik 90

Agent-based Information Systems Built on Read-Write Linked Data

Tobias Käfer 92

Old Agents, New Environments: Open Issues and Unexpected Opportunities

Stefano Mariani 93

Pervasive Autonomous Systems: So Much to Learn from One Another

Simon Mayer 94

Federated Learning in Goal-oriented WoT Environments

Mahda Noura 96

Integrating Multi-Agent Systems within Web-based Microservices Architectures – Multi-Agent MicroServices (MAMS) <i>Eoin O’Neill</i>	97
Human-like AI for Artificial General Intelligence <i>Dave Raggett</i>	99
Ontologies with Temporal Logics Allow Hypermedia Agents to Make Plans <i>Daniel Schraudner</i>	100
Rethinking Agents and Meaning over the Web of Things <i>Munindar P. Singh</i>	101
Challenges for Query Agents on the Decentralized Web <i>Ruben Taelman</i>	105
Signifying Affordances for Effective Interaction of Agents on the Web <i>Danai Vachtsevanou</i>	106
Agents for the Decentralized Web of Data <i>Jan Van den Bussche</i>	108
Socio-Technical Networking Platforms for Multi-Agent Collaboration on the Web <i>Antoine Zimmermann</i>	109

Overview of Demonstrators

Developing Multi-Agent Systems Using JaCaMo-REST <i>Cleber Jorge Amaral, Jomi Fred Hübner, Timotheus Kampik</i>	110
Test-Driven Development for Agent-Oriented Programming <i>Cleber Jorge Amaral, Jomi Fred Hübner, Timotheus Kampik</i>	111
Robust JaCaMo Applications via Exceptions and Accountability <i>Matteo Baldoni, Cristina Baroglio, Roberto Micalizio, Stefano Tedeschi</i>	113
Agent-Oriented Visual Programming for the Web of Things <i>Samuele Burattini</i>	115
Interaction-Oriented Programming <i>Amit K. Chopra, Samuel H. Christie V, Munindar P. Singh</i>	116
Hypermedia Multi-Agent Systems <i>Olivier Boissier, Andrei Ciortea, Fabien Gandon, Simon Mayer, Alessandro Ricci</i>	116
Hypermedia MAS Simulation <i>Rem Collier</i>	117
Building Multi-Agent Microservices with ASTRA <i>Eoin O’Neill</i>	118
Web-based Demos of Cognitive AI <i>Dave Raggett</i>	119
Using MOSAIK for a Decentralized Transportation System <i>Daniel Schraudner</i>	121
Querying over Solid Pods via Link Traversal, with Comunica <i>Ruben Taelman</i>	122

Working Group Reports

Decentralized Hypermedia Ecosystems <i>Olivier Boissier, Pierre-Antoine Champin, Amit K. Chopra, Nicoletta Fornara, Fabien Gandon, Ruben Taelman, Jan Van den Bussche</i>	123
Abstractions for Agents on the Web <i>Andrei Ciortea, Catherine Faron, Jomi Fred Hübner, Eoin O’Neill, María Poveda-Villalón, Alessandro Ricci, Antoine Zimmermann</i>	126
Representation Formats Supporting Goal-oriented Decision-Making on the Web <i>Cleber Jorge Amaral, Samuele Burattini, Jean-Paul Calbimonte, Rem Collier, Matthias Kovatsch, Mahda Noura, Daniel Schraudner, Danai Vachtsevanou</i>	131
Course Check and Revision Strategies for Autonomous Hypermedia Navigation <i>Cleber Jorge Amaral, Samuele Burattini, Jean-Paul Calbimonte, Rem Collier, Andreas Harth, Matthias Kovatsch, Brian Logan, Simon Mayer, Mahda Noura, Sebastian Schmid, Daniel Schraudner, Danai Vachtsevanou</i>	138
Learning and Reasoning with Behavioural Norms <i>Mehdi Dastani, Kemal A. Delic, Jérôme Euzenat, Stefano Mariani, Dave Raggett, Munindar P. Singh</i>	145
Challenges and Opportunities in Governing Agents on the Web <i>Matteo Baldoni, Victor Charpenay, Andrei Ciortea, Stephen Cranefield, Julian Padget, Munindar P. Singh</i>	148
Agents on the Web: Specifying and Implementing Testbeds <i>Cleber Jorge Amaral, Jean-Paul Calbimonte, Jomi Fred Hübner, Timotheus Kampik, Tobias Käfer</i>	156
Participants	162

3 Overview of Position Statements

3.1 How to Ensure Agents on the Web are Behaving as Expected?

Cleber Jorge Amaral (*Universidade Federal de Santa Catarina, BR*)

License © Creative Commons BY 4.0 International license
© Cleber Jorge Amaral

Agents have access to an abundance of resources and interactions on the Web. At present, complex Web agents such as trading bots and personal assistants utilize a combination of data-driven and symbolic techniques to develop their reasoning capabilities [1]. For the Multi-Agent Systems (MAS) community, Agent-Oriented Programming (AOP) languages based on the BDI model of agency, which offers abstractions for knowledge, intentions, objectives, and plans, forms an excellent foundation for combining different symbolic, stochastic, and sub-symbolic AI techniques [2]. However, it can be difficult to ensure the behavior of autonomous and proactive software artifacts is reliable given the diversity of techniques and environment richness.

Indeed, as the Web offers an unpredictable and dynamic environment and agents are usually part of a system that contains several agents, it is tough to reproduce and control interactions of agents with other agents and humans [3]. Besides, it can be hard to test and even to understand agents as they may present behaviors that have no formal specification and they can have several alternative plans for achieving a particular goal. Other aspects such as scaling to numerous autonomous software artifacts running relatively freely in the environment that can also contain malign agents can be even more problematic [4].

Given the aforementioned difficulties, the following questions can be raised:

- How to ensure teleo-reactive agents on the Web are behaving as expected considering a multiplicity of course of actions that they can take for achieving a particular goal?
- How can it be ensured that a new version of a certain MAS is strengthening rather than weakening in the many scenarios to which the preceding version was subjected?
- How can be create reproducible Web environments that are rich in resources, interactions, and dynamic that can be used as a development sandbox and to compare various techniques?

References

- 1 Tarek R. Besold and Artur d'Avila Garcez and Sebastian Bader and Howard Bowman and Pedro Domingos and Pascal Hitzler and Kai-Uwe Kuehnberger and Luis C. Lamb and Daniel Lowd and Priscila Machado Vieira Lima and Leo de Penning and Gadi Pinkas and Hoifung Poon and Gerson Zaverucha. *Neural-Symbolic Learning and Reasoning: A Survey and Interpretation*. 2017. <https://arxiv.org/abs/1711.03902>
- 2 Bordini, Rafael H. and El Fallah Seghrouchni, Amal and Hindriks, Koen and Logan, Brian and Ricci, Alessandro. *Agent programming in the cognitive era*. Autonomous Agents and Multi-Agent Systems, 2020. <https://doi.org/10.1007/s10458-020-09453-y>
- 3 Linz, Tilo. *Testing autonomous systems*. The Future of Software Quality Assurance, Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-29509-7_5
- 4 Koopman, Philip and Wagner, Michael. *Challenges in autonomous vehicle testing and validation*. SAE International Journal of Transportation Safety, 2016. <https://www.jstor.org/stable/26167741>

3.2 Swarm Intelligence for Agents in Smart Edge Environments

Jean-Paul Calbimonte (HES-SO Valais – Sierre, CH)

License  Creative Commons BY 4.0 International license
© Jean-Paul Calbimonte

Motivation

Smart devices, sensors, autonomous robots, drones, and other similar instruments have increasingly become pervasive in industrial and home environments, for numerous use-cases and scenarios. These devices are not limited to the acquisition of data as mere observers of their surroundings, but they are also capable of actuation, and of potentially complex processing of rapidly produced data flows. Given their limitations in terms of computational, storage, and power capabilities, some of these processing features can be delegated to intermediate edge nodes in a network of smart devices, integrated with cloud interactions when required. While in simple cases a manual and explicit set up of this infrastructure is possible, the reality is that more and more use-cases are subject to highly dynamic changes in the number and type of devices, as well as the nature of their interactions. Under these conditions it becomes necessary to use higher-level abstractions that allow the self-organization of smart devices, bringing intelligence to the edge nodes, and allowing the collaborative distribution of their computational tasks through swarm-inspired behavioral patterns.

Challenges and open issues

Considering the potential of swarm agents for self-organizing edge nodes and smart devices, we identify a number of relevant challenges:

- Heterogeneous agent knowledge: Multiple schemas and ontologies exist for representing sensor and IoT devices and data. Are these enough to allow self-organization? [1, 5, 2]
- Are there swarm-inspired models that can be propagated to smart edge agents? Can these act as mediators among devices according to their capabilities? [4]
- What is the role of social coordination of agents in a swarm environment? Can the Web serve as a common place to enable the establishment of inter-deployment knowledge exchanges? [8]
- Negotiation in the context of swarm agents can bring both advantageous optimization patterns, as well as delays and potential conflicts. It may become crucial to identify ways of reducing the risks, and to find compromises that adapt to changing situations and redefinition of goals. [7, 6]
- In this context the risks of manipulation and mischievous behaviors is more than plausible. The conception of integrity, confidence, and transparency mechanisms need to be integrated into this research road-map.
- Edge and sensing devices are often required to acquire and handle sensitive data, which should be subject to strict privacy constraints, while keeping the computation and processing goals. [3, 9]

Opportunities

Addressing these challenges, swarm intelligent agents for smart edge may open a number of research opportunities:

- Semantic representation of shared goals, organization patterns, and behaviors have the potential to facilitate agent coordination and negotiation in swarm smart edge environments.
- Implementation of low-code solutions can help abstracting the complexity of smart edge agents, and their deployment under heterogeneous conditions.
- Smart edge agents necessarily need to go well beyond the Web as a means to interact. Nevertheless, the Web offers a solid and standard mechanism to enable the interoperability, discovery, and accessibility among different swarms.
- Fully decentralized swarms of agents can make use of existing approaches for self-organization that have been successfully used in previous works in other contexts.
- Domain-specific deployments can help drive the requirements and implementation of this idea on different areas including automation, self-driving vehicles, robotics, domotics, eHealth, etc.
- Federated learning and decentralized processing can be adapted to run on swarm agents for smart edge environments, further expanding their applicability.

References

- 1 M. Bermudez-Edo, T. Elsaleh, P. Barnaghi, and K. Taylor. Iot-lite: a lightweight semantic model for the internet of things. In *International IEEE conferences on ubiquitous intelligence & computing*, pages 90–97. IEEE, 2016.
- 2 V. Charpenay and S. Käbisch. On modeling the physical world as a collection of things: The w3c thing description ontology. In *The Semantic Web: 17th International Conference, ESWC 2020, Heraklion, Crete, Greece, May 31–June 4, 2020, Proceedings 17*, pages 599–615. Springer, 2020.
- 3 L. Chen, S. Fu, L. Lin, Y. Luo, and W. Zhao. Privacy-preserving swarm learning based on homomorphic encryption. In *Algorithms and Architectures for Parallel Processing: 21st International Conference, ICA3PP 2021, Virtual Event, December 3–5, 2021, Proceedings, Part III*, pages 509–523. Springer, 2022.
- 4 C. Guéret, S. Schlobach, K. Dentler, M. Schut, and G. Eiben. Evolutionary and swarm computing for the semantic web. *IEEE Computational Intelligence Magazine*, 7(2):16–31, 2012.
- 5 K. Janowicz, A. Haller, S. J. Cox, D. Le Phuoc, and M. Lefrançois. Sosa: A lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics*, 56:1–10, 2019.
- 6 N. Kouka, R. Fdhila, and A. M. Alimi. Multi objective particle swarm optimization based cooperative agents with automated negotiation. In *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part IV 24*, pages 269–278. Springer, 2017.
- 7 F. Lorenzi, D. S. dos Santos, and A. L. Bazzan. Negotiation for task allocation among agents in case-base recommender systems: a swarm-intelligence approach. In *2005 International workshop on multi-agent information retrieval and recommender systems*, pages 23–27, 2005.
- 8 S. Mokarizadeh, A. Grosso, M. Matskin, P. Kungas, and A. Haseeb. Applying semantic web service composition for action planning in multi-robot systems. In *2009 Fourth International Conference on Internet and Web Applications and Services*, pages 370–376. IEEE, 2009.
- 9 B. Zhao, X. Liu, A. Song, W.-N. Chen, K.-K. Lai, J. Zhang, and R. H. Deng. Primpso: A privacy-preserving multiagent particle swarm optimization algorithm. *IEEE Transactions on Cybernetics*, 2022.

3.3 On Information and Interactions on the Web

Victor Charpenay (*Mines Saint-Étienne, FR*)

License  Creative Commons BY 4.0 International license
© Victor Charpenay

Information on the Web is (primarily) symbolic

In the past few years, most spectacular advances in research have been made possible by the Web, to a large extent. The field of computer vision owes much to ImageNet, a dataset compiled with images found on the Web³. In natural language processing, large language models such as BERT or GPT are trained on large corpora found on the Web, including Wikipedia and CommonCrawl⁴ data.

Yet, the Web holds value in itself not because it makes it easy to publish and access large dumps of data (most open data is hard to use because it is heterogeneous and lacks contextual information) but because pieces of information coming from different sources can easily be interlinked.

That makes the Web a source of information that is essentially relational. The Resource Description Framework (RDF) captures that essence by restricting its data model to triples of the form $\langle resource, relation, resource \rangle$, which generalizes hyperlinks. As a consequence, Web agents *should* be able to deal with relational data, if not RDF data.

Information on the Web is (sometimes) approximate

It is legitimate to ask oneself whether the Web would bring anything new to Multi-Agent System (MAS) design. Because of the genericity of its architectural principles, any MAS can be implemented on the Web [3]. But is there a MAS that can *only* be implemented on the Web? If there is one, it will most likely be at a large scale (both with respect to the size of the environment and to the complexity of agent behaviors).

As soon as data is exchanged at a large scale, there is a long-tail phenomenon when it comes to its quality. This phenomenon is well-known among Semantic Web researchers, who have accepted the fact that information found on the Web is sometimes approximate: few publishers will spend a significant amount of time on the quality of the data they publish but most publishers won't. This was the starting point for the Pedantic Web group⁵, that aimed at increasing data quality with low effort (mostly via better tooling).

Accepting approximations requires including a certain amount of statistical inference and learning in agent design. With the coming of Knowledge Graphs (KG), large collections of well-known facts [4], it is possible to learn vector representations of resources and relations, such that an agent can estimate the truth value of unknown $\langle resource, relation, relation \rangle$ statements. Web agents *should* be able to leverage KG latent representations.

Interactions on the Web are (strictly) discrete

To act on the Web, agents must fill in forms and send individual requests. This uniform layer for both agent-to-environment and agent-to-agent interactions requires to describe actions in terms of input, output, preconditions and effects (IOPE). It also requires that agents have discrete, symbolic models of perception and action, very different from e.g. models based on physical simulations.

³ <https://image-net.org/about.php>

⁴ <https://commoncrawl.org/>

⁵ <https://pedantic-web.org/>

Such as statement summons “orthodox” approaches in MAS research, whose common objective has been to find a logical language that was both easy to understand and expressive enough to describe any MAS. One can mention e.g. Michael Wooldridge’s Logic of Rational Agents (*LORA*) [2] and Michael Fisher’s METATEM [1]. These two languages have not been widely adopted for specifying agent behaviors but they remain highly relevant for verifying the behavior of agents that use non-symbolic latent representations.

Both *LORA* and METATEM rely on temporal logics. The Web Ontology Language (OWL), which should be the language of choice to model information on the Web, does not align well with temporal logics. It does, however, easily capture other useful modalities to express beliefs, knowledge, morality, etc. Web agents *should* be able to interpret a mixture of temporal and OWL specifications and have rational (i.e. logical) behaviors with respect to these specifications.

References

- 1 M. Fisher, An Introduction to Practical Formal Methods Using Temporal Logic. Chichester, UK: John Wiley & Sons, Ltd, 2011.
- 2 M. J. Wooldridge, Reasoning about rational agents. Cambridge, Mass: MIT Press, 2000.
- 3 V. Charpenay, T. Käfer, and A. Harth, “A Unifying Framework for Agency in Hypermedia Environments,” in Engineering Multi-Agent Systems, Springer International Publishing, 2022.
- 4 A. Hogan et al., “Knowledge Graphs,” ACM Compututer Surveys, 2021.

3.4 Towards Decentralized Applications based on Protocols, Norms, and Accountability

Amit K. Chopra (Lancaster University, GB)

License  Creative Commons BY 4.0 International license
© Amit K. Chopra

Web applications were envisaged to support social processes, to facilitate interactions between their users. However, a Web application is a conceptually centralized entity that mediates and, as exemplified by popular social media platforms, influences and controls their interactions. Noting this loss of control for users, efforts such as Berners-Lee led Solid are part of the growing trend toward decentralized data in the Web. However, current efforts do not go far enough because they ignore the larger issue of how to accommodate the autonomy of the users. Accommodating autonomy requires not only data decentralization, but also the representation and decentralization of decision making by users [1].


We advocate thinking of applications fundamentally in terms of decentralized multiagent systems (MAS), where each agent represents a user and embodies not only control over the user’s data but also their decision making. We advocate modeling a MAS in terms of social abstractions such as the norms and protocols that apply to interaction between users. The norms could pertain to business interactions, privacy, and in general, governance. Norms give grounding for accountability relationships between users, which are crucial to understanding reasons for norms violations by users. The big question is how to realize such a decentralized multiagent system using Web technologies such the HTTP and HATEAOS. Data semantics promotes interoperability, but we also need interaction semantics, which is where MAS research shines and provides a basis for logical decentralization.

References

- 1 Amit K. Chopra and Mumindar P. Singh *From Social Machines to Social Protocols: Software Engineering Foundations for Sociotechnical Systems*. In Proceedings of the 25th International Conference on World Wide Web, pages 903–914, Montreal, 2016.

3.5 Principled Design of Web-based Multi-Agent Systems

Andrei Ciortea (*Universität St. Gallen, CH*)

License  Creative Commons BY 4.0 International license
© Andrei Ciortea

There has been significant work on engineering Web-based multi-agent systems (MAS), but several points remain open. Several ideas from the involved research communities appear fundamental for deriving a principled way forward.

Environment as a First-class Abstraction

In MAS, multiple agents interact in a shared environment – and thus require a supporting infrastructure to discover other agents, to exchange messages, to coordinate, etc. This led to the view of the *environment as a first-class abstraction* in MAS [12] – and even as a *programming abstraction* [9] that allows agents to shape their environment such that it better fits their needs. The separation of concerns between agents and the environment abstraction is fundamental for engineering Web-based MAS. First, it simplifies integrating architectures for MAS with the Web Architecture. Second, it allows agents and components in their environment to be developed, deployed, and evolve independently from one another. Third, it ensures forward compatibility in long-lived Web-based MAS: the environment abstraction can reify other existing or future conceptual dimensions in MAS. Together with colleagues, we have been developing the idea of hypermedia-based environments as a first-class abstraction in Web-based MAS – and as a conceptual bridge between MAS and the Web [4, 2, 1, 3].

Asynchronous Interaction

It is widely accepted that synchronous interactions are insufficient for engineering MAS (e.g., see [10]). The Web was originally built on synchronous interactions between components – but since then, a plethora of methods and protocols have been developed to support asynchronous interaction on the Web. Such extensions address requirements beyond the classic Web Architecture and it is not always obvious how they fit within a REST-style Web. The reason is that REST, as a coherent closed set of architectural design decisions, was defined in the mid-90s to meet the needs of the Web at that time – and asynchronous interaction was not among them [5]. Several generations of researchers have continued to extend the insights of REST to meet requirements beyond the classic Web Architecture [6]. For example, one such extension is *Asynchronous REST* [7], which allows clients to observe resources. A similar method is implemented by CoAP [11].⁶

⁶ The Constrained Application Protocol (CoAP) is a Web protocol for constrained devices and networks developed in the context of the Web of Things [11].

Uniform Knowledge-Level Abstraction

The Web defines a hypermedia-driven, uniformly-accessible, knowledge-level abstraction of the world.⁷ If we consider the environment as a first-class abstraction in Web-based MAS, where the environment can also reify all the other conceptual dimensions in the MAS, then we have all the elements we need to provide agents with a knowledge-level abstraction of their system that they can also interact with [1]. What remains is to ensure that agents are able to both *perceive* and *act on* their hypermedia-based environment in a uniform way.

Situatedness and Embodiment

The dominant view in AI research is that intelligent behavior is closely related to the environment an agent occupies and is not disembodied [13]. This view emerged in the late '80s in close relationship with research on intelligent robots [8], which are naturally *situated* and *embodied* in a physical environment. On the Web, there is no implicit support for situatedness and embodiment: e.g., multiple agents can browse the same website without being aware of other agents. However, there are also no inherent limitations: if we consider the environment as a first-class abstraction in Web-based MAS, then the hypermedia-based environment can be designed to provide agents with various levels of support for situatedness and embodiment – it is merely a design choice.

References

- 1 Andrei Ciortea, Olivier Boissier, and Alessandro Ricci. 2019. Engineering World-Wide Multi-Agent Systems with Hypermedia. In *Engineering Multi-Agent Systems*, Danny Weyns, Viviana Mascardi, and Alessandro Ricci (Eds.). Springer International Publishing, Cham, 285–301.
- 2 Andrei Ciortea, Olivier Boissier, Antoine Zimmermann, and Adina Magda Florea. 2017. Give Agents Some REST: A Resource-Oriented Abstraction Layer for Internet-Scale Agent Environments. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (São Paulo, Brazil) AAMAS '17*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1502–1504.
- 3 Andrei Ciortea, Simon Mayer, Fabien Gandon, Olivier Boissier, Alessandro Ricci, and Antoine Zimmermann. 2019. A Decade in Hindsight: The Missing Bridge Between Multi-Agent Systems and the World Wide Web. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (Montreal QC, Canada) AAMAS '19*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1659–1663.
- 4 Andrei Ciortea, Antoine Zimmermann, Olivier Boissier, and Adina Magda Florea. 2016. Hypermedia-Driven Socio-Technical Networks for Goal-Driven Discovery in the Web of Things. In *Proceedings of the Seventh International Workshop on the Web of Things (Stuttgart, Germany) WoT '16*. Association for Computing Machinery, New York, NY, USA, 25–30.
- 5 Roy Thomas Fielding. 2000. Architectural styles and the design of network-based software architectures. Ph. D. Dissertation. University of California, Irvine.
- 6 Roy T. Fielding, Richard N. Taylor, Justin R. Erenkrantz, Michael M. Gorlick, Jim Whitehead, Rohit Khare, and Peyman Oreizy. 2017. Reflections on the REST Architectural Style and “Principled Design of the Modern Web Architecture” (Impact Paper Award).

⁷ <https://www.w3.org/DesignIssues/Abstractions.html>

- In Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering (Paderborn, Germany) ESEC/FSE 2017. Association for Computing Machinery, New York, NY, USA, 4–14.
- 7 R. Khare, and R.N. Taylor. 2004. Extending the Representational State Transfer (REST) architectural style for decentralized systems. In Proceedings of the 26th International Conference on Software Engineering. 428–437.
 - 8 Pattie Maes. 1993. Modeling Adaptive Autonomous Agents. *Artificial Life*, 1(1_2), 135–162.
 - 9 Alessandro Ricci, Michele Piunti, and Mirko Viroli. 2011. Environment programming in multi-agent systems: an artifact-based perspective. *Autonomous Agents and Multi-Agent Systems* 23(2), 158–192.
 - 10 Onn M Shehory. 1998. Architectural properties of multi-agent systems.
 - 11 Zach Shelby, Klaus Hartke, and Carsten Bormann. 2014. The Constrained Application Protocol (CoAP). RFC 7252 (Proposed Standard). <http://www.ietf.org/rfc/rfc7252.txt>.
 - 12 Danny Weyns, Andrea Omicini, and James Odell. 2007. Environment as a first class abstraction in multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 14(1), 5–30.
 - 13 Michael Wooldridge. 2000. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence (Second Edition)*, Gerhard Weiss (Ed.). MIT press.

3.6 Towards a Low Entry Barrier for Agents on the Web

Stephen Cranefield (University of Otago, NZ)

License  Creative Commons BY 4.0 International license
© Stephen Cranefield

Introduction

The Web was explicitly designed to allow sharing of information in an open, extensible and scalable way. When considering how agents might enhance the Web, or vice versa, it is useful to revisit Hewitt’s requirements for open systems [1]: **(1)** *continuous change and evolution*, **(2)** *arm’s-length relationships and decentralized decision making*, **(3)** *accommodation of perpetual inconsistency among knowledge bases*, **(4)** *the need for negotiation among system components*, and **(5)** *recognition of the inadequacy of the closed-world assumption*. These challenges have long been tackled by MAS research, which raises the question of what new challenges must be addressed to make agents on the Web a viable technology.

Fielding [2, Ch. 4] presents and motivates the following requirements of the Web architecture: **(a)** *a low entry barrier*, **(b)** *extensibility*, **(c)** *a distributed hypermedia model* (in which application control information [is] embedded within, or as a layer above, the presentation of information”), and **(d)** *the ability to work at internet scale*, implying a need for **(i)** *anarchic scalability* (essentially graceful handling of unanticipated situations) and **(ii)** *independent deployment* of architectural elements in a partial, iterative fashion”. In the MAS context, the dynamic application control that emerges from the proactivity and autonomy of agents can be seen as a natural generalisation of the traditional Web’s distributed hypermedia model.

I believe that requirements **b**, **c** (in its generalised form) and **d(ii)** have been well addressed in MAS research (outside the specific context of the Web), but requirements **a** (low entry barrier) and, to a lesser extent, **d(i)** have lacked significant attention by MAS researchers.

What does a low entry barrier look like for agents on the Web?

It is my observation that over time, much MAS research has shifted away from a peer-to-peer model of agent interaction, in which coordination mechanisms and social intelligence are located within the agents, towards a combination of agent-level and *system-level* design, where the desired properties of an MAS are supported by specific system architectures, services and protocols. While this is entirely appropriate for specific closed application areas (e.g. smart traffic control), I am sceptical that this approach will provide a low entry barrier to making agents a viable and accepted technology (in general) on the Web. Furthermore, we need to make it easier for non-specialist programmers to create and deploy agents on the Web, and limit the requirement to learn specialist agent programming languages and middleware. To this end, I make the following recommendations:


- Provide convenient interfaces between the agent model and conventional programming languages, information models and coordination tools (e.g. [3, 4]).
- Avoid creating MAS-specific middleware and instead use mainstream or at least domain-independent alternatives (e.g. [5, 6]).
- If really necessary, make any newly proposed “Web agent” standards as generic and simple as possible.
- Where possible, locate social intelligence within agents without reliance on “the system”, but ...
- provide individual agent reasoning components as application- and domain-independent libraries or services (e.g. [7]).

References

- 1 Carl Hewitt. The challenge of open systems. In Derek Partridge and Yorick Wilks, editors, *The Foundations of Artificial Intelligence*, pages 383–395. Cambridge University Press, 1990. doi: 10.1017/CBO9780511663116.038.
- 2 Roy Thomas Fielding. *Designing the Web Architecture: Problems and Insights*. PhD thesis, University of California, Irvine, 2000. URL https://www.ics.uci.edu/~fielding/pubs/dissertation/web_arch_domain.htm.
- 3 Stephen Cranefield and Martin K. Purvis. UML as an ontology modelling language. In *Proceedings of the IJCAI-99 Workshop on Intelligent Information Integration, Sixteenth International Joint Conference on Artificial Intelligence*, volume 23 of *CEUR Workshop Proceedings*. CEUR-WS.org, 1999. URL <http://ceur-ws.org/Vol-23/cranefield-ijcai99-iii.pdf>.
- 4 Stephen Cranefield and Surangika Ranathunga. Embedding agents in business processes using enterprise integration patterns. In Massimo Cossentino, Amal El Fallah Seghrouchni, and Michael Winikoff, editors, *Engineering Multi-Agent Systems*, volume 8245 of *Lecture Notes in Computer Science*, pages 97–116. Springer, 2013.
- 5 Stephen Cranefield. Reliable group communication and institutional action in a multi-agent trading scenario. In Frank Dignum, Rogier M. van Eijk, and Roberto A. Flores, editors, *Agent Communication II*, volume 3859 of *Lecture Notes in Computer Science*, pages 258–272. Springer, 2006.
- 6 Stephen Cranefield. Enabling BDI group plans with coordination middleware: semantics and implementation. *Autonomous Agents and Multi Agent Systems*, 35(2):45, 2021. doi: 10.1007/s10458-021-09525-7.
- 7 Surangika Ranathunga, Stephen Cranefield, and Martin K. Purvis. Integrating expectation monitoring into BDI agents. In Louise A. Dennis, Olivier Boissier, and Rafael H. Bordini, editors, *Programming Multi-Agent Systems – 9th International Workshop, ProMAS 2011*, volume 7217 of *Lecture Notes in Computer Science*, pages 74–91. Springer, 2011.

3.7 Society = Autonomy + Adaptation

Jérôme Euzenat (INRIA Grenoble Rhône-Alpes, Saint Ismier, FR & Univ. Grenoble-Alpes, FR)

License  Creative Commons BY 4.0 International license
 © Jérôme Euzenat

Main reference Jérôme Euzenat: “A map without a legend”, *Semantic Web*, Vol. 11(1), pp. 63–68, 2020.

URL <http://dx.doi.org/10.3233/SW-190376>

What makes a true lively society is the capability of their members to autonomously adapt to others. It is not a set of norms cast in iron, be they programming norms or ‘legal norms’. It is a set of beings trying to behave with others.

This behaviour may lead to explicit norms that make explicit what should not have/need to be reinvented, but they may well remain implicit, hence continuously adapted.

We should design software agents so that they are able to elaborate what drives their (social but not only) behaviours. They should be allowed to try, to make mistakes, and to transmit what they know. This is the ground on which evolution may happen. This capacity is what should be built in agents in order for them to behave without breaking too many things.

The goal is not to reach a static equilibrium: in an open-ended agent space there are always opportunities to learn new things, meet new people and visit new places. Hence, rather than the state reached by agents, this is their ability to surf a dynamic disequilibrium that must be sought.

This statement is somewhat made for triggering reactions within the seminar. It reacts to the apparent loss of autonomy of agents. It also extends the one I did for the previous seminar.

3.8 Multi-Agent Systems on the Web as a Special Kind of Knowledge Graphs Ecosystems

Catherine Faron (Université Côte d’Azur – Sophia Antipolis, FR)

License  Creative Commons BY 4.0 International license
 © Catherine Faron

I am interested in the management of knowledge graphs ecosystems on the Web, i.e. the modelling and exploitation of the relationships between not only resources within knowledge graphs, but also the particular resources that are knowledge graphs themselves. By nature, a knowledge graph represents and organises within it descriptions of resources of all types: factual knowledge (from data), procedural knowledge (rules, business processes), domain knowledge (schemas, thesauri, ontologies). At the same time, these knowledge graphs are themselves resources, which have links between each other and need to be described for their management and exploitation. The same models and languages can be used for intra- and inter-graph knowledge management.

My assumption is that Web agents and their interactions within Multi-Agents Systems can be represented and managed as a special kind of knowledge graphs ecosystems, relying on the Solid Protocol specification.

References

- 1 Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel. IndeGx: A Model and a Framework for Indexing RDF Knowledge Graphs with SPARQL-based Test Suits. *Journal of Web Semantics*, 2023, <https://hal.inria.fr/hal-03946680>
- 2 Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel. KartoGraphI: Drawing a Map of Linked Data. *ESWC 2022 – 19th European Semantic Web Conferences*, May 2022, Hersonissos, Greece. Springer. <https://hal.inria.fr/hal-03652865>

3.9 Governing Communities of Autonomous Agents and People on the Web Using Social Norms

Nicoletta Fornara (University of Lugano, CH)

License  Creative Commons BY 4.0 International license
© Nicoletta Fornara

Social norms and policies are fundamental for governing communities of autonomous agents and people on the Web. Various research communities have developed various models designed to formalise different types of norms and policies and to offer various types of automatic reasoning services on these norms. The MAS community has been focused on proposing models for the formal specification of social norms and contracts able to express different types of normative concepts, e.g. obligations, prohibitions, and permissions and on the definition of frameworks for reasoning on norms. An important connection between the MAS and the Semantic Web communities is represented by those norm models that use Semantic Web technologies for defining some components of the norms and for defining their operational semantics, like the OWL-POLAR framework [4] and the T-Norm model [1, 2]. Another policy language based on semantic web technologies is ODRL 2.2⁸ which is a W3C Recommendation since 15 February 2018. It is a policy expression language that can be used to represent permitted, prohibited, and obliged actions over a certain asset. ODRL 2.2 has no formal semantics and I am currently cooperating with other members of the ODRL Community Group for proposing a Formal Semantics⁹ for the language.

All these studies can be a starting point for the engineering and development of mechanisms for governing the interactions of autonomous agents and people on the Web.

It is probably unthinkable that we would reach an agreement on a common model for specifying norms and policies for all web applications, perhaps it is more reasonable to take the path of studying which norms can be formalised with which languages and how to translate norms from one formalisation to another. An initial study in this direction is [3].

It is also difficult to imagine that people not specifically trained could formalise the norms and policies governing the use of their data or of their interactions, so methods have to be devised that can assist such formalisation or vice versa that can describe in natural language the content of norms written in a formal language so that human beings can understand them.¹⁰

⁸ <https://www.w3.org/TR/odrl-model/>

⁹ <https://w3c.github.io/odrl/formal-semantics/>

¹⁰ Funded by the SNSF (Swiss National Science Foundation) grant no. 200021_175759/1.

References

- 1 N. Fornara, S. Roshankish, and M. Colombetti. A framework for automatic monitoring of norms that regulate time constrained actions. In A. Theodorou, J. C. Nieves, and M. D. Vos, editors, *Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems XIV – International Workshop, COINE 2021, London, UK, May 3, 2021, Revised Selected Papers*, volume 13239 of *Lecture Notes in Computer Science*, pages 9–27. Springer, 2021.
- 2 N. Fornara and M. Sterpetti. An architecture for monitoring norms that combines OWL reasoning and forward chaining over rules. In E. M. Sanfilippo et al., editors, *Proceedings of the Joint Ontology Workshops 2021 Episode VII: The Bolzano Summer of Knowledge co-located with the 12th International Conference on Formal Ontology in Information Systems (FOIS 2021), and the 12th International Conference on Biomedical Ontologies (ICBO 2021), Bolzano, Italy, September 11-18, 2021*, volume 2969 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2021.
- 3 S. Roshankish and N. Fornara. A methodology for formalizing different types of norms. In D. Baumeister and J. Rothe, editors, *Multi-Agent Systems*, pages 348–363, Cham, 2022. Springer International Publishing.
- 4 M. Sensoy, T. J. Norman, W. W. Vasconcelos, and K. P. Sycara. OWL-POLAR: A framework for semantic policy representation and reasoning. *Journal of Web Semantics*, 12:148–160, 2012.

3.10 Finding the Critical MAS(S): Resources and Representations Needed for Weaving a Web that Hosts Linked Multi-Agent Systems

Fabien Gandon (INRIA – Sophia Antipolis, FR)

License  Creative Commons BY 4.0 International license
© Fabien Gandon

A small but viral first hMAS. Linked Multi-Agent Systems (MAS) is a vision where MAS are linked through their resources and representations on the Web to allow us to turn the Web into an architecture supporting a network of interoperable MAS. One of the hardest tasks for Tim Berners-Lee in the early 90s was to make people imagine a world with a fully deploy Web and that we have the same cold-start problem with the Agents on the Web [1]. I still believe this to be a strategic question in 2023. In Hypermedia Multi-Agent Systems (hMAS), agents operate on an homogeneous hypermedia fabric that interconnects heterogeneous resources. The Web is the distributed hypermedia that has become the primary software architecture for applications on the Internet but not for Multi-Agent Systems, at least not yet. We are getting closer but we are not there and, in my opinion, a key step is to identify at least one use case that, not only could demonstrate hMAS but, more importantly, that has the potential to reach a critical mass of usage, the tipping point of a network effect like it was the case in the past successes of the Web [2].

The right metadata soil for growing hMAS. A lot of effort has been invested in knowledge acquisition and knowledge publication. We participated to that effort, with methods to index (IndeGx) [4], visualize (Kartographi) and annotate (Metadatamatic) linked data(sets)[5]. But for a first viral hMAS to happen we need the right breeding ground, we need to target metadata that have an impact on targeted adopters, largest users' community, etc. Moreover, with native mechanisms such as conneg (content negotiation over HTTP) the Web can do

much more and support adaptative knowledge exchange of customized representations for human and software agents. The Web has the potential to support profile-based knowledge negotiation for AI methods to obtain or contribute the type of knowledge they can process. Such an open negotiation could also be a key enabler for forward compatibility in open hMAS.

The position of hMAS in the effort of (re)decentralization. Even more than AI, it is distributed AI that has a rendezvous with the Web [3]. By nature, the distributed artificial intelligence paradigm of MAS can participate to the (re-)decentralization of applications and their architectures in general, and on the Web in particular with hMAS. This requires positioning hMAS w.r.t. other initiatives like SOLID. Decentralization in general, has to consider many fronts (architecture, applications, data, schemata) at the same time and hMAS, in particular, will have to consider all of them too.

Extend and make easier the existing solutions. hMAS must be conceived as an extension of the Web and not as another Web or as a Web apart. The futures of the Web must be suited both for “software- and human-agents” [6] and it may reopen the discussion on the relation between humans and agents in MAS both conceptually and practically. This may also echo discussions on other current trends in computer science such as the topic of digital twins (for artefacts, for users, etc.). Finally I would like to conclude on a “meta-risk”: In the history of computer science, we have many examples of a language, an architecture or a tool being reinvented because the previous ones had become too complex for newcomers, for simple use cases, etc. In what we will propose we must thrive to avoid the Déjà Vu of a technology stack that becomes too complex with an adoption cost too high and making the bed for the temptation to create something initially seemingly simpler on the side but that will end up being the first layer of a new technology stack continuing the cycle of reinventing languages, architectures, formats, etc. again and again.

References

- 1 Fabien Gandon. Merry hMAS and Happy New Web: A Wish for Standardizing an AI-Friendly Web Architecture for Hypermedia Multi-Agent Systems. Dagstuhl-Seminar 21072 : Autonomous Agents on the Web, Dagstuhl Schloss, Feb 2021, Dagstuhl, Germany. pp.3. <https://hal.inria.fr/hal-03960471>
- 2 Fabien Gandon, Wendy Hall. A Never-Ending Project for Humanity Called “the Web”. WWW 2022 – ACM Web Conference, Apr 2022, Lyon, France. <https://hal.inria.fr/hal-03633526>
- 3 Fabien Gandon. Distributed Artificial Intelligence and Knowledge Management: Ontologies And Multi-Agent Systems For A Corporate Semantic Web. PhD Thesis, INRIA, Université Nice Sophia Antipolis, November 2002. <https://tel.archives-ouvertes.fr/tel-00378201>
- 4 Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel. IndeGx: A Model and a Framework for Indexing RDF Knowledge Graphs with SPARQL-based Test Suits. Journal of Web Semantics, 2023, <https://hal.inria.fr/hal-03946680>
- 5 Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel. KartoGraphI: Drawing a Map of Linked Data. ESWC 2022 – 19th European Semantic Web Conferences, May 2022, Hersonissos, Greece. Springer. <https://hal.inria.fr/hal-03652865>
- 6 Bettina Berendt, Fabien Gandon, Susan Halford, Wendy Hall, Jim Hendler, Katharina E Kinder-Kurlanda, Eirini Ntoutsi, Steffen Staab. Web Futures: Inclusive, Intelligent, Sustainable The 2020 Manifesto for Web Science. Dagstuhl Manifestos, 2021. <https://hal.inria.fr/hal-03189474>

3.11 About the Place of Agents in the Web

Jomi Fred Hübner (Federal University of Santa Catarina, BR)

License  Creative Commons BY 4.0 International license
© Jomi Fred Hübner

In an integration perspective, we can face two problems when considering agents and the web:

1. how agents can use web resources and
2. how web applications can use agents.

While the former problem is addressed by some proposals, the latter case is mostly ignored. Of course, these problems can not be addressed considering the web simply as a kind of transport layer. A proper integration requires a better conceptualization. These problems are related to the third research question of the seminar: “How to design, represent, and reason about inter-actions among autonomous agents, people, and any other resources on the Web?”

An important aspect regarding the second problem is how to present agents to the web keeping their main property: *autonomy*. Moreover, it would be interesting to keep agents as cognitive entities and interact with them such as. This problem brings thus some initial questions:

- Is an agent web interface (like REST API) enough?
- Is it possible to include an agent in the web presenting it by means of a Thing Description?
- What kind of ‘resource’ is an autonomous agent?
- Can we conceive and develop tools (on the agent side) that can help this integration?

Ideally, we can imagine a scenario where an ordinary web developer is able to reuse available agents on the web while developing an application. We may require that he/she knows what an agent is. However, we should not require that he/she transform his/her application into an agent.

My preference to approach this problem is from an engineering perspective and based on application scenarios.

My assumption is that agents will be useful for the web only if they can be easily used by the web. Moreover, I would avoid a reductionist approach (e.g., transforming agents into services).

3.12 Towards an Analysis-oriented Perspective on Agent-oriented Abstractions for the Web

Timotheus Kampik (Umeå University, SE & SAP Signavio – Berlin, DE)

License  Creative Commons BY 4.0 International license
© Timotheus Kampik

When studying the engineering of multi-agent systems, the notion of an agent is typically seen as a *design* abstraction, *i.e.*, as a tool that directly facilitates implementation. However, one may argue that in the context of complex, large-scale socio-technical systems such as the Web, the suitability of an abstraction as a facilitator of run-time *analysis* is just as or even more important than design-time advantages: after all, it is only during run-time that a software artifact is fully exposed to the technical and social complexities of the larger system, often in unanticipated ways.

Generally, this idea is not new and has, indeed, informed fundamental research on the foundations of knowledge representation and reasoning that multi-agent systems rest upon. For example, in a seminal 1985 paper, the principle of cautious monotonicity in expert systems was motivated by the need to determine whether a system whose domain-specific purpose is unknown or has been obscured still behaves “logically” (in the colloquial sense) [4]. From an application-oriented perspective, recent research results have provided first evidence that the notion of an agent can potentially aid the discovery of symbolic models from system logs, and in particular facilitate the mining of Petri nets that represent behavioral models of agents’ actions and interactions [6, 7], following the established research direction of *agent mining* [3].

However, the design abstractions that the academic multi-agent systems community advocates for appear to be primarily based on philosophical concepts of reasoning, such as the well-known belief-desire-intention model [1, 2], and not on real-world requirements for facilitating run-time interpretability of behaviors in socio-technical systems. Consequently, we consider the following research question as particularly promising in the context of “agents on the Web” research:

How can agent-oriented abstractions facilitate the run-time analysis of complex socio-technical systems on the Web?


For example, the notions of *beliefs* and *belief revision* [5] can potentially be used to explain misaligned behaviors of several agents (caused by inconsistent beliefs), as well as apparently inconsistent behavior by one particular agent (caused by belief change) over time. Yet, beliefs are rarely used as abstractions in software engineering. Although much has been written about the mathematical foundations of belief revision, no comprehensive body of works that puts beliefs into the context of “mainstream”, large-scale software engineering seems to exist. This apparent gap in the literature may warrant further investigation.

References

- 1 Sameera Abar, Georgios K. Theodoropoulos, Pierre Lemarinier, and Gregory M.P. O’Hare. Agent based modelling and simulation tools: A review of the state-of-art software. *Computer Science Review*, 24:13–33, 2017.
- 2 Michael Bratman. *Intention, Plans, and Practical Reason*. Cambridge: Cambridge, MA: Harvard University Press, 1987.
- 3 Longbing Cao, Gerhard Weiss, and Philip Yu. A brief introduction to agent mining. *Autonomous Agents and Multi-Agent Systems*, 25(3):419–424, 2012.
- 4 D. M. Gabbay. Theoretical foundations for non-monotonic reasoning in expert systems. In Krzysztof R. Apt, editor, *Logics and Models of Concurrent Systems*, pages 439–457, Berlin, Heidelberg, 1985. Springer Berlin Heidelberg.
- 5 Peter Gärdenfors and Hans Rott. Belief revision. In *Handbook of logic in artificial intelligence and logic programming (Vol. 4) epistemic and temporal reasoning*, pages 35–132. Oxford University Press, 1995.
- 6 Andrei Tour, Artem Polyvyanyy, and Anna Kalenkova. Agent system mining: Vision, benefits, and challenges. *IEEE Access*, 9:99480–99494, 2021.
- 7 Andrei Tour, Artem Polyvyanyy, Anna Kalenkova, and Arik Senderovich. Agent miner: An algorithm for discovering agent systems from event data (pre-print), 2022.

3.13 Agent-based Information Systems Built on Read-Write Linked Data

Tobias Käfer (KIT – Karlsruhe Institut für Technologie, DE)

License  Creative Commons BY 4.0 International license
© Tobias Käfer

The web architecture has scaled to a global information exchange infrastructure, but in contrast to other architectures, it comes with peculiar constraints, summarised in the REST architectural style [8]. Augmented with knowledge representation using semantic technologies, thus forming Read-Write Linked Data [1], the web presents us with a substrate for integration on the component interaction and data level, i. e. a substrate for interoperability. As such, Read-Write Linked Data is built on standards that are simple yet powerful, such as HTTP and RDF [7, 5]. The underpinnings, their simplicity and formal properties allow for layering approaches for behaviour on top, such as [10, 11].

As the scalability and the adoption of the web has been attributed to its basic standards [8], I would argue that if we want to bring agents to the web, the community should only cautiously extend the underpinnings by inventing extensions and alternatives to the standards for data transmission (HTTP) and data description (RDF), and rather try to embrace the power of what exists and built agent-based systems on top:

Alternatives to HTTP data transmission include SPARQL over HTTP [6], SPARQL with web preemption [12], Linked Data Fragments [14], RDF streaming [13]; RDF is currently being extended to RDF-star [9]. Instead, Linked Data Notifications [3] work in a RESTful manner avoiding the need for streaming for notifications about updates, and the proposal of the data interoperability panel [2] allows for describing subdivisions of data that are accessible in a RESTful manner, instead of querying for triple fragments.

With a cleanly defined substrate, and I prefer Read-Write Linked Data, let us go and connect techniques from agent systems and multi-agent systems intelligently with the web architecture [4].

References

- 1 T. Berners-Lee. *Read-Write Linked Data*. Design Issues. 2009. URL: <http://www.w3.org/DesignIssues/ReadWriteLinkedData.html>.
- 2 J. Bingham, E. Prud'hommeaux and elf Pavlik. *Solid Application Interoperability*. Editor's Draft. 2023. URL: <https://solid.github.io/data-interoperability-panel/specification/>.
- 3 S. Capadisi and A. Guy. *Linked Data Notifications*. Recommendation. W3C. 2017. URL: <https://www.w3.org/TR/ldn/>.
- 4 A. Ciortea, S. Mayer, F. L. Gandon, O. Boissier, A. Ricci and A. Zimmermann. "A Decade in Hindsight: The Missing Bridge Between Multi-Agent Systems and the World Wide Web". In: *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. 2019.
- 5 R. Cyganiak, D. Wood and M. Lanthaler. *RDF 1.1 Concepts and Abstract Syntax*". Recommendation. W3C. 2014. URL: <https://www.w3.org/TR/rdf11-concepts/>.
- 6 L. Feigenbaum, G. Williams, K. Clark and E. Torres. *SPARQL 1.1 Protocol*. Recommendation. W3C. 2013. URL: <http://www.w3.org/TR/sparql11-protocol/>.
- 7 R. Fielding and J. Reschke. *Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing*". RFC 7230 (Proposed Standard). Internet Engineering Task Force, 2014. URL: <http://www.ietf.org/rfc/rfc7230.txt>.

- 8 R. Fielding. “Architectural Styles and the Design of Network-based Software Architectures”. PhD thesis. University of California, Irvine (CA), USA, 2000.
- 9 O. Hartig, P.-A. Champin, G. Kellogg and A. Seaborne. *RDF-star and SPARQL-star*. Community Group Report. W3C. 2021. URL: <https://w3c.github.io/rdf-star/cg-spec/2021-12-17.html>.
- 10 T. Käfer and A. Harth. “Specifying, Monitoring, and Executing Workflows in Linked Data Environments”. In: *Proceedings of the 17th International Semantic Web Conference (ISWC)*. 2018.
- 11 T. Käfer, B. Jochum, N. Aßfalg and L. Nürnberg. “Specifying and Executing User Agents in an Environment of Reasoning and RESTful Systems using the Guard-Stage-Milestone Approach”. In: *Journal on Data Semantics* 10.1-2 (2021).
- 12 T. Minier, H. Skaf-Molli and P. Molli. “SaGe: Web Preemption for Public –SPARQL” Query Services”. In: *Proceedings of the 28th World Wide Web Conference (WWW)*. ACM, 2019.
- 13 E. D. Valle, S. Ceri, F. van Harmelen and D. Fensel. “It’s a Streaming World! Reasoning upon Rapidly Changing Information”. In: *–IEEE” Intelligent Systems* 24.6 (2009).
- 14 R. Verborgh, M. V. Sande, P. Colpaert, S. Coppens, E. Mannens and R. V. de Walle. “Web-Scale Querying through Linked Data Fragments”. In: *Proceedings of the 7th International Workshop on Linked Data on the Web at the 23rd International World Wide Web Conference (WWW)*. 2014.

3.14 Old Agents, New Environments: Open Issues and Unexpected Opportunities

Stefano Mariani (University of Modena, IT)

License © Creative Commons BY 4.0 International license
© Stefano Mariani

Main reference Alessandro Ricci, Angelo Croatti, Stefano Mariani, Sara Montagna, Marco Picone: “Web of Digital Twins”, *ACM Trans. Internet Techn.*, Vol. 22(4), pp. 101:1–101:30, 2022.

URL <https://doi.org/10.1145/3507909>

Hypermedia provides for a new infrastructure and paradigm to design agents’ *virtual* environments, as well as (mediated) interactions. The Web of Things pushes this further to *physical* environments, through actuators encapsulated as hypermedia resources. Should the web bear responsibility for enabling, mediating, and controlling access to them? Probably not. *Digital Twins* could, in turn providing services as hypermedia resources. But can they cope with the web *openness* and *dynamism*? Can agents, while preserving *autonomy*? Probably yes, provided they can *learn*.

Acting through hypermedia

Quoting from the seminar “manifesto”:

“The latest standards allow clients not only to browse and query, but also to observe and act on this hypermedia fabric.”

In a Web of Things perspective what does it even mean for agents to *act* in a hypermedia environment? What can agents *expect* as action feedback and outcomes? Are there *implications* for physical resources? What are the *commitments* agents should hold to when providing hypermedia-accessible resources and services? What is the impact on agents’ architecture, if any?

Openness + dynamism = autonomy + _?

Given that:

“[...] the dynamic and open nature of these systems requires components to be deployed and to evolve independently from one another.”

“Autonomy is [...] agent’s ability to operate on its own, without the need of direct intervention of other people or other agents.”

“[...] hard-coding rules into agents [...] would be impractical in an environment as open and complex as the Web.”

what is the missing factor in the heading equation, if any?

Digital Twins, Agents’ other half

Digital Twins have already been framed within MAS as a novel abstraction to engineer virtual environments deeply bonded with the physical one. Hypermedia can fit in the picture by standardising the means by which such environment is distributed, accessed to, and operated on.

Learning to the rescue

Openness + dynamism = autonomy + *learning*. Do we agree that learning is the best way to deal with open and dynamic environments in *autonomous* way? If so, has hypermedia any value to add to agents’ learning experience? Hypermedia can be the *shared learning environment* where agents learn (i) the structure of the world, (ii) how do they capabilities match the environment affordances, (iii) how to achieve their own goals given so, (iv) whether other agents can help or interfere. Most importantly, agents can make the learnt knowledge, behaviours, and decision making policies available to others as hypermedia resources and services, thus communicating via the very same environment they live within, in a stigmergic way.

Conclusion

I hope that the role of mediation between agents, an hypermedia environment, and the physical world can be discussed and clarified, especially with respect to the Digital Twin abstraction. I also hope that learning, as the cornerstone form of adaptation, is not left out of the picture, as it would be a missed opportunity.

3.15 Pervasive Autonomous Systems: So Much to Learn from One Another

Simon Mayer (Universität St. Gallen, CH)

License  Creative Commons BY 4.0 International license
© Simon Mayer

I would like to share a particularly stimulating finding that could act as an example and also as a motivator in our attempt to find a common language to identify, discuss, and solve the issues on our path to Internet-scalable communities of autonomous agents and people who work together to seamlessly allocate concerns and reach their objectives. To support more

efficient agent-environment interaction, we recently proposed that the environment might contain signification that is *personalized* at run time to each agent that roams it [9, 5]. This was inspired by affordance theory [3] and computer scientists, philosophers, and psychologists who built on top of it (e.g., [7, 6, 1], and the basic idea is visible all around us: in (well-designed) signs, furniture, and, famously, door handles [6]. When designed well, signifiers become environmental cues that can be *intuitively and reliably* discovered and interpreted to provide guidance to agents who roam an environment about *what* are the possible behaviors and *how* these behaviors can be performed [9]. However, importantly, while in classical ecology signifiers are assumed to be run-time static and can only be designed a priori and with respect to agent *stereotypes* rather than actual agent features that are measured at run time, we may drop these constraints in virtual environments and, more and more, in *physical environments as well*. For this reason – and driven by advancements from material science to human-computer interaction – I expect that the findings of our Dagstuhl community for supporting autonomous agents have a high potential to be re-applied to human-computer interaction and, using HCI as a vehicle, to classical ecology. We expect that already in the near future, personalized content will be delivered *to humans* not only through Web browsers and other digital mediation (e.g., mobile apps) but through technologies that mediate individual or group experiences of *physical* reality as well. This delivery may happen through projected or head-worn Mixed Reality, but also through other sensory modalities such as audio [12], haptics [4], or vestibular stimulation [8]; objects in our physical environments with communication and processing abilities [10] might also alter experienced realities directly – e.g., in the context of self-balancing bicycles [11] where it has been shown that users prefer to experience artificially decreased tilting when turning – agent-personalized physics?

Regarding the cognitive dimension, some of these concerns are certainly best left to statistical approaches, in particular if abundant training data is available and can be processed relatively efficiently; others are to be allocated to symbolic systems, not only if more easily accessible explanations of the system’s behavior are desired; we should further support neurosymbolic combinations – to this end, I find the current developments in the field of semantic scene understanding (cf. [2]) particularly well-accessible and transferable; and finally, possibly, there will be tasks that humans are well-suited to solve, and that we also enjoy solving. Let us together work on an architecture that will permit the integration of such heterogeneous systems, including means for environments to support autonomous behavior in agent-agent and agent-environment interaction, means to design and govern communities of autonomous entities towards achieving organizational goals, and means to foster the adoption potential of our approaches into more real use cases.

Acknowledgments

Leading to this and other insights, I have very much enjoyed the discussions with members of my research team and beyond, where I’d like to name specifically Jérémy Lemée, Jannis Strecker, Danai Vachtsevanou, Ganesh Ramanathan, Samuele Burattini, Kenan Bektaş, Kimberly Garcia, and Andrei Ciortea.

References

- 1 A. Chemero and M. Turvey. Complexity, Hypersets, and the Ecological Perspective on Perception-Action. *Biological Theory*, 2(1):23–36, 2007.
- 2 Y. Cong, M. Y. Yang, and B. Rosenhahn. RelTR: Relation Transformer for Scene Graph Generation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–16, 2023.

- 3 J. J. Gibson. *The Ecological Approach to Visual Perception: Classic Edition*. Psychology press, 2014.
- 4 T. Grosshauser and T. Hermann. Augmented Haptics – An Interactive Feedback System for Musicians. In *Proceedings of the 4th International Conference on Haptic and Audio Interaction Design*, HAID '09, pages 100–108, Berlin, Heidelberg, Sept. 2009. Springer-Verlag.
- 5 J. Lemée, D. Vachtsevanou, S. Mayer, and A. Ciortea. Signifiers for Affordance-driven Multi-Agent Systems, 2022. Paper presented at the International Workshop on Engineering Multi-Agent Systems (EMAS) at the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS).
- 6 D. Norman. *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books, 2013.
- 7 D. A. Norman. The Way I See It: Signifiers, Not Affordances. *Interactions*, 2008.
- 8 M. Sra, A. Jain, and P. Maes. Adding Proprioceptive Feedback to Virtual Reality Experiences Using Galvanic Vestibular Stimulation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–14, New York, NY, USA, May 2019. ACM.
- 9 D. Vachtsevanou, A. Ciortea, S. Mayer, and J. Lemée. Signifiers as a First-class Abstraction in Hypermedia Multi-Agent Systems. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2023.
- 10 M. Weiser. The Computer for the 21st Century. *Scientific American*, 1991.
- 11 P. Wintersberger, A. Shahu, J. Reisinger, F. Alizadeh, and F. Michahelles. Self-Balancing Bicycles: Qualitative Assessment and Gaze Behavior Evaluation. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*, MUM '22, page 189–199, New York, NY, USA, 2022. Association for Computing Machinery.
- 12 J. Yang, A. Barde, and M. Billingham. Audio Augmented Reality: A Systematic Review of Technologies, Applications, and Future Research Directions. *Journal of the Audio Engineering Society*, 70(10):788–809, Nov. 2022.

3.16 Federated Learning in Goal-oriented WoT Environments

Mahda Noura (TU Chemnitz, DE)

License  Creative Commons BY 4.0 International license
© Mahda Noura

Joint work of Mahda Noura, Martin Gaedke

Goal-oriented user interaction is a promising new interaction paradigm for collaborative environments of heterogeneous distributed human, software, and hardware agents in the context of IoT. Unlike traditional programming-based approaches, which necessitates knowing and identifying available and potentially transient agents as well as explicitly giving them commands, goal-oriented approaches automatically identify a solution to fulfill a human user's goal [2, 4]. The goal can be expressed in a multimodal fashion through channels such as voice, visual programming, gestures or haptic interaction [5].

There are several categories of approaches used for determining an appropriate sequence of actions for the available agents to fulfill the user's goals such as *semantic reasoning*, symbolic AI *automated planners*, or *reinforcement learning*. Planners perform well when they have a complete understanding of how the world evolves as a result of various actions. However, the deployment of planning for dynamic WoT environments is algorithmically complex and can result in unacceptable response delays, especially when executed on energy-efficient edge

devices [3]. The Reinforcement Learning paradigm, on the other hand, does not require prior knowledge and instead involves exploration of the state-action space [1]. In WoT application scenarios actions of physical devices have an effect on the physical environments, therefore applying reinforcement learning is often impossible as this would result in unexplainable system behaviors via arbitrary device activations.

A novel idea would be to consider exploration not only in isolated multi-agent environments, but also in large-scale distributed multi-agent environments, allowing for exploration to be distributed and knowledge to be shared across them. This would reduce the negative impact on individual environments while also allowing for the optimization of complex decision making via large-scale distributed agent collaboration. Furthermore, due to the decentralized and distributed nature of the learning process, it would help to preserve privacy and data security, especially important in applications involving sensitive data.

Exploration from environments that are different could cause the inferred decisions to be misjudged. However, we argue that for environments like smart homes which are characterized by many different homes providing a large amount of learning data, a limited number of device types, and a limited number of device instances. Due to these characteristics, the chances of goals overlapping are higher, enabling autonomous learning with heuristics. Some research challenges that need to be considered are data aggregation on edge devices, variation of the environments, semantic interoperability, noisy and incomplete data, and data anonymization.

An idea worth discussion in the seminar would be identifying dimensions of the suitability of the different methods across three targeted application areas.

References

- 1 Jiménez, S., De La Rosa, T., Fernández, S., Fernández, F., Borrajo, D. A review of machine learning for automated planning. *The Knowledge Engineering Review* 27(4), 433–467 (2012).
- 2 Mayer, S., Verborgh, R., Kovatsch, M., Mattern, F. Smart configuration of smart environments. *IEEE Transactions on Automation Science and Engineering* 13(3), 1247–1255 (2016).
- 3 Noura, M., Gaedke, M. An automated cyclic planning framework based on plan-do-check-act for web of things composition. In: *Proceedings of the 10th ACM Conference on Web Science*. pp. 205–214 (2019).
- 4 Noura, M., Heil, S., Gaedke, M. Growth: Goal-oriented end user development for web of things devices. In: *International Conference on Web Engineering*. pp. 358–365. Springer (2018).
- 5 Noura, M., Heil, S., Gaedke, M. Natural language goal understanding for smart home environments. In: *Proceedings of the 10th International Conference on the Internet of Things*. pp. 1–8 (2020).

3.17 Integrating Multi-Agent Systems within Web-based Microservices Architectures – Multi-Agent MicroServices (MAMS)

Eoin O’Neill (University College Dublin, IE)

License © Creative Commons BY 4.0 International license
© Eoin O’Neill

In [2], Hubner discussed the need for not only agents interacting with services, but also the ability for services to be able to interact with agents. This is the vision of the Multi-Agent MicroServices [4] architecture. Although interaction between agents and Web services is not

a new concept, with [3] being an example of such a system, the key distinction is that the agents in those systems are external to their environment and are not directly accessible entities. Our goal with the MAMS architecture was to introduce the idea of immersing agents within the environment of the Web and allowing them to expose individual aspects of their agent “body” as directly queryable *virtual resources*. By integrating Multi-Agent Systems (MAS) with Microservices (MS) the MS community benefits from the plethora of research that has been done on fully encapsulated entities interacting and how to handle the interactions between individual, fully-encompassed entities, while the MAS community benefits by enabling agents to utilise the ecology of tools built to support the MS community.

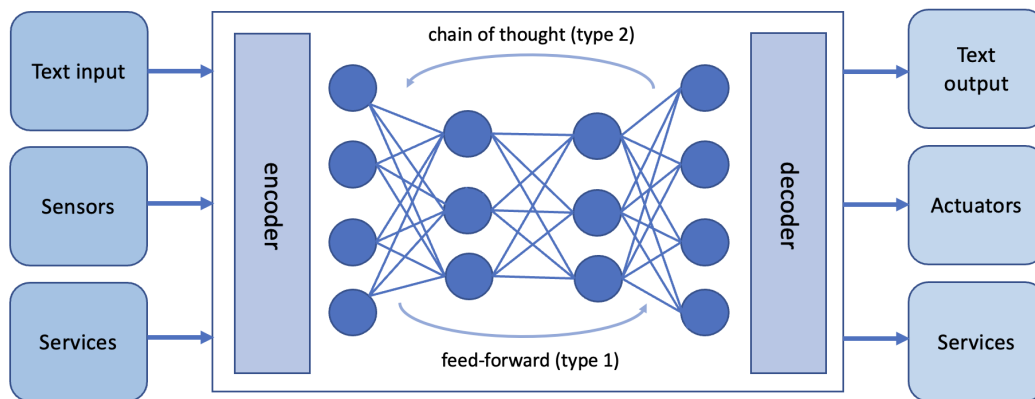
One of the key inspirations for this vision is the HATEOAS principle set forth by Roy Fielding in his description of the REST architectural paradigm [1]. In order to adhere to this principle, each *virtual resource* that the agent exposes, must have a hypermedia representation with actionable links available in order to facilitate state transitions of the resource. By providing a hypermedia body for each resource of the agent, we are embodying the agent within the Web environment, making it identifiable to the other entities that inhabit it.

Each *virtual resource*, an example of which could be to display a history of interactions with other entities as a log in order to represent a public persona to other entities in the environment, or an inbox resource that entities can use to communicate with the agent. In order to immerse agents within a Web-based environment, a hypermedia representation is necessary. The current industry standard in microservices architectures for providing a hypermedia representation of the possible actions for a microservice is the OpenAPI specification. However, by using Linked Data representations, such as the Hydra specification [5], service descriptions can provide a level of detail and ontological context that you cannot achieve by strictly using data representations such as XML or JSON. The Hydra specification allows users to define the inputs and outputs of a resource, the HTTP operations that are accepted in its current state.

The importance of this level of detail is apparent when you start to envision components built as microservices, agent enabled or not, with a description of what it can provide that is consumable by other entities within the Web environment. This can allow interactions with Web-based microservices to occur in an ad-hoc fashion, with all the necessary information for interaction being available to construct requests being directly available in the hypermedia. Additionally, by utilising Linked Data compliant hypermedia representations, the MAS community can make use of the abundance of tools developed for the Semantic Web community to store and manipulate Linked Data. By enabling autonomous agents with the ability to explore resources, reason about their capabilities, see how they line up with their goals and then utilise them if necessary all through the use of hypermedia, we are allowing machines to utilise the Web in the same manner as humans. This is a step towards the serendipitous use of the Web from the perspective of machines.

References

- 1 Fielding, R.T.: Architectural styles and the design of network-based software architectures. University of California, Irvine (2000)
- 2 Hubner, J.: About the place of agents in the web. vol. Web p. 44
- 3 Minotti, M., Santi, A., Ricci, A.: Developing web client applications with jaca-web. In: WOA (2010)
- 4 W. Collier, R., O’Neill, E., Lillis, D., O’Hare, G.: Mams: Multi-agent microservices. In: Companion Proceedings of The 2019 World Wide Web Conference. pp. 655–662 (2019)



■ **Figure 2** architecture for neurosymbolic cognitive agents.

3.18 Human-like AI for Artificial General Intelligence

Dave Raggett (W3C – United Kingdom, GB)

License © Creative Commons BY 4.0 International license
© Dave Raggett

According to Arthur C. Clarke, any sufficiently advanced technology is indistinguishable from magic, and this increasingly applies to the remarkable advances in the field of artificial intelligence, in particular, to large language models and image generators. There are lots of opportunities for the Intelligent Web of Things as agents that can communicate, learn and reason like we do, supporting human-machine collaboration to boost productivity and standards of living.

Symbolic AI is falling behind relative to neural networks, but both have strengths and weaknesses. Symbolic AI is hard to scale up due to a dependency on expensive hand-crafted knowledge. Neural networks, by contrast, can be easily scaled using larger datasets. Neurosymbolic AI combines the strengths of both approaches, yielding better explainability and enabling semantic interoperability for inter-system communication. There are plenty of opportunities for scaling across the computing continuum from the network edge to the cloud.

Stanovich’s tripartite model of mind distinguishes the autonomous mind, the algorithmic mind and the reflective mind. The first is associated with type 1 cognition, which is fast, automatic and opaque. The second and third use type 2 cognition, which is slow, deliberative and open to introspection. This is formed by linking together type 1 processes using working memory to form a chain of thought.

Neurosymbolic systems combine neural networks with symbolic approaches, including back-end information technology systems, sensors and actuators. A natural next step would be to integrate neural networks with cognitive databases that combine symbolic and sub-symbolic approaches.

The Plausible Knowledge Notation (PKN) is an example of how to provide richer semantics compared to W3C’s Resource Description Framework (RDF). PKN supports properties, relations and implications, combined with qualitative metadata and scopes; imprecise concepts and quantifiers, along with analogical reasoning.

Plausible reasoning is a form of argumentation for everyday knowledge, i.e. knowledge that is often uncertain, imprecise, incomplete and inconsistent. In place of logical proof, we have multiple lines of argument for and against the premise in question, just as in the courtroom.

The large language model ChatGPT is remarkably good at plausible reasoning, but can make human-like errors due to careless assumptions. Reinforcement learning with human feedback is proving effective in training large language model for specific tasks, e.g. to solve a range of math and physics problems.

Some outstanding research challenges include:

- Faster and smarter learning by mimicking human learning using a combination of type 1 and type 2 cognition.
- Taming catastrophic interference when learning new tasks.
- Earning trust through citing provenance and avoiding careless mistakes.
- Integrating episodic memory and continual learning through instruction, observation and experience.
- Reflective cognition and support for theory of mind for self and others.
- Acquisition and use of implicit, and explicit, behavioural norms.
- Shared cognitive databases for enabling hive minds where knowledge gained by one agent is immediately available to all other agents in the hive.
- Agent-Client confidentiality and policy-based data sharing.

Researchers should be encouraged by Jeremy Howard and Sylvain Gugger (FastAI), who say that breakthrough work in deep learning absolutely does not require access to vast resources, elite teams or advanced math!

Human-like AI will be hugely disruptive to web search, improved personal privacy and ecosystems of services. This includes personal agents that safeguard your privacy, help you with your health, financial affairs, education and so forth; policing fake news and malicious posts; agents communicating with other agents to find and provide services in open decentralised ecosystems; agents on the Web, in the Metaverse, as robots and embedded in other devices, including cars; in short, the intelligent Web of Things!

For more details, see: <https://www.w3.org/2023/02-Raggett-towards-AGI.pdf/>

3.19 Ontologies with Temporal Logics Allow Hypermedia Agents to Make Plans

Daniel Schraudner (Universität Erlangen-Nürnberg, DE)

License  Creative Commons BY 4.0 International license

© Daniel Schraudner

Joint work of Daniel Schraudner, Andreas Harth

Hypermedia agents are agents that are situated in a hypermedia environment, like e.g. the Web. These agents can interact with artifacts using HTTP methods on resources that are managed by Web servers and interconnected via hypermedia.

Possible interactions for agents are mediated by affordances which in the Web usually are links and forms. I.e. hypermedia agents can follow links to get information about artifacts and can submit forms in order to change the current state of an artifact.

In hypermedia graphs like the web the number of affordances for an agents usually gets quite large very fast. Agents can find out which links to follow in order to get certain information by utilizing ontologies. An ontology could e.g. have the following two axioms:

Coffee \sqsubset Beverage $\sqcap \exists$.madeOutOf.CoffeeBean

CoffeeBean $\sqsubset \exists$.hasSort.{arabica, robusta}

This allows an agent that knows the resource of a coffee to infer that it has to follow the links of types `madeOutOf` and `hasSort` in order to find out the bean sort of the coffee.

Manipulating the state of artifacts, however, is not that straight-forward. A coffee Web service e.g., will need an agent to place a valid order by changing the state of a resource through a form. Afterwards the artifacts will carry out actions that again change their state, e.g. by creating a finished coffee resource (reactivity of artifacts).

Getting a finished coffee resource is usually the goal of an agent placing a coffee order – however, hypermedia agents have no possibility to infer that a valid coffee order will eventually lead to a finished coffee resource by using an ontology.

To solve this problem, we suggest using Temporal Description Logics (ontologies enriched with Temporal Logics) in order to give agent information about how to manipulate the state of artifacts in order to get the wanted reaction of the artifacts. Temporal Description Logics offer the possibility to define temporal concepts, like e.g.:

$$\diamond\text{Coffee} \equiv \text{OrderedCoffee} \sqcap \exists\text{hasPayment}.\text{ValidPayment}$$


This axiom defines that a coffee order that has a valid payment will eventually lead to a finished Coffee resource (as per reaction of the artifact).

An agent that has the goal to get a coffee can now utilize the ontology to infer how it has to manipulate its environment using forms in order to reach its goal.

Having a method to reason about the consequences of their actions will allow hypermedia agents to combine multiple actions to make sophisticated plans in order to achieve complex goals (actually getting a coffee already might require a two-step plan of an agent as it might need to create a `OrderedCoffee` and a `ValidPayment` resource by using different forms).

3.20 Rethinking Agents and Meaning over the Web of Things

Munindar P. Singh (North Carolina State University – Raleigh, US)

License  Creative Commons BY 4.0 International license
© Munindar P. Singh

This position paper takes a fresh look at how we ought to understand autonomy and meaning on the web in relation to the needs of flexibility and governance and with respect to system architecture.

Positions

I take two positions that are well-aligned with the objectives of this Dagstuhl Seminar but nevertheless are departures from what the communities represented in this seminar have primarily pursued.

First, we ought to find new ways to unite the distinct concerns of autonomy, interaction, meaning, and sensing to produce well-governed systems of agents on the web. This position is that our conceptions of these elements must be rethought (and refactored and reconstructed) with awareness of the kinds of usage scenarios that we wish to support.

Second, we should dispense with the traditional conception of a layered system architecture. Black-box layering is dead, or ought to be.

Concepts

Many of the common approaches to multiagent systems (MAS) are conventional in their outlook: they adopt traditional programming abstractions for the most part albeit with a dose of AI, such as through logic programming (with facts interpreted as knowledge and inference rules as plans). This MAS-lite attitude is intended to facilitate adoption of MAS technologies by practitioners versed in traditional methods. The attitude has produced mixed success at best. We would be better served by emphasizing rather than hiding our strengths.

Autonomy

The common realization of autonomy in MAS allows interference from organizational constraints and from the infrastructure. Agents can be blocked from violating norms by the so-called governors in organizations. The infrastructure is set up to control the order in which an agent observes events even when there is no physical reason for such ordering (e.g., because the events have occurred and information about them has already arrived). The net result of focusing on and adopting traditional software assumptions is that autonomy, which would be the main contribution of the MAS field, is sacrificed. Once it is sacrificed, the traditional approaches are perfectly adequate: why would anyone include any MAS at all?

Further, the increasing realization of the importance to ethics in AI provides a challenge and an opportunity. Our current techniques don't facilitate support moral autonomy in its general Kantian sense [9, 11] nor do they correspond to human behavior (beliefs and desires are in general not precursors to behavior).

Meaning

This seminar's description and indeed much of computer science talks of semantics as the only kind of meaning. But classically, that is not so [5, 8]. *Semantics* is the aspect of the meaning of a representation that can be computed from the representation. But what matters more is *pragmatics*, which is the meaning of a representation that relies upon the context [8]. Even something as simple as understanding pronouns thus involves pragmatics.

Semantics is valuable when we can get it. Semantics gives us reusability and easier standardization but at the cost of flexibility. In particular, in multiagent systems on the web, a large part of the context is built by the agents themselves. How language is used and how its usage and meaning evolve fall in the realm of pragmatics.

For agents on the web, meaning would include practices that they establish (e.g., norms and conventions) or which emerge through their interactions. For example, if a bank processes deposits to your account each night before withdrawals, the outcome would be different than if it did so in the reverse order or in the order in which deposits and withdrawals arrived.

What we need are methods to induce semantics from the pragmatics, which would give us gains in transparency and interpretability. Such methods would apply continually to account for the evolution of a sociotechnical system.

Interaction

Conventionally, interaction is modeled from the perspective of its initiator: e.g., as the caller of a method or the sender of a (synchronous or asynchronous) message.

But what an interaction is depends on how it is construed. An agent may act morally, violate social norms, suffer a loss, inflict pain or pleasure by taking an action in a particular setting. Moreover, two or more agents may jointly invent a new language or introduce

or deprecate some meanings by their collaborative actions. Thus, how we understand interaction interplays with autonomy and meaning. Also, interaction can provide a basis for conceptualizing more elaborate agents out of ensembles of agents based on how they interact with effects on their abilities [6] and intentions [7], to mention two works from the early days of the field.

Architecture

A key guiding principle in system design is the end-to-end argument. If a functionality is potentially needed at one layer, it ought not be supported from a lower layer: either the functionality is superfluous and never used or incomplete and the upper layer needs to redo it.

Time and again, software library developers and even standards groups find themselves contravening the end-to-end argument. They give arguments such as the need for performance of some shape or form. I am reminded here of Don Knuth's dictum on premature optimization being the root of all evil.

But perhaps these seemingly unprincipled developers (and groups) are on to something. Taking as a given that the needs that motivate these developers are legitimate as is the end-to-end argument, we should move away from the notion of layers that are fixed across application environments. That is, if developers find it necessary to break the layer encapsulation, maybe that wasn't such a great abstraction to begin with.

We need architectures in which policies addressing different aspects of interaction can be composed. Let's consider retrying transmissions as an example. An interaction may (or may not) be retried upon failure depending on the pragmatics of the application and knowledge of the state of the infrastructure. For example, an action of a physician, patient, and pharmacy to refill a prescription may be retried by just the patient and pharmacy (if the physician's component had progressed sufficiently). Or, the action may be retried with the physician's assistant on the physician's behalf. Or, the retry may be dropped if the patient has met their need through another source. That is, high-level reasoning may be used to handle something as simple as retrying. This view contrasts with many communication protocols that hardcode this functionality, thereby violating the end-to-end argument.

Traditionally, the layers corresponded to different potential businesses. The physical layer is about wires and radios to get a signal across; the network layer is about routing packets; and so on. Thus the layers were standardized and different businesses could build products for them. That is a desirable feature. But we could potentially support greater transparency into each layer as well as greater control. For example, visibility into the nonfunctional properties of a lower layer may help an upper layer adapt better in itself and control the lower layer better.

I expect that these challenges are being addressed in the networking community at the lower layers but we need analogous and open solutions with respect to MAS and governance. Along these lines, we might consider new composition operators for protocols [4] that support additional details to be interposed where appropriate, going beyond existing approaches to contextualize a protocol [2].

Governance

I think of governance of a sociotechnical system as combining the above concerns. Traditional computer scientists in general have produced strong capabilities in the technical tier. The MAS community has investigated flexibility in the social tier through the study of interaction and autonomy and abstractions based on cognition, norms, and economic behavior. Other

communities talk of social meaning but do so in a largely informal way or with low-level abstractions and in the aggregate, such as in social networks. Thus current work suffers from limitations but provides directions to address the flexibility we desire.

But the most interesting part of a sociotechnical system is how the social and technical tiers interplay. Computer scientists of all stripes leave this interplay largely ad hoc with legendary fiascos in the case of technologies such as blockchain [1].

An essential direction therefore is to capture varieties of this interplay to be able to govern a sociotechnical system effectively, e.g., by encoding the social tier explicitly [3] and providing ways to specify expectations rigorously while enabling them to be violated (a crucial requirement of autonomy) [10].

Meta Motivation

A potential counterargument is that we have our hands full dealing with the concepts as they are. Suppose we were more ambitious? Wouldn't that be even harder?

My meta position is that it would be easier. For inspiration, I draw your attention to the *Inventor's Paradox*. This is a situation arising commonly enough in problem solving and was identified by George Pólya. The “paradox” simply is that the more general problem is easier to solve.

In essence, by generalizing a problem, we can bring forth structure in the problem that is otherwise invisible. That structure leads to a more natural solution becoming available.

The prevalent narrow conceptions of autonomy, interaction, semantics on the one hand, and rigid architectures of the other restrict outcomes.

Acknowledgments

I have benefited from discussions with Amit Chopra and Samuel Christie on related topics. Thanks also to the US National Science Foundation (grant IIS-1908374) for supporting this research.

References

- 1 Vitalik Buterin. Critical update re: DAO vulnerability, June 2016. URL: <https://blog.ethereum.org/2016/06/17/critical-update-re-dao-vulnerability/>.
- 2 Amit K. Chopra and Munindar P. Singh. Contextualizing commitment protocols. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 1345–1352, Hakodate, Japan, May 2006. ACM Press. doi:10.1145/1160633.1160884.
- 3 Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh. DESEN: Specification of sociotechnical systems via patterns of regulation and control. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 29(1):7:1–7:50, February 2020. doi:10.1145/3365664.
- 4 Ashok U. Mallya and Munindar P. Singh. An algebra for commitment protocols. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 14(2):143–163, April 2007. doi:10.1007/s10458-006-7232-1.
- 5 Charles Morris. *Foundations of the Theory of Signs*. University of Chicago Press, Chicago and London, 1938.
- 6 Munindar P. Singh. Group ability and structure. In Yves Demazeau and Jean-Pierre Müller, editors, *Decentralized Artificial Intelligence, Volume 2*, pages 127–145. Elsevier/North-Holland, Amsterdam, 1991. Revised proceedings of the *2nd European Workshop on Modeling*

- Autonomous Agents in a Multi Agent World (MAAMAW)*, St. Quentin en Yvelines, France, August 1990.
- 7 Munindar P. Singh. The intentions of teams: Team structure, endodeixis, and exodeixis. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI)*, pages 303–307, Brighton, United Kingdom, August 1998. John Wiley.
 - 8 Munindar P. Singh. The pragmatic web. *IEEE Internet Computing (IC)*, 6(3):4–5, May 2002. Instance of the column *Being Interactive*. doi:10.1109/MIC.2002.1003124.
 - 9 Munindar P. Singh. Consent as a foundation for responsible autonomy. *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, 36(11):12301–12306, February 2022. Blue Sky Track. doi:10.1609/aaai.v36i11.21494.
 - 10 Munindar P. Singh and Amit K. Chopra. Computational governance and violable contracts for blockchain applications. *IEEE Computer*, 53(1):53–62, January 2020. doi:10.1109/MC.2019.2947372.
 - 11 Roger J. Sullivan. *An Introduction to Kant's Ethics*. Cambridge University Press, Cambridge, UK, 1994.

3.21 Challenges for Query Agents on the Decentralized Web

Ruben Taelman (Ghent University, BE)

License © Creative Commons BY 4.0 International license
© Ruben Taelman

While the Web was originally envisioned as a decentralized information space, it has evolved to a place where the majority of data is flowing towards isolated *data silos*. Since these data silos are in the hands of large companies and organizations, this leads to issues related to privacy and vendor lock-in.

To counter these problems, decentralization initiatives such as Solid are gaining popularity, which allows users to store any kind of data in their own personal data vault, which they fully control. These data vaults are personal Knowledge Graphs that can be represented as collections of Linked Data documents and more expressive query APIs.

The presence of such data vaults results in a large-scale distribution of data, where applications involving multiple individuals' data require accessing thousands or even millions of APIs across different data vaults on the Web. These applications cannot effectively be built today due to the lack of querying techniques that can handle the requirements of decentralized environments like Solid.

A promising paradigm to tackle this query problem is *Link Traversal Query Processing (LTQP)*, which is based on the *follow-your-nose principle* of Linked Data where query execution is done over a continuously growing range of documents by autonomously following hyperlinks between documents.


In order to apply LTQP effectively to decentralized environments, several open challenges need to be tackled:

1. To enable clients to select APIs with the most suitable query capabilities, there is a need to **represent the capabilities of query APIs using self-descriptive hypermedia descriptions**. Current approaches have limited expressivity.
2. To enable clients to select and discover APIs matching their query needs, there is a need to **describe the contents of the data exposed through query APIs via hypermedia**, which could be cardinality-based, shape-based, or approximate summaries. Current approaches lack the discovery and privacy-preservation.

3. To enable efficient client-side query execution, there is a need for **algorithms that provide (adaptive) query planning and execution over the heterogeneity of query APIs** in terms of capabilities, contents, and client-side policies. Current techniques suffer from query termination problems caused by unselective link following techniques and inefficient query plans caused by static non-adapting query planning.

3.22 Signifying Affordances for Effective Interaction of Agents on the Web

Danai Vachtsevanou (Universität St. Gallen, CH)

License  Creative Commons BY 4.0 International license
© Danai Vachtsevanou

Having declarative specifications of interaction in Web-based Multi-Agent Systems (MAS) could support autonomous agents in discovering and reasoning about affordances¹¹ on the Web towards effectively achieving their goals [2]. One possible path forward is to consider *signifiers* as a first-class abstraction in Web-based MAS for modelling *cues and signs that reveal information about how to exploit affordances* offered to agents in hypermedia environments [8]. Signifiers enable agents to interact with shared artifacts and other resources in a hypermedia-driven manner, and recommend under which circumstances interaction should take place based on an agent's *abilities* and the agent-environment *context*. As a result, the exploitability and relevance of an affordance can be dynamically evaluated to guide interaction even in affordance-rich and open environments, for instance, based on the agent's ability to reason and act using specific methods and mechanisms, the agent's ability to handle abstractions and processes of a given domain, the agent's goals or the artifact's state.

Further, evaluating interaction relevance could be dynamically distributed among agents, and other entities in the environment that *manage the exposure of signifiers* [6, 8]. Therefore, interaction guidance would remain effective (despite the heterogeneity of components' context and features¹²) across the spectrum from Web-like environment-driven opportunism to complex agent-driven reasoning and planning about action: From the Web perspective, this offers a bridge between local and global hypermedia-driven guidance, e.g. by enabling the step-by-step provisioning of selected interaction specifications (e.g., of W3C Web of Things Interaction Affordances¹³) with respect to valid state transitions and agents' goals. For example, a BDI agent [3] that desires to pick and place an item, may perceive the signifier for moving the gripper of a robotic arm only if the gripper is empty and in range of the item. From the MAS perspective, signifiers provide a bridge between hypermedia-driven affordance exploitation and methods for reasoning about action as examined in MAS research, since hypermedia are defined in terms of abstractions relevant to autonomous agents (e.g. action preconditions and postconditions, agent goals and roles etc.). For example, an agent capable

¹¹ Affordances in this context denote interactions that become possible for agents to enact based on the run-time context and agents' abilities. The term is inspired from affordance theory [1]

¹² For example, signifier exposure could be adjusted with respect to artifacts that implement one native application logic, or the Thin Server architecture [4], and agents capable of simple reflex actions, or automated planning.

¹³ A W3C Web of Things Interaction Affordance is metadata that shows how agents can interact with physical devices modelled as Things [5].

of planning how to pick and place may perceive signifiers for most of the affordances of the robotic arm (including information about action preconditions and postconditions) towards enriching its planning domain.

Given that declarative interaction specifications aim to support *interaction effectiveness* in hypermedia environments (e.g. through the contextual run-time exposure of signifiers), it is interesting to investigate how agents could evaluate such effectiveness. To begin with, addressing the latter requires that interaction specifications capture the full complexity of agents' interactions, which should be realized and, thus, evaluated within an (partially) *observable stateful system* and within a *social context*. For example, agents should be able to reason about how they can *interact for perception*, i.e. with the purpose of affecting their percepts. Percepts are essential for supporting agents' goal deliberation and decision-making processes, including deciding when and how to act towards modifying their perceived environment, or where to focus for evaluating whether their actions succeeded and contributed to their goals. To this end, and given the large size and the dynamicity of the state space, interaction specifications should reveal information about how agents can interact for directing their perception scope, and how interactions for perception relate to interactions for modifying the environment state. Upon exposure of such information, agents would perceive the environment state, and, then, reason about how well their expectations and intentions have been met ¹⁴.

Similarly, agents should be able to discover and interpret how to interact towards perceiving and modifying the current *social state*, for instance, in the context of an organization or an interaction protocol. For example, based on [8], signifiers could be exposed based on agent abilities that derive from the roles that agents hold within an organization. However, agents still cannot reason about policies and norms that capture the behavioral and social expectations that relate to affordance exploitation on the Web, e.g. reason about what organizational goals to adopt. Upon discovering interaction specifications, agents should be able to interpret information for perceiving the social state, acting towards modifying the social state in a regulated environment, and finally, reasoning about how well the behavioral and social expectations have been met.

Overall, towards defining and evaluating interaction effectiveness in Web-based MAS, it is interesting to investigate the following:

- How to model and represent interaction specifications that reveal information about how to interact for perception, and how interaction for perception relates to interaction for modifying the environment state in Web-based MAS.
- How to model and represent interaction specifications that reveal information about how to interact within a social context in Web-based MAS, e.g. based on the norms that apply in the scope of an organization or an interaction protocol.
- How to define and evaluate interaction effectiveness with respect to the expected or desired changes on the environment and social state in Web-based MAS.

References


- 1 A. Chemero and M.T. Turvey. *Complexity, Hypersets, and the Ecological Perspective on Perception-Action*. *Biological Theory*, 2(1), 2007.
- 2 A. Ciortea, S. Mayer, O. Boissier and F. Gandon. *Exploiting Interaction Affordances: On Engineering Autonomous Systems for the Web of Things*. In 2nd W3C Workshop on the Web of Things The Open Web to Challenge IoT Fragmentation, 2019.

¹⁴Per Norman, determining how well expectations and intentions have been met upon affordance exploitation amounts to “crossing the Gulf of Evaluation” [7].

- 3 M.P. Georgeff and A.L. Lansky. *Reactive Reasoning and Planning*. In Proceedings of the 6th National Conference on Artificial Intelligence (AAAI), 1987.
- 4 M. Kovatsch, S. Mayer and B. Ostermaier. *Moving Application Logic from the Firmware to the cloud: Towards the Thin Server Architecture for the Internet of Things*. In 2012 6th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2012.
- 5 M. Kovatsch, M. McCool, S. Käbisch, V. Charpenay and T. Kamiya. *Web of Things (WoT) Thing Description. W3C Recommendation*. W3C, 2020. Available online: <https://www.w3.org/TR/2020/REC-wot-thing-description-20200409/>.
- 6 J. Lemée, D. Vachtsevanou, S. Mayer and A. Ciortea. *Signifiers for Affordance-driven Multi-Agent System*. Paper presented at the International Workshop on Engineering Multi-Agent Systems (EMAS) at the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2022.
- 7 D.A. Norman. *The Design of Everyday Things*. BasicBooks, 2013.
- 8 D. Vachtsevanou, A. Ciortea, S. Mayer and J. Lemée. *Signifiers as a First-class Abstraction in Hypermedia Multi-Agent Systems*. In Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2023.

3.23 Agents for the Decentralized Web of Data

Jan Van den Bussche (Hasselt University, BE)


License  Creative Commons BY 4.0 International license
© Jan Van den Bussche

On the decentralized Web, personal data is managed in so-called “pods”. The idea (as seen, e.g., in the Solid project) is that a pod is an online data store that can give access to resources to specific parties. To manage the various data sharing contracts that exist with many different parties, one could use a “Web agent”. Such an agent can not only keep track of which parties have which access to which resources, but can provide mediator services for finer-grained data sharing. The mediator can map the data we are willing to share with a party to a virtual or materialized RDF graph, directly in a format that is most useful for the party. Methods from the field of information integration, such as schema mappings and query rewriting, can be applied in this context. Yet, the adaptation of these methods to the RDF setting presents new challenges, such as reasoning about conformance of SPARQL queries to shape (e.g., SHACL) schemas.

A separate function that the Web agent can fulfill is to keep parties up-to-date about changes in the personal data. If the shared resource is virtual, this leads to interesting reasoning tasks about independence of queries from updates. These tasks have been investigated in the field of database systems, but much less in the SPARQL context.

3.24 Socio-Technical Networking Platforms for Multi-Agent Collaboration on the Web

Antoine Zimmermann (*Ecole des Mines – St. Étienne, FR*)

License  Creative Commons BY 4.0 International license
© Antoine Zimmermann

A lot of collaboration happening on the Web between people consists in exchanging written messages. In fact, for years, it was the only way people could interact via the Web. Now, since the Web has become a software platform through which we can build and propose any kind of applications, collaborations happen in many different ways, and people on the Web have an open-ended set of possibilities of interaction with the Web and with other people. The long-term goal of the Agents on the Web effort should be that artificial agents can deal with all these possibilities, with sufficient guidelines provided by explanatory resources like metadata, ontologies, etc. However, a framework that would allow all agents to exploit any and all possible interactive resources on the Web in their generality would be as difficult to design as a general AI.

I propose that we instead start the effort of bringing agents to the Web by devising *socio-technical networking platforms* (STN platforms)¹⁵ where agents can easily:

- retrieve specific written messages posted by other agents and isolate them from other content like navigation, titles, dates, authors, etc. (e.g. discussion threads, blog posts, comments, replies, forum messages, wiki discussions, Git issues, reviews, Q&As);
- distinguish different types of messages (e.g., distinguish an original post from a response to another message);
- register to the platform autonomously (possibly with validation by a human person that could be accountable for the agent's actions);
- post messages, either as original posts or as replies to others;
- connect with other agents on the platform, so as to filter which content matters vs what's ignored.

Devising a framework that would allow agents to autonomously make use of this limited set of operations, in a way that is generic and can be instantiated on any social networking platform, is attainable in a near future. As a next step, there should be regulation put in place with the same objective of genericity: devise a framework where agents can autonomously determine what permissions/prohibitions there are on the operations listed above, and have certain agents playing the roles of policy enforcement agents.

When all this will be in place, then we can consider more general modes of interactions on the Web, such as making use of arbitrary software artifact that functions on the Web, or use real time discussion tools.

¹⁵ STN platforms already exist, e.g., Wikipedia where humans and bots collaborate to either ensure better content or to help regulate the platform.

4 Overview of Demonstrators

4.1 Developing Multi-Agent Systems Using JaCaMo-REST

Cleber Jorge Amaral (Federal Institute of Santa Catarina – São José, BR)

Jomi Fred Hübner (Federal University of Santa Catarina, BR)

Timotheus Kampik (Umeå University, SE & SAP Signavio – Berlin, DE)

License © Creative Commons BY 4.0 International license

© Cleber Jorge Amaral, Jomi Fred Hübner, Timotheus Kampik

URL <https://github.com/jacamo-lang/jacamo/tree/master/demos/integration>

In recent years, researchers have started to explore the integration of Multi-Agent Oriented Programming (MAOP) paradigm [1] and resource-oriented web architecture (REST) [2, 3, 4]. The present demo further advances this research line, presenting a novel approach for setting up the integration between a MAS and a range of different kinds of web resources. The demonstration of this approach and the steps to execute it are available at <https://github.com/jacamo-lang/jacamo/tree/master/demos/integration>. In particular, we demonstrate the use of the routing and mediation tool Node-Red (<https://nodered.org/>), for setting up JaCaMo-REST, a resource-oriented web-based abstraction for the Multi-Agent System (MAS).

JaCaMo-REST defines endpoints and provides web infrastructure for exposing MAS entities as web resources. Node-Red provides an integrating engine and an interface for defining flows to wire-up heterogeneous web resources. Its interface allows users to define data flow and transformation sequences from an input endpoint to an output endpoint.

Among the features of JaCaMo-REST, we highlight *proxy agents* and *proxy artifacts*. They allow MAS application agents to access web resources considering them as either agents or artifacts. Thus agents are able to interact with web resource in usual agent-to-agent or agent-to-environment interaction. In this sense, according to the external resources that the MAS must be integrated with, it is possible to define abstractions to interface web resources and MAS internal entities. The referred *proxy* abstractions can be defined using Node-Red flows. Indeed, Node-Red handles protocol transformations through its extensive library of connectors and encapsulates integration specifications such as the address of the endpoints and specific data transformation flows.

In the demonstration, two examples are exploited. In the first example, an external resource that is available via an MQTT *topic* is represented in the MAS as an *proxy agent*. An MAS agent, named *Bob*, can thus send messages to the resource using its Agent Communication Language (ACL). In this example, *Bob* can interact with the external resource without knowing, for instance, its Unique Resource Identifier (URI) or communication protocol, showing that an MAS application can be defined without integration concerns. In another example, *Bob* interacts with a proxy artifact (perceiving and acting upon it).

References

- 1 Olivier Boissier and Rafael H. Bordini and Jomi F. Hübner and Alessandro Ricci and Andrea Santi. *Multi-agent oriented programming with JaCaMo*. Science of Computer Programming, 78, 6, p. 747-761, 2013. <https://doi.org/10.1016/j.scico.2011.10.004>
- 2 Amaral, Cleber Jorge and Hübner, Jomi Fred and Kampik, Timotheus. *Towards Jacamo-rest: A Resource-Oriented Abstraction for Managing Multi-Agent Systems*, 2020. <https://arxiv.org/abs/2006.05619>

- 3 Ciortea, Andrei and Zimmermann, Antoine and Florea, Adina Magda. *Give Agents Some REST: Hypermedia-driven Agent Environments*, Springer, 125-141, 2018. https://doi.org/10.1007/978-3-319-91899-0_8
- 4 W. Collier, Rem and O'Neill, Eoin and Lillis, David and O'Hare, Gregory. *MAMS: Multi-Agent MicroServices*, Association for Computing Machinery, 655-662, 2019. <https://doi.org/10.1145/3308560.3316509>

4.2 Test-Driven Development for Agent-Oriented Programming

Cleber Jorge Amaral (Federal Institute of Santa Catarina – São José, BR)

Jomi Fred Hübner (Federal University of Santa Catarina, BR)

Timotheus Kampik (Umeå University, SE & SAP Signavio – Berlin, DE)

License © Creative Commons BY 4.0 International license

© Cleber Jorge Amaral, Jomi Fred Hübner, Timotheus Kampik

URL <https://github.com/jacamo-lang/jacamo/tree/master/doc/tutorials/tdd>

Agent-Oriented Programming (AOP) [1] is a paradigm that provides first-class abstractions for instilling autonomous behavior into software systems. While this paradigm has not yet been widely adopted by the software engineering mainstream, from an academic perspective, the technology ecosystem can be considered thriving [2, 3].

In modern software engineering, developers commonly apply Test-Driven Development (TDD) approaches, in which a large portion of the tests is written during or even ahead of the implementation of the actual program code. The assumption is that specifying the exact desired behavior of a software component before or alongside the implementation facilitates a more rigorous assessment of the component and ensures testing is not cut short because of time shortage caused by incorrect or imprecise estimations. This applies in particular in the context of autonomous and distributed agent-oriented systems, where reliable governance is a key concern [4] and implementation or operation errors can have disastrous consequences [5]. Although first works that are concerned with the development of QA-related capabilities for AOP have recently emerged [6, 7], so far, no testing utilities for any agent-oriented programming language appear to exist.

This demonstration shows how to specify and perform tests for JaCaMo agents using the AgentSpeak programming language [9, 10]. It is derived from the tutorial is available at: <https://github.com/jacamo-lang/jacamo/tree/master/doc/tutorials/tdd>, as well as from a recent AAMAS demonstration paper [8]. The approach uses a novel goal-oriented testing feature that enables an agent-oriented perspective on automated software testing and test-driven development.

The developed testing tool provides facilities to assert whether an expected outcome is being produced at the unitary level of an agent's inference rules, to test plans for the achievement of the agent's desired goals, and to test agents' integration, in which the facility uses the testing pipeline to assure that the interactions among agents are occurring as expected. The tool also allows for the instantiation of tester agents, which, besides inheriting the features of the agent under test, can run tests to assert whether the rules, plans, and beliefs of the agent under test are as expected. It is possible, for instance, to test if the agent behaves as expected while a particular plan is performed and changes the agent's internal state (e.g., belief adoption). When the main concern is the agent's logic, it is possible to use mock artifacts to avoid interaction with the environment. Additionally, mock agents can be instantiated to be part of a system to test agents' interactions.

The testing tool has several advantages compared to the common practice of using debug messages on production code: (i) the test code can be written separately from the production code, which results in cleaner code and a clearer separation of concerns; (ii) the expected outcomes of agents' functionalities are easier to understand; and (iii) it facilitates the automation of the tests, for instance, using Continuous Integration (CI) tools that enhance collaboration and quality assurance of projects developed by multiple engineers.

References

- 1 Shoham, Yoav. *Agent-oriented programming*. Artificial Intelligence, 60, 1, p. 51-92, 1993. [https://doi.org/10.1016/0004-3702\(93\)90034-9](https://doi.org/10.1016/0004-3702(93)90034-9)
- 2 Mascardi, Viviana and Weyns, Danny and Ricci, Alessandro and Earle, Clara Benac and Casals, Arthur and Challenger, Moharram and Chopra, Amit and Ciorcea, Andrei and Dennis, Louise A. and Díaz, Álvaro Fernández and El Fallah-Seghrouchni, Amal and Ferrando, Angelo and Fredlund, Lars-Åke and Giunchiglia, Eleonora and Guessoum, Zahia and Günay, Akin and Hindriks, Koen and Iglesias, Carlos A. and Logan, Brian and Kampik, Timotheus and Kardas, Geylani and Koeman, Vincent J. and Larsen, John Bruntse and Mayer, Simon and Méndez, Tasio and Nieves, Juan Carlos and Seidita, Valeria and Teze, Baris Tekin and Varga, László Z. and Winikoff, Michael. *Engineering Multi-Agent Systems: State of Affairs and the Road Ahead*. Association for Computing Machinery, New York, NY, USA, 44, 1, p. 18-28, 2019. <https://doi.org/10.1145/3310013.3322175>
- 3 Pal, Constantin-Valentin and Leon, Florin and Paprzycki, Marcin and Ganzha, Maria. *A review of platforms for the development of agent systems*. 2020. arXiv preprint arXiv:2007.08961
- 4 Kampik, Timotheus and Mansour, Adnane and Boissier, Olivier and Kirrane, Sabrina and Padget, Julian and Payne, Terry R. and Singh, Munindar P. and Tamma, Valentina and Zimmermann, Antoine. *Governance of Vol. Web: Challenges and Opportunities*. Association for Computing Machinery, 22, 4, 2022. <https://doi.org/10.1145/3507910>
- 5 Kampik, Timotheus and Amaral, Cleber Jorge and Hübner, Jomi Fred. *Developer Operations and Engineering Multi-agent Systems*. Engineering Multi-Agent Systems, Springer International Publishing, p. 175-186, 2022. https://doi.org/10.1007/978-3-030-97457-2_10
- 6 Amaral, Cleber Jorge and Hübner, Jomi Fred. *Jacamo-Web is on the Fly: An Interactive Multi-Agent System IDE*. Engineering Multi-Agent Systems, Springer International Publishing, p. 246-255. 2020. https://doi.org/10.1007/978-3-030-51417-4_13
- 7 Amaral, Cleber Jorge and Kampik, Timotheus and Cranefield, Stephen. *A Framework for Collaborative and Interactive Agent-oriented Developer Operations*. Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems, Auckland, New Zealand. 2020. <https://dl.acm.org/doi/10.5555/3398761.3399086>
- 8 Amaral, Cleber Jorge and Hübner, Jomi Fred and Kampik, Timotheus. *TDD for AOP: Test-Driven Development for Agent-Oriented Programming* (to appear). Proceedings of the 22nd International Conference on Autonomous Agents and MultiAgent Systems, London, UK. 2023.
- 9 Boissier, Olivier and Bordini, Rafael H and Hubner, Jomi and Ricci, Alessandro. *Multi-agent oriented programming: programming multi-agent systems using JaCaMo*. MIT Press, Cambridge, Massachusetts, USA. 2020.
- 10 Bordini, Rafael H. and Hübner, Jomi Fred and Wooldridge, Michael. *Programming Multi-Agent Systems in AgentSpeak Using Jason (Wiley Series in Agent Technology)*. John Wiley & Sons, Inc., Hoboken, NJ, USA. 2007.

4.3 Robust JaCaMo Applications via Exceptions and Accountability

Matteo Baldoni (University of Turin, IT)

Cristina Baroglio

Roberto Micalizio

Stefano Tedeschi

License  Creative Commons BY 4.0 International license
© Matteo Baldoni, Cristina Baroglio, Roberto Micalizio, Stefano Tedeschi

Robustness, “the degree to which a system or component can function correctly in the presence of invalid inputs or stressful environmental conditions” – generally called *perturbations* [11], is a crucial requirement of distributed software systems [12, 13, 10, 9]. Multi-Agent Systems (MAS) [15] are an effective approach to realize distributed systems by means of heterogeneous, and autonomous agents. Agent *organizations* (MAO), in particular, provide abstractions for modularizing code spread over many components, and orchestrate their execution by way of norms. JaCaMo [8] is one of the best-known platforms for implementing MAOs, but it focuses on providing the means for capturing the normal, correct behavior of the system and lacks of structural mechanisms allowing agents to exchange and propagate information (feedback) when they face perturbations. As in [1], the availability of *feedback* about perturbations is crucial to build robust distributed systems. Also MAS robustness should ground on the ability to convey feedback about perturbation to the agents that can handle it. But since agents generally are peers, and are not related by relationships like caller-callee or parent-child, the realization of robustness should occur through the definition of distributions of responsibilities among the agents, that become part of the MAO. This demonstration presents two extensions to the JaCaMo platform which allow building robust agent organizations. The first borrows from software engineering the concepts of *exception* and *exception handling*, while the second relies on the notion of *accountability*. Exception handling is suitable for treating perturbations anticipated at design time (i.e., exceptions) by activating handlers, that are also specified at design time. The details of the implementation can be found in [14, 6, 2, 3], source code available at <http://di.unito.it/moiseexceptions>.

Accountability, instead, defines feedback “channels” that agents can use at runtime to gain situational awareness about what occurred and then take actions. Raising and handling exceptions as well as asking and returning for an account will be tasks, under the responsibility of specific agents. The two extensions provide the means for representing such tasks as goals, and for distributing the responsibilities of such goals to the capable agents. Note that each such goal can be assigned to many agents, specifying the minimum and maximum cardinality of how many agents need to achieve the goal (as standard in a JaCaMo organization specifications). Moreover, we allow specifying many raising and handling goals for each exception, and many requesting, accounting and treatment goals (possibly involving many agents) for each accountability. Concretely, we leveraged the model presented in [4] and the formalization from [5, 7], source code available at <http://di.unito.it/moiseaccountability>.

References

- 1 David L. Alderson and John C. Doyle. Contrasting views of complexity and their implications for network-centric infrastructures. *IEEE Tr. on Sys., Man, and Cyber.*, 40(4), 2010.
- 2 Matteo Baldoni, Cristina Baroglio, Olivier Boissier, Roberto Micalizio, and Stefano Tedeschi. Demonstrating Exception Handling in JaCaMo. In Frank Dignum, Juan M. Corchado, and Fernando De La Prieta, editors, *Advances in Practical Applications of Agents, Multi-Agent Systems, and Social Good. The PAAMS Collection – 19th International Conference, PAAMS*

- 2021, Salamanca, Spain, October 6-8, 2021, *Proceedings*, volume 12946 of *Lecture Notes in Computer Science*, pages 341–345. Springer, 2021.
- 3 Matteo Baldoni, Cristina Baroglio, Olivier Boissier, Roberto Micalizio, and Stefano Tedeschi. Distributing Responsibilities for Exception Handling in JaCaMo. In Ulle Endriss, Ann Nowé, Frank Dignum, and Alessio Lomuscio, editors, *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '21, pages 1752–1754. International Foundation for Autonomous Agents and Multiagent Systems, 2021.
 - 4 Matteo Baldoni, Cristina Baroglio, Roberto Micalizio, and Stefano Tedeschi. Reimagining Robust Distributed Systems through Accountable MAS. *IEEE Internet Computing*, 25(6), 2021.
 - 5 Matteo Baldoni, Cristina Baroglio, Roberto Micalizio, and Stefano Tedeschi. Robustness Based on Accountability in Multiagent Organizations. In Ulle Endriss, Ann Nowé, Frank Dignum, and Alessio Lomuscio, editors, *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '21, pages 142–150. International Foundation for Autonomous Agents and Multiagent Systems, 2021.
 - 6 Matteo Baldoni, Cristina Baroglio, Roberto Micalizio, and Stefano Tedeschi. Exception handling as a social concern. *IEEE Internet Computing*, 26(6):33–40, 2022.
 - 7 Matteo Baldoni, Cristina Baroglio, Roberto Micalizio, and Stefano Tedeschi. Accountability in multi-agent organizations: from conceptual design to agent programming. *Autonomous Agents and Multi-Agent Systems*, 37(1):1–37, 2023.
 - 8 Olivier Boissier, Rafael H. Bordini, Jomi Hübner, and Alessandro Ricci. *Multi-agent oriented programming: programming multi-agent systems using JaCaMo*. MIT Press, 2020.
 - 9 Samuel H. Christie, Amit K. Chopra, and Munindar P. Singh. Bungie: Improving fault tolerance via extensible application-level protocols. *Computer*, 54(5):44–53, 2021.
 - 10 Samuel H. Christie, Amit K. Chopra, and Munindar P. Singh. Mandrake: multiagent systems as a basis for programming fault-tolerant decentralized applications. *Autonomous Agents and Multi-Agent Systems*, 36(1), 2022.
 - 11 ISO/IEC/IEEE. Systems and software engineering – Vocabulary. *24765:2010(E) – ISO/IEC/IEEE International Standard*, 2010.
 - 12 Anuj K Jain, Manuel Aparico IV, and Munindar P Singh. Agents for process coherence in virtual enterprises. *Communications of the ACM*, 42(3):62–69, 1999.
 - 13 Anup K. Kalia and Munindar P. Singh. Muon: designing multiagent communication protocols from interaction scenarios. *Autonomous Agents and Multi-Agent Systems*, 29(4):621–657, 2015.
 - 14 Stefano Tedeschi. *Exception Handling for Robust Multi-Agent Systems*. PhD thesis, Università degli Studi di Torino, Dipartimento di Informatica, Torino, Italy, 2021.
 - 15 Michael Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, 2009.

4.4 Agent-Oriented Visual Programming for the Web of Things

Samuele Burattini (University of Bologna, IT)

License © Creative Commons BY 4.0 International license
© Samuele Burattini

Joint work of Samuele Burattini, Simon Mayer, Alessandro Ricci, Danaï Vachtsevanou, Jérémy Lemée, Angelo Croatti

Overview

This demo presents the outcome of the work on *Agent-Oriented Visual Programming* [1] carried out in the context of the European project IntelliIoT¹⁶.

The demo is available on GitHub¹⁷ and shows how a multi-agent system controlling two autonomous tractors that expose Web of Things (WoT) Thing Descriptions¹⁸ can be defined and managed through a visual programming interface on a Web application.

More details about the demo can be found on the repository [README.md](#).

Agent-Oriented Visual Programming

The Agent-Oriented paradigm, allows the definition of software systems using high-level abstractions (i.e. software agents). Cognitive architectures for software agents are inspired to the mental abilities of humans, hence we argue that, as it is easy for developers to describe the world in terms of objects in Object-Oriented Programming, it should be even easier to model autonomous behaviour in terms of agents.

In the Belief-Desire-Intention (BDI) model, agents are described in terms of mental qualities that are inspired by *human practical reasoning*[3]. We argue that it is possible to exploit this alignment between human practical reasoning and the BDI architecture to create systems that people can effectively use to program MAS since they more closely match their everyday experience in interacting with artifacts and other agents.

Our endeavor is furthermore motivated by two concrete issues experienced in an industrial scenario based on the Web of Things (WoT). The first one is the ever-increasing interest in forms of end-user programming that shall enable not only experienced programmers but ideally domain experts without programming experience to create or modify software systems of different complexity. The second one is the need to create or modify solutions featuring different degrees of autonomy of software components in performing tasks in a flexible way, dealing with open, dynamic, distributed WoT environments. From this angle, at the same time, the use of semantic-web technologies allows to discover high-level actions at run-time, which promotes the serendipitous creation of applications in such environments – given a proper level of abstraction for exploiting them.

Then, the question is how to design an interface that could reinforce the alignment of agent and human reasoning, hiding the technicalities, to enable individuals without experience in programming to express the behaviour of software agents in the most natural way.

Our approach is built on a blocks-based visual programming environment to create Multi-Agent Systems (MAS) that are then executed on top of the JaCaMo[2] platform and can interact with Web of Things environments.

¹⁶ www.intelliott.eu

¹⁷ <https://github.com/samubura/dagstuhl-demo-wot-autonomous-tractors>

¹⁸ <https://www.w3.org/TR/wot-thing-description/>

References

- 1 Burattini, S., Ricci, A., Mayer, S., Vachtsevanou, D., Lemee, J., Ciortea, A. and Croatti, A., 2022. *Agent-Oriented Visual Programming for the Web of Things*.
- 2 Boissier, O., Bordini, R.H., Hübner, J.F., Ricci, A. and Santi, A., 2013. *Multi-agent oriented programming with JaCaMo*. *Science of Computer Programming*, 78(6), pp.747-761.
- 3 Bratman, M.E., Israel, D.J. and Pollack, M.E., 1988. *Plans and resource-bounded practical reasoning*. *Computational intelligence*, 4(3), pp.349-355.

4.5 Interaction-Oriented Programming

Amit K. Chopra (Lancaster University, GB)

Samuel H. Christie V

Munindar P. Singh (North Carolina State University – Raleigh, US)

License © Creative Commons BY 4.0 International license

© Amit K. Chopra, Samuel H. Christie V, Munindar P. Singh

Main reference Amit K. Chopra, Samuel H. Christie V., Munindar P. Singh: “Interaction-Oriented Programming: Intelligent, Meaning-Based Multiagent Systems”, in Proc. of the 2023 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2023, London, United Kingdom, 29 May 2023 – 2 June 2023, pp. 3041–3043, ACM, 2023.

URL <https://doi.org/10.5555/3545946.3599166>

Interaction-Oriented Programming (IOP) is a novel approach for designing decentralized multiagent systems based on models of interactions. Currently, IOP supports specifying multiagent systems via norms (e.g., commitments) and protocols. It includes tools that can be query databases for norm states, tools for verifying protocols, and programming models based on protocols that can be used to implement agents. IOP’s strength is that it supports flexible decision making by agents based on the meaning of interactions. For more details, see our AAMAS Demo paper.

Our software repository is here: <https://gitlab.com/masr/>.

Some software is here: <https://github.com/akchopr/Cupid/>.

4.6 Hypermedia Multi-Agent Systems

Olivier Boissier (Ecole des Mines – St. Étienne, FR)

Andrei Ciortea (Universität St. Gallen, CH)

Fabien Gandon (INRIA – Sophia Antipolis, FR)

Simon Mayer (Universität St. Gallen, CH)

Alessandro Ricci (Università di Bologna, IT)

License © Creative Commons BY 4.0 International license

© Olivier Boissier, Andrei Ciortea, Fabien Gandon, Simon Mayer, Alessandro Ricci

Main reference Andrei Ciortea, Olivier Boissier, Alessandro Ricci: “Engineering World-Wide Multi-Agent Systems with Hypermedia”, in Proc. of the Engineering Multi-Agent Systems – 6th International Workshop, EMAS 2018, Stockholm, Sweden, July 14-15, 2018, Revised Selected Papers, Lecture Notes in Computer Science, Vol. 11375, pp. 285–301, Springer, 2018.

URL https://doi.org/10.1007/978-3-030-25693-7_15

We are working towards defining a new class of Web-based Multi-Agent Systems (MAS) that can inherit the architectural properties of the Web (scalability, heterogeneity, evolvability, etc.), preserve the architectural properties of MAS (adaptability, openness, robustness, etc.), and are human-centric (usable, transparent, accountable, etc.). Our aim is to leverage

the full potential of the Web as a middleware in MAS. In our approach, we consider the environment as a first-class abstraction in MAS – and we design the agents' environment as a distributed hypermedia application guided by the design rationale of the Web architecture [1]. The distributed hypermedia environment provides agents with a uniform, knowledge-level abstraction of the system that they can navigate, query, observe, and act upon. The hypermedia-based design rationale reduces coupling and enhances the scalability, openness, and evolvability of the MAS. We refer to Web-based MAS that follow this design rationale as Hypermedia MAS.

During the seminar, we showed a demonstrator in which Belief-Desire-Intention (BDI) agents are situated in an open and evolvable hypermedia-based environment distributed across two locations: Schloss Dagstuhl (Germany) and the University of St.Gallen (Switzerland). Given a set of semantic models and a single entry URI into the system, the agents are able to achieve their design objectives by navigating the hypermedia environment to discover, create, perceive, and act on artifacts: the agents are able to create and use a digital counter for coordination, and they can control robotic arms located in St.Gallen. The distributed hypermedia environment is represented in RDF and agents are able to subscribe to triple patterns to receive fine-grained notifications of environmental changes. The demonstrator was built using Yggdrasil¹⁹, JaCaMo²⁰, Corese²¹, and the hardware infrastructure provided by the Chair for Communication- and Interaction-based Systems at the University of St.Gallen²².

References

- 1 Andrei Ciordea, Olivier Boissier, and Alessandro Ricci. 2019. Engineering World-Wide Multi-Agent Systems with Hypermedia. In *Engineering Multi-Agent Systems*, Danny Weyns, Viviana Mascardi, and Alessandro Ricci (Eds.). Springer International Publishing, Cham, 285–301.

4.7 Hypermedia MAS Simulation

Rem Collier (University College Dublin, IE)

License  Creative Commons BY 4.0 International license
© Rem Collier

Agent-Based Modelling (ABM) is an approach to implementing population-based simulations. A standard ABM simulation consists of two key types of entity: a set of agents (representing the population) and an environment (that provides a shared context for interaction between the agents). A range of toolkits are available for implementing ABM simulations covering both small scale desktop approaches and large scale cluster and cloud based approaches [1].

This talk introduces a new approach to implementing simulations using ABM that builds on the Multi-Agent MicroServices (MAMS) architectural style [2]. The approach, known as Hypermedia MAS Simulation [3] decomposes the environment part of an ABM into a set of sub-environments that are implemented as web resources that are hosted by microservices. The state of each sub-environment is then exposed using REpresentational State Transfer using Linked Data representations. The hyperlinks embedded in the linked data representations capture the relationships between sub-environments (and potentially other relationships too).

¹⁹<https://github.com/interactions-hsg/eggdrasil>

²⁰<https://github.com/jacamo-lang/jacamo>

²¹<https://github.com/wimmics/corese>

²²<https://interactions.ics.unisg.ch/>

A key novelty of the approach is that the linked data representations, when specified using Semantic Web languages, form a knowledge graph of the environment. Agents enter the environment by joining one of the sub-environments. Where appropriate, the agents transition between sub-environments and receive periodic updates of the state of the sub-environment they are currently in. These updates include the URL of that sub environment, which the agent can use as an entry point into the environment knowledge graph. All agent-environment interaction is realised through HTTP.


In the talk, the approach is illustrated using a simple worked example of a road network that is decomposed into streets and junctions. For simplicity, a single standard model of a street and a junction is defined and the corresponding environment is implemented using two microservices: one for streets and one for junctions. Hyperlinks connect the start and end of streets with associated junctions. Agents enter the environment by connecting to an junction (their starting point) and navigate the road network based on internal goals.

References

- 1 Abar, Sameera and Theodoropoulos, Georgios K and Lemarini er, Pierre and O’Hare, Gregory MP.: Agent Based Modelling and Simulation tools: A review of the state-of-art software. In: Computer Science Review 24, pp. 13–33, Elsevier (2017)
- 2 Collier, Rem, O’Neill, Eoin, Lillis, David, O’Hare, Gregory: Mams: Multi-agent microservices. In: Companion Proceedings of The 2019 World Wide Web Conference. pp. 655–662 (2019)
- 3 Collier, Rem and Russell, Se an and Ghanadbashi, Saeedeh and Golpayegani, Fatemeh: Towards the Use of Hypermedia MAS and Microservices for Web Scale Agent-Based Simulation, SN Computer Science 3(6), Springer (2022)

4.8 Building Multi-Agent Microservices with ASTRA

Eoin O’Neill (University College Dublin, IE)

License  Creative Commons BY 4.0 International license
  Eoin O’Neill

Joint work of Rem Collier, Eoin O’Neill

Integrating Multi-Agent Systems into larger software systems is difficult. Often, the result is to attempt to re-engineer the system in a way that makes it agent-centric. Such approaches are received negatively by the software engineering community because: (1) it requires the adoption of technology stacks that are not well understood or widely used, and (2) most systems already exist in some form and the cost of transforming it to be *agent-ready* is simply too high.

The emergence of microservices architectures [2] has changed this. Among the many benefits of microservices is the notion of polyglot computing – the construction of systems using multiple languages / technology stacks. Polyglot computing works well with microservices because of other principles, such as the discrete boundaries principle (microservices should be independently deployable). When combined with other architectural styles, such as REpresentational State Transfer (REST), the view emerges (from a deployment perspective) of a microservice as a black box with a uniform interface that can simply be connected into the larger system architecture. Details of the technology stack used become less important, so long as they only impact on the developer of the microservice rather than its users.

In response to this, the Multi-Agent MicroServices (MAMS) architectural style has been proposed [4] and a prototype implementation developed using the ASTRA agent programming language [1] and CArtAgO [3]. The MAMS architectural style promotes the idea of an agent

being represented using a hypermedia body, composed of a set of virtual resources that it chooses to expose. Each virtual resource, accessible at an endpoint embodies an element of the agent, such as its inbox. Based on the HATEOAS constraint of the REST architectural style, the state of the agent can determine what virtual resources are exposed and when. Any and all changes are, at all times represented in the hypermedia representation of the agent.

In this demonstration, a number of simple applications are presented that have been built using the MAMS prototype. In the first example we demonstrate an implementation of a Vickrey Auction using the MAMS architectural style. The Agent Oriented Micro Service (AOMS) exposed two virtual resources that allowed external entities to issue HTTP POST requests to, in order to register as a bidder or to sell an item. Internal agents that represent an Auctioneer and a Bidder agent per client then perform the auction and the bidder is informed by a response issued to the web hook they provide.

In the second example we demonstrate how a MAMS agent can interact with an external microservice that implements a single Tic-Tac-Toe board. The board is represented using hypermedia that the agent perceives in order to play the game. This level of loose-coupling and hypermedia-driven interaction between agents and Plain Old Micro Services (POMS) is the goal of our architectural style.

References

- 1 Collier, R.W., Russell, S., Lillis, D.: Reflecting on agent programming with agentspeak (1). In: PRIMA 2015: Principles and Practice of Multi-Agent Systems: 18th International Conference, Bertinoro, Italy, October 26-30, 2015, Proceedings 13. pp. 351–366. Springer (2015)
- 2 Dragoni, N., Giallorenzo, S., Lafuente, A.L., Mazzara, M., Montesi, F., Mustafin, R., Safina, L.: Microservices: yesterday, today, and tomorrow. Present and ulterior software engineering pp. 195–216 (2017)
- 3 Ricci, A., Piunti, M., Viroli, M., Omicini, A.: Environment programming in cartago. Multi-agent programming: Languages, tools and applications pp. 259–288 (2009)
- 4 W. Collier, R., O’Neill, E., Lillis, D., O’Hare, G.: Mams: Multi-agent microservices. In: Companion Proceedings of The 2019 World Wide Web Conference. pp. 655–662 (2019)

4.9 Web-based Demos of Cognitive AI

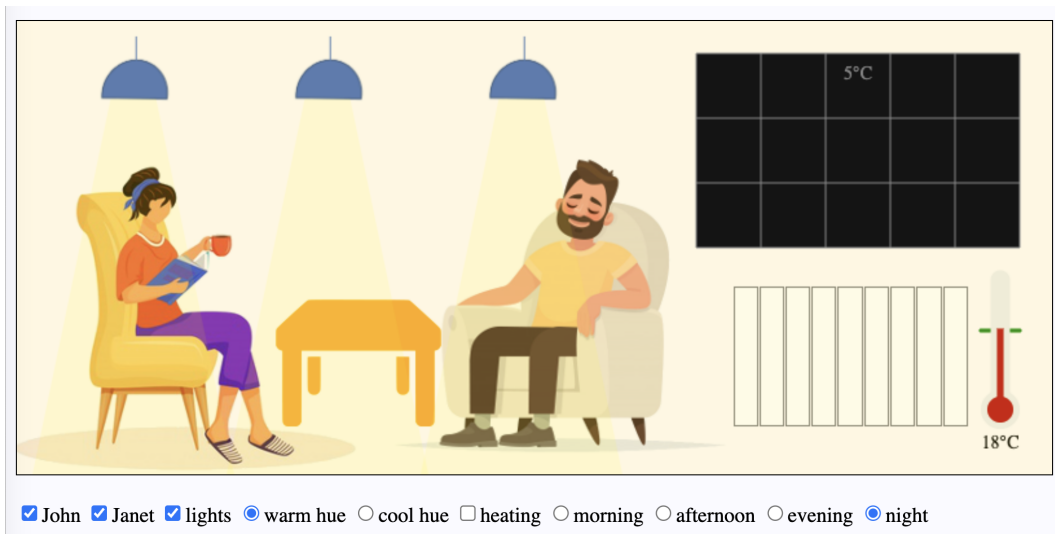
Dave Raggett (W3C – United Kingdom, GB)

License © Creative Commons BY 4.0 International license
© Dave Raggett

I presented three web-based demos, two of which use chunks and rules, inspired by John Anderson’s cognitive architecture named ACT-R, and another demo introducing plausible reasoning.

A chunk is a collection of properties along with a type and a chunk identifier. Chunk rules have a conjunction of conditions expressed as chunks that must match designated cognitive buffers where each buffer holds a single chunk. A rule is stochastically selected from the set of matching rules, and its actions applied. Actions either synchronously update the buffers or invoke asynchronous operations that indirectly update the buffer analogous to CRUD operations with HTTP. Chunk memory mimics human memory, e.g. the forgetting curve, spreading activation and stochastic recall.

See: <https://w3c.github.io/cogai/>



■ **Figure 3** screenshot of smart home demo.

Smart Homes

This includes a real-time model of heat flow between a room and the outside world, using a cognitive agent to manage the lighting and heating according to the preferences of the room's current participants. Janet prefers a warm room and warm lighting, whilst John prefers a cooler room and bluer lighting. If both Janet and John are present at the same time, then Janet has precedence for the lighting and John for the temperature.

This demo features default reasoning using chunk rules, triggered by events such as people entering or leaving the room, a change in the time of day, and the room temperature rising above or falling below the target temperature. The preferences are expressed as facts.

See: <https://www.w3.org/Data/demos/chunks/home/> and Figure 1.

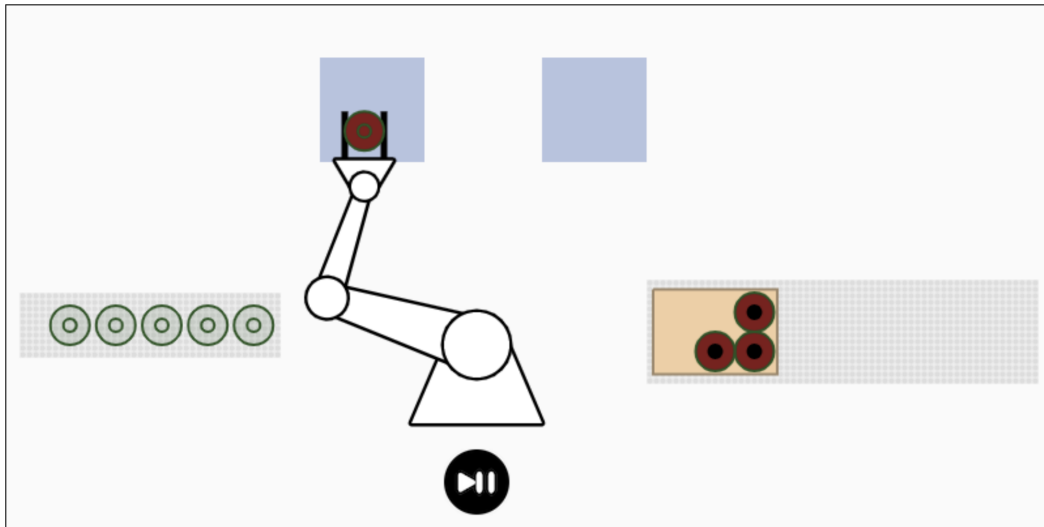
Manufacturing Cell

This features a cognitive agent that controls a robot, two conveyor belts, a bottle filling station, and a bottle capping station. Empty bottles are taken from the first conveyor belt, then filled and capped before being packed into a box, on the second conveyor belt, that takes six bottles. Chunk rules are used to express event driven behaviour. The rules also specify the target state for the robot gripper, delegating real-time control to the robot which smoothly accelerates and decelerates each actuator as required. The demo features matching sound effects.

See: <https://www.w3.org/Data/demos/chunks/robot/> and Figure 2.

Plausible Reasoning

This demo explores the potential for plausible reasoning with imperfect knowledge, i.e. knowledge that may be uncertain, imprecise, incomplete and inconsistent. Logical proof is replaced by plausible arguments for and against a premise, drawing upon work by a long line of philosophers on argumentation theory going all the way back to Ancient Greece, and as used in courtrooms and everyday reasoning. The plausible knowledge notation (PKN) features properties, relationships and implications, along with qualitative metadata that



■ **Figure 4** screenshot of robot demo.

reflects prior knowledge. PKN embraces plausible inferences based upon relationships and implications, fuzzy reasoning, fuzzy quantifiers, qualitative reasoning and analogical reasoning. PKN can be considered as a knowledge representation at a level above RDF with additional semantics for imperfect knowledge.

See: <https://www.w3.org/Data/demos/chunks/reasoning/>

4.10 Using MOSAIK for a Decentralized Transportation System

Daniel Schraudner (Universität Erlangen-Nürnberg, DE)

License © Creative Commons BY 4.0 International license
© Daniel Schraudner

Joint work of Andreas Harth, Sebastian Schmid, Daniel Schraudner

URL <https://github.com/wintechis/mosaik-runtime-documentation>

We investigate possibilities for implementing the decentralized control of transporters with Semantic Web agents to fulfill a given transportation task. We present our demo of the MOSAIK framework [2] as a system to build and simulate simple reflex agents [1] to control transporters using stigmergy for indirect communication via the environment [3] and self-organize based on local decisions. We use Semantic Web technologies as the communication paradigm of stigmergy directly maps to the REST constraints of the application architecture of the web.

In MOSAIK, we discern active components, agents, and reactive components, called artifacts that respond to agents' actions, and form together the agents' environment. As demo scenario, we present a square shop floor with distinct floor tiles and four stations, each with one of four different colors. A station accepts only products of its own color and randomly produces a product of a different color. Three transporters have to bring colored products to the correct station of the same color without being able to perceive more than only their adjacent fields. The transporters, as artifacts, are each controlled by our software agents that communicate with each other via stigmergy: by placing information about the

position of stations in the environment, on the shop floors, the agents help each other to build gradients that lead directly to the stations. With this, no global knowledge of the environment is necessary and the adjustment of the system emerges from local interactions.

As tools, we use BOLD (see <https://github.com/bold-benchmark/bold-server>) and Linked-Data Fu (ldfu) [4]. BOLD is a simulation environment that implements artifacts with multiple RDF graphs, collected in a single named graph representing the environment state. Agents can read and manipulate artifacts via HTTP requests to HTTP resources under the URI of the named graph, which BOLD updates via defined SPARQL INSERT / DELETE queries. The data processing system ldfu can retrieve, process and modify Linked Data based on logical rules and production rules in Notation3. We use N3 to implement the condition-action rules describing the behaviour of the agent which controls the transporters, and interact with the artifacts via RESTful interfaces.

The decentralized transportation system achieves self-organization by implementing a combination of simple reflex web agents that coordinate using web resources as environment for stigmergy.

References

- 1 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 2009, Prentice Hall Press, USA
- 2 Victor Charpenay, Daniel Schraudner, Thomas Seidelmann, Torsten Spieldenner, Jens Weise, René Schubotz, Sanaz Mostaghim, and Andreas Harth, *MOSAİK: A Formal Model for Self-Organizing Manufacturing Systems*, IEEE Pervasive Computing, 2021
- 3 Francis Heylighen, *Stigmergy as a universal coordination mechanism I: Definition and components*, Cognitive Systems Research, 2016
- 4 Andreas Harth and Tobias Käfer, *Linked Data Techniques for the Web of Things: Tutorial*, 2018, Proceedings of the 8th International Conference on the Internet of Things, Santa Barbara, California, USA

4.11 Querying over Solid Pods via Link Traversal, with Comunica

Ruben Taelman (Ghent University, BE)

License  Creative Commons BY 4.0 International license
© Ruben Taelman

Main reference Ruben Taelman, Ruben Verborgh: “Evaluation of Link Traversal Query Execution over Decentralized Environments with Structural Assumptions”, CoRR, Vol. abs/2302.06933, 2023.

URL <https://doi.org/10.48550/arXiv.2302.06933>

In this demonstration, we showed a technique to enable SPARQL query execution over one or more Solid pods in a decentralized environment using the client-side Comunica query engine. The method is a form Link Traversal Query Processing (LTQP) where the structural properties of the Solid ecosystem are taken into account [1] to provide earlier and more complete results.

The demonstrator is based on SolidBench, which is a benchmark that provides a large number of simulated Solid pods containing social network data, and queries over this data. Some of the demonstrated queries cover just one pod, while others span multiple pods.

The demo indicates that a traversal-based query method is an effective way for querying over Solid pods, without having to centralize all of this data beforehand. However, more work is needed to optimize more complex queries are queries spanning multiple pods.

The demo remains available online at https://comunica.dev/research/link_traversal/, together with the open-source implementation of the query engine.

References

- 1 Taelman, Ruben, and Ruben Verborgh. “Evaluation of Link Traversal Query Execution over Decentralized Environments with Structural Assumptions.” arXiv preprint arXiv:2302.06933 (2023).

5 Working Group Reports

5.1 Decentralized Hypermedia Ecosystems

Olivier Boissier (Ecole des Mines – St. Étienne, FR)

Pierre-Antoine Champin (INRIA – Sophia Antipolis, FR)

Amit K. Chopra (Lancaster University, GB)

Nicoletta Fornara (University of Lugano, CH)

Fabien Gandon (INRIA – Sophia Antipolis, FR)

Ruben Taelman (Ghent University, BE)

Jan Van den Bussche (Hasselt University, BE)

License © Creative Commons BY 4.0 International license

© Olivier Boissier, Pierre-Antoine Champin, Amit K. Chopra, Nicoletta Fornara, Fabien Gandon, Ruben Taelman, Jan Van den Bussche

Inspired by the recent focus of decentralized data for Web applications, as supported in infrastructures such as SOLID, this group discussed the idea of decentralized hypermedia applications in general. In particular, this working group focused on understanding how to integrate the technologies and models used and proposed by three areas of research targeted by the seminar.

1. The notion of Hypermedia, Hypermedia as the Engine of Application State (HATEOAS), and ecosystems developed by the Web of Things community;
2. The notion of Linked Data Platform (a W3C recommendation) together with open specifications such as SOLID proposed by the Semantic Web and Linked Data community.
3. Decentralized abstractions studied in multiagent systems. These include the concept of agents, norms or policies, protocols, artifact, workspace, and organization together with processes of regimenting and enforcing norms.

This integration was investigated by first discussing a list of crucial terms used in each of the communities involved, in order to achieve a good level of mutual understanding. Subsequently, by discussing and specifying a scenario in which all the technologies and concepts proposed by each of the communities involved can be used at different levels of abstraction.

Crucial Terms

- **Hypermedia:** a non-linear multimedia medium where resources are linked by hyperlinks i.e. a digital reference to data that can be followed to discover new data
- **Hypermedia as the Engine of Application State (HATEOAS):** is a constraint of the REST application architecture where a client interacts with a network application whose application servers provide information dynamically through hypermedia.
- **Decentralized System:** is a type of system where every node makes its own decision, there is not a central coordinating or governing unit or server.

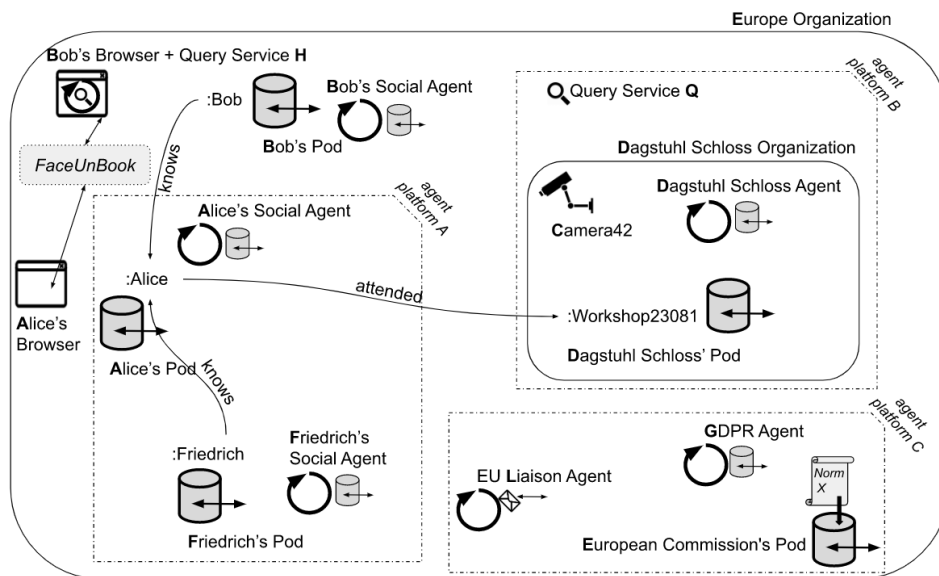
- **SOLID**: Solid is a set of standards that enables a decentralization of data, with interoperable applications over this decentralized data.
- **SOLID ecosystem**: is a specific deployment of a community of users and agents using a given set of applications on top of their Solid pods, and complying with a set of technical and social norms.
- **SOLID pods**: Solid is a specification that lets people store their data securely in decentralized personal online data stores called pods. Pods are like secure personal web servers for data. When data is stored in someone's Pod, they control which people and applications can access it.
- **Pody**: is a term forged during the Dagstuhl Seminar to identify the Pod of an agent that embodies some of its features. This approach is referred to as the empodiment of the agent.
- **Web ID**: URI that identify a person / an agent
- **Agents** : autonomous entities able to perceive, act and interact with other agents, making decisions, subject to norms.
- **Norms**: Capture social expectations, they are by definition violable. Norm types include commitment (obligation), prohibition, permission and so on.
- **Artifacts**: non autonomous entities.
- **Workspace**: a way to organize the artifacts, the artifacts are situated in workspace, agents join workspaces, workspace scope the observability for an agent.
- **Enforcing**: applying sanctions (rewards or fines) when norms are fulfilled or not fulfilled
- **Regimenting**: hindering the violations of norms by agents.
- **ODRL**: the Open Digital Rights Language (ODRL) is a W3C recommendation defining a policy expression language. ODRL policies are used to represent permitted and prohibited actions over a certain asset, and obligations required to be met by stakeholders.

Scenario: Social Media Sharing Application in a Decentralized Hypermedia Ecosystem

We designed and described a plausible decentralized architecture that broadly combines above themes with reference to a social media sharing application. In contrast to the traditional approach of realizing an application as a Web service that mediates interactions between its users, here each user is represented by its own agent who interacts with other agents based on an understanding of applicable norms.

The Scene

Alice is a person sharing her photos on the social network FaceUnBook running this app in her browser and using her pod to host her data and her agent to manage her photo sharing preferences (among other tasks). Bob is a friend of Alice and is following her activity stream on FaceUnBook. His agent is in charge of collecting the activity of all his friends, which will be summarized in the pod of Bob. This summary can then improve discovery when Bob uses the FaceUnBook app, which is backed by the personal query engine of Bob. It combines all search capabilities it can identify to plan and solve this aggregation query. Alice attended the seminar 23081 at Schloss Dagstuhl. Among other things the agent of that organization enforces the norms of the organization that are stored in the organization's pod. The agent also collects pictures taken by the Camera42. This happens in Europe which is an organization with norms. The agent GDPR agent is in charge of enforcing the GDPR norms for every member of Europe. To make all members in Europe findable by the GDPR agent, Europe has a norm saying that all pods in Europe must be registered to Europe's



pod. Friedrich is a kid accompanying Alice in Dagstuhl. The agents, pods and resources in general are hosted on different platforms but discover and interact with each other through the hypermedia fabric and its standards. To support the discovery of resources, pods and their contents, **Q** is a query service available in the environment as one possibility to read and write within the hypermedia fabric, used by either software agents, or by human agents through client-side apps. This scenario follows the empowerment principle i.e. the agents communicate through their pods (pody) both through the inbox container of the pody or through other containers of the pody.

An Interaction to Share Photos May Process as Follows:

Alice took photos at Dagstuhl, in which herself and Friedrich are depicted. She runs the FaceUnBook application in her browser, and uses it to share the photo with her friends. The application posts the photo in Alice's pod, and (HTTP) posts a notification in Bob's inbox (a specific container in his pod). Dagstuhl publishes on its pod ODRL policies about photo sharing that apply to all seminar participants. Alice tags the photo with the tag "Dagstuhl Seminar 23081", then the FaceUnbook application also reads the policy and asks Alice to confirm that she complies with that policy (including the rule that no photo of a children must be published without consent from a caretaker – in that case, Alice is the caretaker and consents). Then the application (HTTP) posts a notification in Dagstuhl's inbox. Bob likes Alice's photo; he improves the contrast in an image editor, and republishes it on his own pod (notifying Bob's friends, including Alice). Alice's agent notices that the republishing violates Alice's preferences and (HTTP) posts a warning to the pods of Alice and Bob.

Moving Forward

People in the group have independently worked on different aspects of the above architecture. A way forward would be to actually realize this application as broadly outlined above, see how things fit together and identify the gaps. Although in the above description Alice acts in accordance with the norms, in another situation, she may act in violation of the norms, e.g., if she posts a photo of a child who needs medical attention. FaceUnbook must not prevent

such sharing but must support recording her violation, based on which the child’s parent, for example, may demand an explanation from her. This brings in accountability and a richer notion of social processes than currently possible.

5.2 Abstractions for Agents on the Web

Andrei Ciortea (Universität St. Gallen, CH)

Catherine Faron (Université Côte d’Azur – Sophia Antipolis, FR)

Jomi Fred Hübner (Federal University of Santa Catarina, BR)

Eoin O’Neill (University College Dublin, IE)

María Poveda-Villalón (Technical University of Madrid, ES)

Alessandro Ricci (Università di Bologna, IT)

Antoine Zimmermann (Ecole des Mines – St. Étienne, FR)

License © Creative Commons BY 4.0 International license

© Andrei Ciortea, Catherine Faron, Jomi Fred Hübner, Eoin O’Neill, María Poveda-Villalón, Alessandro Ricci, Antoine Zimmermann

This document reports on the findings of a Working Group at the Dagstuhl Seminar on *Agents on the Web* relative to what abstractions must be considered when designing multi-agent systems on the Web.

We first consider the case of a single agent that makes use of the Web, consuming or producing information. Then, we examine the case of multiple agents that work together as a system where the Web plays the central role of the environment in which they are situated. Finally, we provide preliminary ideas on how to instantiate the elicited abstractions in practice, relying on Semantic Web models and technologies and the Solid Protocol for Social Linked Data.

5.2.1 Introduction

Agents that are (inter)acting autonomously on the Web and with the Web must make use of the information available in this environment to gather knowledge and make decisions. Different kinds of information correspond to different types of abstractions that we want to formally identify such that they can be represented explicitly and systematically. With such representations of relevant abstractions, artificial agents can replicate activities on the Web that have been so far reserved to human agents: navigating hypermedia environments with a purpose, choosing what next links to follow based on promising signifiers, and collaborating with one another via collaborative Web platforms.

In what we consider here in the context of agents on the Web, an abstraction is a notion that can be defined as a formal structure (such as a mathematical object) which is relevant for a system of autonomous agents to function adequately. If the agents are humans, it is usually not necessary to formalise these abstractions because people have many ways of getting an intuitive understanding of ways to cooperate. For instance, when a community of people discusses issues on a Web forum, the fact that there is a coloured rectangle with the words “New thread” is sufficient to alert humans that this is a button for starting a new textual discussion. When one sees a blue underlined word that says “Reply” under the last message of a thread, one easily understands that this is a link that allows participants to reply to the thread, or to the last post. On the contrary, artificial agents may not be able to interpret these visual cues. If we would like to allow such agents to participate in a

discussion, it is possible to hard code the interactions, e.g., a developer translates the actions needed to read and post messages on a specific forum. However, this is very limiting in terms of autonomy. Instead, we would like to define more abstractly what generic notions have to be known by the agent in order to choose its interactions with Web resources. As an example, online discussions can be characterised by the fact that there are discussion threads, each composed of individual messages that one can respond to. Messages are authored by someone (or something) at a date and time. These concepts are common to any Web forum, to blogs and microblogging platforms, to Git issues, chats, etc.

Once identified, the concepts that pertain to one type of abstractions can be formally defined as terms of an ontology, further described and constrained by logical axioms. In turn, a web platform can make use of such terms to describe itself for autonomous agents. The aim of this document is to report on the findings of a Working Group at the Dagstuhl Seminar on *Agents on the Web* relative to what abstractions must be considered when designing multi-agent systems on the Web. Consequently, this can serve as a starting point to decide what ontologies may be used to describe where, what, how resources are to be used by autonomous agents, for themselves or for a collective goal.

To do this, we start by considering the case of a single agent that makes use of the Web, consuming or producing information. In this first case, the Web is only used as a tool that is assumed to evolve independently of all agents (Section 5.2.2). Then, we examine the case of multiple agents that work together as a system where the Web plays the central role of the environment in which they are situated (Section 5.2.3). Then, we provide preliminary ideas on how to instantiate the abstractions in practice (Section 5.2.4).

5.2.2 Single Agents Interacting with the Web

Before considering the more arduous challenges of coordinating multiple agents on and via the Web, we first tackle the case where a single agent makes use of the Web as if it was a large warehouse full of documents and tools. The agent may exploit the Web in getting information from it: from the Web of documents, the Web of data, or the Web of knowledge. From this, it can derive or update its own beliefs. Or the agent may act upon resources leading to altering the state of the Web itself (e.g., posting a new message adds text to a Web page), or altering the state of the physical world itself (e.g., moving a robotic arm that offers a Web interface using Web of Things technologies).

To be fully autonomous on these types of tasks, Web agents must start their exploration from somewhere. At this point, there must be indications of where to find relevant information. This should lead them to online platforms where they may ask themselves: *What can I do on this platform? How can I do it?* Since the Web is evolving, agents may also be interested in changes occurring, without constantly checking for differences. An agent can be partly or fully guided by hard-coded knowledge provided at implementation time, such as the URIs of a preselected set of resources. However, in general, the agent has to interact with an entry point that serves as a hub towards any relevant resources. Such hub can be a search engine, a query endpoint, etc. that provides, at the minimum, links to resources that match the agent's request. Usually, the resource that the agent is looking for is not atomic: it corresponds to a compound resource that gives multiple options to the agent. For instance, the agent may be looking for online forums. Each online forum is a resource that provides discussion threads, search functionalities, posting abilities, etc. We call the abstraction of such compound resource a *platform*. The functionalities enabled by a platform and that the agent can use are affordances, which can be indicated to the agent via signifiers [10].

In summary, some competency questions must be answerable using the right abstractions, as follows:

- *Where do I find relevant resources?* Entry hubs, links to platforms.
- *What can I do on this platform?* Signifiers, links to affordances.
- *How can I do it?* Action possibilities, links to Manual.
- *How can I know what changes?* Notifications.

5.2.3 Multiple Agents Interacting on the Web

In the case of a single agent interacting with the Web, all online resources are considered as artifacts or tools that can evolve according to an independent life cycle. When agents need to collaborate, cooperate, or at least behave with awareness of each other on the Web, additional abstractions need to be introduced.

The concept of situatedness with regards to Web agents in a *multiagent* context can be defined in relation to the scope of environment that they exist within. As defined in Section 5.2.2, any tools that an agent utilises in its goal-directed behaviour provides a level of situatedness as each interaction will be between two entities that can be identified using the global, unique naming scheme provided by the use of IRIs. This also pertains to agent-to-agent interaction as when agents collaborate and coordinate in order to achieve system-wide, or individual, goals they interact via the Web, with each agent being identifiable by a unique IRI.

In order for collaboration and coordination between Web agents to be a possibility in a multiagent context, the embodiment of an agent within a hypermedia environment is a crucial abstraction that needs to be defined. This allows for the agent to maintain a presence in a Web context which enables other agents to be aware of its existence. However, agents can be embodied in multiple Web contexts, the same way a human would have multiple profiles on different social media platforms. This embodiment of an agent in a particular Web context can represent an identity of a Web agent. The agent's profile can be a certain persona that this particular agent wants to portray within that context. This abstraction can contain the agent's preferred methods of interaction via its *communication interface*, which can be defined as or subject to a set of *norms*. These *norms* can be defined within the hypermedia environment, as resources, that can be directly associated with the Web agent's embodiment.

When discussing multiple Web agents in a hypermedia environment, the scope of the environment is a relevant topic due to the computational limitations of the agents themselves. As a result, an abstraction for specific collections of resources can be viewed as an aggregation of related resources. This abstraction can be viewed as an *area* within a Web-based environment. If we take for example a set of related resources that represents elements of a building (Rooms, Floors etc.), these can be aggregated into an *area*. This example makes logical sense as the relation can be easily visualised and the relations envisioned, however there can be Web resources that are related, but in a less explicit manner and so this abstraction can provide a general way of aggregating resources to provide agents with scope as to the environments they inhabit.

5.2.4 Instantiating the Abstractions

The Semantic Web relies on the Resource Description Framework (RDF), a graph model to structure data by expressing relations between entities, and on RDF Schema and the Web Ontology Language to represent the ontologies used in RDF knowledge graphs, thus providing

semantics to them. The Linked Data principles are a set of best practices to publish RDF data on the Web, namely: 1) use URIs to name things; 2) use HTTP URIs so that names can be looked up; 3) describe things using standards (RDF) so useful information is provided for URIs; and 4) include links to other URIs in things descriptions. Ontologies and linked data together provide the means by which an agent can reliably interpret resources described on the Web, whether they are digital resources or real-world resources. Additionally, with links, a web resource leads to other resources, and so forth, so as to make agents aware of the environment that the Web constitutes.

Besides, the Solid Protocol [2] aims at decentralising personal data management in such a way that Web users regain ownership and control over their data. At the core of it, there is the Solid *pod* (**personal online data store**) that hosts the user’s data and is implemented as a Linked Data Platform [9] with access control on top of it. Solid pods can host any kind of data but are designed in particular to easily manage RDF datasets with fine-grained read/write operations. Overall, the Solid Protocol specifies authentication, storage, access control, and interactions that must be implemented by Solid pods and Solid platforms in order to interoperate with each other and with applications that build on them.

Solid pods containing RDF knowledge graphs can be used to implement the abstractions introduced in Section 5.2.3. In order to represent the different levels and types of knowledge defining those abstractions it is necessary to combine a number of ontologies. For example the Web of Things model [7] might be used to describe the affordances while the Open Digital Rights Language (ODRL [6]) would be useful to define access policies to data or resources. Other existing ontologies might be extended or specialised to represent particular needs in the agents domain like the Organization ontology [8] or sensor and actuator related ontologies as SSN&SOSA [5] and SAREF [4]. Finally, specific domain ontologies related to the agent’s tasks should be developed or reused. The Solid protocol states that “an agent is a person, social entity, or software identified by a URI; e.g., a WebID denotes an agent”. We then assume that such a URI would dereference to an entry point for the data pod of the agent, where an **Agent Description** would be provided as an RDF graph, in addition to the mandatory credentials for authenticating the agent. We call the Solid pod implementing an agent’s Web body a *pod* [12].

5.2.5 Conclusion

The dominant view in AI research is that intelligence is defined in relation to the environment an agent occupies [11]. The agents we have on the Web today also reflect the nature of their environment: they mostly solve the problem of finding and curating information in a Web of Documents. The Web, however, was designed to provide different levels of abstraction – going from computer networks to a knowledge-level representation of the world [1]. With recent standardisation efforts for Linked Data and subsequently the Web of Thing, the Web now extends to the physical world and provides agents with a uniform, knowledge-level hypermedia fabric that allows them “to browse and manipulate reality” – a vision that can be traced back to the early days of the Web.²³ This evolution unlocks new practical use cases for more intelligent agents on the Web, but such agents have to be provided with a proper level of abstraction that allows them to discover and interact with one another – and with the world in general. This report presents an initial proposal for such a set of abstractions with a focus on *situatedness* and *embodiment*. We invite the research community to join us in further investigating and developing these abstractions.

²³ See the keynote of Sir Tim Berners-Lee at the First International Conference on the World Wide Web (WWW’94): <https://videos.cern.ch/record/2671957>

Acknowledgements

Andrei Ciortea and Antoine Zimmermann were supported by funding from the HyperAgents project (grant from French national research agency ANR-19-CE23-0030-01 and from Swiss National Science Foundation No. 189474). María Poveda-Villalón was supported by the Spanish National project KnowledgeSpaces: Técnicas y herramientas para la gestión de grafos de conocimientos para dar soporte a espacios de datos (PID2020-118274RB-I00). Eoin O’Neill is supported by the Science Foundation of Ireland Strategic Partnerships Programme (16/SPP/3296) and is co-funded by Origin Enterprises Plc.

References

- 1 Tim Berners-Lee. Levels of Abstraction: Net, Web, Graph. Design Issues, 2007. <https://www.w3.org/DesignIssues/Abstractions.html>
- 2 Sarven Capadisli, Tim Berners-Lee, Ruben Verborgh and Kjetil Kjernsmo. Solid Protocol, W3C Solid Community Group working draft, 2022. <https://solidproject.org/TR/2021/protocol-20211217>
- 3 Sarven Capadisli and Amy Guy. Linked Data Notification. W3C Recommendation, 2 May 2017. <https://www.w3.org/TR/2017/REC-ldn-20170502/>
- 4 Laura Daniele, Raúl García-Castro, Maxime Lefrançois and María Poveda-Villalón. SAREF: the Smart Applications REFerence ontology. ETSI TS 103 264, v3.1.1, 2020. <https://saref.etsi.org/core/v3.1.1/>
- 5 Armin Haller, Krzysztof Janowicz, Simon Cox, Danh Le Phuoc, Jamie Taylor and Maxime Lefrançois. Semantic Sensor Network Ontology, W3C Recommendation, 2017. <https://www.w3.org/TR/2017/REC-vocab-ssn-20171019/>
- 6 Renato Ianella and Serena Villata. ODRL Information Model 2.2, W3C Recommendation, 2018 <https://www.w3.org/TR/2018/REC-odrl-model-20180215/>
- 7 Sebastian Kaebisch, Michael McCool and Ege Korkan. Web of Things (WoT) Thing Description 1.1, W3C Candidate Recommendation Snapshot, 2023. <https://www.w3.org/TR/2023/CR-wot-thing-description11-20230119/>
- 8 Dave Reynolds. The Organization Ontology, W3C Recommendation, 2014. <http://www.w3.org/TR/2014/REC-vocab-org-20140116/>
- 9 Steve Speicher, John Arwe and Ashok Malhotra. Linked Data Platform 1.0, W3C Recommendation 2015. <https://www.w3.org/TR/2015/REC-ldp-20150226/>
- 10 Danai Vachtsevanou, Andrei Ciortea, Simon Mayer, and Jérémy Lemée. Signifiers as a First-Class Abstraction in Hypermedia Multi-Agent Systems. In Proc. of the 22nd International Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2023.
- 11 Michael Wooldridge. Intelligent Agents. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, MIT Press, 2000. p27–72.
- 12 Antoine Zimmermann, Andrei Ciortea, Catherine Faron, Eoin O’Neill and María Poveda-Villalón. Pody: a Solid-based Approach to Embody Agents in Web-based MAS. To appear in Proc. of Engineering Multiagent Systems, EMAS 2023.

5.3 Representation Formats Supporting Goal-oriented Decision-Making on the Web

Cleber Jorge Amaral (Universidade Federal de Santa Catarina, BR)

Samuele Burattini (University of Bologna, IT)

Jean-Paul Calbimonte (HES-SO Valais – Sierre, CH)

Rem Collier (University College Dublin, IE)

Matthias Kovatsch (Siemens Schweiz AG – Zug, CH)

Mahda Noura (TU Chemnitz, DE)

Daniel Schraudner (Universität Erlangen-Nürnberg, DE)

Danai Vachtsevanou (Universität St. Gallen, CH)

License © Creative Commons BY 4.0 International license

© Cleber Jorge Amaral, Samuele Burattini, Jean-Paul Calbimonte, Rem Collier, Matthias Kovatsch, Mahda Noura, Daniel Schraudner, Danai Vachtsevanou

5.3.1 Overview

One of the key outcomes from this working group was a consensus on the importance of identifying a general representation format to enable both interoperability between different Multi-Agent System (MAS) implementations and adaptability to unexpected differences in the environment for agents that can engage in goal-oriented decision-making on the Web. In the first three days of the seminar, this working group focused on aligning the necessary terminology, investigating the definition of high-level goals, identifying different use cases and decision-making procedures, and establishing the required abstractions for supporting plan execution. On the fourth day, this working group further merged with working group 5 which was focusing on the synthesis of MAS, with a particular focus on strategy identification for agents on the Web. A brief summary of the broad range of topics discussed is provided in the following.

5.3.2 Terminology Alignment

Early in the working group discussions, it became apparent that there was a divergence in the terminology used. This section outlines key terms along with definitions that were agreed by the group.

- *Environment*: Generally viewed as the space that an agent inhabits. Environments can be physical, virtual, or a combination of the two. In MAS, environments are seen as being first-class abstractions in that they are as important as the agents that inhabit them [2]. In this working group, given the scope of the seminar, the environment was considered to be a set of resources that are published on the Web; are accessible through the Web; and are linked to one another (where relevant) via hyperlinks. Those resources could relate to physical (e.g. sensors, actuators) or virtual (e.g. databases, services) resources.
- *Workspace*: In general, a workspace is part of an environment that is oriented towards a specific (set of) tasks or activities. Workspaces are an intermediary layer that adds structure and organisation to the environment. Agents typically operate across one or more workspaces. For the working group, a workspace was viewed as a set of resources that are somehow related (e.g. the resources relating to a smart room or building) and should be monitored by agents operating in that workspace.
- *Situatedness*: This refers to the relationship between an agent and its environment. Situated agents are able to interact with their environment via sensors and actuators. Situated agents are considered to be closely coupled to their environment, and their

behaviours can often be defined in terms of their effect on the environment. In terms of the working group, agents are situated in the Web environment described above. Agents are able to identify resources through Uniform Resource Identifiers and are able to interact with them using HTTP operations.

- *Artifact/Thing*: *Artifacts* are an abstraction of the non-agent entities that exist within an environment. The concept arises from the Agents & Artifacts meta-model [2]. Artifacts typically adhere to a well-defined interface that can be used by agents. Artifacts can be anything (e.g. an diary, a shared whiteboard, or a web browser). In terms of the Web, one possible application of artifacts is to use them as abstractions of resources [3], providing an enabling medium for agent-resource interaction. In such an approach, the artifact would be responsible for exposing the state of the resource and providing the necessary functionality to interact with the resource. In contrast, a *Thing* is a specific type of environment entity; typically a physical sensor / actuator device. The term comes from the *Internet of Things (IoT)* community. One refinement of this view comes from the *Web of Things (WoT) community*, where things are exposed on the Web as a resource. Access to sensor data or control functions is exposed through the Web, and the interface is usually described by an associated *Thing Description*. In this sense, a Thing can be viewed as a stylised Web resource whose representations are constrained by agreed standards.
- *Action/Operation*: Agents perform *actions* in the environments they inhabit. They do so by invoking actuators that implement the primitive mechanisms for affecting the environment. Agents choose the action they wish to perform through a variety of approaches, such as stigmergy, reinforcement learning or symbolic (logical) reasoning. In the context of the working group, actions were taken to generally refer to the submission of HTTP Requests to Web resources. An *Operation* is a type of actuator that is associated with the Agents & Artifacts meta-model. It is part of the well-defined interface described earlier. Each artifact has an associated set of operations that represent the different ways that an agent can interact with the artifact. In terms of the working group, it was agreed that agents perform actions by invoking operations on artifacts. Further the set of operations associated with an artifact representing a Web resource would be the set of valid HTTP Requests that can be used with that resource.
- *Effects/Pre-/Post-Conditions*: In order for some types of agent to decide what action to perform next, they need a description of the context in which an action can be performed and possibly description of the impact/effect of that action. The terminology for describing these descriptions is commonly drawn from early work on planning and they are often defined in terms of logical formulae, but this is not a requirement. The context in which an action may be performed is often defined as a partial description of the state that the environment must be in for the action to be applicable. The action context is often defined as a set of *pre-conditions* that is part of an action description that also includes unique identifier for each action. In addition to providing a context in which the action is applicable, many action descriptions also provide additional information about the *Effect* of the action. This is often described in terms of the changes that the action makes to the environment (both addition and removal). Effects are also known as *Post-conditions*. The action descriptions are developed for specific scenarios as part of the domain model. In the context of this work package, pre- and post- conditions are expected to be defined in terms of the current state of one or more resources and the changes that a given HTTP Request will have on that (set of) resources. How to express a state of multiple web resources was not completely clear.

- *Goals*: A goal is generally viewed as a future state of affairs that the agent (or system) is trying to achieve. Achieving a goal involves the creation of a plan of action. A plan is a set of actions whose cumulative effect is the achieve the goal. The task of constructing a plan of action to achieve a goal is known as planning. Goals that identify future states that the agent wishes to achieve are broadly known as declarative goals. This type of goal contrasts with the idea of procedural goals – goals that encapsulate some behaviour that is to be achieved. For example, the task of paying a bill assigned to an agent can be expressed declaratively via the goal <paid bill> or procedurally via the goal <pay the bill>. The former refers to the state of having paid the bill and the later refers to the procedure of paying the bill. Generally, planning involves the use of declarative goals as it allows the planning problem to be expressed as a search through a space of possible states. States are represented as nodes, and actions are associated with edges that transition the system from one state (node) to another (node) as defined by the pre- and post- conditions of the action description. In terms of this working group, much of the discussion focused on goals in a more abstract sense or around declarative goals.
- *Global Guidance*: Global guidance refers to the creation of high-level plans to achieve domain specific goals. It is primarily concerned with goal refinement and is expected to utilise planning techniques in conjunction with any domain models that have been developed for the scenario. Largely, global guidance is concerned with the strategic decision-making activities of the deployed system. The idea of global guidance is refined further in section 5.3.4.
- *Local Guidance*: Local guidance refers to the refinement of high-level plans based on the context in which the agent is operating. It is primarily concerned with the reification of the abstract domain model derived plans with the actual configuration of the environment. In contrast with global guidance, local guidance is primarily concerned with tactical decision-making and is also discussed further in section 5.3.4.
- *Protocols*: Protocols define structured interactions between agents. They define a set of messages that should be sent between two or more agents with a specified ordering that can be global or partial in nature. Protocols are a building block for defining valid patterns of interaction between agents. In the context of this work package, the existence of protocols was assumed, but was not considered in detail.
- *Organisation*: In Multi-Agent Systems research, an organisation is broadly defined as a collection of agents that work together within an agreed context that is based around a set of well defined roles with clear lines of communication and responsibilities. Organisations apply structure to a Multi-Agent System and agents can play different roles within that structure at different times depending on their own goals and objectives. In the context of this working group, the function of an organisation was not discussed extensively.
- *Norms*: Norms define the standard behaviour of a system (e.g. all cars going in the same direction drive on the same side of the road). In societies, norms often emerge from the behaviour of the individuals within that society (everybody else is doing it, so should we). Norms can be formalised as laws, with an associated penalty, that can be used to enforce compliant behaviour. Like organisations, norms can be used to provide additional structure around the behaviour of agents. In some respect, norms can be seen as a less rigid than organisations because that provide a kind of social guidance that each agent can either adhere to or not. In the context of the working group, there was little discussion around the use of norms.

The provided definitions enabled us to discuss primary abstractions that are commonly encountered in a MAS. Such abstractions have been identified as potential candidates for representation in a common format towards facilitating the interoperability and adaptability of agents on the Web.

5.3.3 Use Case

In this section, we present a scenario to motivate the definition and representation of proper abstractions to support agents' interactions in hypermedia environments. Given that agents can execute a variety of tasks that call for coordination and communication among them, and that the group was interested in challenges posed by a real-world scenario with agents interacting through the Web with an evolving physical environment, a building automation scenario can provide a suitable environment for testing a concept for modelling abstractions of MAS.

In such a scenario, a human user may have a set of preferences relating to environmental conditions stored in a personal Solid Pod²⁴. Agents in the MAS can use the data from the Pod to adopt goals towards increasing the user's comfort with respect to the given set of preferences. In the most simple scenario, we can imagine having a single agent managing a room in the smart building, identifying users, retrieving information from the Pods of the identified users, and striving to implement the desired environmental preferences. Of course, the agent might also need to adapt its behavior based on constraints about energy consumption, or towards supporting users' productivity or health and safety, and so on.

Given a generic specification of the user comfort level as a goal, the agent will be delegated to navigate the hypermedia environment and find the WoT Thing Descriptions of the things it can interact with and adjust the conditions accordingly. Table 1 lists some things and services the agent might need to interact with to achieve its goal.

■ **Table 1** Things and services that allow to modify the comfort level in a smart room scenario.

Thing	Description	Illustrative use
Temperature sensor	Provides the temperature of the room (e.g. in Celsius Degrees)	An agent can find in the thing's description how to read the temperature sensor of the room, monitor it, and possibly actuate over things such as the air conditioning to achieve a target temperature.
Humidity Sensor	Provides the relative humidity of the room	In order to attain a target humidity, an agent can find instructions in the thing's description for how to read the room's humidity sensor, monitor it, and possibly actuate devices like the air conditioner.
Illuminance sensor	Provides the illuminance of the room (e.g. in Lumens)	In the thing's description, an agent can find how to read the room's illuminance sensor, monitor it, and possibly act on items like the lighting and window blinds to reach a desired level of illumination.
Lighting Bulbs	Actuator that affects the illuminance of the room	The power over this device can be adjusted over a range of values by an agent to increase or reduce its effectiveness. This will change the amount of light in the room.
Window Blinds	Actuator that can cover or uncover a window affecting the illuminance and visibility of the room	This actuator can be set in a range of values by an agent for changing the lighting and visibility in the room, opening or closing the covering.
Weather service	Provides the forecast for temperature and humidity in the region in which the room is located	This service can provide the forecast to the next hours. This can be used by an agent that has a threshold in energy consumption which may decide to use natural light for some time even reducing the human comfort (e.g. due to sun glare)

²⁴<https://solidproject.org>

Of course, there are other ways to implement this use case without having a single central agent capable of managing everything. We could for example have different personal agents representing different users, that will try to negotiate the best environmental condition for each user. For instance, for finding a common sense for different preferences of multiple human representative agents can adopt an existent protocol for auctions [1]. Also, an agent representing the room's provider might be present, and have its own preferences (e.g. a certain threshold in energy consumption) in conflict with the comfort level thus they should be negotiated and prioritized according to some extra definitions.

Another possibility may be to also have multiple agents controlling certain room conditions. In each room there might be agents in charge of adjusting the temperature, humidity, and illumination of a room, then the task would no longer require an agent to navigate the hypermedia environment directly looking for things, but instead to navigate the MAS to find which agents are responsible and interact with them. For instance, if the current room temperature is not as a human prefers, its representative agent can ask the agent responsible for controlling the room temperature to adjust it accordingly.

When the room is occupied by more than one agent, the interactions and coordination issues can be even more complex. In this scenario setup, the agents can interact with each other to properly coordinate their actions over the things of the room.

To monitor and control things, and to negotiate a suitable condition for the room, the agents must employ common abstractions to describe their knowledge, intentions, goals, and plans. Table 2 shows possible interactions between agents and things and among agents, as well as necessary abstraction that they must share.

■ **Table 2** Kind of interactions and necessary abstractions.

Interaction	Abstraction description	Examples of circumstances
Agent representing a human / Environment described with Thing's Description	The agent must understand what is described in the Thing's Description.	(i) For understanding the value of the temperature, the agent must know in which unit it is being represented. (ii) For setting up the temperature the agent must know if a particular action would increase or decrease the power over a cooler or a heater.
Agent representing a human / Agents controlling room's conditions	The agents that are interacting must have a common understanding of perceptions, actions, desires and intentions.	(i) The agent that controls the temperature may increase the power of the heater to achieve a request made by an agent representing an human in the room. (ii) The agent that controls the humidity refuses the request of a human representative after realising that achieving the requested humidity conflicts with a higher priority target that is already set.
Agent representing a human / Agent representing another human	The agents that are interacting must negotiate the final desired outcome. Therefore, they must have a common understanding of relevant interaction protocol.	(i) The agent that controls the temperature may reply to a second human representative agent that it will set the temperature for the average of these two, or they must negotiate another target.

5.3.4 Strategic and Tactical Planning

For this section, we assume that we already have a way of expressing goals for agents as well as actions, which agents can take, and their effects on the environment at an abstract level. We will call this a domain model. An example of an abstract representation of an agent's

goal in the domain of building automation would be “I want to make the room brighter”. The domain model would then comprise all possible actions and their effects, e.g. “I can set the power of light bulbs to adjust the brightness” or “I can lift the blinds of a window in order to adapt the brightness of the room to the outside brightness”.

Agents that want to achieve such a goal in hypermedia environments like the Web usually have to make plans on two different levels. We call them strategic planning and tactical planning. Strategic plans are plans on a high level of abstraction, and they are made a priori, have a global scope, and are based on a domain model (which could e.g. be captured with an OWL ontology). A strategic plan in the given domain for the goal “I want to make the room brighter” could, for example, be “I need to find a light bulb in the room, then I can adjust its power”. We make no assumptions about the mechanics of the plan creation or acquisition here (it could be classical PDDL planning, but it could also be a neural network generating the plan).

To actually carry out a plan, an agent has to interact with its environment. In the case of hypermedia environments, these interactions usually are either following a link or submitting a form to a server. This means that agents have to refine their strategic plan to be able to execute it. Agents employ tactical planning to refine their strategic plans. Tactical plans, as opposed to strategic plans, are made at run time, have a local scope, and are based on what the agent found out about the environment (e.g. based on WoT Thing Descriptions that have been found). An example of a tactical plan would be “I send an HTTP POST request with the payload ‘<> rdf:value 10 .' to <http://ex.org/room1/bulb4/properties/power>”. A comparison between strategic and tactical plans can be seen in Table 3.

We need the separation between the two levels of plans because hypermedia environments are usually highly dynamic environments and (strategic) plans that are too detailed get obsolete very fast.

The environment can help agents in creating tactical plans by providing metadata about the available links and forms that the agent can understand (e.g. “If you follow this link, you will get a list of all devices in the room”). The environment could even take a more active role in guiding an agent by presenting them only links and forms which the environment knows that they are useful to the agent (e.g. “I know you are an agent for handling the room brightness, thus I will only show you links to the devices that can influence brightness”).

What parts of a plan should belong to the tactical side and what to the strategic side is not fixed but lies on a continuum. The exact position on the continuum depends on the use application scenario. At one end of the continuum, plans in agents would be hard-coded: the agent already has a very detailed “strategic” plan of what exact steps to carry out in which order (first send this HTTP request, then that HTTP request) based on the domain model before it starts executing it. This works well in a static and deterministic environment, and it is very efficient there, but hard-coded plans are not flexible and cannot adapt to unforeseen changes in the environment at all.

On the other end of the continuum, agents do their best effort link following: the agent has no strategic plan at all (maybe because it does not have knowledge about the domain) and just follows the links and submits the forms that at the moment seem most useful to them; it just uses tactical plans. This approach is very flexible (as it makes no assumptions about the environment) but usually also not very efficient.

For the design of Hypermedia MAS, it clearly is a challenge for system designers to determine where on the described continuum the different agents should be located. It will also be important to find out, how the transitions between the two planning tactics work and how they influence each other.

■ **Table 3** Comparison of the different aspects of strategic and tactical planning.

	Strategic Planning	Tactical Planning
Time	A priori	At run time
Scope	Global	Local
Based on	Domain model	Environment
Focus on leads to	High efficiency, low flexibility	Low efficiency, high flexibility

5.3.5 Designing MAS for Strategic and Tactical Planning

The discussion about strategic and tactical planning presented in the previous section moved our focus from the external data that should be made available to the agent to the internal decision-making processes. We believe that a better understanding of the patterns that agents could use to reason and adapt to changes in their environment is crucial to also identifying how to sufficiently describe such environments with a suitable representation.

We then decided to join our efforts with the working group focused on synthesis in MAS since both groups shared an interest in the activities of the other and felt they were going towards similar objectives.

We ended up with a shared view of the main abstractions that are needed to design a MAS in order to account for many possible strategies to refine tactical planning. We named such strategies *Course Check and Revision Strategies* (CCRS) which are further described in our shared report.

5.3.6 Conclusion and Planned Work

Our initial focus on examining representation formats for describing goals and environments led us to explore ways to enable heterogeneous agents to collaborate effectively on the Web and to enhance agents' autonomous decision-making through tactical planning. Eventually, we shifted our attention towards designing a system that accounts for divergences between the design time abstraction of an environment and its run-time implementation.

The collaboration between our team and the working group on the synthesis on MAS has proven to be fruitful. We are pleased with the outcomes of our joint efforts and intend to pursue this research direction further. Specifically, we aim to investigate the use of CCRS in agent architectures, and how these strategies can be supported by detailed descriptions of the hypermedia environment, the goals of a MAS, and the other agents present within the system.

References

- 1 Fabio Belfemine and Agostino Poggi and Giovanni Rimassa. *Developing multi-agent systems with a FIPA-compliant agent framework*. Software: Practice and Experience, 2001
- 2 Andrea Omicini and Alessandro Ricci and Mirko Viroli. *Artifacts in the A&A meta-model for multi-agent systems*. Autonomous Agents and Multi-Agent Systems, 17, 432-456, 2008.
- 3 Andrei Ciortea, Olivier Boissier, Alessandro Ricci. *Engineering World-Wide Multi-Agent Systems with Hypermedia*. International Workshop on Engineering Multi-Agent Systems, 285-301, Springer, 2018.

5.4 Course Check and Revision Strategies for Autonomous Hypermedia Navigation

Cleber Jorge Amaral (Universidade Federal de Santa Catarina, BR)

Samuele Burattini (University of Bologna, IT)

Jean-Paul Calbimonte (HES-SO Valais – Sierre, CH)

Rem Collier (University College Dublin, IE)

Andreas Harth (Fraunhofer IIS – Nürnberg, DE)

Matthias Kovatsch (Siemens Schweiz AG – Zug, CH)

Brian Logan (University of Aberdeen, GB & Utrecht University, NL)

Simon Mayer (Universität St. Gallen, CH)

Mahda Noura (TU Chemnitz, DE)

Sebastian Schmid (Universität Erlangen-Nürnberg, DE)

Daniel Schraudner (Universität Erlangen-Nürnberg, DE)

Danai Vachtsevanou (Universität St. Gallen, CH)

License © Creative Commons BY 4.0 International license

© Cleber Jorge Amaral, Samuele Burattini, Jean-Paul Calbimonte, Rem Collier, Andreas Harth, Matthias Kovatsch, Brian Logan, Simon Mayer, Mahda Noura, Sebastian Schmid, Daniel Schraudner, Danai Vachtsevanou

5.4.1 Overview

This working group report sums up the activities of the joint sessions between working group 4 (Representation Formats) and working group 5 (MAS synthesis). The groups merged due to similar interests in the relationship of design-time plans and run-time adaptation to real hypermedia environment for autonomous software agents. An extended version of this work has been submitted to the 2023 ACM Hypertext Conference.

5.4.2 Introduction and Background

Supported by Web browsers and online services such as search engines and recommender systems, many people today excel at efficiently navigating the Web’s hypermedia structure towards achieving their goals. The essence of navigating hypermedia is that a user is able to match a discovered – local – hypermedia control with their – global – application goal. For instance, the control `HTTP PUT /cart { "isbn" = "978-1452654126" }` should be matched with the goal to *buy the book “The Design of Everyday Things”*). Users accomplish this matching using information in the representation of the current Web resource – in the above example, this might be a button labelled *Add to Shopping Cart!* or a cart icon. The user hence needs to be able to find out that a specific local hypermedia control is indeed suitable to advance towards their goal; for buying a book, this is true for most users because the online shop has been modelled according to how people shop in the physical world, hence it is *natural* that the book should be “put into the cart” for enabling the user to, eventually, buy it.

From a technical perspective, this process is on the Web enabled by a mechanism that is part of the Uniform Interface constraint of the REST architectural style [1]: *Hypermedia as the Engine of Application State* (HATEOAS) constrains REST systems so that clients only perform hypermedia state transitions on resources using actions that are provided within the hypermedia that the server delivers [2]. With HATEOAS the only implemented capability of a hypermedia client should be the way to select which hypermedia control is useful to achieve the goal. This avoids tight coupling between clients and servers, and permits a hypermedia environment and its clients to evolve independently from one another.

While central to the way humans navigate the Web, the REST HATEOAS constraint is today not emphasized in the design of hypermedia environments that are roamed *by machines*. However, following larger academic emphasis on this topic in the wake of the Web of Things movement [3, 4, 5], industry and standardization groups have been gaining interest in HATEOAS over the past decade, for instance at the World Wide Web Consortium’s (W3C) Web of Things (WoT) Working Group.²⁵ This rising interest is driven by an increasing requirement on machine clients to more autonomously navigate hypermedia environments, and handle environment dynamics [6].

We argue that not least due to today’s wide availability of powerful AI models that might be used by artificial agents to provide guidance in hypermedia environments, revisiting the role of HATEOAS and the matching process of high-level domain goals of agents with low-level hypermedia controls is required. HATEOAS, together with suitable local traversal strategies, could form a cornerstone to permit artificial agents to roam the Web – and use its services – similar to how people do it.

5.4.3 Strategies for Autonomous Navigation

In-line with research in the field of Autonomous Agents and MAS, we suppose that an artificial agent for a hypermedia environment is created by an agent designer according to a (design-time) model that the designer holds of the environment. This creation process involves the design and programming of the agent’s logic towards achieving the agent’s design objective. At run time, the agent is placed in a hypermedia environment and executes its programmed behavior. In environments such as the Web, it is however likely that the agent will encounter an environment that does not exactly match the expectations of the agent designer. The agent then either needs to *cope with* unexpected violations of the design-time assumptions (e.g., nonexistent hyperlinks or link relations); or it may *optimize* its way to achieve the goal if discovered hypermedia controls suggest better ways to achieve its design objective.

We argue that, to permit this level of adaptation to the environment, the agent should be equipped with strategies of evaluating which of a set of encountered hypermedia controls will enable it to realize its design objective.

We refer to such strategies as *Course Check and Revision Strategies* (CCRS). A CCRS describes what an agent does at run time to verify if the next step of its planned course of action is (still) possible and/or prudent given the available run-time information and the agent’s objective. From an application perspective, CCRS can be either reactive or proactive: *Reactive CCRS* are designed to enable agents to cope with unexpected variations of the run-time environment that violate design-time assumptions. The agent hence is required to deviate from the original course of action; for instance, the agent designer might not have known that an agent would need to be logged in to access a specific resource – the agent in this case requires a CCRS that permits it to discover how to log in to the system at run time. *Proactive CCRS* permit agents to discover opportunities that it may choose to exploit to achieve its design objective in a better (faster/cheaper/more privacy-friendly/etc.) way. Following these opportunities, the agent then actively changes its designed course of action to optimize the path towards its design objective.

²⁵ <https://www.w3.org/WoT/wg/>

■ **Table 4** Categories of CCRS with examples (non-exhaustive) based on the CCRS' main dependencies.

Category	Depends on	Examples
<i>Independent</i>	–	Fail, Random links, Retry, Backtrack
<i>Environment-Oriented</i>	environment features	Affordance-based, Stigmergic
<i>AI-Oriented</i>	computational resources, data availability, model/environment fit	Prediction, Reinforcement Learning, Planning, Reasoning
<i>Socially-Oriented</i>	presence of other agents / communication	Consultation, Delegation, Human-in-the-Loop

When creating an agent to reach an objective in a hypermedia environment, the agent designer may then equip its agent not only with the primary (sometimes even hard-coded) course of action – e.g., a list of URIs to access; a list of link relations²⁶ to follow; a list of W3C WoT Interaction Affordances²⁷ to exploit; etc. – but also with one or several CCRS among the ones presented below.

5.4.4 Course Check and Revision Strategies

In this section, we propose a list of CCRS that an agent designer might select from when creating agents that roam Web hypermedia environments – or other environments where HATEOAS-like feedback is available. We do not claim that this list is exhaustive; rather, it should provide a starting point for the creation of a catalog of CCRS while at the same time further illustrating our approach and connecting it to different areas within artificial intelligence and beyond. The choice of a specific CCRS depends on the amount and type of autonomy and flexibility that an agent designer intends to give to the agent it creates as well as on the specific abstraction of the environment that the agent designer holds at design time.

We group CCRS in four categories: *Independent*, *Environment-Oriented*, *AI-Oriented*, and *Socially-Oriented* which are briefly illustrated in Table 4.

5.4.4.1 Independent CCRS

Independent CCRS do not place any special requirements on the hypermedia environment and do not require specific information about available hypermedia controls.

5.4.4.2 Environment-oriented CCRS

Environment-oriented CCRS depend on specific features of the environment to recover from unexpected situations (when used reactively) or to optimize the hypermedia traversal of an agent (when used proactively).

²⁶ <https://www.iana.org/assignments/link-relations/link-relations.xhtml>

²⁷ <https://www.w3.org/TR/wot-thing-description11/>

5.4.4.3 AI-oriented CCRS

In AI-oriented CCRS, the agent is equipped with access to an AI (sub)system that it consults at run time to reactively resolve unexpected situations or to proactively optimize its traversal of the hypermedia environment. For these strategies to remain efficient, the agent designer needs to select a type of AI that is compatible with the resources in the hypermedia environment and that can readily use the information that is provided by these resources towards eluding the appropriate next hypermedia control to use for achieving the agent's design objective.

5.4.4.4 Socially-oriented CCRS

In socially-oriented CCRS, the agent relies on other agents to support its traversal of hypermedia. The supporting agents are considered to be autonomous entities rather than tools that the agent makes use of (we discuss this in more detail in the following Section 5.4.6). In socially-oriented CCRS, the agent designer needs to have sufficient confidence in the ability and willingness of other agents to support the traversing agent at run time, and such agents need to be present, capable to communicate between each other, and perceivable by the created agent.

5.4.5 CCRS in a Single Agent Scenario

We provide a simple yet sufficiently complex scenario to illustrate our proposal and the interplay between the creation of an agent at design time and its execution at run time. For this scenario, suppose that a designer creates an agent that it intends to task with switching on a lamp in a hypermedia environment and then leaving feedback about its usage of the hypermedia environment. At design time, the designer holds expectations about several aspects of the to-be-encountered run-time environment, and each of these expectations is connected to a confidence:

- E1:** The designer knows with full confidence that the environment's entry point URI is `https://mythings.example.org/room-d-12`.
- E2:** The designer knows that the entry point URI represents a thing directory, i.e. a list of hyperlinks to Web APIs of physical devices that are contained in a room. The designer is sure that there is only a single lamp in this list and believes that the hyperlink to the lamp's API is located on the second page of the directory. While the designer neither knows the URI of that page nor of the lamp's API, the designer believes that the directory exposes `next` link relations to support pagination.
- E3:** It is unknown to the designer what specific lamp the agent will encounter at run time, and hence it does not know about the API. The designer furthermore does not know whether the agent will be required to first log in before it can switch the lamp.
- E4:** The designer has high confidence that all hypermedia resources expose hypermedia controls for leaving textual feedback. The designer knows that the link relation that is exposed to hint at these controls is `leaveFeedback` and they know about the specific hypermedia control required.

The designer creates an agent to achieve the design objective based on this environment model: Based on (E1), it programs this agent to first send a GET request to `https://mythings.example.org/room-d-12`. Based on (E2), it programs the agent to, next, find a `next` link relation and follow it until the returned resource representation contains a hyperlink to the API of a *Lamp*. Based on (E3), it programs the agent to achieve the goal

Lamp On which it specifies declaratively using a suitable language (e.g., RDF²⁸). Finally, based on (E4), the designer programs the agent to find a `leaveFeedback` link relation and use it to send a feedback text.

Because of its little knowledge or low confidence with respect to E2 and E3, the agent designer furthermore equips the agent with several CCRS to locally cope with deviations and/or optimize its traversal of the hypermedia environment:

- To cope with a resource not exhibiting the expected link relations in Step 2 and Step 4 of the process (i.e., `next` and `leaveFeedback`), the designer gives the agent an AI-Oriented CCRS like using a large language model for the ability to cope with unexpected vocabularies.
- For achieving the goal of switching on the lamp, the agent designer equips the agent with a CCRS that uses an AI planning system for locally computing a plan. At run time, this CCRS will succeed if suitable metadata is available, where typical planners require the pre- and post-conditions of using a hypermedia control to be specified; these can then be chained to lead the agent to achieve its goal.
- The agent designer knows that the planning strategy, while remaining a possibility, might fail for several reasons. For this reason, the designer includes another CCRS that uses Consultation/Delegation. The agent is hence programmed to contact a (hard-coded or locally discovered) resource – e.g., a specialized local lamp-switching agent or a human agent – when it requires help with achieving its goal of switching on the lamp.

When the agent now starts executing its program in the given hypermedia environment, it will engage none, any, or all its CCRS after each usage of a hypermedia control to possibly modify its future course of action. In this example, the CCRS depend on implicit or explicit environment support, e.g. on information being available about the environment for the *Planning CCRS*, and on other agents being present and capable for the *Consulting/Delegation CCRS*.

5.4.6 CCRS in a Multi-Agent Scenario

In this section, we investigate how our proposed type of increased local autonomy through CCRS impacts the behavior and design of a Multi-Agent System (MAS). In a MAS, several agents work towards the achievement of their individual goals [7]. The way these individual goals are related to each other defines whether the agents within the MAS are *cooperative* – in this case, a top-level goal is usually decomposed into smaller problems that are solved by agents that share results and communicate with each other. In a non-cooperative scenario, agents might have competing interests, but usually are expected to work rationally and towards achieving an equilibrium [8].

In our exploration we considered only a *cooperative* MAS expanding the scenario from Section 5.4.5 to one with two *interdependent* agents in two roles:

- The **Lamp Operator** is tasked to use hypermedia to turn on a lamp in a repository and then leave feedback.
- The **Heater Operator** is tasked to use hypermedia to turn on a heater and then leave feedback. It holds a key that, when submitted through HTTP POST to the `https://mythings.example.org/authorizeAll` endpoint, will authorize other agents to access the repository resource.

²⁸ <https://www.w3.org/RDF/>

In this situation, the MAS designer will create a *multi-agent plan* that requires the **Heater Operator** to first post the key to unlock the repository, then access the repository, and then search for the heater hypermedia control and actuate the heater. The **Lamp Operator** agent will be programmed to wait until it can access the repository and then proceed to find and use the lamp control. In our approach, the MAS designer equips the two agents with suitable CCRS to permit them to recover from unexpected run-time situations and to identify optimization potential at run time.

However, suppose that, at run time, the **Heater Operator** uses a CCRS to discover that there is a *direct access link* to the heater control in the environment that permits it to access that control *without first posting its key*. Although it might be prudent for this individual agent to optimize its course of action, this behavior would not satisfy the goal of the MAS, since the **Lamp Operator** would in this case not be able to complete its design objective.

5.4.6.1 Dealing with Multi-Agent Coordination

The use of CCRS in MAS is hence constrained in multi-agent plans that require coordination. Problems arise whenever an agent discovers an environmental *shortcut* through a CCRS that might be more rewarding for the agent but that leads to the skipping of a *coordination point* that is required for the overall MAS behavior. This effectively limits the freedom of an agent that cannot use its CCRS without considering the effects on the whole MAS, and the direct consequence is that agents either need to *refrain* from using CCRS or they need to *coordinate* with other agents whenever applying a CCRS that might impact other agents.

Several possibilities might be available, however they will be largely dependent on the concrete scenario, including the abilities of the agents and the characteristics of the environment.

In our MAS scenario, the **Heater Operator** might for example follow one of these strategies upon finding a shortcut through a CCRS:

- If the designer *explicitly designed an organization for the MAS*, the agent may first look up the organizational specification for the MAS. This specification will in this case include the obligation on the agent to post the key, and the agent would be in violation of this if it continued following the CCRS.
- If the agent *can communicate with other agents*, it may follow the CCRS-suggested course of action, but first broadcast the key (and how to use it) to other agents so that they can use the key by themselves if needed. If no *direct* communication is possible but the environment permits creating shared resources among the agents in the MAS, the agent might leave the unlock information in the environment for other agents to find.
- If the designer creates an *explicit synchronization point* in the agent plans where the agents are programmed to meet, the agent may choose the CCRS-provided route to reach this point, but – if the other agent is not reaching the synchronization point – it might backtrack and repeat the part of the original plan that it skipped.

5.4.6.2 Dealing with Heterogeneous Agents

The agents might also have *heterogeneous capabilities* – different actions they are able to perform in the environment, different means of communication with other agents, different cognitive capabilities, etc. – and also *heterogeneous knowledge* about their design objectives and the environment – different vocabularies that they may understand, different data they have access to, etc. If it is not possible to capture these differences at design time in a comprehensive model, which may occur in open systems like the Web, agents should be enabled to adapt to such differences at run time to still be able to coordinate in the MAS.

Considering again the heater-and-lamp multi-agent scenario, the **Heater Operator** might decide to delegate the task of using the key to other agents in the environment since it wants to pursue a more direct plan that was discovered using its CCRS. In order for the agent to consider this delegation strategy, the designer must have confidence that other agents can (i) receive such messages, (ii) understand a shared vocabulary to express goals and plans, and (iii) be able to perform a similar subset of actions.

If any of these conditions are violated *at run time*, the delegating agent might realize that no agent is responding to the task it was trying to delegate and use its CCRS to further adapt to this circumstance.

References

- 1 Fielding, R.T. and Taylor, R.N., 2002. *Principled design of the modern web architecture*. ACM Transactions on Internet Technology (TOIT), 2(2), pp.115-150.
- 2 Ruby, S. and Richardson, L., 2007. *RESTful web services* (p. 454). Sebastopol, CA: O'Reilly.
- 3 Guinard, D., Trifa, V. and Wilde, E., 2010, November. *A resource oriented architecture for the web of things*. In 2010 Internet of Things (IOT) (pp. 1-8). IEEE.
- 4 Mayer, S., Inhelder, N., Verborgh, R., Van de Walle, R. and Mattern, F., 2014, October. *Configuration of smart environments made simple: Combining visual modeling with semantic metadata and reasoning*. In 2014 International Conference on the Internet of Things (IOT) (pp. 61-66). IEEE.
- 5 Mayer, S., Inhelder, N., Verborgh, R., Van de Walle, R. and Mattern, F., 2014, October. *Configuration of smart environments made simple: Combining visual modeling with semantic metadata and reasoning*. In 2014 International Conference on the Internet of Things (IOT) (pp. 61-66). IEEE.
- 6 Ciortea, A., Mayer, S., Gandon, F., Boissier, O., Ricci, A. and Zimmermann, A., 2019, May. *A decade in hindsight: the missing bridge between multi-agent systems and the world wide web*. In Proceedings of the International Conference on Autonomous Agents and Multiagent Systems.
- 7 Shoham, Y. and Leyton-Brown, K., 2008. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.
- 8 Axtell, R.L., 2002, July. *Non-cooperative dynamics of multi-agent teams*. In Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 3 (pp. 1082-1089).

5.5 Learning and Reasoning with Behavioural Norms

Mehdi Dastani (Utrecht University, NL)

Kemal A. Delic (The Open University – Milton Keynes, GB)

Jérôme Euzenat (INRIA Grenoble Rhône-Alpes, Saint Ismier, FR & Univ. Grenoble-Alpes, FR)

Stefano Mariani (University of Modena, IT)

Dave Raggett (W3C – United Kingdom, GB)

Munindar P. Singh (North Carolina State University – Raleigh, US)

License © Creative Commons BY 4.0 International license

© Mehdi Dastani, Kemal A. Delic, Jérôme Euzenat, Stefano Mariani, Dave Raggett, Munindar P. Singh

Societies of agents are constantly confronting a changing world, environment, and society. They have to learn how to behave in such a context and to reason on how to exploit what they have learned. This covers learning actions that they can do (including independently or jointly), learning how others behave, and learning what is expected from them. Learning may take its source from the web, including the web of things, the behaviour of other agents, and combinations thereof. It can take advantage of diverse techniques (e.g. neural networks, embedding, reinforcement learning, Bayesian inference and Markov games, clustering, active learning, and inductive logic programming), some of them specific to agents, e.g. multi-agent reinforcement learning (MARL). Learning and reasoning will provide agents with the ability to more flexibly, dynamically, and better perform their tasks. Some challenges in this context come from articulated learning at the agent level and at the organisation level, dealing with implicit objects (only modelled in other agents) and explicit objects (that may be embodied in the web). As a typical example of this, we choose to focus on norms and more precisely behavioural norms.

This is a summary of discussions in the Working Group on Learning and Reasoning with Behavioural Norms. We chose to focus on norms as a topical area of research to focus our discussions, rather than trying to cover learning and reasoning more broadly.

5.5.1 Motivations and Bridges to other Working Groups

Societies of agents are constantly confronting a changing world, environment, and society. They have to learn how to behave in such a context and to reason on how to exploit what they have learned. Learning and reasoning connects well with the other topics discussed during this Seminar, in particular in connection with the following questions:

- How can we incorporate metadata about the data used for learning and reasoning, including metadata to determine its provenance and trustworthiness?
- How can we store and query the data needed for learning and reasoning in a decentralised manner, as in pods for individual agents?
- How can learning and reasoning methods accommodate models of actuators, including through affordances, in the design of an environment through which an agent can observe and control the environment?
- How can agents create and execute collaborative plans?
- How can agents interact with each other felicitously, through learning and reasoning about norms?
- How may flexible governance be accomplished by learning and reasoning about norms, especially to figure out the boundaries between competing norms in a contextual manner?

Because the topic is so wide, we choose to focus on one particular kind of objects: norms.

5.5.2 Learning and Reasoning about Behavioural Norms

Norms are patterns of behaviour for group members in which compliance or violations of the implicit or explicit rules may lead to sanctioning. Norms vary in their extent of explicitness or implicitness and may be prescriptive (including proscriptive) or descriptive. Prescriptive norms may carry the force of legal regulations and laws. Sanctions may be positive or negative and may apply to an individual or a group and be given by an individual or a group [3]. One example is for drivers of vehicles on public roads, self-driving vehicles and humans, who will all need to follow the same norms to get along smoothly. Some of these norms are explicitly described in the highway code, whilst others remaining implicit, or tacit, must be learned through experience. Sanctions may involve honking the horn or flashing the vehicle's headlights – besides actual fines for violating prescriptive norms, not solely descriptive ones.

On the web, where people and software agents meet, they both have to follow social norms. It is therefore interesting to consider groups involving combinations of people and agents. The agents are intelligent provided they can learn and reason as they interact with their environment, which includes other group members. Reasoning involves making decisions about actions with the environment, or changing the agent's beliefs, desires, and intentions. Beliefs may also cover the agent's understanding of people and other agents (as in the Theory of Mind modelling).

Some questions that arise:

- What is the relationship between norms and governance?
- Are we using agents to model people or are we investigating social norms for artificial agents?
- What are suitable metamodels for explicit norms?
- How can we describe behaviours in a symbolic or sub-symbolic form (e.g. using deep neural networks or other embedding techniques)?
- How do we revise or update the rules of individual agents that describe their behaviour?
- How can we model heterogeneous communities of agents with different social values?
- How can we avoid unfair norms that are discriminatory to certain classes of users and agents?
- How can we deal with hostile or malicious agents?
- How are agents supposed to know and learn explicit and tacit norms?
- How can we evaluate the effectiveness of learning social norms?

One approach to learning social norms uses multi-agent reinforcement learning based upon (negative) sanctioning, where an agent, or a group of agents, makes known its disapproval of another agent, based upon the second agent's behaviour. The disapproval could be communicated one-to-one to the agent deemed to have violated the norms, or publicly to the group as a whole.

5.5.3 Use Cases Involving Norms

We chose use cases that make the connection, at least in some respects, with agents and the web but require learning and reasoning. They show the practical use of the following aspects:

- Understanding the benefit of incorporating norms with respect to agents on the Web
- Learning, revising, and reasoning about behavioural norms
- Working with implicit and explicit norms
- Supporting norms in relation to planning tasks
- Supporting norms in relation to causal reasoning – understanding intent on the basis that most people/agents will follow the norm, and recognising violations

The use cases that we considered were as follows:

- Ride sharing across town balancing personal desires and needs, following norms, as well as continual learning of norms and related metadata
- Social media assistant advising you (in a dialogue) whether you would violate the norms should you go ahead and send your draft post, and/or for directing posts to moderators for possible sanctions. This is in the context of federated social networks such as Mastodon.
- Checking compliance with regulations such as GDPR and determining how they should evolve to reduce violations [2]: tighter or looser?

The ride-sharing use case is as follows: volunteers or paid drivers offer to drive people across town, e.g. to shops, medical appointments, and so on [1]. A car may be shared by multiple passengers with different pick-up and drop-off locations. Some people may be limited in how far they can walk. What are the rewards for the drivers to drive and for the passengers to accommodate each other's needs? How can software learn metadata, norms, and sanctions from the journeys undertaken? How can we match drivers to passengers dynamically? How can we integrate with information sources providing the essential live or static knowledge, e.g. about roadwork? The people who want a ride should be able to offer suggestions and express their preferences regarding the origin, destination, and timing of their journey.

What is the contribution of plausible knowledge and statistical models? How can we unite such knowledge with argumentation theory for assessing and balancing conflicting desires and constraints? How can we employ other approaches, e.g. constraint satisfaction algorithms, to facilitate smooth functioning and collaboration?

What kinds of norms arise in this use case? For example, a societal objective would be that the elderly or people with walking sticks should be dropped off close to their destination, and likewise picked up close to their starting points. Journey times should not be unreasonably long for any passenger. Journeys should be navigated to avoid known problems (road works, traffic jams).

How can this use case be implemented in a decentralised manner, where the information belonging to different users stays with them and in their control and where their decision making is likewise local?

5.5.4 Identified Challenges

We identified the following challenges and places for design decisions.

- Should the treatment of norms be data-driven or model-driven? We expect both in various ways but capturing their interplay is non-trivial.
- What are the tradeoffs between tacit and explicit norms?
- What are suitable standards for expressing explicit norms as a basis for reasoning about norms and the decisions to take in reference to the norms?
- How can agents incorporate Theory of Mind reasoning, involving beliefs, desires, and intentions (BDI) and other folk psychological notions?
- How can agents collectively learn norms in a federated manner without revealing confidential information of their users to others?
- How can learning and reasoning about norms be performed at the subsymbolic level?
- What are constraints on agent-to-agent communications, e.g., those that affect the learning performed individually and jointly by the agents?
- What would be the nature of an approach based upon publicly sharing information on norm violations, with such information sharing improving reasoning and acting as a potential deterrent to violators?

- How does learning and reasoning about norms relate to trust and trustworthiness, and provide a foundation for the construction of trustworthy systems? Agents should learn to behave in a trustworthy manner and should learn to calibrate their trust in another agent with its observed trustworthiness.

5.5.5 Conclusion

Enabling agents to learn and reason about behavioural norms will be increasingly important given the rapid pace of advances in artificial intelligence, and the need to ensure that applications enrich human society. This raises many challenging issues that we only had the opportunity to skim through in this report.

Acknowledgements

JE was partially supported by the French Agence Nationale de la Recherche (grants ANR-17-CE23-0007-01 and ANR-19-P3IA-0003). MPS thanks the US National Science Foundation (grant IIS-1908374) for partial support for this research.

References

- 1 Eleni Bardaka, Leila Hajibabai, and Munindar P. Singh. Reimagining ride sharing: Efficient, equitable, sustainable public microtransit. *IEEE Internet Computing (IC)*, 24(5):38–44, September 2020. doi:10.1109/MIC.2020.3018038.
- 2 Andreea Morris Martin, Marina De Vos, Julian Padget and Oliver Ray. Agent-directed runtime norm synthesis. Proc. 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS), London (UK), 2023.
- 3 Luis G. Nardin, Tina Balke-Visser, Nirav Ajmeri, Anup K. Kalia, Jaime S. Sichman, and Munindar P. Singh. Classifying sanctions and designing a conceptual sanctioning process model for socio-technical systems. *The Knowledge Engineering Review (KER)*, 31(2):142–166, March 2016. doi:10.1017/S0269888916000023.

5.6 Challenges and Opportunities in Governing Agents on the Web

Matteo Baldoni (University of Turin, IT)

Victor Charpenay (Mines Saint-Étienne, FR)

Andrei Ciortea (Universität St. Gallen, CH)

Stephen Cranefield (University of Otago, NZ)

Julian Padget (University of Bath, GB)

Munindar P. Singh (North Carolina State University – Raleigh, US)

License  Creative Commons BY 4.0 International license

© Matteo Baldoni, Victor Charpenay, Andrei Ciortea, Stephen Cranefield, Julian Padget, Munindar P. Singh

This paper summarizes the findings of the working group “Governance” at the Dagstuhl Seminar on *Agents on the Web*, which was held in February 2023. This report seeks to identify research challenges relating to governance and the Web architecture.

5.6.1 Introduction

There are major similarities in the motivations behind multiagent systems (MAS) and the Web. Both disciplines and practices seek to advance decentralization and openness in that ideally there is not a single locus of control and agents can behave and interact broadly autonomously under local control.

This report addresses the interplay of multiagent systems with the Web. Specifically, it concerns how constructs and techniques identified in the study of the governance of MAS can be realized over the Web architecture and how the governance of the Web can be beneficially structured based on constructs and techniques developed for the governance of MAS. In simple terms, it seeks to answer the following two questions:

- *What does the Web offer to support the governance of MAS?*

We anticipate ways to use the scalability and evolvability of the Web to build easy-to-use, widely deployed MAS. Scalability and evolvability are two of the several non-functional requirements that can easily be met if a system's architecture relies on the Web.

- *What does the governance of MAS offer the Web?*

We anticipate approaches for governance that provide flexibility and local control with formal models that support correctness and generality going beyond the typically procedural kinds of governance seen on the Web today. A challenge here is to map abstract models for governance in MAS to Web components, in a way that preserves Web architectural constraints and thereby guarantees the associated non-functional requirements of scalability and evolvability.

The main contribution of this report is the identification of crucial challenges pertaining to the interplay of MAS and the Web, and the formulation of some initial research questions that might guide future research on this topic.

5.6.2 Key Concepts and Considerations

We understand the Web as a collection of resources, identified with uniform resource identifiers (URIs) and supporting hyperlinked representations, along with a computational architecture that supports locating and accessing the identified resources. The computational architecture is based on standard protocols for manipulating resources (e.g., HTTP, CoAP²⁹). We think here of architectural constraints (such as for caching, layering, and uniform interaction) as captured in the original design rationale for the Web Architecture [18], but see also the W3C Recommendation for the Web Architecture [22]. Some of these constraints, especially uniform interaction, are captured by the Linked Data principles [21], which can be supplemented by an ontology specification [34]. Some extensions to the above architecture, such as through the observer pattern (implemented in CoAP [28]) and local state transfer [30], aim at going beyond the “Web of Documents” [19].

There are numerous studies relating the Web and MAS. These exhibit two general trends: some focus on applying RDF and Linked Data to expose agents to hypermedia-driven environments [4, 10, 16, 25], while others combine a formal declarative model of norms [7, 12, 26, 31] to specify social protocols [8]. Such social protocols provide a more thoroughly decentralized conception of Berners-Lee's [1] notion of social machines. Recent W3C Recommendations for the Social Web, including Linked Data Notifications, ActivityStream, ActivityPub and WebSub [20], offer ways to connect these two trends, by implementing social protocols over Linked Data.

²⁹ Constrained Application Protocol: <https://datatracker.ietf.org/doc/html/rfc7252>

Dimensions that are common to all MAS architecture, such as the environment, organization and interaction dimensions, are defined at a higher level of abstraction than Web resources and protocols. Constraints such as caching, layering and uniform interaction apply to components, exposing a certain functionality through ports. Web components may only have client or server ports, exchanging messages in a standard protocol. Components with client ports only are called origin clients, those with server ports only are origin servers and a third kind of component, proxies, have an equal number of client ports and server ports, forwarding requests from clients to servers or vice versa [18].

To be able to analyse the interplay between MAS and the Web, a mapping from MAS abstractions to (more concrete) Web components is necessary. Given the complexity of both fields, there is no trivial mapping and most likely not a unique mapping across the two levels of abstraction. In the following, we perform a case study to help identify some preliminary mappings that would transfer effective governance mechanisms developed in MAS research.

5.6.3 Example Scenario: Organ Allocation

Consider a simplified version of the process for allocating donated organs and tissue to potential transplant recipients called Carrel [33]. The distribution of organs and tissues in Carrel, given its Spanish and Catalan context, would be overseen by the Spanish ONT, together with the Catalan Organització CATalana de Trasplantament (OCATT)³⁰.

Tissue distribution is essentially demand-driven because tissues can be preserved and stored over extended periods with no significant degradation. Organ distribution is essentially supply-driven because the need is known before a suitable part becomes available. For the purposes of this report, we only consider the second case of organ distribution where the need is known before availability.

To achieve its goal, the requesting agent must negotiate with the agents that represent hospitals with potential donors. In the original version of Carrel, participating agents are in effect regimented by so-called governor agents, to prevent non-compliance, but in general, agents may choose to take non-compliant actions. Thus, we assume that agent actions in this contemporary Carrel are all visible to the regulatory bodies, as is the case in the physical world: hospital Transplant Coordination Unit agents (TCUs) communicate their requests to members of ONT/OCATT, who then contact other hospitals to find a donor and select the best match. To participate in Carrel, each agent must hold an authorisation in the form of a certificate.

5.6.4 Relevant Multiagent Systems Approaches

We briefly highlight some relevant multiagent systems approaches.

5.6.4.1 Norms and Roles

Early approaches used agent-mediated electronic institutions [33] to model Carrel by capturing the structure of the interactions between hospitals, tissue banks, and institutional agents managing the process, as well as the norms that govern these interactions and the match between a donor and recipient [17]. Following one approach from the MAS literature [6], we can capture the sociotechnical system requirements in terms of accountability [9] and only then proceed to identify the information exchanges between the agents and hence their individual actions.

³⁰<https://trasplantaments.gencat.cat>

5.6.4.2 Allocation Protocol

Consider a simplified *Organ Donation* protocol for TCU agents. The agent holding a surgeon's request for transplant sends its request to another agent, which responds with a donation offer. The requesting agent then has the choice of confirming or rejecting the offer, presumably depending on the strength of this match against any other offers it may have received from other prospective donor hospitals. To facilitate the modeling of norms, such interactions could be captured declaratively in languages such as the Blindly Simple Protocol Language (BSPL) language [29].

5.6.4.3 Norm Representation and Monitoring

Over the last 30 years, researchers in MAS have proposed many approaches to reasoning about norms as well as computational approaches to monitoring a MAS for norm violations [2, 5, 13]. A crucial aspect of this work is to provide a formal representation of norms (see, e.g., [15] for an overview of approaches). For example, the Expectation Event Calculus (EEC) [14] is applicable to this problem.

5.6.5 Candidate Architectures

On the Web, servers hold the power and may hide information, redirect requests or reject them, thereby reducing the action available to agents. We may think of institutions as components with a server port, receiving requests from agents. Unlike a “user agent” (browser) in the Web architecture, an autonomous agent might have both client and server roles simultaneously and can have one-sided elementary interactions, where they are not awaiting a response. The requirements of autonomous agents are closer to those of servients in the lingo of the W3C Web of Things Architecture [24], which are components with both client and server roles that can interact in a peer-to-peer manner. If the agent is visible to other agents, i.e. if its representation is dereferenceable, the representation can point, for instance, to a Linked Data Notification inbox that receives messages from other agents [3].

An institution may not need to materialize as a Web component: if norms are defined at design time, agents may be guaranteed to behave (in general) as per these norms, and certification may occur at design time, such that an agent either uses a single certificate throughout the system's lifetime or periodically renews its certificate. Such an institution would have no sanctioning power, nor any need for sanctioning. It cannot easily change the applicable norms either. Updating a norm would potentially require modifying the behavior of all agents at the same time. This is a basic, but inflexible solution.

5.6.5.1 Institution as Read-Only Server

The behavior of an agent may be decoupled from the norms that regulate it, though: if norms are exposed by a read-only (origin) server, an agent may dereference the norms from time to time and internalize whatever formal specifications the server returns. In this configuration, the institutional component has the power of dictating and changing norms at run time. The ability of a

5.6.5.2 Institution as Read-Write Server

In order for the institution to gain sanctioning power, another component may manage its real-time state. The state of the institution includes the level of compliance of each participating agent, which is directly derived from the confirmations/rejections they generate.

To be able to maintain its state, the institutional component must be able to observe each agent-to-agent interaction. For instance, the certificate may be signed not for an agent but for a pair (agent, donation offer ID), forcing agents to request a new certificate every time they make a decision. If the certification server stores a history of confirmations/rejections, it effectively becomes an institutional component that decides in real time whether agents violate norms and, if they do, to sanction them by rejecting their certification request.

The institutional server would become a stateful read-write component, as agents, through their certification requests, change the state of the overall institution. Yet, it remains a purely reactive component, with a single server port.

5.6.5.3 Institution as Proxy

In the above configuration, the institutional component has no knowledge of how agents negotiate. If the institution is to be omniscient, another kind of component should be used. On the Web, it is common to use proxy servers to monitor activity.

The main architectural constraint over proxies is that they have a client port and a server port, such that incoming requests (on the server port) are either immediately responded to or forwarded to another server, possibly after a rewriting step. In our working example, the institutional server may be replaced by a proxy without modifying in any way the behavior of agents. Agents send requests to the proxy, which can keep track of negotiations and add a certificate on-the-fly if the requesting agent behaves properly. The proxy may turn a confirmation into a rejection, to sanction any misbehaving agent. The requesting agent receives feedback on the sanction through the other agent (which acknowledges the rejection, instead of the initial confirmation).

5.6.5.4 Institution as Servient

An alternative is to capture the institution as an explicit stakeholder supported by an agent on par with the other parties in the system. The institution becoming both reactive and proactive, hence must include independent client and server ports, and becomes a servient.

In ordinary operations, this component may have little to say beyond conveying institutional norms and facts as in the previous approach. However, by identifying this institutional entity, we make it subject to accountability. A party can also question the institution, for example, if they fail to receive an organ in a timely fashion. This process may result in the institutional facts being disputed and adjudicated [32] and the norms potentially revised.

If the institution is embodied by an agent, its monitoring and sanctioning power does not depend on architectural constraints (at the level of Web components) but on the behavior of other agents (at the level of MAS abstractions).

5.6.6 Hypermedia-driven Interaction

Hypermedia-driven interaction can support autonomous agents to interact with Web resources in a uniform way while being decoupled from the underlying components. To illustrate how this works, an HTML page typically provides the user with a number of actions, such as clicking a hyperlink or submitting a form. Performing any such action transitions the user to a new page and exposes a new set of possible actions. In each step, the user's browser retrieves not only an HTML representation of the current page from a server but also the hypermedia controls required to transition to new pages. Hypermedia-driven interaction reduces coupling between components (e.g., browsers, proxies, origin servers) and allows them to evolve independently from one another; a central feature that allowed the Web to scale up to the size of the Internet.

In the present case, the various actions – retrieving formal specifications of norms, requesting a certificate, or sending messages to other agents – can be made available to the agents through hypermedia controls. Such hypermedia controls would encapsulate all the information required by an agent to interact with the institution, but to use the hypermedia controls in a reliable manner the agent would have to operate on an abstract model of the institution. For example, if an agent is required to obtain a certificate at run time to enact an *Organ Donation* protocol, the agent could discover such an action possibility through hypermedia.

The institutional model could evolve throughout the agents' lifetimes: for example, from using norms and certificates defined at design time to a model based on evolving norms [27] and certificates to be obtained at run time. Such an evolution would be reflected in the hypermedia environment through the action possibilities provided to agents at run time and thus, to cope with this evolution, the agents would adapt to a new course of action that meets their objectives. Some related work investigates the design of agents able to plan and adapt to dynamic hypermedia environments (e.g., see [11, 23]).

5.6.7 Discussion: Research Questions and Challenges

The contribution of this report is in identifying some new research questions that can motivate research on the interface of MAS and Web architectures. Specifically, we propose the following:

1. How should we model the presence of a governance layer in a MAS?
2. What aspects of a web-based deployment of a MAS may be subject to governance policies and included in an institutional environment?
3. How do we map the required properties of an institutional environment (monitoring, reasoning, sanctioning power) to constraints and mechanisms of a web-based deployment and more specifically to hypermedia controls?
4. How can we specify and implement composed and stacked agent environments (e.g. institutional environments that overlay physical ones and institutional environments that extend others)?
5. Can and should a governed MAS on the Web be modelled and implemented as a hypermedia application?

5.6.8 Conclusion

Thinking about MAS and the Web together opens up new opportunities in building large-scale sociotechnical systems. Such systems would take advantage of the flexibility derived from MAS and the scalability and familiarity (to most developers) derived from the Web. The possibilities are promising and we invite the research community to join us in investigating them.

Acknowledgments

The working group participants would like to thank Dame Wendy Hall and Mehdi Dastani for the input they provided at Dagstuhl. Victor Charpenay thanks the French Agence nationale de la recherche (grant ANR-19-CE23-0030-01) for partial support for this research. MPS thanks the US National Science Foundation (grant IIS-1908374) and the Department of Defense (Science of Security Lablet) for partial support for this research. Andrei Ciortea

thanks the Swiss National Science Foundation (grant 189474) for partial support for this research.

References

- 1 Tim Berners-Lee. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*. Harper Business, New York, 1999.
- 2 Guido Boella, Leendert W. N. van der Torre, and Harko Verhagen. Introduction to normative multiagent systems. *Computational & Mathematical Organization Theory*, 12(2-3):71–79, 2006. doi:10.1007/s10588-006-9537-7.
- 3 Sarven Capadislil and Amy Guy. Linked data notifications, 2017. URL: <https://www.w3.org/TR/ldn/>.
- 4 Victor Charpenay, Tobias Käfer, and Andreas Harth. A unifying framework for agency in hypermedia environments. In Natasha Alechina, Matteo Baldoni, and Brian Logan, editors, *Engineering Multi-Agent Systems*, pages 42–61. Springer International Publishing, 2022.
- 5 Amit Chopra, Leendert van der Torre, Harko Verhagen, and Serena Villata, editors. *Handbook of Normative Multiagent Systems*. College Publications, 2018.
- 6 Amit K. Chopra, Fabiano Dalpiaz, F. Başak Aydemir, Paolo Giorgini, John Mylopoulos, and Munindar P. Singh. Protos: Foundations for engineering innovative sociotechnical systems. In *Proceedings of the 22nd IEEE International Requirements Engineering Conference (RE)*, pages 53–62, Karlskrona, Sweden, August 2014. IEEE Computer Society. doi:10.1109/RE.2014.6912247.
- 7 Amit K. Chopra and Munindar P. Singh. Custard: Computing norm states over information stores. In *Proceedings of the 15th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1096–1105, Singapore, May 2016. IFAAMAS. doi:10.5555/2936924.2937085.
- 8 Amit K. Chopra and Munindar P. Singh. From social machines to social protocols: Software engineering foundations for sociotechnical systems. In *Proceedings of the 25th International World Wide Web Conference*, pages 903–914, Montréal, April 2016. ACM. doi:10.1145/2872427.2883018.
- 9 Amit K. Chopra and Munindar P. Singh. Accountability as a foundation for requirements in sociotechnical systems. *IEEE Internet Computing (IC)*, 25(6):33–41, September 2021. doi:10.1109/MIC.2021.3106835.
- 10 Andrei Ciortea, Olivier Boissier, and Alessandro Ricci. Engineering world-wide multi-agent systems with hypermedia. In Danny Weyns, Viviana Mascardi, and Alessandro Ricci, editors, *Engineering Multi-Agent Systems*, pages 285–301, Cham, 2019. Springer International Publishing.
- 11 Andrei Ciortea, Simon Mayer, and Florian Michahelles. Repurposing manufacturing lines on the fly with multi-agent systems for the web of things. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '18*, page 813–822, Richland, SC, 2018. International Foundation for Autonomous Agents and Multiagent Systems.
- 12 Owen Cliffe, Marina De Vos, and Julian Padget. Specifying and reasoning about multiple institutions. In *COIN 2006*, pages 63–81, 2007. doi:10.1007/978-3-540-74459-7_5.
- 13 Rosaria Conte, Rino Falcone, and Giovanni Sartor. Introduction: Agents and norms: How to fill the gap? *Artificial Intelligence and Law*, 7(1):1–15, 1999. doi:10.1023/A:1008397328506.
- 14 Stephen Cranefield. Agents and expectations. In *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*. Springer, 2013. doi:10.1007/978-3-319-07314-9_13.

- 15 Stephen Cranefield, Michael Winikoff, and Wamberto Vasconcelos. Modelling and monitoring interdependent expectations. In Stephen Cranefield, M. Birna van Riemsdijk, Javier Vázquez-Salceda, and Pablo Noriega, editors, *Coordination, Organizations, Institutions, and Norms in Agent System VII*, pages 149–166. Springer Berlin Heidelberg, 2012.
- 16 Oğuz Dikenelli, Oylum Alath, and Rıza Cenk Erdur. Where are all the semantic web agents: Establishing links between agent and linked data web through environment abstraction. In Danny Weyns and Fabien Michel, editors, *Agent Environments for Multi-Agent Systems IV*, pages 41–51. Springer International Publishing, Cham, 2015.
- 17 M. Esteva, J. Padget, and C. Sierra. Formalizing a language for institutions and norms. In Jean-Jules Meyer and Milind Tambe, editors, *Intelligent Agents VIII*, volume 2333 of *Lecture Notes in Artificial Intelligence*, pages 348–366. Springer Verlag, 2001. ISBN 3-540-43858-0.
- 18 Roy T. Fielding and Richard N. Taylor. Principled design of the modern web architecture. *ACM Trans. Internet Technol.*, 2(2):115–150, May 2002. doi:10.1145/514183.514185.
- 19 Roy T. Fielding, Richard N. Taylor, Justin R. Erenkrantz, Michael M. Gorlick, Jim Whitehead, Rohit Khare, and Peyman Oreizy. Reflections on the REST architectural style and “principled design of the modern web architecture” (impact paper award). In *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering, ESEC/FSE 2017*, page 4–14, New York, NY, USA, 2017. Association for Computing Machinery. doi:10.1145/3106237.3121282.
- 20 Amy Guy. Social web protocols, 2017. URL: <https://www.w3.org/TR/social-web-protocols/>.
- 21 Tom Heath and Christian Bizer. Linked Data: Evolving the Web into a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1(1):1–136, February 2011. doi:10.2200/S00334ED1V01Y201102WBE001.
- 22 Ian Jacobs and Norman Walsh. Architecture of the world wide web, volume one. Technical report, World Wide Web Consortium, 2004. URL: <https://www.w3.org/TR/webarch/>.
- 23 Matthias Kovatsch, Yassin N. Hassan, and Simon Mayer. Practical semantics for the internet of things: Physical states, device mashups, and open questions. In *2015 5th International Conference on the Internet of Things (IOT)*, pages 54–61, 2015. doi:10.1109/IOT.2015.7356548.
- 24 Michael Lagally, Ryuichi Matsukura, Michael McCool, and Kunihiko Toumura. Web of things (WoT) architecture 1.1. Technical report, World Wide Web Consortium, 2023. URL: <https://www.w3.org/TR/wot-architecture/>.
- 25 Eoin O’Neill, David Lillis, Gregory M. P. O’Hare, and Rem W. Collier. Delivering multi-agent microservices using cartago. In Cristina Baroglio, Jomi F. Hubner, and Michael Winikoff, editors, *Engineering Multi-Agent Systems*, pages 1–20, Cham, 2020. Springer International Publishing.
- 26 Julian Padget, Marina De Vos, and Charlie Ann Page. Deontic sensors. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 475–481. International Joint Conferences on Artificial Intelligence Organization, 7 2018. doi:10.24963/ijcai.2018/66.
- 27 Bastin Tony Roy Savarimuthu and Stephen Cranefield. Norm creation, spreading and emergence: A survey of simulation models of norms in multi-agent systems. *Multiagent and Grid Systems*, 7(1):21–54, 2011. doi:10.3233/MGS-2011-0167.
- 28 Zach Shelby, Klaus Hartke, and Carsten Bormann. The Constrained Application Protocol (CoAP). Technical Report RFC 7252, Internet Engineering Task Force (IETF), Fremont, California, June 2014. Proposed standard; <https://tools.ietf.org/html/rfc7252>.
- 29 Munindar P. Singh. Information-driven interaction-oriented programming: BSPL, the Blindingly Simple Protocol Language. In *AAMAS-11*, pages 491–498, Taipei, May 2011. IFAAMAS. doi:10.5555/2031678.2031687.

- 30 Munindar P. Singh. LoST: Local State Transfer – An architectural style for the distributed enactment of business protocols. In *Proceedings of the 9th IEEE International Conference on Web Services (ICWS)*, pages 57–64, Washington, DC, July 2011. IEEE Computer Society. doi:10.1109/ICWS.2011.48.
- 31 Munindar P. Singh. Norms as a basis for governing sociotechnical systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):21:1–21:23, December 2013. doi:10.1145/2542182.2542203.
- 32 Pankaj R. Telang, Anup K. Kalia, John F. Madden, and Munindar P. Singh. Combining practical and dialectical commitments for service engagements. In *Proceedings of the 13th International Conference on Service-Oriented Computing (ICSOC)*, number 9435 in Lecture Notes in Computer Science, pages 3–18, Goa, India, November 2015. Springer. doi:10.1007/978-3-662-48616-0_1.
- 33 Javier Vázquez-Salceda, Ulises Cortés, Julian Padget, Antonio López-Navidad, and Francisco Caballero. The organ allocation process: A natural extension of the Carrel agent-mediated electronic institution. *AI Communications*, 16(3):153–165, 2003.
- 34 W3C. OWL 2 Web Ontology Language: Document overview, December 2012. W3C Recommendation. Accessed 2023-03-03. URL: <https://www.w3.org/TR/owl2-overview/>.

5.7 Agents on the Web: Specifying and Implementing Testbeds

Cleber Jorge Amaral (Federal Institute of Santa Catarina – São José, BR)

Jean-Paul Calbimonte (HES-SO Valais – Sierre, CH)

Jomi Fred Hübner (Federal University of Santa Catarina, BR)

Timotheus Kampik (Umeå University, SE & SAP Signavio – Berlin, DE)

Tobias Käfer (KIT – Karlsruher Institut für Technologie, DE)

License © Creative Commons BY 4.0 International license

© Cleber Jorge Amaral, Jean-Paul Calbimonte, Jomi Fred Hübner, Timotheus Kampik, Tobias Käfer

5.7.1 Introduction

Agents are often challenging to test due to their complexity and the unpredictability and dynamism of the environments in which they are situated. It is considered infeasible to test all courses of actions that complex agents can take [1, 2]. Even for one particular goal, agents may have several alternative plans or may apply non-deterministic approaches, for example in the context of data-driven methods such as machine learning. In addition, testing stochastic scenarios and situations that entail interaction with other agents and humans can be challenging to reproduce and control [3, 4]. Ensuring that an agent behaves as desired in a relatively friendly environment is already challenging; assuring this at a scale that includes numerous autonomous software artifacts running relatively freely in the environment is a crucial problem [2].

For the specific case of agents on the Web, it is important to consider the Web as *the* decentralized environment where interactions happen, rendering the governance of agent on the Web a key challenge [5]. Interactions can take places between heterogeneous agents (e.g., using usual agent communication languages) but also between agents and other Web entities, applying various of actuators and sensors. We consider the case where the agent environment is identified by a set of resources that are published and made accessible on the Web. While these resources can be linked to real entities, like sensors, robots, actuators, drones, or smart

devices, they can also be linked to virtual entities and services. The *situatedness* of these agents is an integral part of the Web environment, where resources and other agents are identified and located through URIs.

This chapter reports the main results of the working group “Benchmark and Testbeds”, as discussed during the Dagstuhl Seminar “Agents on the Web”. The different challenges regarding testing and benchmarking of agents on the Web were discussed, and specific scenarios and scenarios types for testing and benchmarking were proposed. Through the proposed set of scenarios and scenario types, we identified potential future work regarding the conception and implementation of testbeds for Web Agents.

5.7.2 Testing and Benchmarking of Agents on the Web

The Web provides numerous opportunities for the use of agents and multi-agent systems that are diverse with respect to use cases and underlying technology stacks. In this section we introduce a list of categories of scenario types that can be used for testing and benchmarking MAS on the Web. We assess the scenario categories in terms of how easy (if possible) it is to apply techniques for testing autonomous software artifacts, such as unit, integration, regression, stress, and beta testing, and how to benchmark different application development approaches. A comparison of the categories considering their strengths and weaknesses is provided in Table 5.

The first scenario categories is based on legacy Web-based applications used in a “production-like” environment, i.e. in an environment that reflects the socio-technical complexity of a system whose operation has critical real-world implications, for example in a business context. The actual environment, which can be virtual or physical, is the “real-world” setting in which applications and devices are operating. Because each system has a distinct set of requirements and constraints, throughout the years several protocols (e.g.: CORBA, OPC, RPC) and architectural styles (e.g.: SOAP, REST, GraphQL) were adopted [6]. Legacy Web-based applications are usually a composition of different technologies that are integrated into a functionally somewhat coherent and cohesive system of high technological heterogeneity. Besides the lack of standardization, legacy web-based applications often have many human-machine interfaces of different maturity levels and frequently systems are hard-wired to each other via system actions on graphical user interfaces (in the case of so-called robotic process automation). Navigating this interface heterogeneity can be a key challenge for agents on the (current) Web. Another characteristic is that such applications have little space for mistakes when deploying new versions since they are usually in day-by-day usage by many users and often perform business-critical transactions. Examples of legacy web-based applications are: ERP systems, industrial automation plants [7] and digital marketplaces. One advantage of adopting this scenario category is that it exposes the system under test to a realistic environment which can facilitate software maturity and resilience, in particular because it reflects real-world requirements and performance indicators, allowing for meaningful benchmarking of the current system as well as of partial or full replacements that are technologically more advanced.

The third scenario category is based on simulated environments. Simulated environments can provide reasonably realistic constraints and intuitive visual interfaces for many scenarios. They are often built using simulation frameworks such as Gazebo³¹ and Morse³² which provide

³¹ Gazebo is available at <https://gazebo.org/>

³² Morse is available at <http://morse-simulator.github.io/>

■ **Table 5** Overview of scenario categories, with strengths and weaknesses.

Scenario category	Strengths	Weakness
Heterogeneous Web-based applications in production-like environments	<ul style="list-style-type: none"> • Realistic conditions and use of agents • Easier identification of system performance and compliance • Easier understanding of user interactions impacts over the system 	<ul style="list-style-type: none"> • Systems are often rigid with not much opportunity to apply artificial agents • It is often hard to reproduce some situations that occurs in the actual environment • It can be time-consuming to test applications in such scenarios • It is often challenging to gain access to such scenarios due to privacy and security concerns • Scenarios are often in critical missions with little space for experimentation
Simulated environments	<ul style="list-style-type: none"> • Usually less complex than production environments and easier to grasp for humans • Testing conditions are easier to control and reproduce • Sets of tests can be performed quickly • Usually more cost-effective than testing than production environments 	<ul style="list-style-type: none"> • May lack important characteristics and interactions that can be found in production environments • Time-consuming to develop a scenario with the necessary accuracy
Programming contest scenarios	<ul style="list-style-type: none"> • Has well-defined performance indicators that facilitates comparisons • Competitions often provide more realistic testing conditions than traditional simulators • Development and improvement is made easier by the fact that such scenarios frequently call for a limited range of actions 	<ul style="list-style-type: none"> • May lack important characteristics and interactions that can be found in real-world environments • Testing conditions are limited to the contest proposal
Testbeds	<ul style="list-style-type: none"> • Realistic conditions and use of agents • Easier identification of system performance and compliance • Has well-defined performance indicators that facilitates comparisons 	<ul style="list-style-type: none"> • Complex to define (requires substantial engineering effort) • It can be time-consuming to test applications in such scenarios

tooling for applying physical constraints, and for drawing and animating representations of the environment, thus facilitating human interpretability. Such frameworks also provide tools for monitoring and debugging systems. There are many possible scenarios that can be simulated, such as transportation and search-and-rescue scenarios. As the scenarios can be simpler than the actual environment, simulated environments can facilitate the development of applications. However, the application is not exposed to all situations that can potentially occur in the real world.

Another scenario category are programming competitions, such as MAPC³³ and RoboCup³⁴. These competitions provide different scenarios that can run either simulated or real-world environments. A distinguishing characteristic of this scenario category is that in order to facilitate competition, the scenarios have well-defined performance indicators, which enables the comparison of the different approaches.

Finally, there are scenarios provided as testbeds. Testbeds are defined for experimenting and comparing approaches [8]. For instance, a scenario may allow agents to enter buildings and rooms therein in order to interact with appliances such as light switches, which in turn affects the state of the environment. This scenario shares characteristics of simulated and competition scenarios as it is defined for a small set of actions; it is not supposed to be in use in production, and its performance requirements and indicators can be easily defined and monitored. Ideally, testbeds share some of their characteristics with production-like scenarios

³³ More information about the MAPC can be found at <https://multiagentcontest.org/>

³⁴ More information about the RoboCup can be found at <https://www.robocup.org/>

as they are usually developed using real devices and run in a (potentially cyber-physical) realistic environment. In the context of the WoT, due to the lack of existing simulated and contest scenarios and limited access to WoT scenarios in real-world environments, testbeds can provide ideal conditions for testing and benchmarking MAS on the Web.

5.7.3 Technical Features of Agents on the Web

In this section, we propose a synthesis of technical terms of related to the engineering of agents & MAS and existing standards and practices from the Web.

5.7.4 Interaction Levels for Agents

Envisioned and existing agents on the Web perform tasks on different levels of interaction with the environment and other agents, which can get categorized into:

1. Data Gathering – such agents crawl the web in order to gather information, and may exhibit a querying interface. That is, they operate in read-only mode with safe HTTP requests.
2. Actuation – such agents may also read information, but their characteristic is that they additionally perform writing operations on the environment using unsafe HTTP requests.
3. Communication – such agents engage in conversation with other agents. To this end, they also perform writing operations, but to specific endpoints of other agents, with the goal of communicating.

Note that in the context of testing and benchmarking, in order to provide means for comparing approaches and MAS applications, the focus is “top-level goals rather than approach-specific goals” [9].

5.7.5 Agent Terminology on the Web

In order to combine MAS and Web approaches, a common view on what can constitute agenthood on the Web needs to be established. Depending on the above interaction levels for agents, it is especially on level 3, where the following aspects are required:

- Identity – an agent needs an identifier, i.e., a URI on the Web.
- Endpoint – an agent needs a way to receive communication, e.g., an LDN inbox.
- Entry point – an agent needs a way to expose information, e.g., an FOAF profile document on the Web.

The other levels 1 and 2 require other aspects:

- Situatedness – the agent needs a reference to where it is. On the Web, this could be a (set of) seed URI(s) for a data gathering agent; or an actuating agent could be constrained in its behavior using a query.
- Norms – to formulate desired agent behavior using norms, a notion of norms for the Web needs to be established. This could be implemented as agent behavior descriptions, for which there are no standards and practices yet.
- Signifiers / possible actions – for an agent to find out what it may do in an environment, descriptions of actuatable items are required; on the Web there are, e.g. `schema:PotentialAction`, which represents a coarse match to the agent terminology.
- Adaptivity – an adaptive agent can act in different environments. On the Web this may mean that the agent leverages reasoning to deal with heterogeneous descriptions.

5.7.6 Running, Testing, and Benchmarking Agents on the Web

With those proposed translations, the following steps may need to be taken to run, test, and benchmark agents:

- Define very simple norms – with the notion of behavior hardly defined for the web, simple condition-action rules may be used to implement a very simple norm.
- Analyze norm and policy compliance – here, log analysis techniques may come in, corresponding metrics need to be defined.
- Enact norm violation punishment – depending on the scenario and the control over the environment, it may be possible to punish agents.

5.7.7 Open Questions and Future Work

Testing and benchmarking are cornerstones of modern software engineering. Still, both topics receive relatively little attention in the context of the (engineering) multi-agent systems research. As reliable and trustworthy behavior is both particularly important and challenging when it comes to Web-scale systems, the research need around testing and benchmarking is even more pronounced here. Some of the open questions that future work can address are the following:

- Which abstractions are central to testing and benchmarking of MAS on the Web?
- What are good practices and pitfalls of Web-scale benchmarking of MAS?
- Does testing MAS on the Web require dedicated tooling support?

Initial works towards the implementation of a benchmark framework for testing Agents on the Web have started to emerge, as in the BOLD server³⁵. Future developments and extensions to this type of framework may be a promising starting point for facilitating norm-based validation, agent behavior testing, assurance of policy compliance, and other aspects as discussed above. This working group has concluded that providing testbed environments such as the ones suggested in this chapter can potentially facilitate the systematic evaluation of approaches to engineering agents on the Web and to place a greater focus on the assurance of quality and reliability.

References

- 1 Winikoff, Michael. *BDI agent testability revisited*. *Auton. Agents Multi Agent Syst.*, 31(5), 1094–1132, 2017, <https://doi.org/10.1007/s10458-016-9356-2>.
- 2 Koopman, Philip and Wagner, Michael. *Challenges in autonomous vehicle testing and validation*. *SAE International Journal of Transportation Safety*, 1(4), 15–24, 2016.
- 3 Linz, Tilo. *Testing autonomous systems*. *The Future of Software Quality Assurance*, 61–75, 2020.
- 4 Timotheus Kampik, Cleber Jorge Amaral and Jomi Fred Hübner. *Developer Operations and Engineering Multi-agent Systems*. *Engineering Multi-Agent Systems – 9th International Workshop (EMAS)*, 2021, https://doi.org/10.1007/978-3-030-97457-2_10
- 5 Timotheus Kampik, Adnane Mansour, Olivier Boissier, Sabrina Kirrane, Julian Padget, Terry R Payne, Munindar P Singh, Valentina Tamma and Antoine Zimmermann. *Governance of Autonomous Agents on the Web: Challenges and Opportunities*. *ACM Transactions on Internet Technology*, 4, 1–31, 2022.
- 6 Richardson, Leonard and Ruby, Sam. *RESTful Web Services*. ISBN 9789896540821, 2007.

³⁵ Bold Benchmark: <https://github.com/bold-benchmark>

- 7 Otávio Arruda Matoso and Luis P. A. Lampert and Jomi Fred Hübner and Mateus Conceição and Sérgio P. Bernardes and Cleber Jorge Amaral and Maicon R. Zатели and Marcelo L. de Lima. *Agent Programming for Industrial Applications: Some Advantages and Drawbacks*. arXiv, 2020.
- 8 Bouron, Thierry and Ferber, Jacques and Samuel, Fabien. *MAGES: A multi-agent testbed for heterogeneous agents*. *Decentralized Artificial Intelligence*, 2, 195–214, 1991.
- 9 Victor Lesser, Charles L. Ortiz and Milind Tambe. *Distributed Sensor Networks: A Multi-agent Perspective*. isbn 9781461350392, 1-376, Springer US, 2003.

Participants

- Cleber Jorge Amaral
Federal Institute of Santa
Catarina – São José, BR
- Matteo Baldoni
University of Turin, IT
- Olivier Boissier
Ecole des Mines –
St. Étienne, FR
- Samuele Burattini
University of Bologna, IT
- Jean-Paul Calbimonte
HES-SO Valais – Sierre, CH
- Pierre-Antoine Champin
INRIA – Sophia Antipolis, FR
- Victor Charpenay
Mines Saint-Étienne, FR
- Amit K. Chopra
Lancaster University, GB
- Andrei Ciortea
Universität St. Gallen, CH
- Rem Collier
University College Dublin, IE
- Stephen Cranefield
University of Otago, NZ
- Mehdi Dastani
Utrecht University, NL
- Kemal A. Delic
The Open University –
Milton Keynes, GB
- Jérôme Euzenat
INRIA Grenoble Rhône-Alpes,
Saint Ismier, FR & Univ.
Grenoble-Alpes, FR
- Catherine Faron
Université Côte d’Azur –
Sophia Antipolis, FR
- Nicoletta Fornara
University of Lugano, CH
- Fabien Gandon
INRIA – Sophia Antipolis, FR
- Wendy Hall
University of Southampton, GB
- Andreas Harth
Fraunhofer IIS – Nürnberg, DE
- Jomi Fred Hübner
Federal University of
Santa Catarina, BR
- Tobias Käfer
KIT – Karlsruher Institut für
Technologie, DE
- Timotheus Kampik
Umeå University, SE &
SAP Signavio – Berlin, DE
- Matthias Kovatsch
Siemens Schweiz AG – Zug, CH
- Brian Logan
University of Aberdeen, GB &
Utrecht University, NL
- Stefano Mariani
University of Modena, IT
- Simon Mayer
Universität St. Gallen, CH
- Mahda Noura
TU Chemnitz, DE
- Eoin O’Neill
University College Dublin, IE
- Julian Padget
University of Bath, GB
- María Poveda-Villalón
Technical University of
Madrid, ES
- Dave Raggett
W3C – United Kingdom, GB
- Alessandro Ricci
Università di Bologna, IT
- Sebastian Schmid
Universität Erlangen-
Nürnberg, DE
- Daniel Schraudner
Universität Erlangen-
Nürnberg, DE
- Munindar P. Singh
North Carolina State University –
Raleigh, US
- Ruben Taelman
Ghent University, BE
- Danai Vachtsevanou
Universität St. Gallen, CH
- Jan Van den Bussche
Hasselt University, BE
- Antoine Zimmermann
Ecole des Mines –
St. Étienne, FR



Resilient Software Configuration and Infrastructure Code Analysis

Jürgen Cito^{*1}, Ruzica Piskac^{*2}, Mark Santolucito^{*3},
Andy Zaidman^{*4}, and Daniel Sokolowski^{†5}

- 1 TU Wien, AT. juergen.cito@tuwien.ac.at
- 2 Yale University – New Haven, US. ruzica.piskac@yale.edu
- 3 Barnard College – New York, US. msantolu@barnard.edu
- 4 TU Delft, NL. a.e.zaidman@tudelft.nl
- 5 Universität St. Gallen, CH. daniel.sokolowski@unisg.ch

Abstract

Errors originating from infrastructure and their configurations are one of the major causes of system failures and system degradation, resulting in security vulnerabilities, application outages, and incorrect program executions. Investigating the root causes of such issues and remedies for them requires insight from different research perspectives, including systems, programming languages, software engineering, and verification. To facilitate progress in this field, this Dagstuhl Seminar brought together experts from academia and industry, enabling synergies between different software systems subareas. The seminar was a forum for cross-disciplinary discussions, bridged communities, and forged new conversations on new approaches. Emerging themes that were revealed during the seminar included a focus on Infrastructure as Code, the similarities and differences between configuration engineering and software engineering, the portability (or lack thereof) of program analysis techniques to configuration analysis, the design space of expressibility of configuration languages, and future challenges of analysis for safety, security, and auditing. The seminar led to new short-term and long-term collaborations and connections, including organizing additional workshops and a joint vision paper.

Seminar February 19–24, 2023 – <https://www.dagstuhl.de/23082>

2012 ACM Subject Classification Security and privacy → Operating systems security; Software and its engineering → Parallel programming languages; Software and its engineering → Compilers; Software and its engineering

Keywords and phrases Computing infrastructure, Configuration, Program analysis

Digital Object Identifier 10.4230/DagRep.13.2.163

1 Executive Summary

Jürgen Cito (TU Wien, AT)

Ruzica Piskac (Yale University – New Haven, US)

Mark Santolucito (Barnard College – New York, US)

Andy Zaidman (TU Delft, NL)

License © Creative Commons BY 4.0 International license
© Jürgen Cito, Ruzica Piskac, Mark Santolucito, and Andy Zaidman

Errors originating from infrastructure and their configurations are one of the major causes of system failures and system degradation, resulting in security vulnerabilities, application outages, and incorrect program executions. Investigating the root causes of such issues and

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Resilient Software Configuration and Infrastructure Code Analysis, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 163–182

Editors: Jürgen Cito, Ruzica Piskac, Mark Santolucito, Andy Zaidman, and Daniel Sokolowski



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

remedies for them requires insight from different research perspectives, including systems, programming languages, software engineering, and verification. From these areas, approaches are emerging to manage the complexity of infrastructure and configuration, covering a breadth of forms, such as domain-specific languages, standalone verification tools, automated learning techniques, specification-based synthesis, security annotation extensions, and configuration optimizers.

The Dagstuhl Seminar on Resilient Software Configuration and Infrastructure Code Analysis brought together experts from different fields to explore new cross-disciplinary approaches to configuration management. The seminar facilitated collaboration between academia and industry and enabled synergies between different subareas of software systems. The seminar was a forum for cross-disciplinary discussions, bridged communities, and forged new conversations between academic and industrial perspectives. The shared knowledge built during the seminar is captured in this report, which we hope can act as a body of knowledge for researchers joining this newly forming community.

Overall, the seminar consisted of 3 tutorial talks, 16 presentations, and 5 group discussions. Emerging themes that were revealed during the seminar included a focus on Infrastructure as Code, the similarities and differences between configuration engineering and software engineering, the portability (or lack thereof) of program analysis techniques to configuration analysis, the design space of expressibility of configuration languages, and future challenges of analysis for safety, security, and auditing. In addition, we had a joint evening session with the parallel seminar “Agents on the Web” (Dagstuhl Seminar 23081), where each organizer presented an overview of their seminar. As a result, we started joint discussions where we investigated the use of formal methods, in particular synthesis, for establishing semantic relations between the data.

Key outcomes of this seminar were evident both in new short-term collaboration and connections, as well as the initiation of longer-term projects. For example, a collection of the participants have connected to host a workshop on configuration languages and analysis called CONFLANG 2023, which will be hosted at SPLASH 2023. Additionally, a vision paper outlining key future directions of the field is being drafted by participants of the event.

After two postponements due to COVID-19, this seminar was a pleasure to hold in person and a great success from both a community and a research perspective. We would like to thank the team of Schloss Dagstuhl for their hospitality and support as well as all the participants for their valuable contributions.

2 Table of Contents

Executive Summary

Jürgen Cito, Ruzica Piskac, Mark Santolucito, and Andy Zaidman 163

Overview of Talks

Improving Infrastructure Security by Analyzing Pre-Deployment Artifacts
Claudia Cauli 167

GLITCH: Automated Polyglot Code Smell Detection in Infrastructure as Code
João F. Ferreira 167

Correctness and Fault Tolerance of Kubernetes Operators
Tianyin Xu 168

Configuration Validation and Testing for Cloud Systems: Research and Practice
Tianyin Xu 169

Configurations Here and There, Configurations Everywhere
Myra B. Cohen 170

Correct and Modular Configuration with Nickel
Yann Hamdaoui 171

The Theory of Real-life Small and Large Configurations
Marcel Van Lohuizen 171

Using CUE to Model Configuration
Marcel Van Lohuizen 172

The Do's and Don'ts of Infrastructure Code: A Systematic Gray Literature Review
Dario Di Nucci 172

Automotive (and Some Other) Configuration Problems
Wolfgang Kuchlin 173

Your Shell Reasoning Toolkit
Michael Greenberg 174

IaC for Architectural Reconstruction
Davide Taibi 174

Decentralizing Infrastructure as Code
Daniel Sokolowski 175

Answer Set Programming: The Powerhouse Technology You've Never Heard Of
Michael Greenberg 175

Transposing Static Analyses from Application to Infrastructure Code: the Curious Case of Ansible
Ruben Opdebeeck and Coen De Roover 176

Application, Orchestration, and Infrastructure, Oh My! Cross Layer Static Analysis Strategies
Mark Santolucito 176

TOSCA Explained!
Damian Andrew Tamburri 177

Static Detection of Silent Misconfigurations with Deep Interaction Analysis <i>Ruzica Piskac</i>	177
Generating Infrastructure Code from System Interactions <i>Jürgen Cito</i>	178
Working groups	
Intersections of Infrastructure as Code (IaC) and Configurable Software <i>Myra B. Cohen, Claudia Cauli, Coen De Roover, Dario Di Nucci, Thomas Durieux, João F. Ferreira, Wolfgang Küchlin, Shane McIntosh, Ruben Opdebeeck, Akond Rahman, Martin Schäf, Davide Taibi, and Marcel Van Lohuizen</i>	179
Architectural Reconstruction and IaC <i>Davide Taibi, Wolfgang Küchlin, Ruben Opdebeeck, Ruzica Piskac, and Damian Andrew Tamburri</i>	179
Reasoning about Code and Infrastructure <i>Martin Schäf, Claudia Cauli, Jürgen Cito, Myra B. Cohen, Coen De Roover, Dario Di Nucci, Thomas Durieux, João F. Ferreira, Michael Greenberg, Shane McIntosh, Akond Rahman, Daniel Sokolowski, Marcel Van Lohuizen, and Tianyin Xu</i>	180
Emerging Trends in Infrastructure as Code (IaC) Research <i>Akond Rahman, Jürgen Cito, Myra B. Cohen, Thomas Durieux, João F. Ferreira, Michael Greenberg, Wolfgang Küchlin, Daniel Sokolowski, Davide Taibi, Marcel Van Lohuizen, and Tianyin Xu</i>	181
Participants	182
Remote Participants	182

3 Overview of Talks

3.1 Improving Infrastructure Security by Analyzing Pre-Deployment Artifacts

Claudia Cauli (Amazon Web Services – London, GB)

License © Creative Commons BY 4.0 International license
© Claudia Cauli

Joint work of Claudia Cauli, Meng Li, Nir Piterman, Oksana Tkachuk

Main reference Claudia Cauli, Meng Li, Nir Piterman, Oksana Tkachuk: Pre-deployment Security Assessment for Cloud Services Through Semantic Reasoning. CAV (1) 2021: 767-780

URL https://link.springer.com/chapter/10.1007/978-3-030-81685-8_36

Over the past ten years, the adoption of cloud services has grown rapidly, leading to the introduction of automated deployment tools to address the scale and complexity of the infrastructure companies and users deploy. Without the aid of automation, ensuring the security of an ever-increasing number of deployments becomes more and more challenging. To the best of our knowledge, no formal automated technique currently exists to verify cloud deployments during the design phase. In this case study, we show that Description Logic modeling and inference capabilities can be used to improve the safety of cloud configurations. We focus on the Amazon Web Services (AWS) proprietary declarative language, CloudFormation, and develop a tool to encode template files into logic. We query the resulting models with properties related to security posture and report on our findings. By extending the models with dataflow-specific knowledge, we use more comprehensive semantic reasoning to further support security reviews. When applying the developed toolchain to publicly available deployment files, we find numerous violations of widely-recognized security best practices, which suggests that streamlining the methodologies developed for this case study would be beneficial.

3.2 GLITCH: Automated Polyglot Code Smell Detection in Infrastructure as Code

João F. Ferreira (INESC-ID – Lisboa, PT)

License © Creative Commons BY 4.0 International license
© João F. Ferreira

Joint work of João F. Ferreira, Nuno Saavedra

Main reference Nuno Saavedra, João F. Ferreira: “GLITCH: Automated Polyglot Security Smell Detection in Infrastructure as Code”, in Proc. of the 37th IEEE/ACM International Conference on Automated Software Engineering, ASE 2022, Rochester, MI, USA, October 10-14, 2022, pp. 47:1–47:12, ACM, 2022.

URL <https://doi.org/10.1145/3551349.3556945>

Infrastructure as Code (IaC) is the process of managing IT infrastructure via programmable configuration files (also called IaC scripts). Like other software artifacts, IaC scripts may contain code smells, which are coding patterns that can result in weaknesses. Automated analysis tools to detect code smells in IaC scripts exist, but they focus on specific technologies such as Puppet, Ansible, or Chef. This means that when the detection of a new smell is implemented in one of the tools, it is not immediately available for the technologies supported by the other tools – the only option is to duplicate the effort.

We present GLITCH, a technology-agnostic framework that enables the automated detection of code smells in IaC scripts. GLITCH allows polyglot smell detection by transforming IaC scripts into an intermediate representation on which different smell detectors can be

defined. GLITCH currently supports the detection of nine security smells and nine design & implementation smells. We compare GLITCH with state-of-the-art smell detectors. For security smells, the results show that GLITCH can reduce the effort of writing security smell analyses for multiple IaC technologies and that it obtains higher precision and recall than the current state-of-the-art tools. For the design & implementation smells, we show that GLITCH has enough information in its intermediate representation to detect technology-agnostic smells supported by state-of-the-art tools.

3.3 Correctness and Fault Tolerance of Kubernetes Operators

Tianyin Xu (University of Illinois – Urbana-Champaign, US)

License © Creative Commons BY 4.0 International license
© Tianyin Xu

Main reference Xudong Sun, Wenqing Luo, Jiawei Tyler Gu, Aishwarya Ganesan, Ramnatthan Alagappan, Michael Gasch, Lalith Suresh, Tianyin Xu: “Automatic Reliability Testing For Cluster Management Controllers”, in Proc. of the 16th USENIX Symposium on Operating Systems Design and Implementation (OSDI 22), pp. 143–159, USENIX Association, 2022.

URL <https://www.usenix.org/conference/osdi22/presentation/sun>

Modern cluster managers like Borg, Omega, and Kubernetes rely on the state-reconciliation principle to be highly resilient and extensible. In these systems, all cluster-management logic is embedded in a loosely coupled collection of microservices called controllers. Each controller independently observes the current cluster state and issues corrective actions to converge the cluster to a desired state. However, the complex distributed nature of the overall system makes it hard to build reliable and correct controllers – we find that controllers face myriad reliability issues that lead to severe consequences like data loss, security vulnerabilities, and resource leaks.

In this talk, I present Sieve, the first automatic reliability-testing tool for cluster-management controllers. Sieve drives controllers to their potentially buggy corners by systematically and extensively perturbing the controller’s view of the current cluster state in ways it is expected to tolerate. It then compares the cluster state’s evolution with and without perturbations to detect safety and liveness issues. Sieve’s design is powered by a fundamental opportunity in state-reconciliation systems – these systems are based on state-centric interfaces between the controllers and the cluster state; such interfaces are highly transparent and enable fully-automated reliability testing. To date, Sieve has efficiently found 46 serious safety and liveness bugs (35 confirmed and 22 fixed) in ten popular controllers with a low false-positive rate of 3.5%.

3.4 Configuration Validation and Testing for Cloud Systems: Research and Practice

Tianyin Xu (University of Illinois – Urbana-Champaign, US)

License © Creative Commons BY 4.0 International license
© Tianyin Xu

Main reference Shuai Wang, Xinyu Lian, Darko Marinov, and Tianyin Xu: “Test Selection for Unified Regression Testing”, In Proceedings of the 45th IEEE/ACM International Conference on Software Engineering (ICSE’23), Melbourne, Australia, May 2023.

URL <https://tianyin.github.io/pub/ctest.pdf>

Configuration management is an integral part of modern DevOps-based cloud system management. Many critical operations are done by updating configurations to change system behavior in production dynamically. Today, large-scale cloud and Internet services evolve rapidly, with hundreds to thousands of configuration changes deployed daily. For example, at Facebook, thousands of configuration changes are committed every day, outpacing the frequency of code changes. Other cloud services from Google and Azure also frequently deploy configuration changes. It is not surprising to hear that the “cloud feels more about configuration management than software engineering.”

With the high velocity of changes, faulty configurations inevitably have become major causes of system failures and service outages. For example, faulty configurations are reported as the second largest cause of service disruptions in a main Google production service. At Facebook, 16% of service-level incidents are induced by configuration changes. Many configuration-induced failures led to catastrophic impacts. For instance, in March 2019, a misconfiguration led to Facebook’s largest outage in terms of duration (14 hours); in June 2021, a seemingly-valid configuration change at Fastly triggered an undiscovered software bug and broke the Internet for an hour.

We argue that continuous testing is a key missing piece of today’s configuration management practice. Despite the “configuration-as-code” movement, there is no widely-used, systematic configuration testing technique, and thus configuration changes are not unit-tested—imagining a world where code changes only go through manual review and static analysis, without regression testing.

We will introduce the idea of configuration testing, a new testing technique that enables configuration changes to be unit-tested in DevOps-based continuous integration/deployment. The basic idea of configuration testing is connecting system configurations to software tests so that configuration changes can be tested in the context of code affected by the changes. We will introduce a new type of tests, termed Ctests, to fill the critical need for configuration testing. Ctests complement static validation (the de facto protection), analogous to how testing complements static analysis:

- Ctests can detect failure-inducing configuration changes where the failure root causes are in the code, e.g., valid configuration value changes that trigger dormant bugs.
- Ctests can detect sophisticated misconfigurations (e.g., those that violate undocumented, hidden constraints) by capturing the resulting unexpected system behavior.

We will demonstrate that ctests can be generated by transforming existing software tests that are abundant in mature software projects. The generated ctests selectively inherit test logic and assertions from the original tests. The inherited assertions hold for all correct configuration values. We have successfully generated 7,974 ctests for five widely-used open-source cloud systems (Hadoop Common, HDFS, HBase, ZooKeeper, and Alluxio).

We will show that the generated ctests are effective and outperform state-of-the-art static configuration validation techniques based on extensive evaluations with real-world failure cases, synthesized misconfigurations, and deployed configuration files in public Docker images.

References

- 1 Runxiang Cheng, Lingming Zhang, Darko Marinov, and Tianyin Xu, Test-Case Prioritization for Configuration Testing, In Proceedings of the 30th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA'21), Virtual Event, July 2021.
- 2 Yuanliang Zhang, Haochen He, Owolabi Legunsen, Shanshan Li, Wei Dong, and Tianyin Xu, An Evolutionary Study of Configuration Design and Implementation in Cloud Systems, In Proceedings of the 43rd International Conference on Software Engineering (ICSE'21), Virtual Event, May 2021.
- 3 Xudong Sun, Runxiang Cheng, Jianyan Chen, Elaine Ang, Owolabi Legunsen, and Tianyin Xu, Testing Configuration Changes in Context to Prevent Production Failures, In Proceedings of the 14th USENIX Symposium on Operating Systems Design and Implementation (OSDI'20), Virtual Event, Nov. 2020.
- 4 Qingrong Chen, Teng Wang, Owolabi Legunsen, Shanshan Li, and Tianyin Xu, Understanding and Discovering Software Configuration Dependencies in Cloud and Datacenter Systems, In Proceedings of the 2020 ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE'20), Virtual Event, Nov. 2020.
- 5 Tianyin Xu and Darko Marinov, Mining Container Image Repositories for Software Configurations and Beyond, In Proceedings of the 40th International Conference on Software Engineering, New Ideas and Emerging Results (ICSE'18, NIER), Gothenburg, Sweden, May 2018.
- 6 Tianyin Xu, Xinxin Jin, Peng Huang, Yuanyuan Zhou, Shan Lu, Long Jin, and Shankar Pasupathy, Early Detection of Configuration Errors to Reduce Failure Damage, In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16), Savannah, GA, Nov. 2016.
- 7 Tianyin Xu, Long Jin, Xuepeng Fan, Yuanyuan Zhou, Shankar Pasupathy, and Rukma Talwadker, Hey, You Have Given Me Too Many Knobs! Understanding and Dealing with Over-Designed Configuration in System Software, In Proceedings of the 10th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE'15), Bergamo, Italy, Aug. 2015.
- 8 Tianyin Xu and Yuanyuan Zhou, Systems Approaches to Tackling Configuration Errors: A Survey, ACM Computing Surveys (CSUR), Vol. 47, No. 4, Article 70, Jul. 2015.
- 9 Tianyin Xu, Jiaqi Zhang, Peng Huang, Jing Zheng, Tianwei Sheng, Ding Yuan, Yuanyuan Zhou, and Shankar Pasupathy, Do Not Blame Users for Misconfigurations, In Proceedings of the 24th ACM Symposium on Operating Systems Principles (SOSP'13), Farmington, PA, Nov. 2013.

3.5 Configurations Here and There, Configurations Everywhere

Myra B. Cohen (Iowa State University – Ames, US)

License © Creative Commons BY 4.0 International license
© Myra B. Cohen

Joint work of Paul Gazzillo, Myra B. Cohen

Main reference Paul Gazzillo, Myra B. Cohen: “Bringing Together Configuration Research: Towards a Common Ground”, in Proc. of the 2022 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software, Onward! 2022, Auckland, New Zealand, December 8-10, 2022, pp. 259–269, ACM, 2022.

URL <https://doi.org/10.1145/3563835.3568737>

Configurable software makes up most of the software in use today. Configurability, i.e., the ability of software to be customized without additional programming, is pervasive, and due to the criticality of problems caused by misconfiguration, it has been an active topic

researched by investigators in multiple, diverse areas. This broad reach of configurability means that much of the literature and latest results are dispersed, and researchers may not be collaborating or be aware of similar problems and solutions in other domains. We argue that this lack of a common ground leads to a missed opportunity for synergy between research domains and the synthesis of efforts to tackle configurability problems. To provide a foundation for addressing these concerns, we suggest how to bring the communities together and propose a common model of configurability and a platform, ACCORD, to facilitate collaboration among researchers and practitioners.

3.6 Correct and Modular Configuration with Nickel

Yann Hamdaoui (Tweag I/O – Paris, FR)

License  Creative Commons BY 4.0 International license
© Yann Hamdaoui

From a distance, Infrastructure as Code should really be called Infrastructure as Configuration. DevOps, SRE and other engineers dealing with infrastructure are mostly managing pure configuration data by writing, editing and auditing JSON, YAML or similar serialization formats.

Purely data-oriented languages like JSON might be fine for managing small and simple infrastructure, but when the size and complexity of a configuration grows, data languages don't seem to be the right tool anymore.

There is no way to reuse data and have a single source of truth, with all the pain and inconsistencies that duplication inevitably brings with time. We can't express data dependencies either. For example, the open ports of a firewall may depend on which services are enabled on a server. In JSON, everything must be hardcoded.


Infrastructure tools have thus incorporated programming features in their languages (Terraform, Ansible, Puppet, etc.). But those are often unplanned and unprincipled, resulting in a complex system.

In the end, Configuration Management looks like the poor sibling of Software Engineering. What about tests, types, LSP integration for real-time feedback, completion, and documentation? What about modularity, code reuse, and abstraction?

In this talk, I will present Nickel, a configuration programming language I am currently working on at Tweag, to help finally enter the era of Configuration as Code. I'll discuss more specifically the approach of Nickel to modularity, which is how to write small and reusable configuration snippets that can be combined into a complex configuration based on a merging operation. I'll talk about type-checking and built-in schema validation as well.

3.7 The Theory of Real-life Small and Large Configurations

Marcel Van Lohuizen (CUE – Zug, CH)

License  Creative Commons BY 4.0 International license
© Marcel Van Lohuizen
URL <https://cuelang.org>

Configuration is in more places than people imagine. Every part of your tech stack – databases, apps, schemas, services, workflows, policy, models, and networking must be configured. Not only that, there are dependencies between these configurations. With cloud, multi-cloud,

IoT, and edge computing, the number of things to configure within growing systems has exploded. The consequences of getting the configuration wrong have only worsened over time.

We are in the middle of a major shift where configuration is becoming a first-class citizen across your stack. New engineering roles dedicated to solving configuration problems have emerged: Platform, Site Reliability, Resilience, Observability, Data, “DevOps” and “YAML” engineers all deal with easing configuration toil. There are many tools and approaches that aim to help developers deal with growing configuration complexity: Configuration Management, Infrastructure as Code, GitOps, Policy as Code, and finally, Infrastructure as Data. The industry needs something to unify all these roles, approaches, and challenges and break down configuration silos. CUE (cuelang.org) does exactly that.

In this presentation, you will hear:

- hard-won insights and experiences of configuration at scale, culminating in the design of CUE,
- how configuration can go wrong,
- the need for testing and validation,
- how CUE is paving the way for a holistic approach to configuration via a language, tooling, and APIs that support a vibrant configuration ecosystem.

3.8 Using CUE to Model Configuration

Marcel Van Lohuizen (CUE – Zug, CH)

License © Creative Commons BY 4.0 International license
 © Marcel Van Lohuizen
URL <https://cuelang.org>

Configuration is inherently cross-cutting. Combing configuration, therefore, requires commutative and associative composition. In this tutorial, we show how to use CUE to model all configuration aspects of a distributed system into a single space. Configuration aspects can be data, schema, policy, validation, and templates, or any combination thereof.

3.9 The Do’s and Don’ts of Infrastructure Code: A Systematic Gray Literature Review

Dario Di Nucci (University of Salerno, IT)

License © Creative Commons BY 4.0 International license
 © Dario Di Nucci
Joint work of Indika Kumara, Martin Garriga, Angel Urbano Romeu, Dario Di Nucci, Fabio Palomba, Damian Andrew Tamburri, Willem-Jan van den Heuvel
Main reference Indika Kumara, Martin Garriga, Angel Urbano Romeu, Dario Di Nucci, Fabio Palomba, Damian Andrew Tamburri, Willem-Jan van den Heuvel: “The do’s and don’ts of infrastructure code: A systematic gray literature review”, *Inf. Softw. Technol.*, Vol. 137, p. 106593, 2021.
URL <https://doi.org/10.1016/j.infsof.2021.106593>

This talk provided an overview of the qualitative analysis we conducted to summarize the industrial gray literature (e.g., blog posts, tutorials, white papers) on IaC.

In particular, it provided a general definition of IaC and a broad catalog outlined in taxonomy consisting of ten and four primary categories for best practices and bad practices, respectively, both language-agnostic and language-specific ones, for three IaC languages of best and bad practices for widely used IaC languages (i.e., Ansible, Puppet, and Chef).

Our findings highlighted that the IaC development and maintenance field is in its infancy and deserves further attention.

3.10 Automotive (and Some Other) Configuration Problems

Wolfgang Kuchlin (Universität Tübingen, DE)

License © Creative Commons BY 4.0 International license

© Wolfgang Kuchlin

Joint work of Wolfgang Kuchlin, Carsten Sinz

Main reference Wolfgang Kuchlin, Carsten Sinz: “Proving Consistency Assertions for Automotive Product Data Management”, *J. Autom. Reason.*, Vol. 24(1/2), pp. 145–163, 2000.

URL <https://doi.org/10.1023/A:1006370506164>

We report on industrial configuration problems outside the field of infrastructure as code. We focus on our work on automotive configuration which goes back to the 1990s and continues until today by providing software to the industry in the context of our Steinbeis Transfer Centre STZ OIT.

Automotive configuration is managed on two levels using (a flavor of) Boolean Algebra. Every equipment or sales option is represented by a Boolean variable x , where $x=true$ represents the presence of the option in a car, and $x=false$ represents the absence, such that a valid car order is given by a complete valuation of all variables. The upper level, model description, comprises the configuration rules for an entire model line of cars (e.g. Mercedes C-Class). The lower level, the Bill-of-Materials, contains the list of all materials needed for the model line. Each material also has a selection formula, and the material is needed to produce a car order if the selection formula evaluates to true under the order.

Problems on the upper level include the computation of forced options, which must be contained in any car, and impossible options, which cannot be contained in any car. Problems on the lower level include the detection of car orders which will select alternative materials or will fail to select a necessary material. These problems are efficiently solved by modern CDCL SAT-Solvers, although the typical variance of the model description is practically infinite (more than a trillion).

Beyond the detection of configuration errors, we proceed to configuration optimization. As an example, we describe the problem of computing the configuration of optimal (minimal) sets of test vehicles on which a given set of testing demands can be carried out.

In addition, we report on some other industrial configuration verification problems whose solutions we published in the past. This includes the detection of misconfigurations of Storage Area Networks (SAN) and of Apache Web servers, the analysis of LINUX Kernel configurations, and the analysis of the completeness of the online help system of a line of computer tomographs.

3.11 Your Shell Reasoning Toolkit

Michael Greenberg (Stevens Institute of Technology – Hoboken, US)

License © Creative Commons BY 4.0 International license
© Michael Greenberg

Main reference Michael Greenberg, Austin J. Blatt: “Executable Formal Semantics for the POSIX Shell”, Proc. ACM Program. Lang., Vol. 4(POPL), Association for Computing Machinery, 2019.

URL <https://doi.org/10.1145/3371111>

Main reference Konstantinos Kallas, Tammam Mustafa, Jan Bielak, Dimitris Karnikis, Thurston H.Y. Dang, Michael Greenberg, Nikos Vasilakis: “Practically Correct, Just-in-Time Shell Script Parallelization”, in Proc. of the 16th USENIX Symposium on Operating Systems Design and Implementation (OSDI 22), pp. 769–785, USENIX Association, 2022.

URL <https://www.usenix.org/conference/osdi22/presentation/kallas>

The shell is a critical part of modern software operations. I have three tools for helping you work with the shell:

- libdash, bindings to the dash parser in OCaml and Python
- smooosh, an executable small-step operational semantics tested against the POSIX spec
- pash-annotations, specifications of common commands

3.12 IaC for Architectural Reconstruction

Davide Taibi (University of Oulu, FI)

License © Creative Commons BY 4.0 International license
© Davide Taibi

Main reference Tomás Cerný, Amr S. Abdelfattah, Vincent Bushong, Abdullah Al Maruf, Davide Taibi: “Microservice Architecture Reconstruction and Visualization Techniques: A Review”, in Proc. of the IEEE International Conference on Service-Oriented System Engineering, SOSE 2022, Newark, CA, USA, August 15-18, 2022, pp. 39–48, IEEE, 2022.

URL <https://doi.org/10.1109/SOSE55356.2022.00011>

The continuous development of new services and the time pressure imposed by the development of new features commonly result in the introduction of non-optimal and temporary solutions. As a result, the software architecture is commonly not compliant with the originally designed one, and nobody has a complete view of the whole system. Static and dynamic analysis methods can help to reconstruct the architecture of the system.

In this talk, we showed some possible techniques for reconstructing the architecture from different perspectives: the infrastructural, the structure of the system analyzed statically or dynamically, and the structure of the organization.

Last we presented a roadmap of possible shared research work towards the development or the extension of tools for IaC architectural reconstruction.

3.13 Decentralizing Infrastructure as Code

Daniel Sokolowski (Universität St. Gallen, CH)

License © Creative Commons BY 4.0 International license
© Daniel Sokolowski

Joint work of Daniel Sokolowski, Pascal Weisenburger, Guido Salvaneschi

Main reference Daniel Sokolowski, Pascal Weisenburger, Guido Salvaneschi: “Decentralizing Infrastructure as Code”, *IEEE Softw.*, Vol. 40(1), pp. 50–55, 2023.

URL <https://doi.org/10.1109/MS.2022.3192968>

Main reference Daniel Sokolowski, Pascal Weisenburger, Guido Salvaneschi: “Automating serverless deployments for DevOps organizations”, in *Proc. of the ESEC/FSE '21: 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, Athens, Greece, August 23–28, 2021, pp. 57–69, ACM, 2021.

URL <https://doi.org/10.1145/3468264.3468575>

DevOps unifies software development and operations in cross-functional teams to improve software delivery and operations (SDO) performance. Ideally, cross-functional DevOps teams independently deploy their services, but the correct operation of a service often demands other services, requiring coordination to ensure the correct deployment order. This issue is currently solved either with a central deployment or manual out-of-band communication across teams, e.g., via phone, chat, or email. Unfortunately, both contradict the independence of teams, hindering SDO performance – the reason why DevOps is adopted in the first place.

We conducted a study on 134 IT professionals, showing that, in practice, they resort to manual coordination for correct deployments even if they expect better SDO performance with fully automated approaches. We find that Infrastructure as Code (IaC) automates deployments for single teams, falling short of decentralized deployments across groups. To enable testing and automation advances for decentralized organizations, we need mature IaC solutions that embrace and consolidate software engineering principles.

To address this issue, we proposed μ s ([mju:z] “muse”), a novel IaC system automating deployment coordination in a fully decentralized fashion, still retaining compatibility with DevOps practice – in contrast to today’s solutions. We implement μ s, demonstrate that it effectively enables automated coordination, introduces negligible definition overhead, has no performance overhead, and is broadly applicable, as shown by the migration of 64 third-party IaC projects.

3.14 Answer Set Programming: The Powerhouse Technology You’ve Never Heard Of

Michael Greenberg (Stevens Institute of Technology – Hoboken, US)

License © Creative Commons BY 4.0 International license
© Michael Greenberg

Main reference Aaron Bembenek, Michael Greenberg, Stephen Chong: “From SMT to ASP: Solver-Based Approaches to Solving Datalog Synthesis-as-Rule-Selection Problems”, *Proc. ACM Program. Lang.*, Vol. 7(POPL), pp. 185–217, 2023.

URL <https://doi.org/10.1145/3571200>

Answer set programming (ASP) is a powerful tool that folks in PL/FM/SE do not know about. I give two case studies applying ASP to synthesis (once for Datalog, once for RBAC policies); I also explain how ASP relates to SAT (its parent) and SMT (its sibling).

3.15 Transposing Static Analyses from Application to Infrastructure Code: the Curious Case of Ansible

Ruben Opdebeek (VU – Brussels, BE) and Coen De Roover (VU – Brussels, BE)

License © Creative Commons BY 4.0 International license
© Ruben Opdebeek and Coen De Roover

Joint work of Ruben Opdebeek, Jonas De Bleser, Coen De Roover, Dario Di Nucci, Camilo Velázquez-Rodríguez, Ahmed Zerouali

Main reference Ruben Opdebeek, Ahmed Zerouali, Coen De Roover: “Smelly Variables in Ansible Infrastructure Code: Detection, Prevalence, and Lifetime”, in Proc. of the 19th IEEE/ACM International Conference on Mining Software Repositories, MSR 2022, Pittsburgh, PA, USA, May 23-24, 2022, pp. 61–72, ACM, 2022.

URL <https://doi.org/10.1145/3524842.3527964>

This talk presents our journey and experiences in transposing static analyses from application to infrastructure code.

We realized the need to analyze infrastructure code while designing the Chaokka automated tester. Chaokka injects cloud-specific faults during the execution of a test suite to assess the resilience of a cloud-native application against these faults. The approach uses delta debugging to speed up the exploration of the corresponding fault space, which we expect to benefit further from architectural insights extracted from the supporting infrastructure code (e.g., how many replicas the application maintains of each service, ...).

The second part of the talk focuses on analyses for infrastructure code in general and Ansible in particular. To this end, we present two static representations of Ansible code. First, we describe a structural model akin to an AST and its applications in empirical research on Semantic Versioning in the Ansible Galaxy ecosystem. Second, we present a program dependence graph representation that captures an Ansible script’s control and data flow. We describe the challenges faced when building such graphs caused by unconventional and undocumented Ansible semantics. Finally, we describe two applications of Ansible program dependence graphs, namely the detection of variable-related smells and a graph pattern mining approach to defect detection.

3.16 Application, Orchestration, and Infrastructure, Oh My! Cross Layer Static Analysis Strategies

Mark Santolucito (Barnard College – New York, US)

License © Creative Commons BY 4.0 International license
© Mark Santolucito

With Infrastructure as Code (Iac), the process of configuring complex infrastructure has been simplified and made more accessible to developers. However, with the new expressive power of IaC and the consequentially more complex cloud infrastructure deployments, understanding this complexity has emerged as an unsolved issue. In particular, estimating the cost of an Infrastructure as Code deployment requires understanding the pricing models of every cloud resource being used, as well as an understanding of the interactions between the resources. Existing work either relies on historical usage metrics to predict cost (which has limited utility for new deployments), or on coarse-grain static cost analysis that ignores interactions between resources. We pose as a challenge to the community the need for fine-grained static cost analysis techniques for IaC that incorporate reasoning about the interactions between resources. We walk through a motivating example of this problem that demonstrates the need for such analysis as well as the complexity of the problem.

3.17 TOSCA Explained!

Damian Andrew Tamburri (TU Eindhoven, NL)

License © Creative Commons BY 4.0 International license
© Damian Andrew Tamburri

Joint work of Paul Lipton, Derek Palma, Matt Rutkowski, Damian A. Tamburri

Main reference Paul Lipton, Derek Palma, Matt Rutkowski, Damian A. Tamburri: “TOSCA Solves Big Problems in the Cloud and Beyond!”, IEEE Cloud Computing, Vol. 5(2), pp. 37–47, 2018.

URL <https://doi.org/10.1109/MCC.2018.022171666>

The anatomy of infrastructure code, as well as all elements and abstractions necessary in writing and maintaining that blueprint, are addressed, among other IaC formats, in the industrial standard for IaC, namely, the OASIS “Topology and Orchestration Specification for Cloud Applications” (TOSCA) industrial standard adopted by as many as 60+ big industrial players worldwide.

Paraphrasing from the standard specification itself, “TOSCA [...] uses [...] service templates to describe cloud apps as a topology template, [...]. TOSCA provides a type system of node types to describe the possible building blocks for constructing a service template and relationship type to describe possible relations”. TOSCA-ready orchestrators such as Cloudify or Apache Brooklyn normally come with a considerable number of reusable TOSCA nodes (e.g., a MySQL DB or a WordPress host), while more are proliferating in both research and practice.

Although several alternatives to TOSCA exist (e.g., HashiCorp Terraform, Ubuntu Juju), for the sake of simplicity, we focus on outlining the anatomy of infrastructure code using the technology-agnostic TOSCA standard notation and its intended levels of abstraction.

3.18 Static Detection of Silent Misconfigurations with Deep Interaction Analysis

Ruzica Piskac (Yale University – New Haven, US)

License © Creative Commons BY 4.0 International license
© Ruzica Piskac

Joint work of Jialu Zhang, Ruzica Piskac, Ennan Zhai, Tianyin Xu

Main reference Jialu Zhang, Ruzica Piskac, Ennan Zhai, Tianyin Xu: “Static detection of silent misconfigurations with deep interaction analysis”, Proc. ACM Program. Lang., Vol. 5(OOPSLA), pp. 1–30, 2021.

URL <https://doi.org/10.1145/3485517>

The behavior of large systems is guided by their configurations: users set parameters in the configuration file to dictate which corresponding part of the system code is executed. However, it is often the case that, although some parameters are set in the configuration file, they do not influence the system runtime behavior, thus failing to meet the user’s intent. Moreover, such misconfigurations rarely lead to an error message or raising an exception. We introduce the notion of silent misconfiguration, which are prohibitively hard to identify due to (1) lack of feedback and (2) complex interactions between configurations and code.

In this talk we present ConfigX, the first tool for the detection of silent misconfigurations. The main challenge is understanding the complex interactions between configurations and the code they affected. Our goal is to derive a specification describing non-trivial interactions between the configuration parameters that lead to silent misconfigurations. To this end, ConfigX uses static analysis to determine which parts of the system code are associated with configuration parameters. ConfigX then infers the connections between configuration parameters by analyzing their associated code blocks. We design customized control- and

data-flow analysis to derive a specification of configurations. Additionally, we conduct reachability analysis to eliminate spurious rules to reduce false positives. Upon evaluation on five real-world datasets across three widely-used systems, Apache, vsftpd, and PostgreSQL, ConfigX detected more than 2200 silent misconfigurations. We additionally conducted a user study where we ran ConfigX on misconfigurations reported on user forums by real-world users. ConfigX easily detected issues and suggested repairs for those misconfigurations. Our solutions were accepted and confirmed in the interaction with the users who originally posted the problems.

3.19 Generating Infrastructure Code from System Interactions

Jürgen Cito (TU Wien, AT)

License  Creative Commons BY 4.0 International license
© Jürgen Cito

Setting up complex, large-scale infrastructure is an iterative process that includes installing dependencies and adjusting parameters within a vast space of configuration options and eventual infrastructure testing in a trial-and-error fashion. The result of this process serves as the underlying base for the deployment and execution of computation defined in the program code. This ad-hoc style of infrastructure setup cannot scale to operations with scalable compute instances (i.e., cloud computing). Infrastructure as code allows to express all infrastructure concerns in the form of code, enabling automation to achieve scalability, transparency, and reproducibility. However, creating code for infrastructure inherently differs from writing program code. The fast feedback cycles required for the iterative setup process outlined above are not feasible when infrastructure code is written in the same way as program code. Thus, infrastructure code is commonly written retrospectively in a cumbersome back-and-forth process querying configuration parameters that are prone to human error. My proposed research aims to intercept lower-level system interactions by experts setting up systems interactively and infer higher-level actions. We generate action sequences guided by the existing infrastructure serving as an oracle goal state.

4 Working groups

4.1 Intersections of Infrastructure as Code (IaC) and Configurable Software

Myra B. Cohen (Iowa State University – Ames, US), Claudia Cauli (Amazon Web Services – London, GB), Coen De Roover (VU – Brussels, BE), Dario Di Nucci (University of Salerno, IT), Thomas Durieux (TU Delft, NL), João F. Ferreira (INESC-ID – Lisboa, PT), Wolfgang Küchlin (Universität Tübingen, DE), Shane McIntosh (University of Waterloo, CA), Ruben Opdebeeck (VU – Brussels, BE), Akond Rahman (Auburn University, US), Martin Schäf (Amazon Web Services – New York City, US), Davide Taibi (University of Oulu, FI), and Marcel Van Lohuizen (CUE – Zug, CH)

License © Creative Commons BY 4.0 International license

© Myra B. Cohen, Claudia Cauli, Coen De Roover, Dario Di Nucci, Thomas Durieux, João F. Ferreira, Wolfgang Küchlin, Shane McIntosh, Ruben Opdebeeck, Akond Rahman, Martin Schäf, Davide Taibi, and Marcel Van Lohuizen

In this breakout, we discussed the intersections between infrastructure as code and traditional configurable software, where configurability models the variability in system behavior rather than defining how the system is composed at a single point in time. Our goal was to find ways we can share techniques, tools, and benchmarks across these two domains. We discussed the fact that configurability in software is usually declarative, while IaC is often imperative. We also discussed the idea that the IaC view can be considered a unit level (or bottom-up view) of the system, while those in traditional configurable systems are using a systems approach (or top-down view). However, we agreed that many challenges are shared, so it would be important to look for places we can merge approaches. Some roadblocks to using a systems approach on IaC are that the systems are often unbounded; hence the complete configuration space is unknown. There is also no uniform modeling approach for IaC. We further discussed the problem that many of the dynamic techniques used in traditional configurable software, such as testing across configurations, may not work. We also noted that it is often too expensive to perform dynamic analyses because many IaC systems are often charged by usage micro-services. We agreed we should start with some example benchmarks, such as train-ticket to further explore the differences (and commonalities).

4.2 Architectural Reconstruction and IaC

Davide Taibi (University of Oulu, FI), Wolfgang Küchlin (Universität Tübingen, DE), Ruben Opdebeeck (VU – Brussels, BE), Ruzica Piskac (Yale University – New Haven, US), and Damian Andrew Tamburri (TU Eindhoven, NL)

License © Creative Commons BY 4.0 International license

© Davide Taibi, Wolfgang Küchlin, Ruben Opdebeeck, Ruzica Piskac, and Damian Andrew Tamburri

During this breakout session, we discussed on pros and cons of techniques for architectural reconstruction. We identified mainly two analysis techniques: static and dynamic analysis techniques. Static analysis techniques scan different software artifacts, including IaC, source code, but also repository activity from git logs. We agreed that static analysis of source code might not be a viable solution, mainly because of the large number of technologies existing on the market. However, analysis of IaC might still be performed, because of the limited number of languages adopted.

Dynamic analysis instead models the system’s behavior at runtime and allows one to understand how the system is performing or structured at runtime.

We highlighted the importance of merging the different visualizations, including the runtime analysis and the visualization of the infrastructure, which might become very important when considering continuum edge-to-cloud systems, where different types of devices and their resources can make the difference.

Security analysis might also be an important research avenue that could benefit from static, dynamic analysis, and their composition.

4.3 Reasoning about Code and Infrastructure

Martin Schäf (Amazon Web Services – New York City, US), Claudia Cauli (Amazon Web Services – London, GB), Myra B. Cohen (Iowa State University – Ames, US), Coen De Roover (VU – Brussels, BE), Dario Di Nucci (University of Salerno, IT), Thomas Durieux (TU Delft, NL), João F. Ferreira (INESC-ID – Lisboa, PT), Michael Greenberg (Stevens Institute of Technology – Hoboken, US), Shane McIntosh (University of Waterloo, CA), Akond Rahman (Auburn University, US), Daniel Sokolowski (Universität St. Gallen, CH), Marcel Van Lohuizen (CUE – Zug, CH), and Tianyin Xu (University of Illinois – Urbana-Champaign, US)

License © Creative Commons BY 4.0 International license

© Martin Schäf, Claudia Cauli, Jürgen Cito, Myra B. Cohen, Coen De Roover, Dario Di Nucci, Thomas Durieux, João F. Ferreira, Michael Greenberg, Shane McIntosh, Akond Rahman, Daniel Sokolowski, Marcel Van Lohuizen, and Tianyin Xu

URL https://docs.google.com/document/d/1S0eN_5tjQ15DU6VO8b5T4oDAMBvOAVikx8KvOpJMfPc

Many of the most prevalent CWE security issues, such as cross-site scripting (xss) or command injection, can be detected statically via data-flow or taint analysis. We simply define what parameters are considered untrusted (e.g., the arguments to a Servlet request in a Java application) and what sinks are considered sensitive (e.g., the members of a Servlet response in Java). This method works great if the application is known to the analysis. However, in a cloud setting, we usually have separate code bases and repositories for each compute component of the system. Consider an AWS application that receives some job identifier via a message queue, process that information via a serverless function, and then has the serverless function pass that job identifier on to another queue once processing is done.

While this is a typical design for processing information with micro-services, from a static, this serverless function has a cross-site scripting vulnerability since it parrots its input without modification. This cross-site scripting vulnerability will be a false alarm unless the input to the function can be controlled by an adversary and the output is sent back to a browser. To decide this, we will need to analyze the code and its infrastructure.

- We see multiple challenges in this space. The static analysis of the code will need to produce a summary that can be re-used by the static analysis of the infrastructure code.
- The mitigation for the code vulnerability may not happen in the code (e.g., xss could be mitigated by a firewall or gateway), so tracking issues, producing counter-examples, or generating patches will have to span over infrastructure code and application code.
- The infrastructure code itself may not be complete and not show the entire system, so analysis of infrastructure and application code will need to be able to report conditional findings (e.g., if message queue X receives untrusted data, then lambda Y will forward it to queue Z).

4.4 Emerging Trends in Infrastructure as Code (IaC) Research

Akond Rahman (Auburn University, US), Jürgen Cito (TU Wien, AT), Myra B. Cohen (Iowa State University – Ames, US), Thomas Durieux (TU Delft, NL), João F. Ferreira (INESC-ID – Lisboa, PT), Michael Greenberg (Stevens Institute of Technology – Hoboken, US), Wolfgang Küchlin (Universität Tübingen, DE), Daniel Sokolowski (Universität St. Gallen, CH), Davide Taibi (University of Oulu, FI), Marcel Van Lohuizen (CUE – Zug, CH), and Tianyin Xu (University of Illinois – Urbana-Champaign, US)

License © Creative Commons BY 4.0 International license

© Akond Rahman, Jürgen Cito, Myra B. Cohen, Thomas Durieux, João F. Ferreira, Michael Greenberg, Wolfgang Küchlin, Daniel Sokolowski, Davide Taibi, Marcel Van Lohuizen, and Tianyin Xu

The objective of this working group is to identify emerging areas in the domain of Infrastructure as Code (IaC). As part of this breakout group, participants from academia and industry both agreed that the domain of IaC research has a lot of potential and till date remains an under-explored area. Through rigorous discussion, the participants suggested the following areas as emerging and therefore need to be addressed: (i) discovering and evaluating representations and reasoning for IaC, (ii) deriving metrics to measure complexity of configurations, (iii) derivation of single source truth for IaC-based infrastructure, (iv) investigate quality assurance for IaC compilers, (v) understand how to combine static and dynamic analysis procedures for IaC, (vi) quantify coverage for IaC-specific tests, such as idempotency detection coverage, and (vii) architectural inference from IaC.

While discussing these emerging areas participants shared real-world examples as well as described their personal experiences when working with industry partners. Participants also mentioned research papers in other CS-related research areas, such as security and software product lines can benefit IaC research. Participants also remarked that the ongoing discussions of the seminar helped them realize the scope and novelty of IaC research.

Participants

- Claudia Cauli
Amazon Web Services –
London, GB
- Jürgen Cito
TU Wien, AT
- Myra B. Cohen
Iowa State University –
Ames, US
- Coen De Roover
VU – Brussels, BE
- Dario Di Nucci
University of Salerno, IT
- Thomas Durieux
TU Delft, NL
- João F. Ferreira
INESC-ID – Lisboa, PT
- Michael Greenberg
Stevens Institute of Technology –
Hoboken, US
- Yann Hamdaoui
Tweag I/O – Paris, FR
- Wolfgang Küchlin
Universität Tübingen, DE
- Anthony W. Lin
RPTU – Kaiserslautern, DE
- Shane McIntosh
University of Waterloo, CA
- Ruben Opdebeeck
VU – Brussels, BE
- Ruzica Piskac
Yale University – New Haven, US
- Akond Rahman
Auburn University, US
- Guido Salvaneschi
Universität St. Gallen, CH
- Mark Santolucito
Barnard College –
New York, US
- Martin Schäf
Amazon Web Services – New
York City, US
- Daniel Sokolowski
Universität St. Gallen, CH
- Davide Taibi
University of Oulu, FI
- Damian Andrew Tamburri
TU Eindhoven, NL
- Marcel Van Lohuizen
CUE – Zug, CH
- Tianyin Xu
University of Illinois –
Urbana-Champaign, US



Remote Participants

- Andy Zaidman
TU Delft, NL

Algorithmic Foundations of Programmable Matter

Aaron Becker^{*†1}, Sándor Fekete^{†2}, Irina Kostitsyna^{†3},
Matthew J. Patitz^{†4}, Damien Woods^{†5}, and
Ioannis Chatzigiannakis^{†6}

- 1 University of Houston, US. atbecker@uh.edu
- 2 TU Braunschweig, DE. s.fekete@tu-braunschweig.de
- 3 TU Eindhoven, NL. i.kostitsyna@tue.nl
- 4 University of Arkansas – Fayetteville, US. patitz@uark.edu
- 5 Maynooth University, IE. damien.woods@mu.ie
- 6 Sapienza University of Rome, IT. ichatz@diag.uniroma1.it

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 23091, “Algorithmic Foundations of Programmable Matter”, a new and emerging field that combines theoretical work on algorithms with a wide spectrum of practical applications that reach all the way from small-scale embedded systems to cyber-physical structures at nano-scale.

The aim of this seminar was to bring together researchers from computational geometry, distributed computing, DNA computing, and swarm robotics who have worked on programmable matter to inform one another about the newest developments in each area and to discuss future models, approaches, and directions for new research. Similar to the first two Dagstuhl Seminars on programmable matter (16271 and 18331), we did focus on some basic problems, but also considered new problems that were now within reach to be studied.

Seminar February 26 – March 3, 2023 – <https://www.dagstuhl.de/23091>

2012 ACM Subject Classification Theory of computation → Models of computation; Theory of computation → Computational geometry; Theory of computation → Distributed algorithms; Computer systems organization → Robotics; Computing methodologies → Artificial intelligence

Keywords and phrases computational geometry, distributed algorithms, DNA computing, programmable matter, swarm robotics

Digital Object Identifier 10.4230/DagRep.13.2.183

1 Executive Summary


Aaron Becker (University of Houston, US)

Sándor Fekete (TU Braunschweig, DE)

Irina Kostitsyna (TU Eindhoven, NL)

Matthew J. Patitz (University of Arkansas – Fayetteville, US)

Damien Woods (Maynooth University, IE)

License  Creative Commons BY 4.0 International license

© Aaron Becker, Sándor Fekete, Irina Kostitsyna, Matthew J. Patitz, and Damien Woods

Algorithmic Foundations of Programmable Matter is an area that aims at designing models and algorithms for materials that can change their physical properties in a programmable fashion or based on external stimuli.

* At the time of the seminar on sabbatical at TU Braunschweig, DE

† Editor / Organizer

‡ Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Algorithmic Foundations of Programmable Matter, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 183–198

Editors: Aaron Becker, Sándor Fekete, Irina Kostitsyna, Matthew J. Patitz, Damien Woods, and Ioannis Chatzigiannakis



DAGSTUHL
REPORTS

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

This was the third successful Dagstuhl Seminar on this topic, following seminar 16271, which brought together a number of algorithmic core topics, and seminar 18331, which helped to identify a number of particular challenges. In this third seminar 23091, we continued and extended these first steps, in particular, by considering systems with many small components (such as in physics), very large structures (such as in space), and very complex structures (such as in biology).

Despite having to deal with a number of pandemic and post-pandemic issues (which led to the cancellation of the originally planned seminar 21091 in March 2021), we were able to benefit from a broader range of interaction, including digital preparation: We ran an online event over two half days (accounting for time zones from Japan to the US), bringing together the originally planned group of attendees with additional interested and motivated colleagues. This not only helped to maintain the community spirit and keep in touch about ongoing developments, but it also helped to identify the most motivated members of the community. This preparatory event was successful in the following ways.

- Participants worked out a refined format for preparing joint work on new research problems.
- Participants of the digital meeting demonstrated their great commitment to attend: At the end of the second day (i.e., after 8h of online interaction spanning 11h of time difference), the number of live participants was still at 43, i.e., basically undiminished.

This turned out to be extremely helpful for the organization of the renewed seminar 23091, in particular with respect to participation: Despite the inevitable late cancellations due to illness or other unforeseeable emergencies, there was a total of 38 in-person participants.

Another key feature of the seminar was to again make use of the interactive electronic tool `coauthor`¹, which allowed exceptionally intensive and interdisciplinary collaboration throughout the week, allowing dynamic formulation and development of research problems, ideas, progress and formulation of results. Thus, `coauthor` greatly facilitated the work done during the seminar, enabling not just identification of, but also dynamic research work on a number of new topics.

On the content side, a number of research areas were brought together, with theoretical connections to the areas of *Distributed Computing*, *Computational Geometry*, and *Self-Assembly*. In addition, a number of application areas provided further directions, including *Swarm Robotics* with their methods for dealing with systems composed of many individual components that together form complex and reconfigurable structures; *Engineering and Physics*, with a variety of technologies and applications for developing flexible and innovative materials and constructions; and *Biomolecular applications* with a spectrum of real-life scenarios.

Overall, we brought together a combination of established experts from the mentioned areas. On the senior level, participants included a number of leading authorities who are established in more than one of the mentioned topics; on the junior level, we had a good selection of highly talented scientists who will continue to advance the field by specific contributions.

Making use of the excellent experiences during the previous seminar, the seminar started with a plenary introduction of all participants, their research areas and their specific challenges and expectations for the seminar. This was followed by a number of plenary sessions, in which experts gave overviews over broad developments and specific open problems.

¹ <https://github.com/edemaine/coauthor/>

- Amira Abdel-Rahman gave an excellent survey over the challenges and perspectives of using self-replicating hierarchical robotic swarms in application areas such as space missions.
- Dan Halperin gave an overview of geometry at the service of robotics, ranging from snapping fixtures to multi-robot coordination.
- Damien Woods gave an introduction to practice and theory of self-assembly using synthetic DNA, and then went on to give an overview of work going on in the wet lab, covering a range of different practical topics.
- Kay Römer described scalable real-time localization for robotic ensembles, providing insights into practical hardware considerations.
- Benoît Piranda gave an overview of ongoing work on both the development of real-world Catoms (as a powerful model for producing programmable matter platforms), as well as their simulation and visualization.
- Tom Peters (with support by both Irina Kostitsyna and Christian Scheideler) described the state of the art and ongoing challenges in the amoebot model.

Further presentations were given by Jo Ellis-Monaghan (on DNA self-assembly), Sándor Fekete (describing work on coordinating large ensembles of particles at extreme dimensions, from space to targeted drug delivery), Timothy Gomez (on two-handed self-assembly), Christian Scheideler (on reconfigurable circuit extension), and Giovanni Viglietta (describing programmable matter from fractal formation to genetic programming).

Extending precedent, we put particular emphasis on sessions and presentations that were focused not only on previous and ongoing work, but also on coming challenges, by running a considerable number of open problem sessions. A variety of open problem sessions were headed by Amira Abdel-Rahman (path planning for assembly and reconfiguration), Yuval Emek (self-stabilizing chemical reaction networks), Dan Halperin (minimum number of control alterations), Tanja Kaiser (open-ended evolution), Maria Kokkou (leader election with faults), Irene Parada (fast parallel reconfiguration), Nicolas Schabanel (DNA origami and experiments), Cynthia Sung (self-folding structures), and Ryuhei Uehara (slide-and-pack puzzles).

2 Table of Contents

Executive Summary

<i>Aaron Becker, Sándor Fekete, Irina Kostitsyna, Matthew J. Patitz, and Damien Woods</i>	183
---	-----

Overview of Talks

Self-Replicating Hierarchical Robotic Swarms <i>Amira Abdel-Rahman</i>	188
Some open questions from DNA self-assembly <i>Jo Ellis-Monaghan</i>	188
Space Ants: Episode II – Coordinating Connected Catoms <i>Sándor Fekete</i>	188
Unique Assembly Verification in Two-Handed Self-Assembly <i>Timothy Gomez</i>	189
From snapping fixtures to multi-robot coordination: Geometry at the service of robotics <i>Dan Halperin</i>	189
Fast communication in amoebot systems <i>Tom Peters and Irina Kostitsyna</i>	190
How to manage the movement of robots with VisibleSim to achieve locomotion and self-reconfiguration <i>Benoît Piranda</i>	190
Scalable Real-Time Localization for Robotic Ensembles <i>Kay Römer</i>	191
Open Problems for the Reconfigurable Circuit Extension of the Amoebot Model <i>Christian Scheideler</i>	191
Reconfigurable Circuit Extension of the Amoebot Model <i>Christian Scheideler</i>	192
Programmable Matter: From Fractal Formation to Genetic-Programming Solutions <i>Giovanni Viglietta</i>	193
DNA based self-assembly and robotics <i>Damien Woods</i>	193

Open problems

Path Planning for the Assembly and Reconfiguration of Structures using Neural Cellular Automata <i>Amira Abdel-Rahman</i>	194
A self-stabilizing chemical reaction network implementation of a modular clock <i>Yuval Emek</i>	194
Minimum number of control alterations (squares in polygon, discs in curved work-space) <i>Dan Halperin</i>	195

Open-ended evolution by minimizing surprise <i>Tanja Katharina Kaiser</i>	195
Leader election in grids with faults <i>Maria Kokkou</i>	196
Fast parallel reconfiguration without scaling <i>Irene Parada</i>	196
DNA Origami: Open Questions on Helix and Scaffold Routing; Feedback From Experiments <i>Nicolas Schabanel</i>	197
When can a structure self-fold? <i>Cynthia Sung</i>	197
Computational Complexity of Recent Slide-and-Pack Puzzles <i>Ryuhei Uehara</i>	197
Participants	198

3 Overview of Talks

3.1 Self-Replicating Hierarchical Robotic Swarms

Amira Abdel-Rahman (MIT – Cambridge, US)

License  Creative Commons BY 4.0 International license
 Amira Abdel-Rahman

Main reference Amira Abdel-Rahman, Christopher Cameron, Benjamin Jenett, Miana Smith, Neil Gershenfeld: “Self-replicating hierarchical modular robotic swarms”. *Communications Engineering*, 1(1), 35, 2022

URL <https://doi.org/10.1038/s44172-022-00034-3>

We introduce the self-replication of hierarchical robotic swarms made by robotic swarms. This is accomplished by discretizing their construction from a feedstock of primitive building blocks that are simultaneously simpler than prior reconfigurable robotic modules and can be composed to create a wide range of functionality. The motivating application is the assembly of high-performance (high specific modulus and strength) cellular structures, for which the swarms function as a combined material-robot system. The discretization significantly simplifies the swarm’s navigation, error correction, and coordination. Such a system can build serially, can grow exponentially by building more robots, and can grow hierarchically by building larger robots. We present an algorithm for this novel path-planning problem to compile a shape into a swarm to construct it, show an experimental design of the building block basis set, and model the scaling tradeoffs.

3.2 Some open questions from DNA self-assembly

Jo Ellis-Monaghan (University of Amsterdam, NL)

License  Creative Commons BY 4.0 International license
 Jo Ellis-Monaghan

Joint work of Nataša Jonoska, Greta Pangborn, Nadrian Seeman, among many others

This talk gives a brief overview of three areas in DNA self-assembly, with emphasis on open problems, including:

- Self-assembly for flexible armed tiles, and also geometric tiles for example in the octet truss.
- Identifying reporter strands for verifying experimental results.
- The problems and possibilities of knotted scaffolding strands in DNA origami.

3.3 Space Ants: Episode II – Coordinating Connected Catoms

Sándor Fekete (TU Braunschweig, DE)

License  Creative Commons BY 4.0 International license
 Sándor Fekete

How can a set of identical mobile agents coordinate their motions to transform their arrangement from a given starting to a desired goal configuration? We consider this question in the context of actual physical devices called Catoms, which can perform reconfiguration, but need to maintain connectivity at all times to ensure communication and energy supply. We demonstrate and animate algorithmic results, in particular a proof of hardness, as well as an algorithm that guarantees constant stretch for certain classes of arrangements: If

mapping the start configuration to the target configuration requires a maximum Manhattan distance of d , then the total duration of our overall schedule is in $O(d)$, which is optimal up to constant factors.

3.4 Unique Assembly Verification in Two-Handed Self-Assembly

Timothy Gomez (MIT – Cambridge, US)

License © Creative Commons BY 4.0 International license
© Timothy Gomez

Joint work of David Caballero, Timothy Gomez, Robert Schweller, Tim Wylie

Main reference David Caballero, Timothy Gomez, Robert T. Schweller, Tim Wylie: “Unique Assembly Verification in Two-Handed Self-Assembly”, in Proc. of the 49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France, LIPIcs, Vol. 229, pp. 34:1–34:21, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.

URL <https://doi.org/10.4230/LIPIcs.ICALP.2022.34>

One of the most fundamental and well-studied problems in Tile Self-Assembly is the Unique Assembly Verification (UAV) problem. This algorithmic problem asks whether a given tile system uniquely assembles a specific assembly. The complexity of this problem in the 2-Handed Assembly Model (2HAM) at a constant temperature is a long-standing open problem since the model was introduced. Previously, only membership in the class coNP was known and that the problem is in P if the temperature is one ($\tau = 1$). The problem is known to be hard for many generalizations of the model, such as allowing one step into the third dimension or allowing the temperature of the system to be a variable, but the most fundamental version has remained open.

In this paper, we prove the UAV problem in the 2HAM is hard even with a small constant temperature ($\tau = 2$), and finally answer the complexity of this problem (open since 2013). Further, this result proves that UAV in the staged self-assembly model is coNP -complete with a single bin and stage (open since 2007), and that UAV in the q -tile model is also coNP -complete (open since 2004). We reduce from Monotone Planar 3-SAT with Neighboring Variable Pairs, a special case of 3SAT recently proven to be NP-hard. We accompany this reduction with a positive result showing that UAV is solvable in polynomial time with the promise that the given target assembly will have a tree-shaped bond graph, i.e., contains no cycles. We provide a $\mathcal{O}(n^5)$ algorithm for UAV on tree-bonded assemblies when the temperature is fixed to 2, and a $\mathcal{O}(n^5 \log \tau)$ time algorithm when the temperature is part of the input.

3.5 From snapping fixtures to multi-robot coordination: Geometry at the service of robotics

Dan Halperin (Tel Aviv University, IL)

License © Creative Commons BY 4.0 International license
© Dan Halperin

Robots sense, move and act in the physical world. It is therefore natural that understanding the geometry of the problem at hand is often key to devising an effective robotic solution. I will review several problems in robotics and automation in whose solution geometry plays a major role. These include designing optimized 3D printable fixtures, object rearrangement by robot arm manipulators, and efficient coordination of the motion of large teams of robots. As we shall see, exploiting geometric structure can, among other benefits, lead to reducing the dimensionality of the underlying search space and in turn to efficient solutions.

3.6 Fast communication in amoebot systems

Tom Peters (TU Eindhoven, NL), Irina Kostitsyna (TU Eindhoven, NL)

License © Creative Commons BY 4.0 International license
© Tom Peters and Irina Kostitsyna

Joint work of Tom Peters, Irina Kostitsyna, Bettina Speckmann
Main reference Irina Kostitsyna, Tom Peters, Bettina Speckmann: “Brief Announcement: An Effective Geometric Communication Structure for Programmable Matter”, in Proc. of the 36th International Symposium on Distributed Computing, DISC 2022, October 25-27, 2022, Augusta, Georgia, USA, LIPIcs, Vol. 246, pp. 47:1–47:3, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
URL <https://doi.org/10.4230/LIPIcs.DISC.2022.47>

The concept of programmable matter envisions a very large number of tiny and simple robot particles forming a smart material that can change its physical properties and shape based on the outcome of computation and movement performed by the individual particles in a concurrent manner. We use geometric insights to develop a new type of shortest path tree for programmable matter systems. Our feather trees utilize geometry to allow particles and information to traverse the programmable matter structure via shortest paths even in the presence of multiple overlapping trees.

3.7 How to manage the movement of robots with VisibleSim to achieve locomotion and self-reconfiguration

Benoît Piranda (FEMTO-ST Institute – Montbéliard, FR)

License © Creative Commons BY 4.0 International license
© Benoît Piranda

Joint work of Benoît Piranda, Julien Bourgeois
Main reference Pierre Thalamy, Benoît Piranda, André Naz, Julien Bourgeois: “VisibleSim: A behavioral simulation framework for lattice modular robots”, Robotics Auton. Syst., Vol. 147, p. 103913, 2022.
URL <https://doi.org/10.1016/j.robot.2021.103913>

This presentation introduces VisibleSim a behavioral simulator for distributed robots placed in a regular 3D lattice. I first present our programming model and then show a video that introduce VisibleSim, presenting the several robots that can be simulated, with different shapes and capabilities (change color, move from one cell of the lattice to another one. . .). VisibleSim is a C++ project that allow to observe the effects of running the same program in all the robots, in order to debug, capture picture and video from the simulation. The second part of the presentation shows step by step how to write an application code (BlockCode) for VisibleSim to displace a line of 4 SlidingCubes. Using our online generator which generate a runnable skeleton of the application from a formular, for your application the code that must be added is very limited to get a quick result. Finally in the third part of my presentation, I have presented an application of self-reconfiguration developed with VisibleSim applied on 3D Catoms robots which are placed in a Face Centered Cubic lattice and are able to roll to neighbor to change their position in the grid. We first define meta-modules made of 10 modules, and the basic process to dismantle, transfer and assemble these meta-modules. Using VisibleSim, we evaluate an algorithm that compute a distributed version of the Max-Flow algorithm to compute a set of separated paths between sources (Meta-module that must be removes) and destinations (connected meta-modules that must be added), and finally moves modules along these paths.

3.8 Scalable Real-Time Localization for Robotic Ensembles

Kay Römer (TU Graz, AT)

License © Creative Commons BY 4.0 International license
© Kay Römer

Joint work of Kay Römer, Bernhard Großwindhager, Michael Stocker, Michael Rath, Mustafa Bakr, Klaus Witrissal, Carlo Boano

URL <https://dependablethings.tugraz.at/>

For controlling the movement of swarms of robots or drones, the positions of the robotic entities need to be tracked at sub-decimeter level. This entails a number of research challenges, (i) minimizing the overhead for the infrastructure needed to support tracking such as the number of reference anchors (ii) scaling to a large number of densely deployed robotic elements (iii) supporting a high update rate for positions in the order of hundreds to thousands of position fixes per second, and (iv) robustness in harsh environments, for example due to obstacles blocking the line of sight between anchors and robotic elements. In this talk we present two results to address these challenges using Ultra-Wide-Band (UWB) communication. In the first solution, a single anchor supports localization by exploiting the reflections of the UWB signals from the walls. In the second solution, multiple anchors submit UWB pulses quasi simultaneously to allow robotic entities to localize themselves with just a single UWB message reception operation, thus supporting scaling to an arbitrary number of robotic elements and more than 1000 position updates per second.

3.9 Open Problems for the Reconfigurable Circuit Extension of the Amoebot Model

Christian Scheideler (Universität Paderborn, DE)

License © Creative Commons BY 4.0 International license
© Christian Scheideler

Joint work of Christian Scheideler, Joshua Daymude, Shlomi Dolev, Michael Feldmann, Irina Kostitsyna, Andreas Padalkin, Andrea Richa, Daniel Warner

In my presentation, I raised a number of open problems for reconfigurable circuit extension of the amoebot model. In this extension, amoebots can set up circuits stretching many amoebots. If at least one amoebots beeps on such a circuit, all amoebots of that circuit will receive it, and otherwise no amoebot will receive a beep. When a beep is received, an amoebot cannot determine the direction it came from or the number of amoebots that beeped.

First, consider the case of rapid shape transformation. We have shown that when performing joint expansions and contractions on top of the circuit model, one can transform a line of n amoebots into a parallelogram in just $O(\log n)$ rounds [1]. However, what is the minimum number of joint expansions and contractions to get from one structure to another? Can this be computed efficiently and transformed into a distributed algorithm?

Second, consider the case of rapid energy transfer or, more concretely, the problem of repairing an amoebot structure by moving a collection of amoebots from places where they should be removed to places where they are needed. Both of these problems can be reduced to a fairly elementary problem of establishing a shortest path from one position of the amoebot structure to another (e.g., [2]). Is there a fast algorithm for constructing a circuit between these two positions representing a shortest path in the circuit model?

Third, consider the case of seed-less shape transformation. All amoebot algorithms that transform an arbitrary amoebot structure in a hexagon are based on a seed (e.g., [3]). Can we also come up with an algorithm that does not need a seed or leader to perform the transformation.

Finally, consider the situation of faulty amoebots (e.g., [4]). How can we design algorithms for shape transformation that are resilient to permanent failures of amoebots?

References

- 1 Michael Feldmann, Andreas Padalkin, Christian Scheideler, and Shlomi Dolev. *Coordinating Amoebots via Reconfigurable Circuits*. *Journal of Computational Biology* 29(4), pp. 317-343, 2022.
- 2 Irina Kostitsyna, Tom Peters, and Bettina Speckmann. *Fast Reconfiguration for Programmable Matter*. In arXiv, <https://arxiv.org/abs/2202.11663>, 2022.
- 3 Joshua J. Daymude, Andrea W. Richa, and Christian Scheideler. *The Canonical Amoebot Model: Algorithms and Concurrency Control*. In DISC 2021, pp. 20:1–20:19, 2021.
- 4 Irina Kostitsyna, Christian Scheideler, and Daniel Warner. *Fault-Tolerant Shape Formation in the Amoebot Model*. In DNA 28, pp. 9:1–9:22, 2022.

3.10 Reconfigurable Circuit Extension of the Amoebot Model

Christian Scheideler (Universität Paderborn, DE)

License © Creative Commons BY 4.0 International license
© Christian Scheideler

Joint work of Christian Scheideler, Andreas Padalkin, Daniel Warner, Shlomi Dolev

Main reference Michael Feldmann, Andreas Padalkin, Christian Scheideler, Shlomi Dolev: “Coordinating Amoebots via Reconfigurable Circuits”, *J. Comput. Biol.*, Vol. 29(4), pp. 317–343, 2022.

URL <https://doi.org/10.1089/cmb.2021.0363>

The Amoebot model has been proposed as a model for programmable matter. However, many important issues are not captured in the basic form of this model, including rapid shape transformation, energy, fault-tolerance, and the 3D case. In my presentation, I focus on rapid shape transformation and propose a reconfigurable circuit extension of the Amoebot model for this. In this extension, the Amoebots have a constant number of so-called pins on each side that they can internally connect via wires. The connected components formed by these wires then establish circuits. Amoebots can send very primitive signals on these circuits called beeps. If an Amoebot beeps on a circuit, every Amoebot connected to that circuit will instantly notice this. However, an Amoebot does not know where the beep came from or the number of beeps. As I show, this circuit extension allows many central problems like leader election, compass alignment, and symmetry checks to be solved significantly faster than in the basic model. This will then provide the basis for synchronization which will be important for rapid shape transformation.

3.11 Programmable Matter: From Fractal Formation to Genetic-Programming Solutions

Giovanni Viglietta (JAIST – Ishikawa, JP)

License © Creative Commons BY 4.0 International license
© Giovanni Viglietta

Main reference Giuseppe Antonio Di Luna, Paola Flocchini, Nicola Santoro, Giovanni Viglietta, Yukiko Yamauchi: “Mobile RAM and Shape Formation by Programmable Particles”, in Proc. of the Euro-Par 2020: Parallel Processing – 26th International Conference on Parallel and Distributed Computing, Warsaw, Poland, August 24-28, 2020, Proceedings, Lecture Notes in Computer Science, Vol. 12247, pp. 343–358, Springer, 2020.

URL https://doi.org/10.1007/978-3-030-57675-2_22

A typical approach to Shape Formation by Programmable Matter is to elect a leader and use it to coordinate and direct the whole system. This approach has some disadvantages, such as the introduction of bottlenecks and the lack of fault tolerance. In this talk, I reviewed this basic approach and I contrasted it with a recent experimental study of Genetic Programming applied to Programmable Matter.

3.12 DNA based self-assembly and robotics

Damien Woods (Maynooth University, IE)

License © Creative Commons BY 4.0 International license
© Damien Woods

Joint work of Cai Wood, Constantine G Evans, Abeer Eshra, Tristan Stérin, Ahmed Shalaby, David Doty, Irina Kostitsyna, Damien Woods



The talk began with an introduction to practice and theory of self-assembly using synthetic DNA, and then went on to give an overview of work going on in our lab on the following topics.

1. DNA robotics: theory and implementation of the Turning Machine model (a sub-model of the Nubot model). [Wood, Kostitsyna, Woods]
2. Control of self-assembly dynamics 1: controlling growth order of tiles in self-assembly. A 32x32 canvas of tiles was given that can be grown in any order in a scaled black-wise fashion. In theory, there are a number of methods that could implement such behaviour and we’ve found holes to stop growth in the wrong direction to work best for us (ongoing work suggests that glue mismatches could also work). [Evans, Shalaby, Doty, Woods]
3. Control of self-assembly dynamics 2: using covers/blockers/poison to tune the rate of nucleation, including giving a method to completely suppress nucleation while allowing for seeded growth. [Rogers, Evans, Woods]
4. Thermodynamically-favoured computation: DNA origami-inspired technique to achieve thermodynamically favoured computation, leading to in-principle low/no errors. Six example finite automata have been implemented (addition, bit-parity, etc.), as well as a bit-copy system who’s answer is flipped between 0 and 1 five times, each time changing the energy landscape and thus thermodynamic favourability. [Eshra, Stérin, Woods]

4 Open problems

4.1 Path Planning for the Assembly and Reconfiguration of Structures using Neural Cellular Automata



Amira Abdel-Rahman (MIT – Cambridge, US)

License  Creative Commons BY 4.0 International license
 Amira Abdel-Rahman

In this open problem presentation, we introduced a method to learn cellular automata (CA) rules to grow and reconfigure optimal structures based on given boundary conditions and objective functions. Now that we have CA local rules to update a voxel state based on only its neighbors' state, we proposed working on the materialization of this method to be used for the assembly and reconfiguration of structures. We proposed first working on implementing distributed path planning algorithms for the swam assembly and reconfiguration of structures using the learned neural cellular automata (NCA) rules. We also proposed working on understanding the theoretical bounds of the geometries can/cannot be built using these algorithms in order to edit/update/constrain the learned NCA rules.

4.2 A self-stabilizing chemical reaction network implementation of a modular clock

Yuval Emek (Technion – Haifa, IL)

License  Creative Commons BY 4.0 International license
 Yuval Emek

This open problem is concerned with a (discrete) CRN task referred to as self-stabilizing modular clock. For this task, let Z_k denote the additive cyclic group over the integers modulo k . We wish to build a CRN over a species set $\Sigma \cup \{E\}$ where E is a designated highly reactive external species, associated with a function $\kappa : \Sigma \rightarrow Z_k$ that maps each species $A \in \Sigma$ to a clock value $\kappa(A) \in Z_k$. A configuration over Σ in which all molecules have the same clock value is said to be synchronized. We are interested in constructing a CRN that satisfies the following properties:

1. From any initial configuration, the CRN is guaranteed to stabilize to a halting synchronized configuration.
2. The external species E reacts with all species in Σ and is not produced by any reaction.
3. If we are in a synchronized halting configuration with clock value $i \in Z_k$ and a “small” number of external molecules (“a drop”) are added to the solution, then the CRN is guaranteed to stabilize to a halting synchronized configuration with clock value $i+1 \pmod k$.

It is trivial to design such a CRN operating under a fair scheduler, assuming the “usual” notion of fairness ([1]), however we are interested in the following weaker versions of fairness: (F1) Every reaction which is applicable infinitely often is scheduled infinitely often (that is, no reaction can be “starved” indefinitely). (F2) Every reaction which is applicable infinitely often is either scheduled or becomes inapplicable infinitely often (that is, no reaction can be “starved” continuously and indefinitely).

Does there exist such a CRN? For which values of k ? Does the task become easier if we omit the requirement of halting configurations? Does the task become easier if we relax the “self-stabilization” requirement so that starting from an (adversarial) initial configuration, the system is guaranteed to go back to “proper operation” after a few rounds of “external species drops”?

References

- 1 Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, and René Peralta. 2006. Computation in networks of passively mobile finite-state sensors. *Distributed Comput.* 18:4, pages 235–253, 2006.

4.3 Minimum number of control alterations (squares in polygon, discs in curved workspace)

Dan Halperin (Tel Aviv University, IL)

License © Creative Commons BY 4.0 International license
© Dan Halperin

Given two axis-aligned unit-square robots translating in a polygonal environment, we wish to find a collision-free motion between free start and target configurations such that at each point in time only one robot moves and that will minimize the number of alternations between the robots. What is the maximum number of alternations needed when the polygonal environment has n vertices? A lower bound of 14 alternations [Andrey Leshchenko] is shown below for a polygon with holes.

4.4 Open-ended evolution by minimizing surprise

Tanja Katharina Kaiser (Universität Lübeck, DE)

License © Creative Commons BY 4.0 International license
© Tanja Katharina Kaiser

Joint work of Tanja Katharina Kaiser, Louis Mayer, Heiko Hamann

Main reference Tanja Katharina Kaiser, Heiko Hamann: “Innate Motivation for Robot Swarms by Minimizing Surprise: From Simple Simulations to Real-World Experiments”, *IEEE Trans. Robotics*, Vol. 38(6), pp. 3582–3601, 2022.

URL <https://doi.org/10.1109/TRO.2022.3181004>

In nature, open-ended evolution has led to the continuous emergence of novel and increasingly complex artifacts. Researchers are trying to replicate this open-endedness in artificial intelligence, but it remains one of the last great challenges. Minimize surprise evolves actor-predictor pairs of artificial neural networks by rewarding only high prediction accuracy (i.e., low surprise) in the optimization process. In previous swarm robot experiments, we have shown that our minimize surprise method [2] leads to the emergence of diverse swarm behaviors and enables robots to adapt to a changing environment. In first preliminary experiments, we explore here the potential of minimize surprise for open-ended evolution in EvoCraft [1], a framework for Minecraft designed to study open-ended algorithms. So far, we have limited ourselves to a “Game of Life”-like 2D cellular automata setting and have been able to evolve different patterns. Our vision is to extend the setting to 3D and to allow the open-ended evolution of diverse entities that are capable of interacting with the environment.


References

- 1 D. Grbic, R.B. Palm, E. Najjarro, C. Glanois, S. Risi. *EvoCraft: A New Challenge for Open-Endedness*. In: Castillo, P.A., Jiménez Laredo, J.L. (eds) *Applications of Evolutionary Computation, EvoApplications 2021, Lecture Notes in Computer Science*, vol 12694, Springer, Cham, 2021

- 2 T. K. Kaiser, H. Hamann. *Innate Motivation for Robot Swarms by Minimizing Surprise: From Simple Simulations to Real-World Experiments*. In: *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3582-3601, Dec. 2022

4.5 Leader election in grids with faults

Maria Kokkou (Aix-Marseille University, FR)

License  Creative Commons BY 4.0 International license
© Maria Kokkou

The leader election problem has been well studied under the Amoebot model for systems without faults. However, for systems that do contain faults the problem remains open. An interesting starting point is to determine the minimum assumptions we need to deterministically elect a unique leader even if the particle configuration contains holes and the grid contains obstacles.

The motivation for this problem comes from two possible faults. First, as a big number of particles is involved, it is possible that some particles have already crashed before the execution of the algorithm, creating obstacles within the system. The second motive comes from the possibility of faults existing in the grid itself, such as missing or damaged nodes which cannot be occupied by particles. Since obstacles restrict the ability of particles to move, an algorithm that depends on movement would need to account for an unknown number of nodes being unavailable or blocked. Furthermore, obstacles can create symmetries that do not exist in the triangular grid, such as six-symmetric configurations where a node is occupied by an obstacle and all its neighbouring nodes are occupied by competing particles.

4.6 Fast parallel reconfiguration without scaling

Irene Parada (UPC Barcelona Tech, ES)

License  Creative Commons BY 4.0 International license
© Irene Parada

The problem we consider is parallel reconfiguration of modular robots preserving connectivity. To guarantee fast parallel reconfiguration, the existing algorithms for modular robots need either meta-modules [1, 2] or a constant scale factor [3] (blowing up the shape size by a constant). The question is whether we can relax these conditions. More precisely, whether requiring a minimum local density of the initial and final configurations allows for an efficient reconfiguration plan.

References

- 1 Ferran Hurtado, Enrique Molina, Suneeta Ramaswami, and Vera Sacristán. *Distributed reconfiguration of 2D lattice-based modular robotic systems*. *Auton. Robot.* 38, 383–413, 2015. <https://doi.org/10.1007/s10514-015-9421-8>
- 2 Irene Parada, Vera Sacristán, and Rodrigo I. Silveira. *A new meta-module design for efficient reconfiguration of modular robots*. *Auton Robot* 45, 457–472, 2021. <https://doi.org/10.1007/s10514-021-09977-6>
- 3 Sándor P. Fekete, Phillip Keldenich, Ramin Kosfeld, Christian Rieck, and Christian Scheffer. *Connected coordinated motion planning with bounded stretch*. *Proc. ISAAC*, 2021. <https://doi.org/10.4230/LIPIcs.ISAAC.2021.9>

4.7 DNA Origami: Open Questions on Helix and Scaffold Routing; Feedback From Experiments

Nicolas Schabanel (*ENS – Lyon, FR*)

License  Creative Commons BY 4.0 International license
© Nicolas Schabanel

Helix routing consists in deciding the path followed by the DNA helices in DNA origami: it is a problem similar to coil wrapping when you want to wrap a surface with non-overlapping ribbons of fixed width. Scaffold routing is to decide the path followed by the scaffold of a DNA origami along the helices. Our experiments show that many parameters have to be taken into account to have the DNA origami fold properly. This raises the question of how to come up with an algorithmic approach to solve this issue.

4.8 When can a structure self-fold?

Cynthia Sung (*University of Pennsylvania – Philadelphia, US*)

License  Creative Commons BY 4.0 International license
© Cynthia Sung

Self-folding is a process through which a 2D sheet autonomously folds into a desired 3D shape. Self-folding fabrication typically involves layering rigid and active materials that allow the sheet to respond to heat, moisture, light, or other signals in a predetermined way. Designing these structures to reliably self-fold into the desired shape repeatedly remains a challenge. Common approaches to “programming” the folding process involve specifying that certain folds are mountain or valley folds. However, it has been shown that a mountain-valley assignment is insufficient to uniquely determine the final 3D shape. Recent models of self-folding have considered the case where precise torques at each of the folds can be specified, but even under this model, it is sometimes impossible to control what the final 3D shape will be. We now consider a new self-folding mode, in which forces can be applied directly at each face and consider the associated questions of what can or cannot be reliably self-folded under these methods.

References

- 1 Tomohiro Tachi and Thomas C. Hull. *Self-foldability of Rigid Origami*. J. Mechanisms Robotics. 9, 021008, 2017. <https://doi.org/10.1115/1.4035558>

4.9 Computational Complexity of Recent Slide-and-Pack Puzzles

Ryuhei Uehara (*JAIST – Ishikawa, JP*)

License  Creative Commons BY 4.0 International license
© Ryuhei Uehara

In the last decade, a series of puzzles in a new category has been developed in puzzle society. We are given a frame and a set of pieces. The goal of the puzzle is packing all the pieces into the frame. It seems that they are similar to the classic popular silhouette puzzles like the tangram. However, in recent puzzles, the frame has small entrance, and we have to find how to pack the pieces into the frame. In 2D, it can be seen as the combination of sliding block puzzle and packing puzzle. It seems that these puzzles are PSPACE-complete in general, but so far, we have no results in this framework.

Participants

- Amira Abdel-Rahman
MIT – Cambridge, US
- Yotam Ashkenazi
Ben Gurion University –
Beer Sheva, IL
- Aaron Becker
University of Houston, US
- Julien Bourgeois
FEMTO-ST Institute –
Montbéliard, FR
- Ioannis Chatzigiannakis
Sapienza University of Rome, IT
- Shantanu Das
Aix-Marseille University, FR
- David Doty
University of California –
Davis, US
- Jo Ellis-Monaghan
University of Amsterdam, NL
- Yuval Emek
Technion – Haifa, IL
- Constantine Evans
Maynooth University, IE
- Sándor Fekete
TU Braunschweig, DE
- Timothy Gomez
MIT – Cambridge, US
- Daniel Hader
University of Arkansas –
Fayetteville, US
- Dan Halperin
Tel Aviv University, IL
- Nataša Jonoska
University of South Florida –
Tampa, US
- Tanja Katharina Kaiser
Universität Lübeck, DE
- Maria Kokkou
Aix-Marseille University, FR
- Irina Kostitsyna
TU Eindhoven, NL
- Dominik Krupke
TU Braunschweig, DE
- Julien Leclerc
University of Houston, US
- Jakub Lengiewicz
University of Luxembourg –
Esch-sur-Alzette, LU
- David Liedtke
Universität Paderborn, DE
- Friedhelm Meyer auf der Heide
Universität Paderborn, DE
- Othon Michail
University of Liverpool, GB
- Irene Parada
UPC Barcelona Tech, ES
- Matthew J. Patitz
University of Arkansas –
Fayetteville, US
- Tom Peters
TU Eindhoven, NL
- Benoît Piranda
FEMTO-ST Institute –
Montbéliard, FR
- Andréa Richa
Arizona State University –
Tempe, US
- Kay Römer
TU Graz, AT
- Nicolas Schabanel
ENS – Lyon, FR
- Christian Scheideler
Universität Paderborn, DE
- Arne Schmidt
TU Braunschweig, DE
- Ahmed Shalaby
Maynooth University, IE
- Cynthia Sung
University of Pennsylvania –
Philadelphia, US
- Ryuhei Uehara
JAIST – Ishikawa, JP
- Giovanni Viglietta
JAIST – Ishikawa, JP
- Damien Woods
Maynooth University, IE



A Human-Computer Interaction Perspective to Drive Change towards Sustainable Future

Susanne Boll^{*1}, Kaisa Väänänen^{*2}, Nicola Bidwell^{*3},
Marc Hassenzahl^{*4}, and Robin Neuhaus^{†5}

1 University of Oldenburg, DE. susanne.boll@uol.de

2 Tampere University, FI. kaisa.vaananen@tuni.fi

3 University of Melbourne, AU & International University of Management, NA.
nic.bidwell@gmail.com

4 University of Siegen, DE. marc.hassenzahl@uni-siegen.de

5 Universität Siegen, DE. robin.neuhaus@uni-siegen.de

Abstract

In our everyday lives, people are constrained by routines, social expectation, and the soft and hard technologies and infrastructures that shape this. The way they approach things, think about things, are expected to be, and are governed is rarely questioned in terms of the finite nature of resources nor impacts of this. The challenge is to change the way people think and behave, and to reshape these very tools and expectations. However, change is exhausting, challenging, confronting, and requires support.

Technology can provide such a support, BUT it would be naïve to assume that this change will happen without friction, without dispute, and without constraints. But on the other hand, most of the conveniences that need to be changed are predicated on a false and falacious assumption that we can go as much, as fast, as high, and as pleasantly as we want without any regard for others. In this workshop, we explored how human computer interaction can facilitate, require, or even enforce the path we should take to use less, do slower, or act differently. In this Dagstuhl Perspectives Workshop we discussed the contribution that HCI can make in light of the SDGs and what role HCI must play in informing and changing the behavior of individuals and collectives.

Seminar February 26 – March 3, 2023 – <https://www.dagstuhl.de/23092>

2012 ACM Subject Classification Human Computer Interaction; Hardware → 3D integrated circuits; Software and its engineering → Software design engineering; Networks → Network performance analysis

Keywords and phrases Dagstuhl Perspectives Workshop

Digital Object Identifier 10.4230/DagRep.13.2.199

* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

A Human-Computer Interaction Perspective to Drive Change towards Sustainable Future, *Dagstuhl Reports*, Vol. 13, Issue 2, pp. 199–241

Editors: Susanne Boll, Nic Bidwell, Kaisa Väänänen, and Marc Hassenzahl



DAGSTUHL Dagstuhl Reports

REPORTS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Susanne Boll (University of Oldenburg, DE)

Kaisa Väänänen (Tampere University, FI)

Nic Bidwell (University of Melbourne, AU & International University of Management – Windhoek, NA)

Marc Hassenzahl (University of Siegen, DE)

License  Creative Commons BY 4.0 International license
© Susanne Boll, Kaisa Väänänen, Marc Hassenzahl, and Nicola Bidwell

Background and motivation

Sustainability from an ecological, social and economic perspective requires more attention than ever. To address and to implement the United Nation's set of sustainability goals is of utmost importance to all scientific communities, including information technology researchers. As a multidisciplinary field located within information technology, Human-computer interaction (HCI) constantly transfers knowledge and skills between the technically-oriented computer science, and psychology, social sciences, design, as well as other disciplines related to sustainability, such as environmental sciences. HCI's focus embeds technology into personal and societal practices and, thus, greatly influences the way technology is used in everyday life. At the same time, each new technology impacts the way we live. It shapes people's behavior and experiences. Through this mutual relationship, technological innovations have a great power to transform the world to be more ecological, resource-efficient, and inclusive. Not only can technologies be more resource-friendly, but also they can offer more ecological and social ways of interacting with the world. As a field, HCI is positioned to conceptualize, design and evaluate digital tools, able to suggest and to support the transformation of current into more sustainable lifestyles. Yet, while in principle HCI has all a field needs to contribute to a more sustainable future, it only hesitantly takes on the challenge of establishing sustainability at the heart of designing interactive information systems.

While sustainability is a global goal and needs many stakeholders from countries, organizations, and companies to make the transformation into sustainable societies happen, there is a large potential of individuals to make behavioral changes and contribute to sustainability as an individual and a collective. Digital interactive technology can play a major role in informing, supporting, reflecting, and acting in their everyday life and their individual societal and cultural context. However, this is not a single behavior change challenge in which we just have to offer a mobile app nudging us and that's it. Making such changes, taking different decisions comes with a large number of tensions such as who decides what the better lifestyle would be, and who could afford that change in their individual context.

A wider view on sustainability – the UN sustainability goals

Transformation requires goals. A good starting point for central requirements of sustainability are the UN sustainability goals. The 2030 Agenda for Sustainable Development, [1], adopted by all United Nations Member States in 2015, provides a shared blueprint for peace and prosperity for people and the planet, now and in the future. This agenda “is a plan of action for people, planet, and prosperity” and forms “universal goals and targets which involve the entire world, developed and developing countries alike.” At its heart are the 17 Sustainable Development Goals (SDGs), which shall provide direction to transform the world

via actionable objectives, from ending poverty to ensuring access to clean water and clean energy, to education and decent work for all. The goals assume that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests.

Digitalization will play an important, even transformative, role in meeting the SDGs, and so can the field of HCI. An open challenge though is to translate these goals into actionable HCI evaluation and design practice, academic and professional, alike.

Participants and workshop approach

This workshop brought together scientists related to HCI and sustainability for this Dagstuhl Perspectives Workshop. We built on the United Nations' Sustainable Development Goals (SDGs) to identify central opportunities for HCI to advance sustainability in research, education, and everyday information systems. Through a series of interactive discussions, we explored the tensions around a more sustainable map of research topics and approaches suitable for advancing sustainability through HCI research. As the main outcome, the workshop will formulate a manifesto and roadmap for Human-Computer Interaction contributing to our common sustainable future. The core questions of the Perspectives Workshop were

- When and where is sustainability negotiable? We can observe that humans, and particularly the most disadvantaged, hardly have a chance of a liveable future. What does this mean for designing interactive systems that support sustainable behavior? Is sustainable behavior actually negotiable or a must for all?
- How do we address conflicts between people's individual goals, business goals, and global sustainability goals?
- What can HCI offer for those who cannot afford a more sustainable lifestyle? Lifestyle HCI sustainabiliteers who can feel better versus an inclusive approach? How do we model the contextual factors of individuals who make / do not make sustainable decisions?
- Can interactive technology help people who want to change to make changes in their lifestyle towards a more sustainable behavior? We contrasted technology that tells people what to do with technology in which persons are involved in the steps of behavior change.
- Can technology outbalance injustice in the situations in which sustainability challenges are caused by a few?
- How can technology help people compensate differently, adequately, and fairly?

Results

The 26 participants from all over the world were bringing in different perspectives to the workshop not only from their scientific and methodological background but also their lens of sustainability challenges in and from different areas of the world. Before the workshop, we collected a reading list of own and related work to set the stage for the discussions. For the workshop we invited three talks (online and in-person) that were setting the stage for the role of HCI for a sustainable future, the tensions of making sustainable and unsustainable individual decisions, and if we can or want to live with much less than today and still be happy. Based on the framework of the United Nations Sustainability Goals we discussed the tensions and opportunities for HCI contributing to sustainability for individual SDGs.

We used speculative explorations of positive sustainable futures both from a technology perspective as well as from a research and practice perspective.

As a major result from the workshop:

1. We selected and discussed SDGs in which HCI can make a contribution to a sustainable future and identified potential and open challenges which we will present in Section 5.
2. We used speculation as a form of analysis of possible futures and developed three design fictions in group works in Section 6.
3. We identified core areas where HCI can make a contribution to a sustainable future which will be elaborated and published in a manifesto.

References

- 1 Department of Economic United Nations and Social Affairs Sustainable Development. Transforming our world: the 2030 agenda for sustainable development, 2015.

2 Table of Contents

Executive Summary

Susanne Boll, Kaisa Väänänen, Marc Hassenzahl, and Nicola Bidwell 200

Keynotes

Community, technology and HCI: Towards sustainable societies?
Jacki O’Neill 205

Digital and sustainable: Shaping the twin transition
Stephan Ramesohl 206

Torn Between Impulses and Reflection. Why Don’t You Just Let Me Carry On as Usual?
Matthias Laschke 206

Abstracts

Towards sustainable multimedia experiences
Katrien De Moor 208

Designing sustainable digiphysical experiences
Markus Fiedler 209

Situating HCI & moving the dial on sustainability
Adrian Friday 209

On the Importance of Policy and Impact for Sustainable HCI
Jason Jacques 210

Sustainable HCI – Focusing on Interactions of Communities and Societies
Kirsikka Kaipainen 210

An encouraging and hopeful exchange of diverse perspectives
Lenneke Kuijer 210

Sustainability and Development
Neha Kumar 211

Designerly Approaches to Behavior Change
Carine Lallemand 211

Justice-led Technology Design for Biodiversity Conservation
Joycelyn Longdon 212

Design for Imagination and Climate Futures
Dan Lockton 212

Crafting Positive Narratives for Sustainable Living
Robin Neuhaus 212

The Complicated Interconnections between Digitalization and Sustainable Development
Thomas Olsson 213

Collaborative HCI for Sustainability: Bridging Divides and Driving Change
Eunice Sari 214

What Is “the Transition” We Should Design for? <i>Shadan Sadeghian</i>	214
Marginalised Perspectives on Interaction Design & Sustainability <i>Katta Spiel</i>	215
Toward Sustainable Ubiquitous Computing-A Human-centered Approach <i>Tengxiang Zhang</i>	215
Working Groups on SDGs	
Working Group on SDG 3: Good Health and Wellbeing <i>Carine Lallemand, Matthias Laschke, Eunice Sari, Kaisa Väänänen</i>	216
Working Group on SDG 4: Quality Education <i>Eunice Sari, Markus Fiedler, Carine Lallemand, Thomas Olsson, Marc Hassenzahl, Robin Neuhaus</i>	218
Working Group on SDG 8: Decent Work and Economic Growth Education <i>Jason Jacques, Shadan Shadeghian, Katta Spiel, Robin Neuhaus, Thomas Olsson, Neha Kumar</i>	218
Working Group on SDG 10: Reduce inequalities <i>Daniel Pargman, Kaisa Väänänen, Neha Kumar, Kirsikka Kaipainen, Jacki O’Neill, Katrien De Moor</i>	219
Working Group on SDG 11: Sustainable cities and communities <i>Shadan Sadeghian, Dan Lockton, Anjali Karol Mohan, Jacki O’Neill, Lenneke Kuijer, Adrian Friday</i>	219
Working Group on SDG 12: Responsible Consumption and Production <i>Susanne Boll, Joycelyn Longdon, Katrien de Moor, Markus Fiedler, Gözel Shakeri, Lenneke Kuijer, Dan Lockton</i>	220
Working Group on SDG 13: Climate Action <i>Alexander Raake, Jacki O’Neill, Anjali Karol Mohan, Adrian Friday, Kirsikka Kaipainen, Daniel Pargman</i>	221
Working Group on SDG 16: Peace, Justice and Strong Institutions <i>Kaisa Väänänen, Anjali Mohan, Katrien de Moor, Robin Neuhaus, Katta Spiel</i>	222
Speculations	
Introduction and Method	223
Four Speculations	223
Reflection: Emerging Themes	227
Conclusions and Actions	
About the organizers	230
Literature overview	233
Participants	241
Remote Participants	241

3 Keynotes

One goal of the Dagstuhl Perspectives Workshop was to focus on the role of interactive digital technology for a more sustainable future. We invited three different talks (online and in-person) that are setting the stage for the role of HCI for a sustainable future, the tensions of making sustainable and unsustainable individual decisions and if we can or want to live with much less than today and still be happy.

- *Sustainability and the role of digitization*: As there is a large field and there is a large body of research about the key threats but also the key potentials digitization offers to this transformation, we set the stage by an invited talk about key facts, but also urban legend and opportunities for digital interactive technologies for a more sustainable future.
- *Sustainability and social injustice*: We will invited a speaker from the context of the Climate in Color platform which explores and discusses the intersection of climate science and social justice and is making climate conversations more accessible and diverse. This will address the tensions of often unequally distributed effects and opportunities between different social groups.
- *Living with less*: Key speakers in the field would argue that for a sustainable future we will need to let go of constant growth and consumption but transform into a post-growth economy. An invited talk will present the challenges and opportunities of a post-growth economy and the role digitization and interactive personal technology can play in this.

3.1 Community, technology and HCI: Towards sustainable societies?

Jacki O'Neill (Microsoft Africa Research Institute – Nairobi, KE,
Jacki.ONeill@microsoft.com)

License © Creative Commons BY 4.0 International license
© Jacki O'Neill

Sustainability has tended to focus on individual rather than collective freedoms; the freedom to buy and consume and behave in certain ways, not the freedom of the world's children and grandchildren to live in a world free of the devastation of climate change. While, many projects, including examples from Africa, use technology to address climate change by, for instance, collecting data, improving predictions or carbon sequestering, technology cannot magically solve climate change on its own. Large scale governmental, community and collective action are needed to really change the supply chains, markets and ways of living and working that support unsustainable consumerism and neoliberal ideas of "choice". Since, democratic governments have little incentive to make changes that will be unpopular and impact perceived quality of life, they must be pushed into action. Technology can help to empower citizens in this conversation, e.g., using satellite imagery for storytelling. However, with some notable exceptions, most of sustainable HCI is driven by the interests and needs of the Global North and is about individual change. This not only misses the 80 percent of the world's middle class who will live in Global South in 2030 but misses the point of addressing the global systemic problem. Framing design for sustainability as the informed choice of individual consumers arises because we conceive a market economy to be a natural, not a social, construction in the neoliberal world. As a result grants and funding cycles focus on interventionist innovations not problem solving and disciplinary boundaries determine the types of things we work on and separate research from practice and action. A possible

answer is to engage with projects and people on-the-ground doing sustainability work, such as climate scientists and activists, government and policy makers, data scientists and engineers. This may require rethinking HCI and reframing its own isolated and individualist approach. *Talk summarized by Nic Bidwell*

3.2 Digital and sustainable: Shaping the twin transition


Stephan Ramesohl (Wuppertal Institut, Research Unit Digital Transformation, DE, stephan.ramesohl@wupperinst.org)

License  Creative Commons BY 4.0 International license
© Stephan Ramesohl

The talk provided insights into the challenges posed by anthropogenic climate change and discussed the role the digital could play in supporting the necessary transformations. Socio-economic and technological innovations must go hand in hand. Together they provide new opportunities for a transformation towards more sustainability. The talk presented two in-depth examples of how technology can support and drive socio-cultural change in the domain of personal mobility and business models for a circular economy. In general, digital technology must be understood as a network of opportunities to shape everyday practices to be sustainable. In the mobility domain, for example, traffic management, sharing platforms, virtualisation (meetings, homeoffice), autonomous driving, digital government solutions, as well as new approaches to town planning together create a the transformation, which enables people to be less mobile, to choose eco-friendly modes of transportation (bike, walking, public transport) as well as to make necessary individual mobility more efficient. *Talk summarized by Marc Hassenzahl*

3.3 Torn Between Impulses and Reflection. Why Don't You Just Let Me Carry On as Usual?

Matthias Laschke (University of Siegen, Interaction Design for Sustainability and Transformation, DE, matthias.laschke@uni-siegen.de)

License  Creative Commons BY 4.0 International license
© Matthias Laschke

In light of the threatening consequences of anthropogenic climate change, it seems paramount to reduce individual energy consumption. For example, within the European Union, as much as 24 percent of the yearly energy is consumed by individual households. This is due to the many energy-intensive daily routines or activities of consumers within their households, such as heating, taking hot showers, or doing several courses of laundry a day. Thus, one way to counteract the consequences of climate change would be to reduce energy consumption of these activities, for example, by reducing the room temperature, showering more quickly, or by reducing the amount of laundry and its energy consumption. Unfortunately, consumers rarely act accordingly. A possible explanation is that climate objectives are very abstract and future-oriented goals and, therefore, not easily put into practice. In addition, the many energy-intensive daily activities offer direct gratification through immediate comfort, which will increase their execution. Therefore, consumers might need external support to follow and

Digital solutions are socio-technological innovations

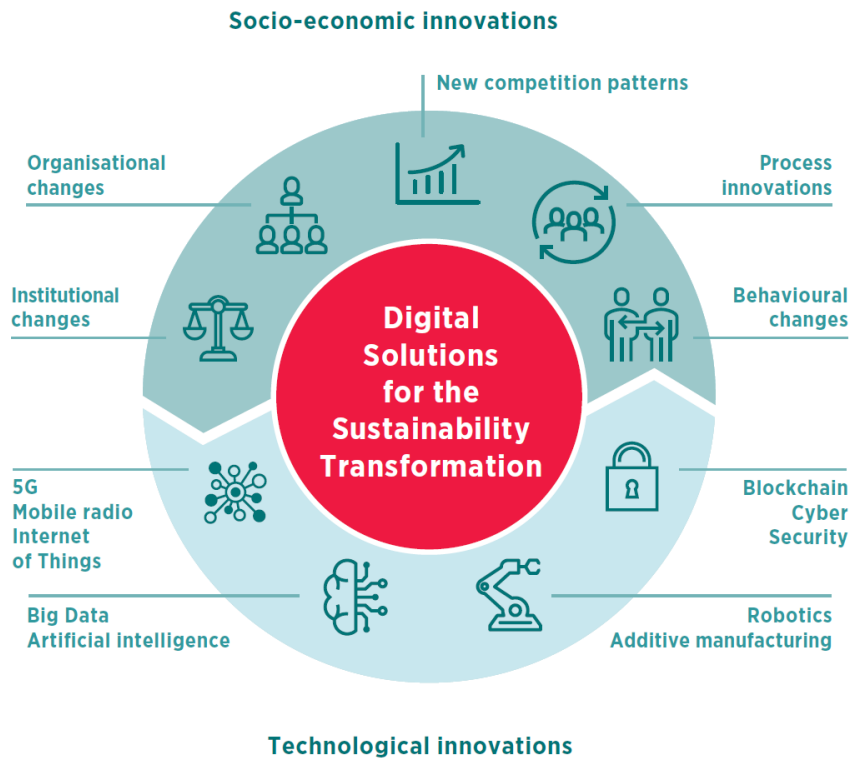


Figure 1 Digital Solutions for the sustainability of transformation.

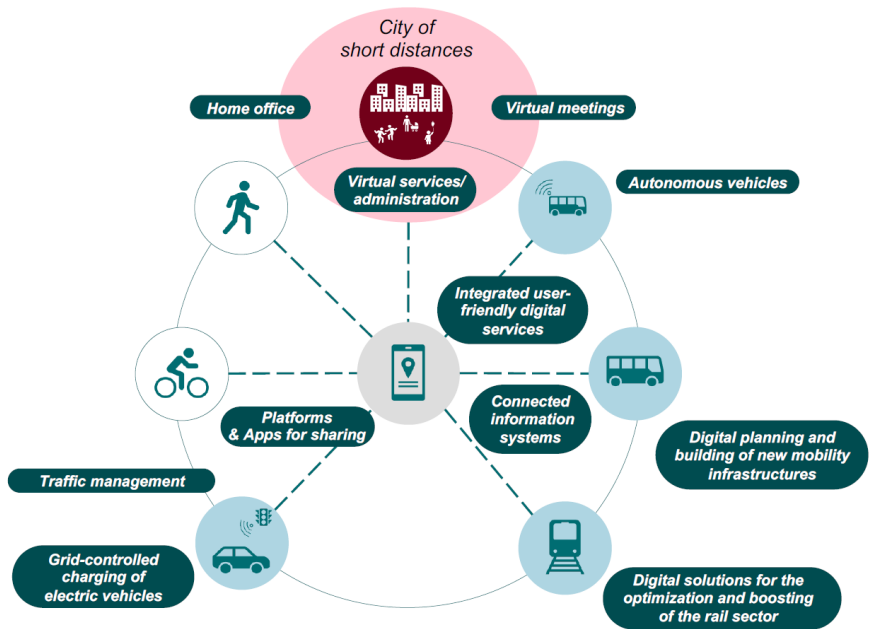


Figure 2 Example.

achieve the abstract goal of reducing individual energy consumption. As most of the energy-consuming activities are linked to or achieved through the use of technology, technology design is a possible lever to foster behavior change and reduce individual energy consumption. For quite a while now, Human-Computer Interaction (HCI) studies how technology could change consumer behavior. Broadly, two different types of approaches are offered: (1) feedback and arguments to persuade consumers to behave differently (e.g., Persuasive Technology) or (2) situational interventions (i.e., “choice architectures”) to prevent impulsive actions and to disturb unconscious unwanted routines. The underlying assumption of persuasion is that users make rational decisions but lack the “right” information and knowledge to do so. From a psychological perspective, this approach is based on the precondition of a reflective system in which behavior emerges as the result of conscious decision-making. In the context of sustainable behavior, technologies designed according to this approach, for example, confront the user with information on how much hot water is consumed during a shower or visualize water as a limited amount that decreases. The underlying assumption of situational interventions is that people primarily do not act rationally but are driven by mostly unconscious impulses, automatic routines, and limited resources to make conscious decisions. In contrast to a reflective system, these approaches build upon an impulsive system in which behavior emerges from previously learned schemata. Technologies that implement this approach, for example, confront users with the choice between a bicycle or a car key to break routines or hide less sustainable choices and even remove them completely.

4 Abstracts

In alphabetical order.

4.1 Towards sustainable multimedia experiences

Katrien De Moor (Norwegian University of Science and Technology – Trondheim, Torgarden, NO, katrien.demoor@ntnu.no)

License  Creative Commons BY 4.0 International license
© Katrien De Moor

My research focuses on human-technology experiences and in particular, multimedia and IoT-technology enabled experiences. My activities in these areas are grounded in human-centered approaches and focus on aspects related to the design, evaluation and use practices of different types of media: how can they be made more sustainable and trigger more sustainable use practices? How can we ensure that they are inclusive and put other ethical considerations at the core? More recently, I’m also working on user awareness (and lack of it) of the “invisible” carbon footprint of digital media. How can we support the exploration and uptake of concrete strategies and mechanisms that may help to trigger critical reflection and more conscious and responsible consumption (both from a well-being and environmental point of view)? And how can we – as researchers and educators – push stakeholders with decision power to take action? This week has given me hope and courage to help addressing these and some of the related big questions and challenges.

4.2 Designing sustainable digiphysical experiences

Markus Fiedler (*Blekinge Institute of Technology – Karlshamn, SE, markus.fiedler@bth.se*)

License © Creative Commons BY 4.0 International license
© Markus Fiedler

Personally, I have been interested in the co-existence of ICT and sustainability for 1½ decades, and it has been interesting to see the engagement of some researchers on one side, and the reluctance of scientific communities to pick on the sustainability challenges coming with ICT. As designers of ICT applications and technologies that allow to bridge between various realities (on the continuum between physical and digital experiences), we have great opportunities to make differences in the ways that people perceive and upon sustainability issues. From own previous experience, quality and energy efficiency can well come together, and our knowledge in those areas can enable “Quality of Experience and Sustainability by Design”. In any case, it is very nice to see how the interest in sustainability issues within ICT is taking up speed, which gives hop for more sustainable solutions that enable more sustainable behaviour.

4.3 Situating HCI & moving the dial on sustainability

Adrian Friday (*Lancaster University, GB, a.friday@lancaster.ac.uk*)

License © Creative Commons BY 4.0 International license
© Adrian Friday

Nearly 15 years ago, I recognised that I had become disatisfied that my research was centred around the creation of novel technologies, without addressing the ways in which these technologies were and are impacting society and the environment. I hadn’t fully appreciated at this time, and perhaps I still don’t, but there’s a deep relationship between technology, economics, society and environment – and also how technology can reinforce and even accelerate unsustainable growth. I went on to study energy use in the home, how we use energy to achieve thermal comfort, how technology can promote or negatively impact environmental and social justice in digitally mediated work, and ICT’s global impacts. This journey has helped me realise that digital technology is a tool for transformation, but it’s also political, economic and embedded almost invisibly in society. What are the roles for technology not just in optimising and tuning an unsustainable way of living; or in plugging supposed information deficits – but in really changing thinking; challenging and making visible the drivers of unsustainability; creating empathy and radical coordinated action; and driving bold leadership and governance on climate. I enjoyed the perspectives from meeting people from across the world variously engaged with climate and technology; plus the solidarity and kinship that we’re working together. This reminded me that the climate change impacts are urgent, global, colonial, and that technology also has a role in connecting us with those most impacted. A lasting impression is our “ecopill” design, which helped me confront the pervasive presumption that many (too many?) of us have the choice to consume at the expense of others. It is discomfiting that in the first world at least, we can only seemingly choose a sustainable future if we can afford and buy our way there as a consumer? How do we give space for other non-techno-centric, non-growth centred, and nature respecting paradigms?

4.4 On the Importance of Policy and Impact for Sustainable HCI

Jason Jacques (University of St. Andrews, GB, jtj2@st-andrews.ac.uk)

License  Creative Commons BY 4.0 International license
© Jason Jacques

My work mainly focuses on the ethical application of labour and how technology influences decisions and lives. My recent work has focused on how our activities affect the environment. In our discipline we often resort to quantitative data before we accept the impact – which is challenging. The climate impact is colloquially settled, but the multifaceted impact is so broad. In many ways technology impacts policy. This has been a story/theme of the workshop. However, equally important is how policy impacts behaviour. I hope that this mutual feedback cycle can be capitalised on, and cause us to develop our technology to feed into this aspect to support a positive impact on sustainability, both for the community in the room – but also more broadly as our work influences colleagues, our institutions, and the wider world.

4.5 Sustainable HCI – Focusing on Interactions of Communities and Societies


Kirsikka Kaipainen (Tampere University, FI, kirsikka.kaipainen@tuni.fi)

License  Creative Commons BY 4.0 International license
© Kirsikka Kaipainen

We need to broaden our understanding of people in different communities with different backgrounds in order to drive sustainable futures. Personally, this is one of the reasons I came to this workshop. Although all participants are working in the academic world, there are various perspectives and lived experiences to learn about. I'm also searching for the next best step to take in my professional world, where to best use my own resources for mitigating climate change and addressing social sustainability challenges. While individual level behavior change is important (and tech can definitely support it), it seems urgent to focus on interaction between different societal levels, such as how HCI could help in deliberative democracy, and also how to determine what kinds of “lifestyles” are okay for people: forming a big picture, looking at and analyzing people's and communities' footprints and handprints, providing informed decision support for decision-makers. Already, this workshop has given me a lot of food for thought.

4.6 An encouraging and hopeful exchange of diverse perspectives

Lenneke Kuijer (TU Eindhoven, NL, S.C.Kuijer@tue.nl)


License  Creative Commons BY 4.0 International license
© Lenneke Kuijer

Ever since my PhD research, my work has situated itself at the touching points of design, domestic energy demand and social practice theories. When I moved to the TU Eindhoven's Industrial Design Department this shifted to a stronger focus on interaction design. I find the HCI community very inspiring due to its diversity of perspectives and generally rigorous

& constructive review culture. This exchange of diverse perspectives and constructive atmosphere is also something that I particularly appreciate about this workshop. Like the others present, I am deeply concerned about the state of the world and feel helpless often with regards to the things I can personally do to change this. I find it inspiring and stimulating to find others struggling with very similar challenges and emotions; it eases some of the pain and offers hope and inspiration. I believe HCI and I personally as a researcher and teacher can make a difference. Systemic understanding translated into concrete, strategic interventions can play a key role here. Thank you very much.

4.7 Sustainability and Development

Neha Kumar (Georgia Institute of Technology – Atlanta, USA, neha.kumar@gatech.edu)

License  Creative Commons BY 4.0 International license
© Neha Kumar

My research lies at the intersection of Human-Computer Interaction (HCI) and sustainable global development. I focus on care infrastructures, from global health informatics and integrating data-driven/AI approaches with care to care work(ers), employing care-based methods, as well as planetary care. Also in my role as serving president of SIGCHI, I work towards global inclusivity and (hopefully) a more sustainable ecosystem of HCI conferences. This workshop has been fantastic in bringing together such an incredible group of conversations on sustainability (and the broad, diffuse scope of this word) as well as the sustainable development goals all day yesterday. I plan to take back some of this energy to my research, teaching, and service responsibilities in my university and beyond. Looking forward to many more stimulating discussions over the next few days. Excellent as this experience has been, it would be fantastic to, next time, have/see more participation from those who live and work in the Global South – in Dagstuhl or elsewhere. Front of mind today, as I listen to Jacki O’Neill’s keynote talk, is how we can approach this work more radically, move towards more hybrid ways of operation, work towards doing more with less, and maybe doing less in some ways as well.

4.8 Designerly Approaches to Behavior Change

Carine Lallemand (University of Luxembourg, LU & TU Eindhoven, NL, c.e.lallemand@tue.nl)

License  Creative Commons BY 4.0 International license
© Carine Lallemand

My research interests are mainly focused on experience design and evaluation methods, with a strong will to develop methods and tools that can be relevant for practitioners. As a scientist, I want to “equip” those who are currently designing technologies and shaping experiences of individuals, groups, communities, society. I also have some experience working on topics related to behavior change for healthy lifestyles, especially sedentary office work or exercising motivation. These projects contribute to both physical and mental vitality. They rely on aesthetics of interaction principles and the use of data as a creative material. To question the status quo, I also enjoy exploring alternative ways of designing, for instance through critical and speculative design, the aesthetics of friction or slow technology. The SDGs, I am most

closely related to are: 3/ good health, and well-being 4/ quality education 8/ Decent work and economic growth I am however excited to spend the week here in Dagstuhl discovering other perspectives and envisioning how we, HCI researchers, educators and citizens, could contribute to the sustainability challenges ahead.

4.9 Justice-led Technology Design for Biodiversity Conservation

Joycelyn Longdon (University of Cambridge, GB, jl2182@cam.ac.uk)

License  Creative Commons BY 4.0 International license
© Joycelyn Longdon

My research interests are focused on co-designed/co-created technologies for tropical forest biodiversity conservation, specifically working with rural forest-fringe communities that are commonly excluded from conservation and technology research. I am currently working with bioacoustic technologies that use environmental sound recordings and machine learning to monitor biodiversity. Through interactive visualisation and participatory design methods, I hope to build an interface that acts as a boundary object between conservationists and forest communities and addresses tensions that arise with the “datafication” of forests.

4.10 Design for Imagination and Climate Futures

Dan Lockton (Eindhoven University of Technology, NL, d.j.g.lockton@tue.nl)

License  Creative Commons BY 4.0 International license
© Dan Lockton

My work as a researcher / educator / designer explores designing tools (of different kinds) for participatory (re-)imagining and futuring in an age of transitions (and crises) in climate, energy, and social inequalities. We need new models – new ways to think and new ways to live – to help us imagine, understand, live, and even what we see as possible in our collective futures. Design and HCI have a role to play, and the urgency of climate crisis makes this really very acute. Design methods can enable people to share their experiences with others, giving voice (or other appropriate metaphor) to underrepresented ideas and hopes, and turning these ideas into experienceable prototypes which we can live with and reflect on, sharing our imaginaries with each other. We can bring plural possible futures to life in the present. I would like the HCI/Sustainability community to explore and value the possibilities of imagination in relation to climate futures, and to value the capabilities to bring alternative models to life.

4.11 Crafting Positive Narratives for Sustainable Living

Robin Neuhaus (University of Siegen, DE, robin.neuhaus@uni-siegen.de)

License  Creative Commons BY 4.0 International license
© Robin Neuhaus

In light of ever accelerating climate change, we are facing a regional and global challenge to change the way we live towards a more sustainable lifestyle, in order to maintain a livable planet for everyone. Sustainability has emerged as a critical objective in all areas of life and

is not bound by disciplines. As designers and practitioners in the field of human-computer interaction (HCI), we must contribute to this common goal and re-evaluate our goals and actions. As tempting as it might be, we should not solely seek to contribute by making things more efficient or to solve immediate problems with more technology. Alternatively, I believe we can leverage HCI's methodological tools and their potential to shape habits and society through technology and applications. One key aspect is the importance of crafting positive narratives for a future sustainable way of living. While the future can be challenging to imagine and change can be intimidating, HCI is accustomed to exploring and visualizing potential future technology, its use, and even its potential societal impacts. By employing speculative methods, HCI researchers and designers can actively engage with different possible futures. Here, approaches such as speculative design and design fiction show promise to depict and negotiate desirable ways of living sustainably, highlighting the positive role technology can play – or where less technology can be beneficial. Speculative prototypes can facilitate discussions about how to rethink practices and aspects of our daily lives to be more sustainable. By visualizing a part of a possible future as if it was real, we can find out what is acceptable or realize how such changes can even spark joy.

4.12 The Complicated Interconnections between Digitalization and Sustainable Development


Thomas Olsson, (Tampere University, FI, thomas.olsson@tuni.fi)

License © Creative Commons BY 4.0 International license
© Thomas Olsson

In terms of sustainable development, digitalization can be viewed through Jacques Derrida's concept of *pharmakon*: a composite of remedy, poison, and scapegoat. This reminds us that (i) different sustainability goals might be sensibly pursued by the means of digitalization, (ii) the processes and products of digitalization may also be harmful and unsustainable, even if the intentions were benign, and (iii) digitalization might be unfairly blamed for the faults elsewhere in our society. I think this is a wonderfully simple, yet powerful framework to regard the complicated relationship between these two megatrends shaping our times. My own contribution to this area has so far focused on developing a course for master students in Human-Computer Interaction, aiming to educate the future UX/interaction designers more critical perspectives, systems thinking, and social scientific viewpoints in general, as well as to underline the importance of sustainability also in this seemingly de-material and apolitical professional field. Additionally, recently we've been trying to consider the digital world through the concept of *digital excess*, and I'll happily tell more about that another time!

4.13 Collaborative HCI for Sustainability: Bridging Divides and Driving Change

Eunice Sari (UX Indonesia – Jakarta, ID & University of New South Wales – Indonesia and Australia, eunice@urindo.com)

License  Creative Commons BY 4.0 International license
© Eunice Sari

As a academic representing both the Global North and South, my pragmatic sustainability perspective addresses the need for increased awareness in the Global South. Climate change impacts have fueled this awareness, leading many young individuals to initiate community-based projects. Sustainability has become integral to my academic and non-academic work in the Global North. However, significant gaps in understanding sustainability persist across contexts. I aim to identify these gaps and foster unity through collaboration. The recent Dagstuhl Perspectives Workshop on A Human-Computer Interaction Perspective to Drive Change Towards a Sustainable Future has been enlightening. Engaging with leading experts in HCI-UX and sustainability, we focus on the well-being of our planet, exploring possibilities, potentials, and challenges to find optimal solutions.

4.14 What Is “the Transition” We Should Design for?

Shadan Sadeghian (University of Siegen, DE, shadan.sadeghian@uni-siegen.de)

License  Creative Commons BY 4.0 International license
© Shadan Sadeghian

Sustainability has been discussed as a prominent goal in legislation, political, and development processes. Specifically with the current environmental crisis, there is an urgent need for action for radical changes towards sustainable development. Sustainability is the result of practices and behaviours that makes the future life of human and non-human beings better on environmental, social, and economical levels. The question is how to make these radical changes and whether people are ready or willing to make those changes? Through this workshop, we discussed that change should be rather systematic and can primarily be induced through legislation and policies. However, the role of technology in changing towards sustainability through use and interaction with it cannot be ignored. The introduction of new technologies per se could imply new cycles of production which could lead to more CO2 emissions. Nevertheless, we cannot put a stop on production, but we can produce more sustainably and for sustainability. One question is how many products we need, and how can HCI that aims for making more usable, enjoyable products avoid over consumption? Of course, omitting technologies comes with changing the practices around their use. For example, removing the car from a family requires them to redefine their daily commutes, and this only will be possible if they have access to required resources such as public transport, or are physically fit to bike. Another way is to design technologies that change the practices through their use, contribute to sustainability while fulfilling other needs (e.g., riding a scooter improves health besides being a way to commute). Designing technologies that are attractive to people and can be integrate to their daily practices can also contribute to economic growth which is one of the main drives of big players in industry, politics, and legislation.

4.15 Marginalised Perspectives on Interaction Design & Sustainability

Katta Spiel (TU Wien, AT, katta.spiel@tuwien.ac.at)

License © Creative Commons BY 4.0 International license
© Katta Spiel

I research marginalised perspectives on technology through a lens of Critical Access, particularly around embodied computing. There, I'm focusing on issues of (minority) genders & disability (specifically neurodivergent and Deaf populations; though I'm hearing). My work hopes to inform the design & development of technologies in ways that allow for reflection, community and access beyond formalised criteria checklists. The workshop here shows me how social & ecological sustainability are modes that co-constitute each other by necessity. We won't address issues of the climate catastrophe without decentering our individual(istic) experiences and attending to the (human & non-human) others in a love-oriented epistemological relationship.

4.16 Toward Sustainable Ubiquitous Computing-A Human-centered Approach

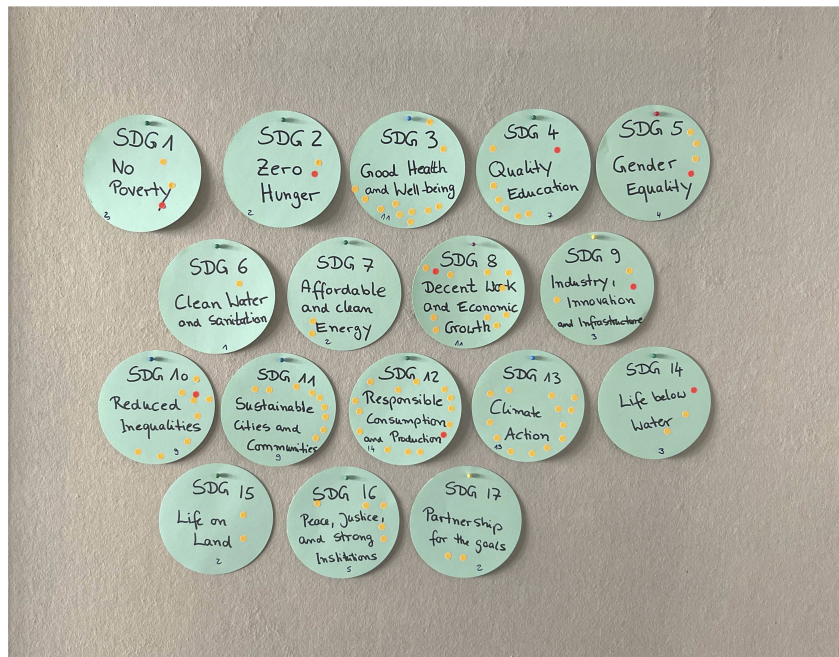
Tengxiang Zhang (Chinese Academy of Sciences – Beijing, CN, tengxiangz@sigchi.org)

License © Creative Commons BY 4.0 International license
© Tengxiang Zhang

The vision of ubiquitous computing, characterized by pervasive sensing, computing, and interactive electronic devices, presents significant sustainability challenges. In his talk, Dr. Zhang discussed his research that tackled these sustainability issues in the context of ubiquitous computing. Using a human-centered approach, Dr. Zhang suggested centralizing both computing power and energy on a few wearable devices on the body. This concept aimed to make off-body ubiquitous devices battery-less and computation-free, contributing to their sustainability. To bring this concept to life, Dr. Zhang developed backscatter-based ultra-low-power sensors and interaction-based power transfer mechanisms. These innovations are steps towards creating a more sustainable ubiquitous computing future.

5 Working Groups on SDGs

Through small group work and plenum discussions, we explored the role of each of the selected SDGs and its targets in current HCI research and practice. This includes substance (What type of interactive products are related to the goal? What type of research questions and interventions are already addressing these goals?) as well as methods (What type of development/research methods are already conducive to the goal?). This resulted in a map of topics and methods in HCI, which already address sustainability and an overview of white spots, that is, future opportunities to engage in sustainability through HCI.



■ **Figure 3** Poster Board from our discussion about the 17 SDGs and potentials we see for HCI.

5.1 Working Group on SDG 3: Good Health and Wellbeing

Carine Lallemand (University of Luxembourg, LU & TU Eindhoven, NL, c.e.lallemand@tue.nl)

Matthias Laschke (University of Siegen, DE, Matthias.Laschke@uni-siegen.de)

Eunice Sari (UX Indonesia – Jakarta, ID & University of New South Wales – Sydney, AU, eunice@uxindo.com)

Kaisa Väänänen (Technical University of Tampere, FI, kaisa.vaananen@tuni.fi)

License © Creative Commons BY 4.0 International license
© Carine Lallemand, Matthias Laschke, Eunice Sari, Kaisa Väänänen

SDG 3 Good Health and Wellbeing “ensures healthy lives and promotes well-being for all at all ages”. In our discussion, we first agreed on a possible approach and structure for our analysis. Here, the division into three forms of knowledge – system knowledge, target knowledge, and transformation knowledge (e.g., according to Pohl, C., Hirsch Hadorn, G. (2006): *Gestaltungsprinzipien für die transdisziplinäre Forschung*, Munich: Oekom Verlag) – seemed to be particularly useful for the analysis. This approach is often applied in the context of sustainability science and involves an iterative process of learning, planning, and implementation involving multiple stakeholders and perspectives. Systems knowledge involves understanding the structure and function of the system under study, while target knowledge involves defining clear goals for sustainability. Transformational knowledge refers to the strategies and actions needed to achieve those goals, including identifying potential trade-offs and developing strategies and actions to facilitate transformative change.

The goal to be achieved was set by the SDG 3 “Good Health and Wellbeing ensures healthy lives and promotes well-being for all at all ages”. Therefore, we first focused our discussion on the existing knowledge (i.e., system knowledge) in HCI regarding the goal.

The following items summarize which HCI topics provide system knowledge related to SDG 3:

- Impact of environment on health and subjective wellbeing: Climate change affects not only the environment but also the health and wellbeing of humans. There are, besides the existence of mankind, also very concrete and short-term interests to stop climate change as soon as possible.
- Access to healthcare (i.e., inclusive healthcare): Universal access to healthcare is almost non-existent worldwide
- Data/AI developments and lack of data literacy: More and more data is available. This could be used for the benefit of humanity and nature. At the same time, few people can process and interpret this data. Data is not even a separate subject in school, although it is central in an information age.
- Intrinsic motivation from people: People and their motivation are deciding factors when it comes to the big transformation.
- Tech-push solutions rather than supporting human needs: The development of new technologies in HCI is still too much based on new technological rather than social possibilities. Despite a certain technological saturation (i.e., phenomena like “good enough”), enhancing well-being is not at first place in innovation development.
- Productivity as a driver or innovation rather than subjective wellbeing: The current economic system, to which HCI belongs and contributes, is focused on growth and productivity. Self-determination and an increase in wellbeing have a subordinate role. They are rather soft factors.

The following items represent topics we covered in our discussion of how HCI’s research efforts already contribute to achieving SDG 3:

- More than one human (design)
- Personal informatics (self-tracking)
- Telemedicine/Homecare
- Explainable/transparent AI
- Positive design/Experience design
- Exploration of meaningful design spaces of new technologies
- Low- or no tech
- Calm/warm/slow technologies
- Workplaces that promote wellbeing
- Physcalisation/visualization of data
- Healthcare for women

5.2 Working Group on SDG 4: Quality Education


Eunice Sari (UX Indonesia – Jakarta, ID & University of New South Wales – Sydney, AU, eunice@uxindo.com)

Markus Fiedler (Blekinge Institute of Technology – Karlshamn, SE, markus.fiedler@bth.se)

Carine Lallemand University of Luxembourg, LU & TU Eindhoven, NL, c.e.lallemand@tue.nl)

Thomas Olsson, (Tampere University, FI, thomas.olsson@tuni.fi)

Marc Hassenzahl (University of Siegen, DE, marc.hassenzahl@uni-siegen.de)

License  Creative Commons BY 4.0 International license

© Eunice Sari, Markus Fiedler, Carine Lallemand, Thomas Olsson, Marc Hassenzahl, Robin Neuhaus

SDG 4 Quality Education (QE) aims to ensure inclusive and equitable quality education, and HCI provides a set of opportunities: 1) Improved accessibility to education for everyone, no matter of location, disabilities, socio-economic conditions, and situations in life. HCI promotes lifelong situated and experiential learning (and assessment) opportunities, as well as stakeholder participation in design and operation. 2) New or improved learning experiences. HCI allows for designing digital tools and interfaces that facilitate positive, engaging and effective learning experiences, social connections between teachers and learners from different places, resilience, lifelong and challenge-based education. 3) Improved organisation. HCI supports the right-scaling of educational resources and joint curriculum development. It helps to remove institutional barriers and enables a holistic end-to-end learning process. However, HCI also brings a set of challenges of relevance for any QE-related design process, for instance: 1) Focus on socio-technical aspects beyond infrastructures; 2) Match people and resources; 3) Realise that technologies make their own policies and realities; 4) Increase digital literacy; 5) Teach HCI beyond universities; 6) Articulate our values; 7) Frame societal issues; and 8) Combine research and activism.

5.3 Working Group on SDG 8: Decent Work and Economic Growth Education

Jason Jacques (University of St. Andrews, GB, jtj2@st-andrews.ac.uk)


Shadan Sadeghian (University of Siegen, DE, shadan.sadeghian@uni-siegen.de)

Katta Spiel (TU Wien, AT, katta.spiel@tuwien.ac.at)

Robin Neuhaus (University of Siegen, DE, jtj2@st-andrews.ac.uk)

Thomas Olsson (Tampere University, FI, thomas.olsson@tuni.fi)

Neha Kumar (Georgia Institute of Technology – Atlanta, USA)

License  Creative Commons BY 4.0 International license

© Jason Jacques, Shadan Shadeghian, Katta Spiel, Robin Neuhaus, Thomas Olsson, Neha Kumar

Our group looked at the targets and indicators and were in consensus that these were too focused on numbers, and failed to capture meaningfulness of work. We discussed how decent work may or may not be meaningful, and how HCI could focus on work becoming more focused on meaning. As an example, we considered academia and how work models in academia target numbers over quality of life, even though “when we are happy we are performing better” arguably. On the topic of growth, we talked about what would entail “sensible growth” or “slowth” and how this might impact our HCI research and practice. We concluded that “capitalism is likely not saving us.”

5.4 Working Group on SDG 10: Reduce inequalities

Daniel Pargman (KTH Royal Institute of Technology – Stockholm, SE, pargman@kth.se)

Kaisa Väänänen (Technical University of Tampere, FI, kaisa.vaananen@tuni.fi)

Neha Kumar, (Georgia Institute of Technology – Atlanta, USA)

Kirsikka Kaipainen, (Tampere University, FI, kirsikka.kaipainen@tuni.fi)

Jacki O’Neill (Microsoft Africa Research Institute – Nairobi, KE,

Jacki.ONeill@microsoft.com)

Katrien de Moor (Norwegian University of Science and Technology – Trondheim, Torgarden, NO, katrien.demoor@ntnu.no)

License © Creative Commons BY 4.0 International license

© Daniel Pargman, Kaisa Väänänen, Neha Kumar, Kirsikka Kaipainen, Jacki O’Neill, Katrien De Moor

We started our discussion on SDG 10 “Reduced inequalities” by an initial reflection on the understanding of inequality and on how “improvement” in this respect may take very different forms. Briefly discussing all goal targets and indicators associated to SDG 10 led to a shared view that (1) the proposed targets are not very radical and that (2) there is a lack of goal targets aiming to reduce growth of those with a higher income as a strategy to address this goal (instead of targeting mainly income growth from below). While we discussed various opportunities for HCI in the context of SDG 10, some of the recurring themes have to do with understanding inequality and barriers to equality at different levels; methods and tools for increasing representation and for giving a voice to those who are currently not heard (based on an understanding of how they want to be voiced) and for ensuring that they are heard; research and practices to unpack bias in technology (e.g., bias in AI) and to transition to more transparent and bias-aware technology. However, there are various barriers to (further) realising the potential role that HCI can play, one of them being how to make people care? How to ensure that relevant activities in this respect trigger action, e.g., at the policy level?

5.5 Working Group on SDG 11: Sustainable cities and communities

Shadan Sadeghian (University of Siegen, DE, shadan.sadeghian@uni-siegen.de)

Dan Lockton (Eindhoven University of Technology, NL, d.j.g.lockton@tue.nl)

Anjali Karol Mohan (INDÉ – Bangalore, IN, anjali@integrateddesign.org)

Jacki O’Neill (Microsoft Africa Research Institute – Nairobi, KE,

Jacki.ONeill@microsoft.com)

Lenneke Kuijer (TU Eindhoven, NL, S.C.Kuijer@tue.nl)

Adrian Friday, (Lancaster University, GB, a.friday@lancaster.ac.uk)

License © Creative Commons BY 4.0 International license

© Shadan Sadeghian, Dan Lockton, Anjali Karol Mohan, Jacki O’Neill, Lenneke Kuijer, Adrian Friday

Our starting point was looking at societal and economic drivers that we were seeing as working against environmental and social sustainability. We believe that systemic change is needed to redress the balance between forces of globalisation, environmental and societal exploitation that are working against social and environmental sustainability. Especially in light of recent geo-political changes highlighting how long, remote or complex supply chains can be disrupted and are neither safe nor resilient under assertions and future impacts of climate change.

How is the flow of money, and digitisation of services with environmental and climate impacts, and of labour driving these systems. What role does HCI and the design of digital technologies in reshaping these to better meet local needs and reduce the externality costs? Ultimately, we felt that an opportunity space led in addressing and supporting pro-social and pro-community activities, and addressing and correcting folk theories and anti-climate misinformation and discourses of delay. We saw this as “digital infrastructuring” of sustainability, and promoting better awareness and connection to technologies, business and governmental failures to act with regards to “theft of the commons” to promote pro-sustainable agency in all its forms. We finished with a discussion about how to transform deeply embedded and market led aspirations of consumerism to provide a notion of alternative values to make being more sustainable aspirational, c.f. “attractive sustainability”. How might we amplify and build community around potentially disparate actors, overcome barriers, and give voice for greater solidarity and drive toward pro-sustainable community?

5.6 Working Group on SDG 12: Responsible Consumption and Production

Susanne Boll (University of Oldenburg, DE, susanne.boll@uni-oldenburg.de)

Joycelyn Longdon (University of Cambridge, GB, jl2182@cam.ac.uk)

Katrien de Moor (Norwegian University of Science and Technology – Trondheim, Torgarden, NO, katrien.demoor@ntnu.no)

Markus Fiedler (Blekinge Institute of Technology – Karlshamn, SE, markus.fiedler@bth.se)

Gözel Shakeri (University of Oldenburg, DE, gozel.shakeri@uni-oldenburg.de)

Dan Lockton (Eindhoven University of Technology, NL, d.j.g.lockton@tue.nl)

License  Creative Commons BY 4.0 International license

© Susanne Boll, Joycelyn Longdon, Katrien de Moor, Markus Fiedler, Gözel Shakeri, Lenneke Kuijer, Dan Lockton

We were aware from the start that a large proportion of HCI (and design) research focused on “sustainability” already arguably addresses SDG 12 – from “design for sustainable behaviour” (e.g. energy use feedback) through applications of social practice theory, to servitisation (product-service systems), etc, it seems as though HCI and design researchers have found the “responsible consumption” aspects of SDG 12 to be something very much within the ambit of their domain. However, we noted that the “production” aspects of SDG 12 were less well-addressed in HCI, at least within the HCI areas that we as a group had experience in. We noted that while aspects like circular economy ideas, sharing economy, alternatives to fast fashion, repairability (and right-to-repair), attention to production and material supply chains (traceability, and the labour practices involved), distribution logistics, and so on, had some currency in design research, they were not so often discussed within mainstream HCI, and yet there were many opportunities, including consideration of how to achieve higher quality digital experiences more sustainably (linking to discussions elsewhere at the Dagstuhl Perspectives Workshop about, for example, how academic conference norms might change) which could have significant impact.

Our discussion included noting that while the focus of HCI on SDG 12 topics is often the individual, a refocus on social groups, policymakers, and industry would also offer potentially transformative opportunities. Even with individual-focused HCI work on consumption, it is often about making technologies (or use) more “efficient”, rather than challenging norms, disrupting consumption patterns, doing more with existing things rather than creating new

things, or embodying alternative economic models. Is it preferable to consume less, or to consume better? The answer is not as simple as the question sounds. In some ways some of this work could be characterised as greenwashing a “wants” economy rather than addressing underlying issues. Can HCI go beyond the assumptions that if only people have more/better information (a deficit model), they will necessarily consume more responsibly? Can HCI be more radical, showing the magnitude of change needed, and enabling new patterns of living? We explored whether visualisation (or “visibilisation”, maybe using AR) or tangibilisation, making parts of systems that are currently invisible, visible (e.g. showing e-waste and whole product lifecycles), or figures that are relevant – including to decision-makers upstream rather than just consumers – would make a difference. We asked whether HCI could engage with some of the sociological issues such as attachment to products (or ownership, or “the new”), the social value of “success” attributed to material goods, phenomena such as taste regimes, conspicuous consumption, mindless consumption, perceptions of norms around consumption (e.g. through practical “speculative” design enacting different ways of living, or presenting alternative realities to stimulate people’s imaginations) We were also aware that much of this work in HCI has traditionally been centred on a very global north conception of consumption practices and needs, and that contexts, societal norms, assumptions around “development”, and relevant issues to address can be very different in the global south – what local practices, or indigenous knowledges, offer different ways of addressing SDG 12?

5.7 Working Group on SDG 13: Climate Action

Alexander Raake (TU Ilmenau, DE, alexander.raake@tu-ilmenau.de)

Jacki O’Neill (Microsoft Africa Research Institute – Nairobi, KE,

Jacki.ONeill@microsoft.com)

Anjali Karol Mohan (INDÉ – Bangalore, IN, anjali@integrateddesign.org)

Kirsikka Kaipainen (Tampere University, FI, kirsikka.kaipainen@tuni.fi)

Adrian Friday (Lancaster University, GB, a.friday@lancaster.ac.uk)

Daniel Pargman, (KTH Royal Institute of Technology – Stockholm, SE, pargman@kth.se)

License © Creative Commons BY 4.0 International license

© Alexander Raake, Jacki O’Neill, Anjali Karol Mohan, *Adrian Friday*, Kirsikka Kaipainen, Daniel Pargman

We saw taking and supporting climate action as addressing an interconnected and systemic problem, “doing the work”, “promoting the salience” of the need for climate action at multiple stages and by many actors, together, and with greater awareness as to the outcomes. We thus framed HCI in the context of working to address systemic issues that prevent climate inaction.

This could mean holding various stakeholders (broadly cast – governments, businesses, leaders, citizens, scientists) to account. Providing tools to gather or communicate evidence as to the positive and negative impacts of policies impacting the climate, and their effectiveness or unintended consequences, to help identify meaningful positive change. Helping share the experience of climate change as it impacts people, society and environment; recognising that the impacts of climate change are not evenly experienced across the world in space and time. Using this experience as a lever to increase empathy and motivate changes in different stakeholder groups. Addressing the information war, supporting clear narratives, accounts and storytelling around action or inaction. Especially overcoming barriers to appreciate the climate science such as dealing with the inherent uncertainty and challenges of communicating

this science, to make it a critical societal topic of debate, where climate inaction is no longer tolerated (c.f. drink driving). Practically, joining and supporting communities to accelerate and infrastructure positive and timely action. And finally, countering discourses of delay and misinformation, avoiding critical misunderstandings and even toxic and subverted discourse.

We saw this as a step change for HCI, from driving more efficient X, or better user experience Y, to deeply engaging with barriers to climate action, positive change, and more transformative and timely action across the systems.

5.8 Working Group on SDG 16: Peace, Justice and Strong Institutions

Nicola Bidwell (University of Melbourne, AU & International University of Management, NA, nic.bidwell@gmail.com)

Joycelyn Longdon (University of Cambridge, GB, jl2182@cam.ac.uk)

Alexander Raake (TU Ilmenau, DE, alexander.raake@tu-ilmenau.de)

Gözel Shakeri (University of Oldenburg, DE, gozel.shakeri@uni-oldenburg.de)

Katta Spiel (TU Wien, AT, katta.spiel@tuwien.ac.at)

Jason Jacques (University of St. Andrews, GB, jtj2@st-andrews.ac.uk)

License  Creative Commons BY 4.0 International license

© Kaisa Väänänen, Anjali Mohan, Katrien de Moor, Robin Neuhaus, Katta Spiel

In looking at Sustainable Development Goal 16, Peace, Justice & String Institutions, we began with an overview of the goal targets, which include a broad array of areas. Some of these targets have with concrete outcomes, such as ending abuse of children (16.2), while others have concrete timelines, such as providing legal identity for all by 2030 (16.9); though there was some discussion about the extent and inclusiveness of this with reference to the target to enforce non-discriminatory laws and policies (16.B). In some respects, the group felt there was scope to strengthen the targets, such as only reducing illicit flows of arms (16.4), while in other aspects there were some concerns raised about how the target might conflict with other personal freedoms, such as with the target to strengthen national institutions (16.A). Much of our discussion was focused on the democratic aspects of the goal and its targets, including reducing corruption (16.5), improving transparency (16.6), broadening participation at local (16.7) and international (16.8) levels, and finally regarding increasing access to information (16.10). We considered these targets in relation to aspects that must reflect the inclusiveness of the goal: cultural context, trust in the community, epistemic injustice, and belief in decision makers. We felt that the goal highlighted several opportunities around improving access to data, visibility, and transparency. These areas are strongly supported by HCI, through techniques such as visualisation, data curation, presentation, and structuring information in a customisable and layered experience. Finally, we briefly considered the tensions that were raised in creating a common understanding and allowing space for multiple meanings to be explored. We also touched upon more opportunistic engagement with the goal, and goals, and how this might create further distortion or tension both by actors and through our own work.

6 Speculations

6.1 Introduction and Method

Transformation requires critical engagement with possible futures. While many treat the future as something to discover, we understand the future as something – at least in part – to be “made”. In this sense, we not only anticipate future problems based on today but imagine desirable futures to explore and better understand the mindsets, methods and technologies necessary to make these futures happen. [20] argued that an utopian science, a science of possible change, requires research into three different types of knowledge: system knowledge, goal knowledge, and transformation knowledge. Traditionally, science focuses on the production of system knowledge, that is descriptions, models, or theories of how things are. However, to change we need to explore and discuss how things ought to be to set desired goals. Futuring and speculation, for example in the form of design fiction, is one way to critically [10] or constructively [9] engage with possible futures (see also [2] [16] [17] [21]). Speculation is a way to analyse and to better understand the intricate entanglement of technology, people, and policies. It is able to question dominant narratives of technology in general and Human-Computer Interaction and provides starting points to rewrite them.

In the workshop, we introduced positive speculation as a form of analysis of possible futures (see [9]). We formed four groups. The instruction was to start with the knowledge encapsulated in some of the Sustainable Development Goals (SDG) discussed on Tuesday (see section 5), however, to acknowledge their limits, to go beyond them, or to ignore them completely. Groups were then asked to imagine “positive” (meaningful, in line with the SDGs) everyday situations (who, where, activities, feelings) in the near future (2028) and to explore how this situation had been and is still shaped by people, policies, technologies. The groups did so through worldbuilding and storytelling, roleplay and a form of anticipatory ethnography (e.g., [16]). In general, we encouraged the groups to make their own futures, pick topics they find engaging, without a strong focus on technology.

The results were presented in the form of 20 minute group performances and discussed in the plenum at the end of the day (16:00 – to 18:00).

6.2 Four Speculations

In the following, we summarize the four speculations as presented by the groups.

6.2.1 GROUP 1: “Global Sugar Act”

Susanne Boll (University of Oldenburg, DE, susanne.boll@uol.de)

Markus Fiedler (Blekinge Institute of Technology – Karlshamn, SE, markus.fiedler@bth.se)

Lenneke Kuijer (TU Eindhoven, NL, S.C.Kuijer@tue.nl)

Neha Kumar (Georgia Institute of Technology – Atlanta, USA, neha.kumar@gatech.edu)

Eunice Sari (UX Indonesia – Jakarta, ID & University of New South Wales – Sydney, AU, eunice@uxindo.com)

Thomas Olsson (Tampere University, FI, thomas.olsson@tuni.fi)

License © Creative Commons BY 4.0 International license

© Kaisa Väänänen, Anjali Mohan, Katrien de Moor, Robin Neuhaus, Katta Spiel

“The United Nations” think-tank for revising the 2030 Integrated Development Goal for Nutrition, Environment, Decent Work and Degrowth began its online meeting. Four participants, four time-zones, four beverages. One participant, with her early morning beverage,



■ **Figure 4** Global Sugar Act Speculation

discusses how it is oat-milk with a novel scent-based taste of coffee [inspired by existing technology: air-up¹; another, at the end of her day, sips a light “beer”, also scent-based taste only, and the others sip scent-based health and chocolate drinks. The group discusses how quickly the scent-based infusions changed the beverage market since the Global Sugar Act was introduced. Originally, the act was implemented to address the implications of sugar as a serious threat to health. While originally focused on the health issue, the Act had multiple additional consequences, economically, environmentally as well as on innovation. First, sugar-producing regions adjusted their agricultural regimes. Naturally sweetened products became rare luxuries, especially chocolate. Cacao producers stepped into other products and professions, improving their labour position in the process. Today’s meeting focuses on translating lessons from this transition to coffee-producing regions, which are experiencing decline because of the more recent switch to scent-based coffee. Observers asked how the Global Sugar Act was introduced in the first place, and why honey, a popular replacement of sugar, was still available in spite of it.”

This speculation prompted reflection on the multifaceted entanglement of policies, technological responses, and further (economical) impact. It depended on a radical policy decision regarding global health, which then had consequences for food production practices across the world, but also on everyday practices such as having breakfast or celebrating a birthday. It highlighted the potential of an existing technology to step into the space created by radical policy and to turn it into a triple win (when optimistically approached).

6.2.2 GROUP 2: “Glow in Care”

Joycelyn Longdon (University of Cambridge, GB, jl2182@cam.ac.uk)

Dan Lockton (Eindhoven University of Technology, NL, d.j.g.lockton@tue.nl)

Kirsikka Kaipainen (Tampere University, FI, kirsikka.kaipainen@tuni.fi)

Gözel Shakeri (University of Oldenburg, DE, gozel.shakeri@uni-oldenburg.de)

Shadan Sadeghian (University of Siegen, DE, shadan.sadeghian@uni-siegen.de)

Nicola Bidwell (University of Melbourne, AU & International University of Management, NA, nic.bidwell@gmail.com)

License © Creative Commons BY 4.0 International license

© Kaisa Väänänen, Anjali Mohan, Katrien de Moor, Robin Neuhaus, Katta Spiel

“Welcome egalitarian council of the Global Majority, and all people and life beings of the former Global Souths”, said the chair. Delegates responded, “Glow in care”. The meeting began with the monthly report on the Global North’s transition to a care-based economy as

¹ * inspired by existing technology (air-up: <https://uk.air-up.com/>)



■ **Figure 5** Glow in Care Speculation.

introduced by the newly formed sovereign coalition of the Global Majority. Under the new economic system people received “glow” when they actively engaged in sharing resources sustainably or acted with an ethic of care (Gilligan 1982). Professor John assessed the progress using the Doughnut Economics model (Raworth, 2017) and summarised population characteristics in a matrix with Egotism-Reciprocity and Resistance-Enthusiasm dimensions. He illustrated these using interviews with people at different stages of transition: a senior woman who resisted the non-exploitative use of resources; an enthusiastic, but tired, former taxi driver whose livelihood now depends on cycling a rickshaw; a “whistle-glower”, who informs about the illegal cartels who were stealing “glow”; and, a millennial enthusiast who felt less anxiety and loneliness under the new system.”

This design fiction sought to frame societies in the Souths from the perspective of socio-psychological abundance, not deficit. It depicted a radical transformation of the global economic system, with “glow” to replace currency to represent actions that contribute to sustainability. This transformation had begun with local movements of avocado and coffee farmers joining together. In order to prompt reflecting on what constitutes the capability for resilience and the need to learn from the Global Majority in managing resources sustainably the performance broke “the fourth wall”.

6.2.3 GROUP 3: “The EcoFLOW Pill”

Daniel Pargman (KTH Royal Institute of Technology – Stockholm, SE, pargman@kth.se)

Carine Lallemand (University of Luxembourg, LU & TU Eindhoven, NL, c.e.lallemand@tue.nl)

Adrian Friday (Lancaster University, GB, a.friday@lancaster.ac.uk)

Jason Jacques (University of St. Andrews, GB, jtj2@st-andrews.ac.uk)

Matthias Laschke (University of Siegen, DE, Matthias.Laschke@uni-siegen.de)

License © Creative Commons BY 4.0 International license

© Kaisa Väänänen, Anjali Mohan, Katrien de Moor, Robin Neuhaus, Katta Spiel

“Activists”, the British Prime Minister said quietly, “were now engineering the EcoFLOW pill and dosing people involuntarily.” The Prime Minister took the issue seriously since it was out of hand and had forced his government to now heavily promote the voluntary use of EcoFLOW. Many citizens had already voluntarily taken the pill, with its effect estimated to last for 12-14 months. Doctors around the country were routinely explaining to patients how the drug leads to negative physical responses to others and their own eco-unfriendly practices, such as taking the car, and to the enhanced experience of engaging in sustainable practices,



■ **Figure 6** EcoFlow Speculation

for example a heightened sense of enjoyment of the nature walk and of cycling to work. To do so, EcoFLOW stimulates the vagus nerve, part of the parasympathetic system. The strong bodily reactions to unsustainable practices has created strong tensions in relationships and in families if members have taken different decisions in regards to consuming the pill. Indeed, activists' practice of setting up serial Tinder dates to slip the pill into the drink was causing widespread disruptions to marriages and parents were justifiably upset when their teenage children slipped them the pill."

Composed of three narrative parts (the manifesto of scientists published on the Dagstuhl website, the British Prime Minister's speech and a doctor's appointment), this speculation prompted reflection on ways to creatively exploit the felt-experience of ecoanxiety. It turned sustainable practices into a disruptive social challenge, rather than framing sustainability as an individual choice. In addition it prompted thoughts on activism, potential illegal activist practices, and the tension between individual freedom and the need to act.

6.2.4 GROUP 4: U4ya – “Community and Water”

Kaisa Väänänen (Technical University of Tampere, FI, kaisa.vaananen@tuni.fi)

Anjali Karol Mohan (INDÉ – Bangalore, IN, anjali@integrateddesign.org)

Katrien de Moor (Norwegian University of Science and Technology – Trondheim, Torgarden, NO, katrien.demoor@ntnu.no)

Robin Neuhaus (University of Siegen, DE, robin.neuhaus@uni-siegen.de)

Katta Spiel (TU Wien, AT, katta.spiel@tuwien.ac.at)

License © Creative Commons BY 4.0 International license

© Kaisa Väänänen, Anjali Mohan, Katrien de Moor, Robin Neuhaus, Katta Spiel

“Remember how it all started: “To promote more efficient use of water, the government is now subsidising the costs for community housing projects to remodel.” We came together as a community, discussed how we can change the way we use water and have fun together. We built models of our remodelled housing complex and slowly everything took shape. Of course, it was a little hard in the beginning to let go of the large bathtubs and spa amenities we used to have in our personal bathrooms. But now many of us greatly cherish the new shared activities that came with it and how we are closer than ever. Enjoying a comfortable bath while exchanging what's new over a beer, or caring for the community garden as the laundry is slowly soaking – I would not want to miss that anymore. A visit every now and then to the museum of Water-‘ways’ that Were (commissioned by the water utilities department), housing the relics of water infrastructure – shower heads, bathtubs, water bottles, big water pipes, laundromats – are all reminders of the wasteful (and individualistic) lifestyle that we “enjoyed” from a “place of ignorance”. The experience does draw expressions of incredulousness from our children and is often a topic of over-dinner conversation for our families. To promote more efficient use of water, the government is now subsidising the costs



■ **Figure 7** EcoFlow Speculation.

for community housing projects to remodel. Community members engaged in play, using miniatures to physically model the design of their infrastructure. Imagining how walkways between shared showering, bathing and washing facilities might open new opportunities for social engagement and care, community members.”

This speculation engaged physical play around built infrastructure. It prompted considering how the need to support more sustainable practices and consuming less resources would not lead to a feeling of missing out, but to a better quality of life. A closely knit diverse community has managed to turn the challenge of a strict rationing of water supply and strict regulations of reusing grey water into positive communal practices. Bathing and laundry have become weekly social rituals that do not only matter for hygiene but also strengthen togetherness. Water is used and reused in planting and gardening projects, which the community takes great pride in. In that, it was relevant to the group to include a community perspective over individualised approaches or system level overwhelmedness. While more individual technological solutions, such as a novel shower designed to only “freshen up” were discussed as well, the design fiction stresses how new community practices can help shape narratives of a more sustainable positive future. Additionally, this fiction makes a point in turning not just to an always far away future that assumes technocentrism in a Global North sense, but rather a sideways consideration of existing practices with geographies that continue with deep-rooted, long standing cultural practices that have the potential to steer sustainable living. It shows that orienting ourselves towards a community requires rethinking our current individualistic practices and creating infrastructures that support differing individual needs within a collective.

6.3 Reflection: Emerging Themes

In this section, we briefly summarize the themes emerging from the speculations itself and the discussions following the presentations of the speculations.

6.3.1 Theme: Role of Community

A predominant narrative of change, especially when it involves technology, is one of individual behavior. In this view, people are “free” to either adopt or ignore more sustainable lifestyles. It remains basically choice, how to be mobile, how to heat, how to eat, or how to be healthy. Other people, that is the social, is framed as a source of pressure, interfering with individual choice. It is the individual, who is responsible for sustainability. Two of speculations, “U4ya” and “Glow in Care”, were about the pursuit practices that produce, reproduce and use the

commons and create communality, also known as of “commoning” [5]. They frame the social, that is, community as a resource. “Glow in Care”, for example, envisions how the cumulated knowledge of communities (here from the Global South) challenges the dominant Western way of exploiting rather than caring. A community based on shared values has the power to change the existing. “U4ya” points at community as a “safety net”, a way to face the manifold challenges posed by necessary radical changes required to live more sustainable. Both views acknowledge the importance of community as a motivator and a support of change. In fact, all the design fictions spoke in various ways to the need for coordinated actions and reframing efforts in HCI from an isolated and individualist approach to sustainability (see also Jacki O’Neil’s keynote in Section 3.1). While not made explicit in the speculation, technology can play a role in facilitating community building, for example, by providing accessible infrastructures for communities, exchange of values, views, and strategies, for example between people in the Global Norths and those on the ground in the Souths (e.g., [6]). The tension between the overwhelmingly individualist frame of HCI, and indeed all of modern life, and collectivist approaches can be creative.

6.3.2 Theme: Positive Practices

Especially in saturated Western societies, sustainability is often understood as a threat to the current lifestyle, a dystopian step back rather than an utopian step forward. Sustainable life is depicted as bleak and undesirable. The speculations paint a different picture. They emphasize the capability of humans to adapt their practices to changes in conditions. This can be preserving, for example, when in the “Sugar Tax” speculation global change in the availability of resources, such as cocoa or coffee, lead to the wide adoption of a new technology (here: scent-based taste) to be able to carry on with beloved rituals and habits (e.g., breakfast). It can also be disrupting, when for example “U4ya” reshapes practices of water use, such as bathing or laundry, from the rather hidden and private into something shared and communal. In both speculations, people quite naturally respond to external demands by developing adapted practices, which are positive and meaningful. First, this can be understood as a plea to simply trust the ability of people to respond positively to changed circumstances. Second, we can support this by providing potential positive narratives of the everyday life after the introduction of external change, such as a new policy (see [13]). Third, we should encourage scientific real world experimentation with change to make alternatives experienceable and to foster people’s ability to positively respond to alternatives. From a technology-perspective this highlights how technology is inextricably entangled with our practices. It should not be thought of as separate and neutral, but as integral and constituting. While in HCI the typical strategy is to adapt technologies to current practices, this entanglement offers the possibility to introduce new sustainability-related technologies first and to rely on the transformative powers of these technologies.

6.3.3 Theme: Policies and Technology

All four speculations tell stories about how technology and its adoption is heavily intertwined not only with the social per se (see theme “the role of community”) but also with policies. In “Sugar Tax”, the banning of sugar had been a starting point for small technological changes, in “Glow with Care” an alternative “regime” defined care to be the ultimate “currency”, or in “U4ya” extensive regulations on water use triggered a social and a technological response. Interestingly, these speculation place policy changes at the beginning of their story, rather than at the end. They see policy changes as a trigger rather than as a response. “ecoFLOW”

is different. Here a “technology” is set free and only after the fact policies emerge as a response. However, no matter the direction, these speculations remind us of the fact that technology should never be seen as independent from rules and regulations. While HCI as a discipline seems well aware of the social embedding of technology, it seem less aware of the politics of the technologies suggested. For example, forms of contextual analysis have been a ubiquitous part of research and practice for decades. A lot of effort is put in understanding people directly affected, their tasks, and the physical as well – albeit to some lesser extent – organizational context they use the technology in. There is no equivalent standard analysis for the political dimension of a new technology. But this seems required especially in the field of change for sustainability. In line with the present speculations, typical questions to answer would be, which policies may support the beneficial effect of any given technology (and which may obstruct) as well as to which new policies a given technology may lead.

6.3.4 Theme: Autonomy and Coercion

In general, technologies for change are often scrutinized from the standpoint of manipulation and coercion. Nudging through technology is often critically seen as ethically questionable. “ecoFLOW” paints a more nuanced picture. In this speculation, a pill creates viscerally negative sensations when the person who took it or other close people behave unsustainably. In one part of the story, people are tricked by activists into swallowing the “ecoFLOW” pill, and as a consequence not only need to behave sustainable against their will but also need to explain under which debatable circumstances they caught the pill. Others however take the pill as a form of self-commitment, because they know that they cannot rely on their willpower alone. In another variant of the story, members of the younger generation slip the pill to their parents, as a form of self-defense, because changes in behavior are imminent to survive, even if some people do not agree. This speculation highlights the many shades along a potential continuum of autonomy and coercion. In fact, it is hard not to feel sorry for the activists or the children because they lose patience with all the people who seem unable to change, even though it seems to be of utmost importance to do so. The “ecoFLOW” speculation reframes unsustainable behavior as a sickness, a pathology, one has to manage. While everybody seems amiable and understanding of the difficult condition people are put in, there is an acceptance of the underlying premises about which behavior is good and bad. To coerce into the good may not be ideal, but as long the result justifies the means it might be acceptable. In this sense, the “ecoFLOW”-pill is akin to a new “social norm” established through technology. From the perspective of the design of technologies for behavior change, it seems crucial to better understand the many shades of autonomy/coercion to find acceptable and effective solutions. On the one hand, technologies, which do not create feelings of friction, that is restrict autonomy, may not be in any way effective in changing beloved routines. On the other hand, too much friction may lead to resistance and reactance, which in turn will reduce effectiveness as well. Finding a balance is the ultimate challenge of behavior change technologies (see Hassenzahl and Laschke, 2015).

6.3.5 Theme: Activism and Research

A final theme emerged from the discussion about the speculations rather than from the speculations themselves, namely the relationship between activism and research. Many scientific disciplines take pride in remaining neutral. In terms of Schneidewind’s (2017) model (see introduction to this section on speculation) they primarily create system knowledge with the intention to factually describe how it is. In technology-oriented disciplines, basic

technologies take the place of system knowledge. HCI, however, is per definition concerned with the way technology is used in everyday life. This was always accompanied with strong normative notions about how technology ought to be (e.g., usable) and how it as to be sensibly designed (e.g., in a human-centered way). In this sense, HCI always had been an utopian science, concerned with making the world better through technology. The same logic that drove the notion of HCI as acting as an “advocate of user’s needs” could be applied to sustainability. HCI could face the responsibility of becoming a lobby for the environment, just as it has already become a lobby for users. For HCI, research and activism doesn’t seem to be as much of a contradiction.

7 Conclusions and Actions

In a last step, we will group the narratives into larger coherent themes. This can be input to a provocative as well as reflective paper/website about possible sustainable futures, the role HCI might play, and the potential transformations needed to achieve this. As a result from the workshop the participants collected a set of statements that characterize the role that HCI should take on making a positive change for sustainability. We will elaborate on these statements in our Dagstuhl Manifesto for this Perspectives Workshop.

The organizers envision a number of follow-up activities from the Dagstuhl Perspectives Workshop. The central venues in the field of behavioral science and computer interaction offer a great platform to propose new workshops and bring together the participants from the Dagstuhl Perspectives Workshop and further researchers in the field. We also envision that the Dagstuhl Perspectives Workshop will initiate collaborations on projects and project proposals and mutual visits of the participants and their PhD students to exchange and continue the discussions that started in Dagstuhl. During the closing days of the Dagstuhl Perspectives Workshop we collected several of the participating researchers which will be implemented in the months to come. Below some of these actions are summarized.

- Education: Participants will work on a curriculum on sustainability that complements our existing courses in Human Computer Interaction. We will aim to educate future generations and equitable access to sustainability.
- Recognizing Sustainability in HCI: Give paper that contribute to sustainability in HCI more visible recognition. Invite authors to add a section to their paper that makes a sustainability statement on their research and encourage journals and magazines to ask for such a section. Also raise awareness for sustainability topics for reviewers.
- Funding and project: Lobby at funding agencies to a strong recognition on sustainability in funding schemes and references to the SDGs as one framework for addressing global goals in our research.

7.1 About the organizers

7.1.1 Nic Bidwell

Nicola Bidwell is inaugural Adjunct Chair for Sustainability on the Executive Committee of the Special Interest Group for Computer Human Interaction (SIGCHI) of the Association for Computer Machinery (ACM); and, for nearly 20 years, has researched in HCI with a focus on the Global Souths and socio-ecological sustainability. Her research encompasses working with Aboriginal people in north Australia and inhabitants of rural Namibia, South Africa,

Mozambique, Kenya, Uganda, Indonesia, Argentina, Mexico and India. She initiated the first panel on Indigenous Led Digital Enterprise at a leading HCI forum in 2008, co-founded the African HCI Conference in 2016 and her ethnographically informed design work with rural collaborators in South Africa received an award for its contribution to social and economic development. Her research is distinguished for combining novel intellectual contributions with measurable and sustained social impact, as illustrated by being chosen as a best-practice case in *Future-proofing: Making Practice-Based IT Design Sustainable*. Nic's analyses of relations between spectrum regulation and involvement in designing and deploying technologies for community networks (CNs) and community radio have also informed policy debate. Her recent studies include experiences of conditional programming in philanthropic donations, interactions with algorithms in ride-sharing platforms, and onto-epistemic translations in designing probabilistic programming languages for marginalised knowledge practices. Nic has taught in universities around the world. She is currently Associate Professor and Head of the Techno-Anthropology research group at Aalborg University, Denmark; however, for the past fourteen years she lived and worked in southern Africa, where she continues as an adjunct Professor at the International University of Management, Namibia. Nic's work has theoretically and practically directly aligned with nearly half of the SDGs. Her support and studies of CNs, for instance, tackle economic marginalisation (e.g., SDG 8) and gender equality (SDG 5) directly; her design and implementation of community-based solar charging in three different projects have tackled affordable clean energy (e.g. SDG 13); and, across her work, from the rainforests of tropical north Queensland to the Kalahari desert, there is a constant theme of considering indigenous people's and local ways of knowing and caring for their ecological environment (e.g. SDG 10).

7.1.2 Susanne Boll

Prof. Dr. techn. Susanne Boll is Professor of Media Informatics and Multimedia Systems in the Department of Computing Science at the University of Oldenburg, in Germany. She serves on the board of the OFFIS–Institute for Information Technology, in Oldenburg. Prof. Dr. Boll earned a doctorate with distinction from the Technical University of Vienna, Austria. She received her Diploma in Computer Science with distinction from the Technical University of Darmstadt, Germany in 1996. Her research field lies at the intersection of human computer interaction and interactive multimedia in which she has an excellent scientific track record. Her research passion is developing interactive technology for people, joining novel innovative technology development with user needs and social acceptance in the center of her research.

Susanne Boll has been debating with the scientific community on what multimedia can do for sustainability [8] and how human computer interaction related to the Sustainable Development Goal 3: Good Health and Well-being [3]. In several scientific research projects she has developed novel technology for personal health and interactive technology in ambulant and stationary care [11]. As a female researcher in computer science she has always been an advocate for gender equality in STEM for many years and has led scientific projects for introducing Computer Science to girls. She has studied how physical computing can facilitate the understanding required for learning the subject matter for children [1] and middle-school girls [19] which relates to Sustainable Development Goal 5: Gender Equality. Her ongoing work looks into acceptable technology for humans in cooperation with automation and artificial intelligence which relates to Sustainable Development Goal 8 to promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all (e.g., [15]). Recently she was the lead author of a white paper on “Mit Künstlicher

Intelligenz zu nachhaltigen Geschäftsmodellen” (With artificial intelligence to sustainable business models [7]. She is also currently building a new theme of interaction design in her group for eco-friendly purchase with Dr. Gözel Shakeri joining her team (c.f. [22]).

7.1.3 Marc Hassenzahl

Marc Hassenzahl is Professor for “Ubiquitous Design” in the Department of Business Computing at the University of Siegen, Germany. With his group of designers and psychologists, he explores the theory and practice of designing pleasurable, meaningful and transforming interactive technologies. He currently leads or participates in five national and international research projects. While the main theme of all projects is to use design as a means to create meaningful technology-mediated everyday experiences, the projects range from the question of how to use AR/VR in the health-domain (HIVE, BMBF) to how to best design assistive robots (e-VITA, EU, Honda Research Grant). He is author of “Experience Design. Technology for all the right reasons” [12], co-author of “Psychologie in der nutzerzentrierten Produktgestaltung. Mensch-Technik-Interaktion-Erlebnis” (“People, Technology, Interaction, Experience”) (Springer, with Sarah Diefenbach). His current h-index (according to Google Scholar) is 56 with a total of 22158 citations. Sustainability and Human-Computer Interaction is a recurrent theme in Marc’s work. His broader focus is on behavioral and attitudinal change and the role technology can play as an intervention to instill transformation of everyday practices. Individual motivation and ability to change already had been a topic in his doctoral dissertation, which addressed intertemporal choice and self-regulation (e.g., willpower) – both typical problems in acting sustainably. Since 2010, he has worked with Dr. Matthias Laschke on a particular approach to technological intervention, called “pleasurable troublemakers” [14]. This led to the successful application of Dr. Laschke for the funding of MOVEN (MOtivationale und VERhaltensändernde Nachhaltigkeitsstechnologien/ Motivational and Behavior Change Technologies for Sustainability) a “Junior Research Group in Social-Ecological Research” (FONA, BMBF) in cooperation with LMU Munich, RWTH Aachen, the Wuppertal Institute as well as Miele. The faculty is about to establish a new Juniorprofessorship “Interaction Design for Sustainability and Transformation”. Sustainability and HCI is a recurrent theme in Prof. Hassenzahl’s teaching, for example, with research courses given in the HCI Master in the winter term 20/21, as well as in 21/22. In his role as dean, he is a member of DECODE, the European Deans Council For Sustainable Development.

7.1.4 Kaisa Väänänen

Kaisa Väänänen is a full professor of user experience (Human-Computer Interaction) in Tampere University, Finland. Kaisa leads the research group of Human-Centered Technology (IHTE) in the unit of Computing Sciences. Kaisa has over 25 years of experience in research related to human computer interaction both in university and industry. In 1995-2004, she worked at Nokia Inc, in leading positions of user needs research and strategic consumer insights. Kaisa’s research interests cover user experience and human-centered design, with emphasis on design research of digital services for advancing sustainability and human-centered AI. Kaisa is an author of >160 peer-reviewed academic publications. She is very active in the international research community, and frequently takes part in organizing conferences related to user experience and human-computer interaction. In 2013-2018, Kaisa served six years as a member of Academy of Finland’s research council for Natural Sciences and Engineering.

Related to SDGs, Kaisa Väänänen's research is addressing themes from both environmental and social sustainability perspectives: 1) Digital societal participation of all kinds of youths (SDGs 3, 10 and 16). In this research, Kaisa's team has studied the design and use of digital tools for societal participation, for example in legislation, mental wellbeing and educational development [23]. 2) Smart persuasive technologies for sustainable behaviour (SDGs 11, 12 and 13). In this research line, the research has focused on social robots and how they can be used as proactive agents and mediators of people's sustainable behaviour [4]. 3) Novel digital technologies for urban environments (SDGs 3, 11 and 13). In this line of research, positive citizen experiences and wellbeing is studied through human-centered design of interactive technology concepts for urban residents [18]. Outcomes of all topics include models, frameworks and design guidelines for human-technology interactions advancing technologies.

8 Literature overview

Before the workshop, we asked the participants to add one paper or one video to the workshop wiki that in any way relates to the topic of the workshop and that they have identified as key paper or video in the field (theory, new directions, and visions; not their own research). This allowed each participant to prepare themselves for the workshop by going through the list of proposed papers and videos to read or watch those which are yet unknown. By this we will also make sure that each of the interdisciplinary fields can share their key publications which might not be so visible to other fields.

Below we present the compilation of this reading list.

- Abowd, G. D. (2020). The Internet of Materials: A Vision for Computational Materials. *IEEE Pervasive Computing*, 19(2), 56–62. <https://doi.org/10.1109/MPRV.2020.2982475>
- Bardzell, J., Bardzell, S., & Light, A. (2021). Wanting To Live Here: Design After Anthropocentric Functionalism. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–24. <https://doi.org/10.1145/3411764.3445167>
- Bates, O., Thomas, V., & Remy, C. (2017). Doing good in HCI: Can we broaden our agenda? *Interactions*, 24(5), 80–82. <https://doi.org/10.1145/3121386>
- Baumer, E. P. S., & Silberman, M. S. (2011). When the implication is not to design (technology). *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2271–2274. <https://doi.org/10.1145/1978942.1979275>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Biørn-Hansen, A., Pargman, D., Eriksson, E., Romero, M., Laaksolahti, J., & Robért, M. (2021). Exploring the Problem Space of CO2 Emission Reductions from Academic Flying. *Sustainability*, 13(21), Article 21. <https://doi.org/10.3390/su132112206>
- Blevis, E., Chow, K., Koskinen, I., Poggenpohl, S., & Tsin, C. (2014). Billions of interaction designers. *Interactions*, 21(6), 34–41. <https://doi.org/10.1145/2674931>
- Bremer, C., Knowles, B., & Friday, A. (2022). Have We Taken On Too Much?: A Critical Review of the Sustainable HCI Landscape. *CHI Conference on Human Factors in Computing Systems*, 1–11. <https://doi.org/10.1145/3491102.3517609>

- Breves, P., & Schramm, H. (2021). Bridging psychological distance: The impact of immersive media on distant and proximal environmental issues. *Computers in Human Behavior*, 115, 106606. <https://doi.org/10.1016/j.chb.2020.106606>
- Caraban, A., Karapanos, E., Gonçalves, D., & Campos, P. (2019). 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3290605.3300733>
- Chauhan, C., Parida, V., & Dhir, A. (2022). Linking circular economy and digitalisation technologies: A systematic literature review of past achievements and future promises. *Technological Forecasting and Social Change*, 177, 121508. <https://doi.org/10.1016/j.techfore.2022.121508>
- Chopra, S., Clarke, R. E., Clear, A. K., Heitlinger, S., Dilaver, O., & Vasiliou, C. (2022). Negotiating sustainable futures in communities through participatory speculative design and experiments in living. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–17. <https://doi.org/10.1145/3491102.3501929>
- Dema, T. (2021). Engaging remote communities in technology design for connecting people to and through nature [PhD, Queensland University of Technology]. <https://doi.org/10.5204/thesis.eprints.211484>
- Döring, N., Moor, K. D., Fiedler, M., Schoenberg, K., & Raake, A. (2022). Videoconference Fatigue: A Conceptual Analysis. *International Journal of Environmental Research and Public Health*, 19(4), Article 4. <https://doi.org/10.3390/ijerph19042061>
- Dourish, P. (2010). HCI and Environmental Sustainability: The Politics of Design and the Design of Politics. In *DIS 2010 – Proceedings of the 8th ACM Conference on Designing Interactive Systems* (p. 10). <https://doi.org/10.1145/1858171.1858173>
- Ferreira, D., Dey, A. K., & Kostakos, V. (2011). Understanding Human-Smartphone Concerns: A Study of Battery Life. In K. Lyons, J. Hightower, & E. M. Huang (Eds.), *Pervasive Computing* (pp. 19–33). Springer. https://doi.org/10.1007/978-3-642-21726-5_2
- Fiedler, M., Popescu, A., & Yao, Y. (2016). QoE-Aware Sustainable Throughput for Energy-Efficient Video Streaming. 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), 493–500. <https://doi.org/10.1109/BDCloud-SocialCom-SustainCom.2016.78>
- Freitag, C., Berners-Lee, M., Widdicks, K., Knowles, B., Blair, G., & Friday, A. (2021). The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. *Patterns*, 2, 100340. <https://doi.org/10.1016/j.patter.2021.100340>
- Fromhold-Eisebith, M., Grote, U., Matthies, E., Messner, D., Pittel, K., Schellnhuber, H.-J., Schiefendecker, I., & Schlacke, S. (2019). *Unsere gemeinsame digitale Zukunft: Zusammenfassung. Wissenschaftlicher Beirat d. Bundesregierung Globale Umweltveränderungen.*
- Gnanasekaran, V., Fridtun, H. T., Hatlen, H., Langøy, M. M., Syrstad, A., Subramanian, S., & Moor, K. D. (2021). Digital carbon footprint awareness among digital natives: An exploratory study. *Norsk IKT-Konferanse for Forskning Og Utdanning*, 1, Article 1.
- Grönwald, L., Weiblen, J., Laschke, M., Christoforakos, L., & Hassenzahl, M. (2023). Sustainability by Design. How to Encourage Users to Choose Energy-Saving Programs and Settings when Washing Laundry. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3544548.3581150>

- Hansson, L. Å. E. J., Cerratto Pargman, T., & Pargman, D. S. (2021). A Decade of Sustainable HCI: Connecting SHCI to the Sustainable Development Goals. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3411764.3445069>
- Hassenzahl, M., Dörrenbächer, J., Laschke, M., & Sadeghian, S. (2022). European Union's Green Smart Directive or How Resource-Conscious Smart Systems Saved the World. *Nordic Human-Computer Interaction Conference*, 1–5. <https://doi.org/10.1145/3546155.3547277>
- Hassenzahl, M., & Laschke, M. (2015). Pleasurable Troublemakers. In S. P. Walz & S. Deterding (Eds.), *The Gameful World* (pp. 167–196). The MIT Press. <https://doi.org/10.7551/mitpress/9788.003.0011>
- Herglotz, C., Robitza, W., Kränzler, M., Kaup, A., & Raake, A. (2022). Modeling of Energy Consumption and Streaming Video QoE using a Crowdsourcing Dataset (arXiv:2210.05444). arXiv. <https://doi.org/10.48550/arXiv.2210.05444>
- Howson, P., & de Vries, A. (2022). Preying on the poor? Opportunities and challenges for tackling the social and environmental threats of cryptocurrencies for vulnerable and low-income communities. *Energy Research & Social Science*, 84, 102394. <https://doi.org/10.1016/j.erss.2021.102394>
- Hsu, Y.-C., & Nourbakhsh, I. (2020). When human-computer interaction meets community citizen science. *Communications of the ACM*, 63(2), 31–34. <https://doi.org/10.1145/3376892>
- Ickin, S., Wac, K., & Fiedler, M. (2013). QoE-based energy reduction by controlling the 3g cellular data traffic on the smartphone. 2013 22nd ITC Specialist Seminar on Energy Efficient and Green Networking (SSEEGN), 13–18. <https://doi.org/10.1109/SSEEGN.2013.6705396>
- Inayatullah, S., Paine, E., & Scharmer, O. (2022). The Navigation System, The Planetary Gardener, and the Prism: Metaphors for Bringing the Future into Being. *Journal of Awareness-Based Systems Change*, 2(2), Article 2. <https://doi.org/10.47061/jasc.v2i2.5047>
- Jacques, J. T. (2020). CHI 2020: Right Here, Right Now? A bottom-up approach to estimating the carbon emissions from more than twenty years of CHI conference travel. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3334480.3381806>
- Junge, I. P. (2021). Single Use Goes Circular—An ICT Proto-Practice for a Sustainable Circular Economy Future. *Journal of Sustainability Research*, 3(1). <https://doi.org/10.20900/jsr20210009>
- Kannabiran, G., & Reddy, A. V. (2022). Exploring Kolam As An Ecofeminist Computational Art Practice. *Proceedings of the 14th Conference on Creativity and Cognition*, 336–349. <https://doi.org/10.1145/3527927.3531452>
- Kasila, K., Vainio, S., Punna, M., Lappalainen, P., Lappalainen, R., Kaipainen, K., & Kettunen, T. (2020). Individual differences in processes of lifestyle changes among people with obesity: An acceptance and commitment therapy (ACT) intervention in a primary health care setting. *Primary Health Care Research & Development*, 21, e12. <https://doi.org/10.1017/S146342362000016X>
- Klapperich, H., Laschke, M., Hassenzahl, M., Becker, M., Cürliş, D., Frackenpohl, T., Köhler, H., Ludwigs, K., & Tippkämper, M. (n.d.). A social practice approach to wellbeing-driven design.
- Kuijer, L., & Robbins, H. (2021). Teaching alternative Paradigms through Critical Design. *Interaction Design and Architecture(s)*, 51, 172–201. <https://doi.org/10.1145/3411764.3445069>

- 55612/s-5002-051-008 Laschke, M., Diefenbach, S., & Hassenzahl, M. (2015). "Annoying, but in a Nice Way": 9(2).
- Lee, J., Kaipainen, K., & Väänänen, K. (2020). Local foodie: Experience design of a mobile augmented reality application for tourists to encourage local food consumption. *Proceedings of the 23rd International Conference on Academic Mindtrek*, 110–119. <https://doi.org/10.1145/3377290.3377298>
 - Lockton, D., Chou, M., Krishnaprasad, A., Dixit, D., La Vattiata, S., Shon, J., Geiger, M., & Zea-Wolfson, T. (2019, December 10). Metaphors and imaginaries in design research for change.
 - Lockton, D., Lallemand, C., & Menheere, D. (2022). Designing qualitative interfaces: Experiences from studio education. *DRS Biennial Conference Series*. <https://dl.designresearchsociety.org/drs-conference-papers/drs2022/researchpapers/196>
 - Longdon, J. (2020). Environmental data justice. *The Lancet Planetary Health*, 4(11), e510–e511. [https://doi.org/10.1016/S2542-5196\(20\)30254-0](https://doi.org/10.1016/S2542-5196(20)30254-0)
 - Magoutas, B., Bothos, E., Schrammel, J., Orji, R., & Mentzas, G. (2017). Exploring the Links Between Persuasion, Personality and Mobility Types in Personalized Mobility Applications: 12th International Conference on Persuasive Technology. *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*, 107–118. https://doi.org/10.1007/978-3-319-55134-0_9
 - Meyer, J., Beck, E., Holdt, K. von, Koppelin, F., Pauls, A., & Boll, S. (2021). ActiThings: Reminders for physical activity exercises in daily lives of inactive older adults. 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI), 386–395. <https://doi.org/10.1109/ICHI52183.2021.00064>
 - Mthoko, H., Adamu, M., & Lazem, S. (2023). Context, Culture, and Fabulations: In Search of a Home for Our Veiled African Design Stories. *Interactions*, 30(2), 30–33. <https://doi.org/10.1145/3581642>
 - Nardi, B., Tomlinson, B., Patterson, D. J., Chen, J., Pargman, D., Raghavan, B., & Penzenstadler, B. (2018). Computing within limits. *Communications of the ACM*, 61(10), 86–93. <https://doi.org/10.1145/3183582>
 - Nielsen, K. S., Clayton, S., Stern, P. C., Dietz, T., Capstick, S., & Whitmarsh, L. (2021). How psychology can help limit climate change. *The American Psychologist*, 76(1), 130–144. <https://doi.org/10.1037/amp0000624>
 - Nyström, T., & Mustaquim, M. M. (2014). Sustainable information system design and the role of sustainable HCI. *Proceedings of the 18th International Academic MindTrek Conference: Media Business, Management, Content & Services*, 66–73. <https://doi.org/10.1145/2676467.2676486>
 - Olsson, T., & Väänänen, K. (2021). How does AI challenge design practice? *Interactions*, 28(4), 62–64. <https://doi.org/10.1145/3467479>
 - Personalization and the Smart Home: Questioning techno-hedonist imaginaries – Kari Dahlgren, Sarah Pink, Yolande Strengers, Larissa Nicholls,
 - Jathan Sadowski, 2021. (n.d.). Retrieved 3 July 2023, from <https://journals.sagepub.com/doi/pdf/10.1177/13548565211036801>
 - Pietilä, I., Meriläinen, N., Varsaluoma, J., & Väänänen, K. (2021). Understanding youths' needs for digital societal participation: Towards an inclusive Virtual Council. *Behaviour & Information Technology*, 40(5), 483–496. <https://doi.org/10.1080/0144929X.2021.1912182>
 - Preist, C., Schien, D., & Bleviss, E. (2016). Understanding and Mitigating the Effects of Device and Cloud Service Design Decisions on the Environmental Footprint of Digital

- Infrastructure. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 1324–1337. <https://doi.org/10.1145/2858036.2858378>
- Root, E., Heuten, W., & Boll, S. (2019). Maker Cards: Evaluating design cards for teaching physical computing to middle-school girls. Proceedings of Mensch Und Computer 2019, 493–497. <https://doi.org/10.1145/3340764.3344445>
 - Royakkers, L., Timmer, J., Kool, L., & van Est, R. (2018). Societal and ethical issues of digitization. *Ethics and Information Technology*, 20(2), 127–142. <https://doi.org/10.1007/s10676-018-9452-x>
 - Sadeghian, S., Wintersberger, P., Laschke, M., & Hassenzahl, M. (2022). Designing Sustainable Mobility: Understanding Users' Behavior. Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 34–44. <https://doi.org/10.1145/3543174.3546833>
 - Sari, E. (2022). Changing the world through HCI and UX. *Communications of the ACM*, 65(3), 7. <https://doi.org/10.1145/3510541>
 - Sari, E., Tedjasaputra, A., Kurniawan, Y., Zulaikha, E., Asfarian, A., Ghazali, M., Sivaji, A., Abu Bakar, J. A., Wong, C. Y., Norowi, N. M., Makany, T., Perera-Schulz, D., Chintakovid, T., Nuchitprasitchai, S., & Ong, E. (2023). HCI in Southeast Asia: The Journey Forward. *Asian HCI Symposium '22*, 48–51. <https://doi.org/10.1145/3516492.3558812>
 - Schien, D., Shabajee, P., Chandaria, J., Williams, D., & Preist, C. (2021). Using behavioural data to assess the environmental impact of electricity consumption of alternate television service distribution platforms. *Environmental Impact Assessment Review*, 91, 106661. <https://doi.org/10.1016/j.eiar.2021.106661>
 - Schmidt, A., Giannotti, F., Mackay, W., Shneiderman, B., & Väänänen, K. (2021). Artificial Intelligence for Humankind: A Panel on How to Create Truly Interactive and Human-Centered AI for the Benefit of Individuals and Society. In C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, & K. Inkpen (Eds.), *Human-Computer Interaction – INTERACT 2021* (pp. 335–339). Springer International Publishing. https://doi.org/10.1007/978-3-030-85607-6_32
 - Scuri, S., Ferreira, M., Jardim Nunes, N., Nisi, V., & Mulligan, C. (2022). Hitting the Triple Bottom Line: Widening the HCI Approach to Sustainability. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 1–19. <https://doi.org/10.1145/3491102.3517518>
 - Shakeri, G., & McCallum, C. H. (2021). Envirofy your Shop: Development of a Real-time Tool to Support Eco-friendly Food Purchases Online. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 1–10. <https://doi.org/10.1145/3411763.3451713>
 - Shukla, R., Kiran, N., Wang, R., Gummeson, J., & Lee, S. I. (2019). SkinnyPower: Enabling batteryless wearable sensors via intra-body power transfer. Proceedings of the 17th Conference on Embedded Networked Sensor Systems, 68–82. <https://doi.org/10.1145/3356250.3360034>
 - Skowronek, J., Raake, A., Berndtsson, G. H., Rummukainen, O. S., Usai, P., Gunkel, S. N. B., Johanson, M., Habets, E. A. P., Malfait, L., Lindero, D., & Toet, A. (2022). Quality of Experience in Telemeetings and Videoconferencing: A Comprehensive Survey. *IEEE Access*, 10, 63885–63931. <https://doi.org/10.1109/ACCESS.2022.3176369>
 - Smids, J., Nyholm, S., & Berkers, H. (2020). Robots in the Workplace: A Threat to – or Opportunity for – Meaningful Work? *Philosophy & Technology*, 33(3), 503–522. <https://doi.org/10.1007/s13347-019-00377-4>

- Soden, R. (2022). Reimagining environmental data. *Interactions*, 29(1), 44–47. <https://doi.org/10.1145/3501302>
- Spiel, K. (2021). "Why are they all obsessed with Gender?" – (Non)binary Navigations through Technological Infrastructures. *Designing Interactive Systems Conference 2021*, 478–494. <https://doi.org/10.1145/3461778.3462033>
- Spiel, K. (2022). Transreal tracing: Queer-feminist speculations on disabled technologies. *Feminist Theory*, 23(2), 247–265. <https://doi.org/10.1177/14647001221082299>
- Stamhuis, S., Brombacher, H., Vos, S., & Lallemand, C. (2021). Office Agents: Personal Office Vitality Sensors with Intent. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–5. <https://doi.org/10.1145/3411763.3451559>
- Strengers, Y. (2014). Smart energy in everyday life: Are you designing for resource man? *Interactions*, 21(4), 24–31. <https://doi.org/10.1145/2621931> The "Artificial" Colleague: Evaluation of Work Satisfaction in Collaboration with Non-human Coworkers | 27th International Conference on Intelligent User Interfaces. (n.d.). Retrieved 3 July 2023, from <https://dl.acm.org/doi/10.1145/3490099.3511128> Visualising Forest Sound: Justice-led Ecoacoustic Data Interaction – CHI '23. (n.d.). Retrieved 27 June 2023, from <https://programs.sigchi.org/chi/2023/program/content/99121>
- Wang, Y., Zhou, J., Li, H., Zhang, T., Gao, M., Cheng, Z., Yu, C., Patel, S., & Shi, Y. (2019). FlexTouch: Enabling Large-Scale Interaction Sensing Beyond Touchscreens Using Flexible and Conductive Materials. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(3), 109:1-109:20. <https://doi.org/10.1145/3351267>
- Williams, K., Kane, B., Clarke, C., & Widdicks, K. (2021). The Environmental Impact of Research Communities: Insights from conference sustainability chairs. *XRDS: Crossroads, The ACM Magazine for Students*, 27(4), 46–51. <https://doi.org/10.1145/3468216>
- Ystgaard, K. F., Atzori, L., Palma, D., Heegaard, P. E., Bertheussen, L. E., Jensen, M. R., & De Moor, K. (2023). Review of the theory, principles, and design requirements of human-centric Internet of Things (IoT). *Journal of Ambient Intelligence and Humanized Computing*, 14(3), 2827–2859. <https://doi.org/10.1007/s12652-023-04539-3>
- Zhang, T., Yi, X., Yu, C., Wang, Y., Becker, N., & Shi, Y. (2018). Touchpower: Interaction-based Power Transfer for Power-as-needed Devices. *GetMobile: Mobile Computing and Communications*, 22(2), 27–31. <https://doi.org/10.1145/3276145.3276156>

References

- 1 Swamy Ananthanarayan and Susanne Boll. Physical computing for children: Shifting the pendulum back to papertian ideals. *Interactions*, 27(3):40–45, apr 2020.
- 2 James Auger. Speculative design: crafting the speculation. *Digital Creativity*, 24(1):11–35, 2013.
- 3 Elke Beck, Kai von Holdt, Jochen Meyer, and Susanne Boll. Sneaking physical exercise into sedentary work life: Design explorations of ambient reminders in opportune moments. In *2019 IEEE International Conference on Healthcare Informatics (ICHI)*, pages 1–7, 2019.
- 4 Nasim Beheshtian, Sina Moradi, Aino Ahtinen, Kaisa Väänänen, Kalle Kähkönen, and Markus Laine. Greenlife: A persuasive social robot to enhance the sustainable behavior in shared living spaces. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, NordiCHI '20, New York, NY, USA, 2020. Association for Computing Machinery.

- 5 Nicola J. Bidwell. Rural uncommoning: Women, community networks and the enclosure of life. *ACM Trans. Comput.-Hum. Interact.*, 28(3), jul 2021.
- 6 Nicola Jane Bidwell. *Keeping the World Alive: An Alternative Story for Funding Community Networks*, pages 99–104. FGV Direito Rio, December 2021. Official Outcome of the IGF Dynamic Coalition on Community Connectivity. Presented at the United Nations Internet Governance Forum. Katowice, Poland, December 2021.
- 7 Susanne Boll and Markus Schnell. Mit künstlicher intelligenz zu nachhaltigen geschäftsmodellen – nachhaltigkeit von, durch und mit ki. Sep 2022.
- 8 Susanne Boll, Hari Sundram, Svetha Venkatesh, Martha Larson, and Mohan Kankanhalli. The world has changed - the world needs to change. what multimedia has to offer for our common digital future. In *Proceedings of the 28th ACM International Conference on Multimedia*, MM '20, page 4796–4798, New York, NY, USA, 2020. Association for Computing Machinery.
- 9 Judith Dörrenbächer, Matthias Laschke, Diana Löffler, Ronda Ringfort, Sabrina Großkopp, and Marc Hassenzahl. Experiencing Utopia. A Positive Approach to Design Fiction. *arXiv e-prints*, page arXiv:2105.10186, May 2021.
- 10 Anthony Dunne and Fiona Raby. *Speculative Everything: Design, Fiction, and Social Dreaming*. The MIT Press, 2013.
- 11 Sarah Forberger, Karin Bammann, Jürgen Bauer, Susanne Boll, Gabriele Bolte, Tilman Brand, Andreas Hein, Frauke Koppelin, Sonia Lippke, Jochen Meyer, Claudia R Pischke, Claudia Voelcker-Rehage, and Hajo Zeeb. How to tackle key challenges in the promotion of physical activity among older adults (65+): The aequipa network approach. *International Journal of Environmental Research and Public Health*, 2017.
- 12 Marc Hassenzahl. *Experience Design: Technology for All the Right Reasons*. Synthesis Lectures on Human-Centered Informatics. Morgan & Claypool Publishers, 2010.
- 13 Marc Hassenzahl, Judith Dörrenbächer, Matthias Laschke, and Shadan Sadeghian. European union's green smart directive or how resource-conscious smart systems saved the world. In *Nordic Human-Computer Interaction Conference*, NordiCHI '22, New York, NY, USA, 2022. Association for Computing Machinery.
- 14 Marc Hassenzahl and Matthias Laschke. Pleasurable Troublemakers. In *The Gameful World: Approaches, Issues, Applications*. The MIT Press, 01 2015.
- 15 Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. Social acceptability in hci: A survey of methods, measures, and design strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–19, New York, NY, USA, 2020. Association for Computing Machinery.
- 16 Joseph Lindley and Paul Coulton. Back to the future: 10 years of design fiction. In *Proceedings of the 2015 British HCI Conference*, British HCI '15, page 210–211, New York, NY, USA, 2015. Association for Computing Machinery.
- 17 Conor Linehan, Ben J. Kirman, Stuart Reeves, Mark A. Blythe, Theresa Jean Tanenbaum, Audrey Desjardins, and Ron Wakkary. Alternate endings: Using fiction to explore design futures. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, page 45–48, New York, NY, USA, 2014. Association for Computing Machinery.
- 18 Jouko Makkonen, Rita Latikka, Laura Kaukonen, Markus Laine, and Kaisa Väänänen. Advancing residents' use of shared spaces in nordic superblocks with intelligent technologies. *AI Soc.*, 38(3):1167–1184, dec 2022.
- 19 Erika Root, Wilko Heuten, and Susanne Boll. Maker cards: Evaluating design cards for teaching physical computing to middle-school girls. In *Proceedings of Mensch Und Computer 2019*, MuC '19, page 493–497, New York, NY, USA, 2019. Association for Computing Machinery.

- 20 Uwe Schneidewind. *Utopische Wissenschaft. Oder: Warum gerade Wirtschaftswissenschaften als Möglichkeitswissenschaft konzipiert sein sollten.*, pages S. 147–159. 01 2017.
- 21 Gözel Shakeri, Frederike Jung, Ferran Altarriba Bertran, Adrian Friday, and Daniel Fernández Galeote. Eco-joy: Imagining sustainable and joyful food eco-label futures. In *Adjunct Proceedings of the 2022 Nordic Human-Computer Interaction Conference, NordiCHI '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- 22 Gözel Shakeri, Frederike Jung, Ferran Altarriba Bertran, Adrian Friday, and Daniel Fernández Galeote. Eco-joy: Imagining sustainable and joyful food eco-label futures. In *Adjunct Proceedings of the 2022 Nordic Human-Computer Interaction Conference, NordiCHI '22*, New York, NY, USA, 2022. Association for Computing Machinery.
- 23 Jari Varsaluoma, Iikka Pietilä, Kaisa Väänänen, and Tiina Rättilä. Experimenting with youth-centred e-participation. *Young People as Agents of Sustainable Society: Reclaiming the Future*, 2023.

Participants

- Nic Bidwell
IUM – Windhoek, NA
- Susanne Boll
Universität Oldenburg, DE
- Katrien De Moor
NTNU – Trondheim, NO
- Markus Fiedler
Blekinge Institute of Technology –
Karlshamn, SE
- Adrian Friday
Lancaster University, GB
- Marc Hassenzahl
Universität Siegen, DE
- Jason Jacques
University of St Andrews, GB
- Kirsikka Kaipainen
Tampere University of
Technology, FI
- Lenneke Kuijer
TU Eindhoven, NL
- Neha Kumar
Georgia Institute of Technology –
Atlanta, US
- Carine Lallemand
University of Luxembourg –
Esch-sur-Alzette, LU &
Eindhoven University of
Technology, NL
- Matthias Laschke
Universität Siegen, DE
- Dan Lockton
TU Eindhoven, NL
- Joycelyn Longdon
University of Cambridge, GB
- Anjali Karol Mohan
INDE – Bangalore, IN
- Robin Neuhaus
Universität Siegen, DE
- Jacki O’Neill
Microsoft Africa Research
Institute – Nairobi, KE
- Thomas Olsson
University of Tampere, FI
- Daniel Pargman
KTH Royal Institute of
Technology – Stockholm, SE
- Alexander Raake
TU Ilmenau, DE
- Shadan Sadeghian Borojeni
Universität Siegen, DE
- Eunice Sari
UX Indonesia – Jakarta, ID
- Gözel Shakeri
Universität Oldenburg, DE
- Katta Spiel
TU Wien, AT
- Kaisa Väänänen
University of Tampere, FI



Remote Participants

- Stephan Ramesohl
Wuppertal Institut, DE
- Tengxiang Zhang
Chinese Academy of Sciences –
Beijing, CN