DAGSTUHL
REPORTS

Volume 12, Issue 12, December 2022

Dagstuhl Reports, Vol. 12, Issue 12 ISSN 2192-5283

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
May, 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`.

Digital Object Identifier: 10.4230/DagRep.12.12.i

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.

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

Report from Dagstuhl Seminar 22491

Cognitive Augmentation

Kai Kunze*[1], Pattie Maes*[2], Florian 'Floyd' Mueller*[3], and Katrin Wolf*[4]

1 **Keio University – Yokohama, JP.** kai@kmd.keio.ac.jp
2 **MIT – Cambridge, US.** pattie@media.mit.edu
3 **Monash University – Clayton, AU.** floyd@exertiongameslab.org
4 **Berliner Hochschule für Technik, DE.** katrin.wolf@bht-berlin.de

Abstract

Mobile phones and other connected wearable systems transformed the way we interact with information, offering access to vast amounts of knowledge at our fingertips. However, the challenge remains on how to make this information more accessible and intuitive. The field of cognitive augmentation aims to enhance our cognitive abilities through technology, allowing us to interact with digital data more naturally and efficiently. This Dagstuhl Seminar brought together experts in neuroscience, psychology, physiology, wearable computing, human-computer interaction, machine perception, and pattern recognition to discuss the possibility of augmenting our cognitive skills and creating new digital senses. The seminar explored the latest findings in these fields and their potential for improving human performance, productivity, and creativity. Ultimately, the goal is to bridge the gap between humans and machines, enabling a more seamless and intuitive interaction between the two. The main discussion topic centered around the possibilities and challenges of digitally augmenting our cognition.

Seminar December 4–9, 2022 – http://www.dagstuhl.de/22491
2012 ACM Subject Classification Human-centered computing → Human computer interaction (HCI); Human-centered computing → Ubiquitous and mobile computing
Keywords and phrases augmented humans, augmented reality, wearable computing
Digital Object Identifier 10.4230/DagRep.12.12.1

1 Executive Summary

Kai Kunze (Keio University – Yokohama, JP)
Pattie Maes (MIT – Cambridge, US)
Florian 'Floyd' Mueller (Monash University – Clayton, AU)
Katrin Wolf (Berliner Hochschule für Technik, DE)

The real and digital worlds are increasingly more interconnected, leaving people to split their attention between tasks in the physical world in an increasing amount of ubiquitous systems and IoT services. We see an increase in accidents related to the usage of digital tools (such as interacting with a smartphone while driving). As governments and healthcare experts around the world call for changing lifestyles in response to the Covid-19 pandemic, the development, and usage of remote communication and touchless technologies are rapidly becoming an essential part of the "new normal". At the same time, the absence of touch and physical contact highlights their critical importance in human life, from school to hospital to care facilities. We need more intuitive, direct ways to interface with technology. Students

* Editor / Organizer

Cognitive Augmentation, *Dagstuhl Reports*, Vol. 12, Issue 12, pp. 1–26
Editors: Kai Kunze, Pattie Maes, Florian 'Floyd' Mueller, and Katrin Wolf
DAGSTUHL REPORTS Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

and educators find it harder and harder to concentrate, and news outlets are already talking about the distraction economy. The seminar focused on people interacting with information from the digital domain, in a minimally disruptive way, creating novel sensory experiences using and extending human perception and ultimately cognition.

The overall objective of the seminar was to foster research, explore, and model new means for increasing human intake of information in order to lay the foundation for augmented cognition, especially through somatosensation: the ability to sense the environment through our body.

Machine Learning has often been used to mimic or surpass some cognitive functions of the human mind (visual object/face recognition, playing chess, etc.). Such efforts appear to put humans and computers in a competitive relationship, as emphasized in AI vs. Human competitions. Once a fear of AIs "replacing" human workers is now taken much more seriously and discussed in the public sphere. This Dagstuhl proposal suggests a different approach to the human-computer relationship by applying a cooperative and empowering framework. One important characteristic of the human mind is that it has significant fluctuations in productivity and capacity. Our mind has ebbs and flows, and is affected by various factors, some of which we do not even realize. These fluctuations manifest in patterns in human behavior and physiological signals (body temperature, eye movements, galvanic skin response, etc.). With this seminar, we aim to discuss technologies that can give us more insights into the ebb and flow of the human mind as a basis for cognitive augmentation.

The participants developed several frameworks and taxonomies for understanding and evaluating different types of cognitive augmentation, based on their goals, methods, and impacts. The frameworks focus on enhancement, compensation, offloading, and replacement, and consider factors such as safety, efficacy, and social impact. There are several publication plans and several participants already agreed to organize conference workshops together (for example at AugmentedHumans 2023 and UbiComp/ISWC 2023).

Overall, the Dagstuhl Seminar on Cognitive Augmentation advanced our research fields by creating a shared understanding of the concept and its implications, promoting interdisciplinary collaboration and communication, and identifying promising directions for future research and development. The outcomes of the seminar are to inform the design, implementation, and evaluation of cognitive augmentation technologies, and contribute to augmenting human cognition by applying ethical principles.

2 Table of Contents

3 Seminar Overview

The concept to use information technology to augment the human intellect goes back to the 1960s [Winograd, Engelbart]. The basic idea is to extend the computational (and other) capabilities of the human mind using technology. While augmenting the human intellect is a long time focus of various research efforts, we think it is necessary to rethink what intelligence amplification means and become more specific towards Cognitive Augmentation [Schmidt]:

- Cognitive scientists and psychologists have now better insights on how perception, cognition and, in general, how our mind works
- Affordable physiological and cognitive activity recognition systems are being developed in the HCI and wearable computing research communities
- Advances in new sensing and actuation technologies (e.g., somatosensory and odor actuation, wearable EEG) that can be integrated with our senses
- Advances in real-life tracking of physical, cognitive and emotional states

The ongoing technical progress in these key areas will enable fundamentally new approaches to amplify the human intelligence. The technological advances mentioned above already started to fundamentally transform how we communicate with each other and will do this even further. The seminar was split into topics tackling state-of-the-art, future directions, and potential issues and perils:

- Discussion of existing and emerging technology for cognitive augmentation approaches and methodologies between the different communities.
- Discovering compelling application cases
- Exploring new sensing and actuation modalities to amplify human intelligence
- Exploring technology for more effective skill sharing and learning Discussing privacy and ethical implications

In contrast to more traditional Dagstuhl Seminars, we followed a more active approach using an ignite talk format to introduce participants and reduce the time of talks and to leave the rest of the week for hands-on sessions, group work, deeper discussions, and socializing.

We also invited the participants to bring demonstrations of their research to Dagstuhl, so participants can experience the research on their own body. We believe that this is crucial in the field of cognitive augmentation, as our mind is embodied and simple discussion or written reports cannot completely convey the effectiveness and experience of some technologies.

The seminar started with very short ignite talks on the first day, so all participants could get to know the interests of each other. This session was followed by a clustering of discussion items, ideas and interests for the participants to foster spontaneous discussions and an exchange of ideas during breaks. The second day, we had a special full day workshop on “Human Augmentation Design”, a detailed description will follow in the next section. The third day was more speculative dealing with human futures and ethical implications. The forth day focused on human cognitive attributes (e.g. attention, agency etc.) and included a special workshop on memory augmentation (details also given later). The last day was used for wrap up and for discussions on collaborations after the seminar (e.g. joint publications, workshop organizations, book chapters etc.).

The participants organized and participated in several workshops that focused on different aspects of cognitive augmentation. In the rest of this report, we will highlight two workshops (one with a hands-on-prototyping session): A Human Augmentation Design and a Memory Augmentation workshop. We conclude with the list of ignite talks from the participants.

4 Workshops

In the following we will go into details about two workshops, the Human Augmentation Design and the Memory Augmentation workshop held as part of the seminar.

4.1 Human Augmentation Design: Immediate Issues

Steeven Villa (LMU München, DE), Matthias Hoppe (LMU München, DE), Thomas Kosch (HU Berlin, DE), and Katrin Wolf (Berliner Hochschule für Technik, DE)

4.1.1 Introduction

Human augmentation is a new Human-Computer Interaction (HCI) field. Subsequently, cognitive augmentation proliferated as a research area into the human augmentation field. HCI professionals and academics spent the last decades to show how the new wave of ubiquitous technologies, such as AI, AR, and VR, can be used to improve human cognition. However, with the creation of new technologies allowing people to improve their natural abilities and the addition of new users, a whole new set of challenges arises. In this one-day workshop, we compile and discuss the meta-issues that human augmentation, and especially cognitive augmentations, will face in the coming years. In the activities performed in the workshop, such as planning, prototyping, evaluating, and discussing augmentation technologies, participants focused on the validity of research methodologies, the definition of human augmentation used, the societal challenges included, and ethical implications of their research. The overarching objective of this workshop is to raise awareness of the immediate conceptual, methodological, and social challenges that face the field of cognitive augmentation.

4.1.2 Agenda

- **Part one – Introduction:**
  - Welcome and introduction – Thomas Kosch
  - Keynote: "Human Augmentation Design: Immediate Issues" – Steeven Villa
- **Part two – Hands-on-session:**
  - Hands-on Session: Prototyping and Evaluating Cognitive Augmentation – Steeven Villa, Thomas Kosch, Matthias Hoppe, All attendees
  - Result Presentations – All attendees
- **Part three – Discussion:**
  - Walk and Talk: Concepts, Implications, and Ethics of Cognitive Augmentation – All attendees
  - Open Discussion of Results – All attendees

4.1.3 Outcomes – Part One: Introduction

In the first part, a keynote by Steeven Villa was presented (see Figure 1). During the keynote, the points of society's perceptions of human augmentation, the impact of human augmentation on self-perception, and risk-taking behavior were addressed. In the last part of the keynote, methodological recommendations for evaluation validity were presented. The session was followed by a round of questions and discussion.

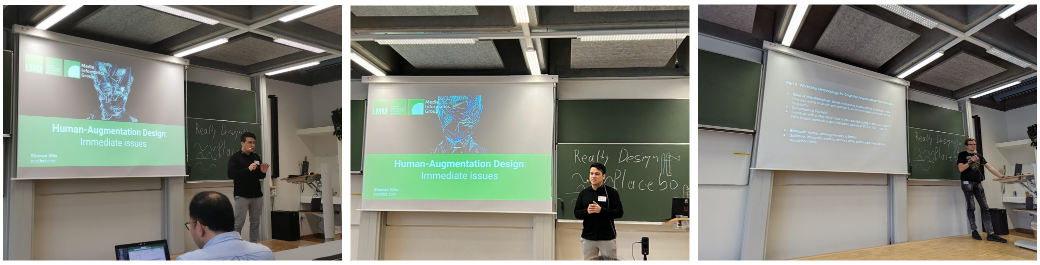


Figure 1 Pictures from the Intro.

Figure 2 Pictures from the Hands-On Session.

4.1.4 Outcomes – Part Two: Hands-On Session

During the hands-on part of the workshop, the participants were put into groups and asked to brainstorm and make prototypes of cognitive enhancements. We instructed participants to develop hypothetical and conceptual lo-fi prototypes. This allowed participants to freely explore and express their ideas without being restricted by physical constraints.

The groups were then instructed to develop a cognitive augmentation prototype and a plan for its evaluation and analysis, as well as its long-term effects on the user. The prototype idea then had to be presented in the form of a user story, showing how the prototype solves a problem and talking about how it will affect society in 20, 50, and 100 years.

After the prototyping session, four groups presented their results, the concepts are presented below:

“Life Rerouting”: while one might wonder how their life would have changed if they had taken a different route, with an AI-simulated life map, you can find answers to these questions and help yourself think and make decisions. “Life Rerouting” not only lets you stay in different timelines, but it also helps you figure out how the path you choose now will affect your future.

“Umm-less”: The system helps a speaker that might be stuck in a conversation and cannot find the right words by using Auto-Complete Talking that shows world-anchored AR cues. The system recognizes the conversation and context and suggests next possible worlds by combining transcription of the conversation with auto-complete suggestions.

“TransCap”: Every person has their strengths and weaknesses in different skills. Learning a new skill can take time and effort. TransCap is an augmentation device that can transport human capabilities. Instead of learning a new skill just to be used once, one can borrow it from another person who already knows the desired skill.

“Collective Intelligence”: How can human beings understand each other and create social empathy? By being exposed to different cultures, communities, or even species, collective intelligence is shaped. However, by fading information via AR (hiding other people or

Figure 3 More Pictures from the Hands-On Session.

Figure 4 Pictures from the Walk and Talk and discussions.

blocking out fears), this exposure is changed. Therefore, exposure can change the bigger picture and cause dangerous impacts on society caused by the management of companies or social groups.

4.1.5 Outcomes – Part Three: Discussion

In the final part, we encouraged participants to engage in an open discussion about the points raised during the keynote and the prototyping session. We invited participants to walk in groups around the facilities of Dagstuhl. The participants then spent the remaining time of the session discussing these topics outside. After the session outside, some participants returned to the room to visualize and summarize the concepts discussed during the day.

4.1.6 Conclusion

While the creation and evaluation of human augmentation technologies are still challenging, the human augmentation field is growing. In a keynote, the current challenges of human augmentation technologies were presented and discussed by the participants. Subsequently, the participants presented low-fidelity prototypes of future human augmentation technologies while denoting their challenges. A final discussion of the implications of human augmentation technologies showed how evaluation methods need improvement and a design space for such technologies needs to be created. Overall, the workshop is considered a success, building the next cornerstone for the human augmentation community.

4.2 Memory Augmentation Workshop

Samantha W.T. Chan (MIT – Cambridge, US)

Co-authors/Organizers/Facilitators: Samantha Chan, Pattie Maes, Evangelos Niforatos, Nathan Whitmore, Gayathri Subramanian, Jiawen Han, Rakesh Patibanda, Florian 'Floyd' Mueller
Aim: The workshop aimed to introduce the topic of Memory Augmentation in HCI through a series of five short talks. Workshop participants (mainly from the Cognitive Augmentation Dagstuhl Seminar) are led by facilitators to discuss taxonomies, open questions, opportunities, and ideas for future interventions in the area.
Materials: The workshop organizers prepared a document with a rough outline of the taxonomy and a spreadsheet listing existing systems for memory augmentation and their relation to the taxonomy. A Miro board and Zoom room link were also prepared beforehand.

4.2.1 Agenda

Five short talks are given to set the stage (8-min talk + 5 min Q&A for each speaker).

- Pattie Maes, MIT Media Lab (Remote) Introduction to workshop
  Introduce the purpose of the session. "7 sins of memory" talk about some prior HCI work
- Evangelos Niforatos, TU Delft (Remote)
  Episodic memory / Contextual Lifelogging & Affect / Physiological sensing (AI+HMDs)
- Samantha Chan, MIT Media Lab (In-Person)
  Needs of the Elderly, Augmenting Prospective Memory
- Nathan Whitmore, MIT Media Lab (Remote)
  Role of Sleep in Memory and the use of Targeted Memory Reactivation
- Gayathri Subramanian, NorthWestern University (Remote)
  Entrainment and Brain Stimulation for Memory Improvements

Breakout sessions are done to work on a memory augmentation taxonomy.

- Pattie Maes – Intro to breakout sessions Breakout groups (45 mins)
- 3 in-person groups (Facilitators: Samantha, Rakesh, Jiawen)
  Facilitators take notes using whiteboards and pin-boards and take pictures afterwards or use Miro board.
- 1 remote group (Facilitator: Pattie) – uses Miro board.
- Guiding Questions for the Breakout Groups:
  1. Which memory augmentation systems/apps do you know?
  2. What remains to be tackled in the field of Human Memory Augmentation (challenges/opportunities)?
  3. What are interesting ideas for memory augmentation interventions? Use Schacter's 7 sins of memory as the framework to guide the discussion.
- Facilitators report back on group discussions.

4.2.2 Feedback on Taxonomy

Workshop organizers developed and added two taxonomy categories to the Miro board for participants to discuss and provide feedback on. The resulting outlines are shown below, the items in italics are what were added to the taxonomies during the discussions.

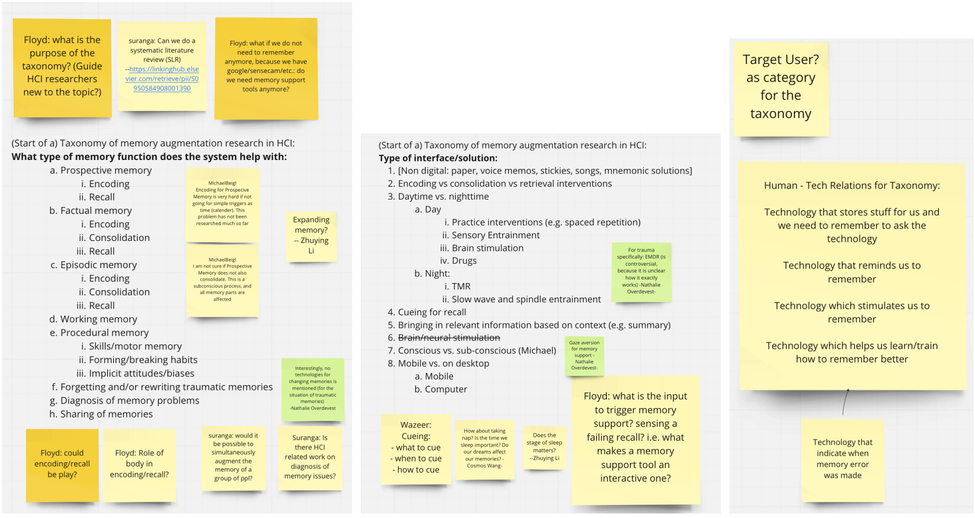
1. Taxonomy Category: What type of memory function does the system help with:
   - Prospective memory: Encoding / Recall
   - Factual memory: Encoding / Consolidation / Recall
   - Episodic memory: Encoding / Consolidation / Recall
   - Working memory
   - Procedural memory: Skills/motor memory, Forming/breaking habits, Implicit attitudes/biases
   - Forgetting and/or rewriting traumatic memories: Diagnosis of memory problems / Sharing of memories
2. Taxonomy Category: Type of interface/solution:
   - Non-digital: paper, voice memos, stickies, songs, mnemonic solutions
   - Encoding vs consolidation vs retrieval interventions
   - Day
     - Practice interventions (e.g. spaced repetition)
     - Sensory Entrainment
     - Brain stimulation
     - Drugs
   - Night
     - Targeted Memory Reactivation (TMR) during sleep
     - Slow wave and spindle entrainment
   - Cueing for recall
   - Bringing in relevant information based on context (e.g. summary)
   - Conscious vs. sub-conscious
   - Mobile vs. on desktop

General Feedback: It was suggested to clearly define the purpose of the taxonomy, that it is meant to guide HCI researchers who are new to the topic. A question was raised on whether we need memory support tools anymore since we already have tools like Google and SenseCam.

Feedback on taxonomy based on memory function: It was argued that there is still a need for tools to support aspects of memory function such as prospective memory, changing memories, and diagnosis of memory issues (See Figure 5).

Feedback on taxonomy based on type of interface/solution: Many participants described non-digital solutions such as sticky notes and paper memos. Some of these non-digital solutions can be digitalized. We should clarify that the taxonomy should only include digital interfaces. It was mentioned that we should consider: What are the inputs to trigger memory support? Could it be when sensing a failing recall? What makes a memory support tool an interactive one? (see 5).

Potential new taxonomy categories: Participants suggested that the interfaces could also be categorized based on target users and in terms of human-technology relations, as proposed below (see also Figure 5).



Floyd: what is the purpose of the taxonomy? (Guide HCI researchers new to the topic?)

suranga: Can we do a systematic literature review (SLR) --https://linkinghub.elsevier.com/retrieve/pii/S0950584908001390

Floyd: what if we do not need to remember anymore, because we have google/sensecam/etc.: do we need memory support tools anymore?

(Start of a) Taxonomy of memory augmentation research in HCI:
What type of memory function does the system help with:

a. Prospective memory
   i. Encoding
   ii. Recall
b. Factual memory
   i. Encoding
   ii. Consolidation
   iii. Recall
c. Episodic memory
   i. Encoding
   ii. Consolidation
   iii. Recall
d. Working memory
e. Procedural memory
   i. Skills/motor memory
   ii. Forming/breaking habits
   iii. Implicit attitudes/biases
f. Forgetting and/or rewriting traumatic memories
g. Diagnosis of memory problems
h. Sharing of memories

MichaelBeigl
Encoding for Prospecive Memory is very hard if not going for simple triggers as time (calender). This problem has not been researched much so far

MichaelBeigl
I am not sure if Prospective Memory does not also consolidate. This is a subconscious process, and all memory parts are affected

Expanding memory? -- Zhuying Li

Interestingly, no technologies for changing memories is mentioned (for the situation of traumatic memories) -Nathalie Overdevest

Floyd: could encoding/recall be play?

Floyd: Role of body in encoding/recall?

suranga: would it be possible to simultaneously augment the memory of a group of ppl?

Suranga: Is there HCI related work on diagnosis of memory issues?

(Start of a) Taxonomy of memory augmentation research in HCI:
Type of interface/solution:

1. [Non digital: paper, voice memos, stickies, songs, mnemonic solutions]
2. Encoding vs consolidation vs retrieval interventions
3. Daytime vs. nighttime
   a. Day
      i. Practice interventions (e.g. spaced repetition)
      ii. Sensory Entrainment
      iii. Brain stimulation
      iv. Drugs
   b. Night:
      i. TMR
      ii. Slow wave and spindle entrainment
4. Cueing for recall
5. Bringing in relevant information based on context (e.g. summary)
6. ~~Brain/neural stimulation~~
7. Conscious vs. sub-conscious (Michael)
8. Mobile vs. on desktop
   a. Mobile
   b. Computer

For trauma specifically: EMDR (is controversial, because it is unclear how it exactly works) -Nathalie Overdevest-

Gaze aversion for memory support - Nathalie Overdevest-

Wazeer:
Cueing:
- what to cue
- when to cue
- how to cue

How about taking nap? Is the time we sleep important? Do our dreams affect our memories? - Cosmos Wang-

Does the stage of sleep matters? --Zhuying Li

Floyd: what is the input to trigger memory support? sensing a failing recall? i.e. what makes a memory support tool an interactive one?

Target User? as category for the taxonomy

Human - Tech Relations for Taxonomy:

Technology that stores stuff for us and we need to remember to ask the technology

Technology that reminds us to remember

Technology which stimulates us to remember

Technology which helps us learn/train how to remember better

Technology that indicate when memory error was made

Figure 5 Miro boards from the taxonomy based on memory function, type of interface and suggestions on new taxonomy categories.

Taxonomy Category: Human-Technology Relations.

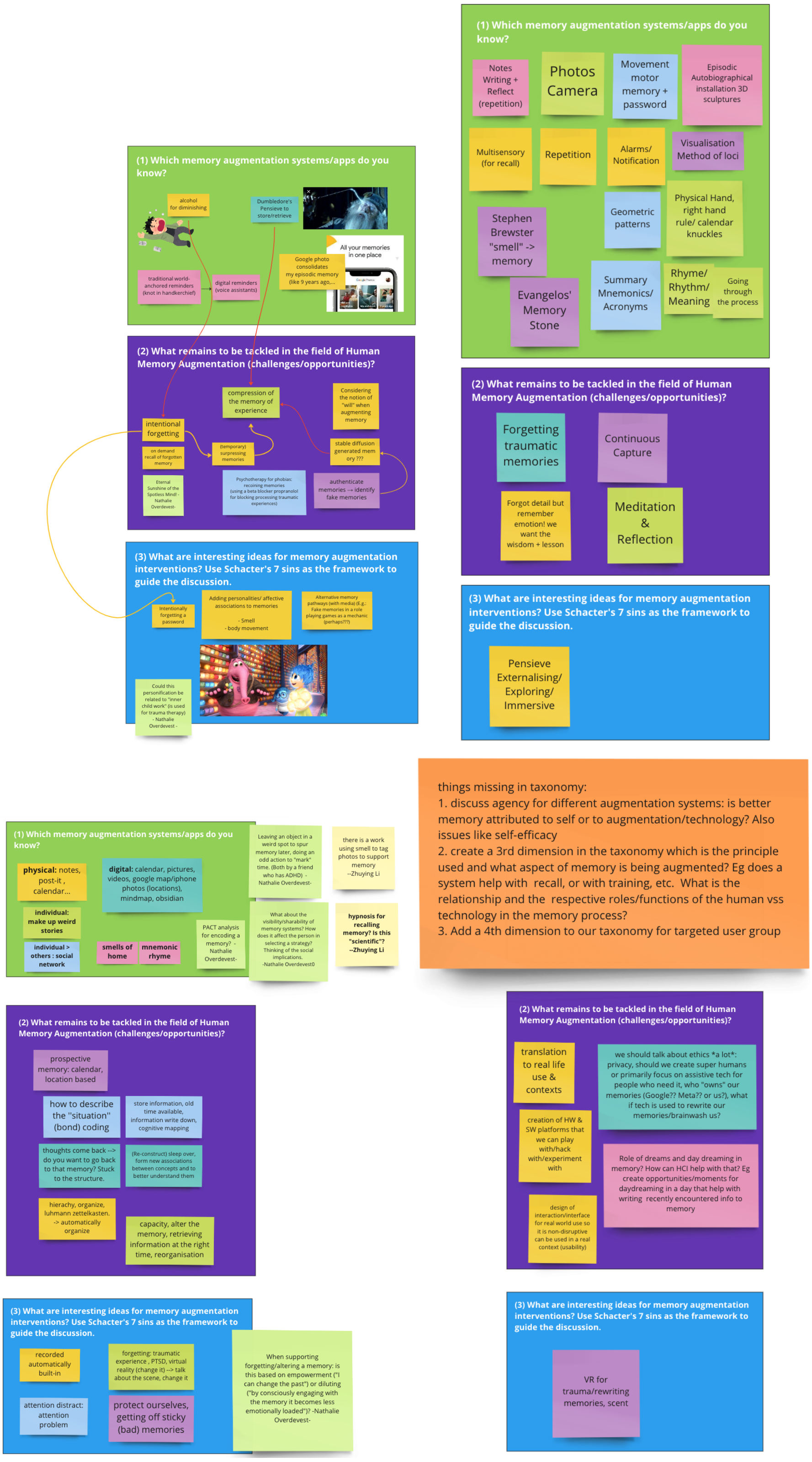
- Technology that stores stuff for us, and we need to remember to ask the technology
- Technology that reminds us to remember
- Technology which stimulates us to remember
- Technology that indicate when memory error was made

4.2.3 Responses to Guiding Questions

The four Breakout Groups discussed answers to the three guiding questions. The summary of the responses to each question are as follows. Figures 4 to 7 show the question responses on Miro board for each group respectively.

1. Which memory augmentation systems/apps do you know? Many participants discussed mnemonic devices and strategies that they have used or seen before. These include "traditional world-anchored reminders (knot in handkerchief)" (Group 1), gestures or memories mapped to our physical hands like the right-hand rule and calendar month lengths using knuckles (Group 3), making up weird stories, and leaving objects in weird locations (Group 2). A few participants talked about movement-related or kinesthetic strategies for remembering, such as doing an odd action to mark time (Group 2), and using geometric patterns and motor memory for remembering passwords (Group 3). Mnemonic rhymes (Group 2, Group 3), rhythm, summary, acronyms, and the method of loci were also mentioned (Group 3). Another strategy included receiving reminders from other people (Group 2).

Existing physical tools are used to record and externalize memories, such as notes (Group 2, Group 3), post-it, calendar (Group 2), the act of writing down, reflecting and repetition were thought to assist in remembering (Group 3). Digital tools were also mentioned. All groups talked about using digital media (photos and videos) and albums "to consolidate episodic memory" (Group 1) with tools like Google Photos (Group 1), Google Maps, and iPhone photos (Group 2). Participants also pointed out digital reminders from voice assistants (Group 1), alarms and notifications (Group 3) on computers or phones. A participant recalled seeing a prototype that "stored" episodic, autobiographical memories in an installation with 3D sculptures (Group 3). The Olfoto interface [1], which used smell (olfactory cues) to tag photos/memories (Group 2, Group 3), was mentioned by participants from two groups.



(1) Which memory augmentation systems/apps do you know?

alcohol for diminishing

Dumbledore's Pensieve to store/retrieve

traditional world-anchored reminders (knot in handkerchief)

digital reminders (voice assistants)

Google photo consolidates my episodic memory (like 9 years ago,...

All your memories in one place

(2) What remains to be tackled in the field of Human Memory Augmentation (challenges/opportunities)?

compression of the memory of experience

Considering the notion of "will" when augmenting memory

intentional forgetting

on demand recall of forgotten memory

(temporary) surpressing memories

stable diffusion generated mem ory ???

Eternal Sunshine of the Spotless Mind! - Nathalie Overdevest-

Psychotherapy for phobias: recoining memories (using a beta blocker propranolol for blocking processing traumatic experiences)

authenticate memories → identify fake memories

(3) What are interesting ideas for memory augmentation interventions? Use Schacter's 7 sins as the framework to guide the discussion.

Intentionally forgetting a password

Adding personalities/ affective associations to memories

- Smell
- body movement

Alternative memory pathways (with media) (E.g.: Fake memories in a role playing games as a mechanic (perhaps???)

Could this personification be related to "inner child work" (is used for trauma therapy) - Nathalie Overdevest -

(1) Which memory augmentation systems/apps do you know?

Notes Writing + Reflect (repetition)

Photos Camera

Movement motor memory + password

Episodic Autobiographical installation 3D sculptures

Multisensory (for recall)

Repetition

Alarms/ Notification

Visualisation Method of loci

Stephen Brewster "smell" -> memory

Geometric patterns

Physical Hand, right hand rule/ calendar knuckles

Evangelos' Memory Stone

Summary Mnemonics/ Acronyms

Rhyme/ Rhythm/ Meaning

Going through the process

(2) What remains to be tackled in the field of Human Memory Augmentation (challenges/opportunities)?

Forgetting traumatic memories

Continuous Capture

Forgot detail but remember emotion! we want the wisdom + lesson

Meditation & Reflection

(3) What are interesting ideas for memory augmentation interventions? Use Schacter's 7 sins as the framework to guide the discussion.

Pensieve Externalising/ Exploring/ Immersive

(1) Which memory augmentation systems/apps do you know?

physical: notes, post-it , calendar...

digital: calendar, pictures, videos, google map/iphone photos (locations), mindmap, obsidian

individual: make up weird stories

individual > others : social network

smells of home

mnemonic rhyme

PACT analysis for encoding a memory? - Nathalie Overdevest-

Leaving an object in a weird spot to spur memory later, doing an odd action to "mark" time. (Both by a friend who has ADHD) - Nathalie Overdevest-

there is a work using smell to tag photos to support memory --Zhuying Li

What about the visibility/sharability of memory systems? How does it affect the person in selecting a strategy? Thinking of the social implications. -Nathalie Overdevest0

hypnosis for recalling memory? Is this "scientific"? --Zhuying Li

(2) What remains to be tackled in the field of Human Memory Augmentation (challenges/opportunities)?

prospective memory: calendar, location based

how to describe the "situation" (bond) coding

store information, old time available, information write down, cognitive mapping

thoughts come back --> do you want to go back to that memory? Stuck to the structure.

(Re-construct) sleep over, form new associations between concepts and to better understand them

hierachy, organize, luhmann zettelkasten. -> automatically organize

capacity, alter the memory, retrieving information at the right time, reorganisation

(3) What are interesting ideas for memory augmentation interventions? Use Schacter's 7 sins as the framework to guide the discussion.

recorded automatically built-in

forgetting: traumatic experience , PTSD, virtual reality (change it) --> talk about the scene, change it

attention distract: attention problem

protect ourselves, getting off sticky (bad) memories

When supporting forgetting/altering a memory: is this based on empowerment ("I can change the past") or diluting ("by consciously engaging with the memory it becomes less emotionally loaded")? -Nathalie Overdevest-

things missing in taxonomy:
1. discuss agency for different augmentation systems: is better memory attributed to self or to augmentation/technology? Also issues like self-efficacy
2. create a 3rd dimension in the taxonomy which is the principle used and what aspect of memory is being augmented? Eg does a system help with recall, or with training, etc. What is the relationship and the respective roles/functions of the human vss technology in the memory process?
3. Add a 4th dimension to our taxonomy for targeted user group

(2) What remains to be tackled in the field of Human Memory Augmentation (challenges/opportunities)?

translation to real life use & contexts

we should talk about ethics *a lot*: privacy, should we create super humans or primarily focus on assistive tech for people who need it, who "owns" our memories (Google?? Meta?? or us?), what if tech is used to rewrite our memories/brainwash us?

creation of HW & SW platforms that we can play with/hack with/experiment with

Role of dreams and day dreaming in memory? How can HCI help with that? Eg create opportunities/moments for daydreaming in a day that help with writing recently encountered info to memory

design of interaction/interface for real world use so it is non-disruptive can be used in a real context (usability)

(3) What are interesting ideas for memory augmentation interventions? Use Schacter's 7 sins as the framework to guide the discussion.

VR for trauma/rewriting memories, scent

Figure 6 Miro board response summaries from all groups (Group 1–4 starting from left to right, top to bottom).

2. What remains to be tackled in the field of Human Memory Augmentation (challenges/opportunities)? Many participants highlighted that there remains a lack of HCI systems for "intentional forgetting", temporarily suppressing memories (Group 1) and forgetting traumatic memories (Group 3). Another key area was in altering memories (Group 2), be it situations for re-coining memories via psychotherapy for phobias (Group 1) or perhaps reconstructing memories during sleep to form new associations between concepts and to better understand them (Group 2). Along with these aspects, participants also emphasized the importance of future systems being able to authenticate memories or identify false and altered memories (Group 1).

Another common theme that emerged from the discussions was the need for better "compression" of memory and experience (Group 1). Systems should be able to automatically organize memories (Group 2), and summarize wisdom and lessons learned from the memories (Group 3). Future tools should help users to easily code and describe the situation and information related to the memory (Group 2). Retrieving information at the right time and prospective memory related reminder systems (e.g., calendar, location based) still remain as key issues to address (Group 2).

In terms of designing tools for memory augmentation, there remain opportunities for the creation of hardware and software platforms to experiment with and good translation of tools from research settings to real life use and contexts such that they are non-disruptive and usable (Group 4).

The topic of ethics and privacy in using tools for memory augmentation remains to be addressed. Participants in Group 4 posed open questions on the topic, for example, "Should we create superhumans or focus on those who need it?", "Who [would] "own' our memories?", "What if technologies are used to rewrite our memories and brainwash us?"

3. What are interesting ideas for memory augmentation interventions? Use Schacter's 7 sins of memory as the framework to guide the discussion. The "sin of memory persistence" was discussed the most. Participants thought of ways to protect and get rid of sticky (bad) memories (Group 2). In one scenario, participants talked about remembering old passwords instead of the current one and wondered if there could be a system to help in intentionally forgetting a password (Group 1). A few participants raised the concept of forgetting traumatic experiences in virtual reality (VR) settings (Group 2, Group 4) where the system could enable the user to talk about the experience and then change it. This change or alteration of memories could be based on empowerment ("I can change the past") or diluting ("by consciously engaging with the memory [...] it becomes less emotionally loaded") (Group 2). A related idea was to have a digital "Pensieve" (a fictional item in the Harry Potter book series for exploring memories) to assist users in externalizing and exploring memories in an immersive setting (Group 3) which could also be in VR.

The "sin of absent-mindedness" often refers to the lack of attention when encoding memories. A participant suggested that there could be "[a] wearable that "caresses' the wearer when they are encoding [memories]". The "sin of misattribution" (i.e., having the wrong perception of the memory's source) might be addressed with "quizzes to check your memories – [like a] family pub quiz".

Other ideas that might not be directly related to the 7 sins of memory included adding personalities or affective associations to memories. This could be via smell and body movement, and could be related to the "inner child" concept in trauma therapy (Group 1). The idea of exploring alternative memory pathways with media was also pitched, for example, introducing fake memories or backstories to explore and reflect on different choices as if you are playing a role-playing game (Group 1). Another concept was to create opportunities for daydreaming to help with encoding recent memories (Group 4).

References

1 Brewster, S., McGookin, D., Miller, C. (2006, April). Olfoto: designing a smell-based interaction. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 653-662).

5 Overview of Talks

5.1 Cognitive Augmentation: New sensory experiences and memories need new technology

Michael Beigl (KIT – Karlsruher Institut für Technologie, DE)

Cognitive augmentation is the use of technology to enhance our cognitive abilities, such as to enhance our skills or memory or to improve problem-solving. To fulfill that promise, new technologies need to be developed that are able to augment and enhance these abilities. Creating new sensory experiences and memories go together. In our research, we address cognitive augmentation technologies to create entirely new sensory experiences that then impact the way we remember. For example, we develop technologies that allow users to experience sensations that are not possible because we lack the senses, but which are at the same time pleasant or useful. This has the potential to open new avenues for exploration and discovery, and could then fundamentally change the way we think about sensory experiences and open up new applications. These enhancements to our sensory experiences can also have a profound impact on our cognitive abilities. By expanding our range of sensory input, we can not only improve our ability to experience the physical world around us, but find new ways to learn, remember, and make decisions based on that new information. We conduct research in the area of creating new sensory experiences mainly through haptic displays and investigate new ways of learning new skills and adding new facts to memory through passive haptic learning with applications ranging from psychological treatment to sports, music, and industry. A major focus of our group is the development of (open) hardware and software for sensory and cognitive augmentation, such as OpenEarable, Tactile Interfaces and EdgeML.

As a grain of salt, I find it impossible to prospectively define Grand Challenges. Some Challenges might be:

1. Development of effective, open and easily reproducible technologies (both in hardware and software) to enhance and enhance human senses
2. Overcoming the technical and biological barriers to implementing sensory augmentation. This will probably be mainly joint work with ongoing health activities.
3. Dealing with the social and psychological effects of sensory augmentation: Finally, there are also significant social and psychological challenges associated with sensory augmentation.

5.2 Seamless integration of virtuality as cognitive augmentation

Michael D. Bonfert (Universität Bremen, DE)

While the term metaverse is heavily debated and defined in various ways, the concept starts taking shape in practice. Virtual and online worlds will be incorporated more and more into our physical world, making them ubiquitous. To access and control digital information, HCI needs to explore future interfaces with virtual content that move from device-centric 2D interactions to reality-anchored, pervasive 3D interactions. These interfaces might involve multimodal access and multisensory feedback for integrating virtuality into everyday activities.

As a preliminary step, we need to understand how to design interactions in the purely virtual, as the following examples illustrate. With Get a Grip! we explored mapping physics-based object handling known from reality to dexterous manipulation of virtual objects. The controller Triggermuscle creates a sense of weight for virtual objects. In the case study, Seeing Faces Is So Important, we explored holding meetings on social VR platforms and compared it to videoconferencing during the Covid-19 pandemic. More generally, with the Interaction Fidelity Model, we aim to describe how closely VR interactions reproduce interactions from reality by distinguishing between distinct aspects of fidelity and providing precise terminology.

Conceiving future XR interactions that integrate virtuality into physical reality, we must achieve a careful balance. While digital information is seamlessly incorporated into the real world, it should not distract from reality. While we might want pervasive access to virtuality, we must achieve this for everyone, with high accessibility across modalities. While the virtual content can be personal, it should not be intrusive and must remain secure. The realities should ideally augment each other and the users, rather than compete. These trade-offs will be a challenge for the next decades.

5.3 Augmenting Human Memory

Samantha W.T. Chan (MIT – Cambridge, US)

Memory is a cognitive ability used in our everyday lives. However, it eventually declines and its adaptive nature results in memory troubles. Most of our lapses in memory are due to prospective memory troubles related to forgetting to perform future intended tasks, for example, forgetting to take medication or bring items when leaving the house. With the advancement of technologies and the growth of the aging population, new requirements and opportunities arise. Many technologies available today lack the means to understand our complex everyday situations to support prospective memory. There is a need to provide support through enhancing our memory and assisting with memory tasks, as well as to encourage our receptivity to these interventions. My work aims to investigate the ways to augment prospective memory through technologies that are aware of our cognitive contexts and integrate with our perception and behaviors. I introduce how we can enhance memory through memory training by digitally mediating a memory strategy. The user studies show that this improves users' performance on prospective memory tasks and self-reported memory. Technologies can encourage receptivity to memory training by sensing cognitive contexts through physiological signals (bio-signals) and suggesting training sessions during moments of low emotional arousal and cognitive load (calm moments). I unveil user perceptions of voice reminders which use the voices of friends and family, and discuss how this might benefit user receptivity. Cognitive understanding of the user and cognitive influence can be combined for implicit interactions to augment human memory.

5.4 Group Physiology in Cognitive Augmentation

Jiawen Han (Keio University – Yokohama, JP)

My talk is about how to apply physiological data to augment group interaction and collective cognition. With the advent of wearable sensing, individuals' behavioral and physiological data could be tracked and analyzed almost in real-time. Unlike behavioral data, physiological signals could hardly be observed by the naked eye. But analysis and interpretation afterward could help us reflect and recall past live group events.

I shared some of my works on quantifying a group's physiological data and relating to collective cognition triggered by external stimuli. Moreover, I introduce one of my works to use physiological data as input to predict life experiences and be later shared with others. In that case, group physiology could be not only reflected but also augmented with feedback systems. However, we also need to think about the balance or the trade-off between augmentation and distraction. What kind of system or feedback shall we provide or design to keep the essence of the natural process, with cognitive ability augmented unconsciously?

References

1 Jiawen Han, George Chernyshov, Moe Sugawa, Dingding Zheng, Danny Hynds, Taichi Furukawa, Marcelo Padovani, Kouta Minamizawa, Karola Marky, Jamie A Ward, and Kai Kunze. 2022. Linking Audience Physiology to Choreography. ACM Trans. Comput.-Hum. Interact. Just Accepted (August 2022). https://doi.org/10.1145/3557887

2 Yan He, George Chernyshov, Jiawen Han, Dingding Zheng, Ragnar Thomsen, Danny Hynds, Muyu Liu, Yuehui Yang, Yulan Ju, Yun Suen Pai, Kouta Minamizawa, Kai Kunze, and Jamie A. Ward. 2022. Frisson Waves: Exploring Automatic Detection, Triggering and Sharing of Aesthetic Chills in Music Performances. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 6, 3, Article 118 (September 2022), 23 pages. https://doi.org/10.1145/3550324

5.5 Being embodied in abstract and multiple bodies

Matthias Hoppe (LMU München, DE)

My research is exploring the use of Virtual Reality (and Mixed Reality in general) to go beyond what is possible in the real world. VR allows a user to slip into other bodies in various worlds. The self-perception of one's body can be altered by presenting the user by varying the look of one's body, or change how others are perceived by a simple touch gesture that utilizes social touch [1]. However, VR enable to present the user with a different viewpoint as well. In previous research, we presented the existence of a perspective continuum in VR [1]. Here, the viewpoint of the experience is altered to enable an out-of-body experience for the users, as they have control over a character via motion control. However, the users see themselves from their outside, therefore resulting in an experience where they are embodied in "two bodies", one is the character, the other is the camera. How this experience and embodiment of the two bodies is perceived, highly depends on the users. While some feel like "they are the character" while looking at it from the outside, others have the feeling of only being an observer, or even being embodied in two bodies at the same time. This poses the question of the limits of human embodiment if one person can be present and embodied in multiple bodies at the same time? Can these bodies also take on abstract forms? Can humans be augmented to not only be present in multiple bodies, but also multiple times?

References

1 Hoppe, Matthias, et al. "There Is No First-or Third-Person View in Virtual Reality: Understanding the Perspective Continuum." CHI Conference on Human Factors in Computing Systems. 2022. Hoppe, Matthias, et al. "A human touch: Social touch increases the perceived human-likeness of agents in virtual reality." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020.

5.6 Thinking Social Acceptability alongside with Cognitive Augmentation

Marion Koelle (OFFIS – Oldenburg, DE)

The emergence of cognitive augmentation technologies promises a range of unique benefits: they may improve accessibility and help create equitable opportunities for everybody, while also supporting people in developing, even extending, their natural abilities and skills. Yet, new social dynamics, needs, and challenges may dynamically emerge when a novel technology is "released" into-the-wild. Cognitive augmentation technology may (and likely will) create social tensions or intensify existing individual or societal vulnerabilities. Therefore, it is crucial to research social, societal and ethical implications alongside technology-driven advances of cognitive augmentation.

In my work, I ask how we might operationalize social acceptability and related social and societal phenomena as a core component in HCI research. My past research focused on social acceptability issues with body-worn cameras (see Figure 7) that can serve as assistive devices for the visually impaired or improve indoor navigation (1) but face significant social acceptability issues, e.g., due to concerns about bystander privacy (2). More recently, I expanded my research focus towards the design and fabrication of on-body interfaces, e.g., electrode sleeves (3) or hair interfaces (4), focusing on their perceptual, social and ethical aspects and their wearability in social contexts. With our current and future research, my freshly founded research group and I aim to contribute theories and models that help articulate the role technology plays in social contexts, empirical studies that test and refine these theories, and new tools and methods that help to design interfaces that meet social and societal needs.

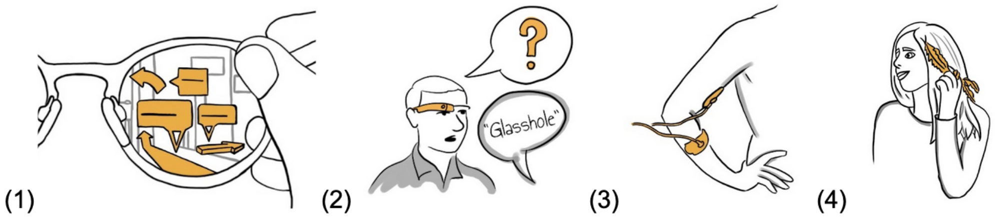


Figure 7 Potential social acceptability issues.

5.7 Cognitive Augemtation using Assistive Devices using Robotics Technology

Yuichi Kurita (Hiroshima University, JP)

I enjoyed my first stay in Dagstuhl to the fullest. The discussions with researchers related to cognitive augmentation were very stimulating and suggested what I should do next.

I develop assistive devices using robotics technology (mainly pneumatic artificial muscles and teleoperated robots). In our field, improving exercise and work performance are important evaluation axes. However, through our discussions at Dagstuhl, we strongly recognized that these affect cognition simultaneously. Conversely, cognitive augmentation will also affect physical activity. We need to determine whether this impact will be positive or negative.

In addition, I felt that social action is also important to ensure that these technologies reach a wider audience on an equal basis. Human augmentation technology should not be a technology for a limited number of people. Let's bring augmentation technology to all people. I feel the important thing is to design the future. We must imagine what we want to be and what we need to do to make it happen.

5.8 Craftsmanship Augmentation: Learn, Master, and Retain Delicate Skills

Jie Li (EPAM Systems – Hoofddorp, NL)

As people are increasingly relying on technology to complete tasks that used to be done by hand. We are losing certain traditional skills such as calligraphy, especially handwriting in Chinese. Character amnesia is a frequent phenomenon that experienced Chinese speakers forget how to write certain characters previously well-known to them. The same goes with other delicate craftsmanship that is a tangible manifestation of the cultural heritage but is on the verge of extinction (e.g., Sichuan embroidery).

In addition to dependency on technology, younger generations find it demanding to learn to write Chinese characters or these delicate skills that often require a lengthy apprenticeship. To ensure that the knowledge and skills are passed onto future generations, cognitive augmentation technology could be used to help people learn and master these skills more efficiently and easily.

By placing sensors on the human body, technology such as brain-computer interfaces (BCIs) or electromyography (EMG) have the potential to translate electrical activities of the brain or the muscles into commands that the computer or device can understand and execute. People can be augmented to control a digital pen or brush using their thoughts or using programmed muscle memories, even if they are unable to physically perform these tasks due to injury or illness. This could help to prevent these skills from being lost due to lack of practice.

Grand challenges regarding craftsmanship augmentation include that:

1. Can we develop augmented craftsmanship that can efficiently help us learn, master and retain a sense of control and precision that is comparable to or even much better than a skilled craftsman?
2. How can we ensure that these skills are still appreciated when we can speed up the learning and mastering process?

5.9 Cognitive augmentation for attunement

Zhuying Li (Southeast University – Nanjing, CN)

My research interests lie in developing cognitive augmentation technology to help people be more attuned to themselves, others, and the environment. To explore this, we have built a series of design works such as HeatCraft, a wearable system that can generate thermal stimuli based on one's body temperature sensed by an ingestible sensor. The stimuli's temperature and body temperature change reversely, so when body temperature raises, the stimuli's temperature drops down. Our study shows that this device could influence how people perceive and use their body. For example, a participant chose to drink some icy water to get the heat in winter. Another example is our recent work in progress GoChirp, an AI-powered wearable device that can continuously sense the existence of surrounding birds via recognizing bird songs, and provide haptic sensations when birds are detected. The device was designed for enhancing people's awareness of wild lives and engaging them with nature in urban lives. In general, we hope the future workaround cognitive augmentation is not only about transhuman and augmenting one's capability, but also considers the experiential perspective. We hope to develop more humanized cognitive augmentation technology to better support attunement.

5.10 Enhancing human capabilities and experiences

Stephan Lukosch (University of Canterbury – Christchurch, NZ)

My research focuses on human augmentation to enhance our capabilities and experiences. I study the effect and impact of human augmentation in different domains such as sports, games, health, safety & security, or engineering. Cognitive augmentation is a form of human augmentation that aims to create novel experiences using and extending human perception and cognition. In previous work, we have explored how a location-based information system in combination with augmented reality impacts work processes, awareness, and workload for hotspot policing [1]. While the location-based information system enhanced human perception and cognition, the timing of information delivery, the adaptation of content to situational characteristics and the modes of information delivery are of interest for further investigation. In more recent work, we have investigated the use of everyday objects to interact in augmented reality [2], how to provide augmented feedback for the correct execution of strength exercises, how to bridge the spatial gap between local and remote players in location-based games, how to provide collaboration awareness during complex co-located collaborative tasks, or how to support the mental imagery practice of elite athletes in preparation for competitions. While all projects are embedded in different domains, all have aimed at enhancing cognitive capabilities and experiences and provided initial insights into a deeper understanding of human and cognitive augmentation. Still, further research is necessary to identify use cases for cognitive augmentation or to determine benefits of cognitive augmentation. Can cognitive augmentation be used to facilitate presence, awareness, decision-making, knowledge recall, skill acquisition, or empathy? What kind of sensors and input do we need to create cognitive augmentation? From an ethical and societal perspective, it is further necessary to explore if we want to create new cognitive capabilities or whether we want to enhance/augment existing ones? What are possible benefits, drawbacks and ethical dilemmas for cognitive augmentation? Should, e.g., others be aware of someone using cognitive augmentation? These questions are just some of those that future work on cognitive augmentation will need to address.

References

1 Engelbrecht, H. & Lukosch, S. G., Dangerous or Desirable: Utilizing Augmented Content for Field Policing, International Journal of Human Computer Interaction (IJHCI), 2020, 36, 1415-1425

2 Greenslade, M.; Clark, A. & Lukosch, S., User-Defined Interaction Using Everyday Objects for Augmented Reality First Person Action Games, IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2022, 842-843

5.11 Cognitive Enhancement

Pattie Maes (MIT – Cambridge, US)

While today’s devices put the world’s information at our fingertips, they do not help us with some of the cognitive skills that are arguably more important to leading a successful and fulfilling life, such as attention, grit, motivation, creativity, memory and learning, mindful decision-making, and emotion regulation. Building upon insights from psychology and neuroscience, my research group at MIT creates AI systems and interfaces for enhancing human cognition. Our designs range from assistive technologies that help people overcome disabilities to technologies that help people further develop their natural abilities and skills.

One of the areas we are particularly interested in lately is opportunities for enhancing cognition during sleep. Good sleep is crucial not just for health, but also daytime cognitive performance. There is an opportunity for HCI researchers to venture into systems and experiences that interface with and influence the sleeping mind. Commercially available devices for sleep such as Oura ring, Fitbit and Smartwatches primarily focus on monitoring and quantifying sleep. Instead, our work looks at how stimuli presented during sleep, including scent, sound and electrical stimulation issued at specific moments of the sleep cycle, can impact memory (strengthening & weakening), time to sleep onset, and depth of sleep.

5.12 Immersive Accessibility

Roshan Lalintha Peiris (Rochester Institute of Technology, US)

I currently co-direct the En-Ability Lab at RIT. The objectives of the En-Ability Lab is to enable, empower and enhance abilities through technology and design. At the En-Ability Lab, one of my main research directions is "Immersive Accessibility" that looks at immersive technologies (AR/VR and Haptics) and accessibility. Under Immersive Accessibility, we conduct research under two main themes, 1) Making Immersive Technologies Accessible and 2) Immersive Technologies for Accessibility Applications. Under theme 1, we primarily identify accessibility limitations of existing immersive technologies and explore methods to make them more accessible. Here, one example project is SoundVizVR where we examine new ways to visualize sounds in VR using mini-maps and indicators to improve the accessibility of VR for Deaf or Hard of Hearing (DHH) users. Under the 2nd theme, we explore new ways to use immersive technologies for accessibility. For example, in the OneButtonPIN research, we examine using haptics to make the PIN code entry process more secure and accessible for Blind and Low Vision participants while the Haptic Captioning project aims to enhance captions and improve speaker indication for DHH caption viewers.

With Cognitive Augmentation, my main goal is to examine ways in which we can re-imagine existing assistive technologies to be more implicitly integrated with users.

5.13 Sensing Visual Attention in Remote Education to Support Learning

Tobias Wagner (Universität Ulm, DE)

Due to the increasing integration of cameras in everyday devices (e.g., smartphones, smartwatches, head-mounted displays) and the continuous progress of hardware and software in eye-tracking, there is a possibility of capturing users' visual attention continuously and in real time. Sensing the human eye allows the assessment of users' cognitive states and processes, providing crucial information for systems that aim to augment users' cognitive abilities. We were interested in using gaze-based information in remote education to enhance meaningful learning, improve communication and collaboration [1, 2]. Here, challenges such as accurately sensing the visual attention of a crowd of distributed users in real-time [3] and the meaningful augmentation of learners' cognitive abilities remain. In the future, we will investigate how educational systems can provide gaze-based feedback to support learners' cognitive processes and abilities during learning and perfecting skills.

References

1 Teresa Hirzle, Marian Sauter, Tobias Wagner, Susanne Hummel, Enrico Rukzio, and Anke Huckauf. 2022. Attention of Many Observers Visualized by Eye Movements. 2022 Symposium on Eye Tracking Research and Applications. https://doi.org/10.1145/3517031.3529235

2 Marian Sauter, Tobias Wagner, and Anke Huckauf. 2022. Distance between gaze and laser pointer predicts performance in video-based e-learning, independent of the presence of an on-screen instructor. 2022 Symposium on Eye Tracking Research and Applications. https://doi.org/10.1145/3517031.3529620

3 Marian Sauter, Teresa Hirzle, Tobias Wagner, Susanne Hummel, Enrico Rukzio, and Anke Huckauf. 2022. Can Eye Movement Synchronicity Predict Test Performance With Unreliably-Sampled Data in an Online Learning Context? 2022 Symposium on Eye Tracking Research and Applications. https://doi.org/10.1145/3517031.3529239

5.14 Diving into Virtual Worlds

Po-Yao (Cosmos) Wang (National Taiwan University – Taipei, TW)

I envision a future where people can dive into game worlds. To make this happen, I previously pursued my Master's degree in Computer Science and Information Engineering at National Taiwan University. To expand game genres, I published Game Illusionization: A workflow for applying optical illusions to video games to tell people how to make games with illusions. People can follow our workflow step by step to create their own illusion games. We offer an illusion database for people to search their desired illusions. And we also provide editors to help people build their illusion games in Unity Game Engine. I hope this work can inspire people to apply illusions to interactions more, and also make the game industry more thriving. Apart from illusions, I also research in Virtual Reality. However, VR headsets are not immersive enough. So, I think maybe dreaming is one of the solutions, where we have full sensations, and we think everything in the dreams are so realistic. Especially, I would like to research in the lucid dream. Lucid dream means we can control our dream content, just like playing video games, but in a more immersive way. Through explorations, I believe one day we could build immersive experiences for gamers.

5.15 Augmenting Liveness

Jamie A. Ward (University of London, GB)

What does it mean to be live and to experience liveness? The experience of live performance, be it theater, dance, music, comedy, or even inspirational talks, is a unique multi-sensorial phenomena that connects people in a particular moment. But the concept of liveness defies clear definition. It might be connected to co-presence in a particular space, yet equally it can transcend space and be experienced across many spaces at once, such as during live broadcast events. It might be the social connection forged in a shared experience, but can equally apply to an individual experiencing some event happening for the first time. This would suggest a link to the uniqueness of the performance, enhancing the feeling that those involved, those who get to experience it, are privy to something special. One intriguing take is that liveness is related to jeopardy, that there is always the possibility that things might go wrong – that, even in the most choreographed performance, no-one knows what is going to happen next. This, for me, is the true spirit of what it means to be live: a shared exploration by audience and performers of where they are headed and what they can discover together.

In my work, I ask how we can measure liveness, and how we can recreate it during situations that are not live – how might we augment our experiences to be more live? To do this I bring together my research in wearable computing, human activity recognition and social neuroscience, and combine this with my experience as a professional actor working in theater. I use the methodology of “theater as a laboratory” to try to uncover the social signals and behaviors that occur during live performance, and uncover what these tell us more broadly about the human condition. My work uses wearable motion sensors, wearable neural hyper-scanning and physiological sensors to study the brains and bodies of audiences and performers during live theater and dance. From a technical perspective, this work aims to solve problems related to multi-person sensing, streaming, and signal processing. From a user-experience perspective, the work aims to solve ways in which such data might be used to enhance performances – to augment our experience in some way. And from a wider perspective, the work aims to uncover what it means to be live.

5.16 Coding Our Cognition

Kai Kunze (Keio University – Yokohama, JP)

I’m interested in how we use computer programming principles to systematically change our cognitive processes. “Coding your cognition” refers to the intentional and systematic process of modifying or improving one’s cognitive abilities, such as memory, attention, and creativity, through computational means. This concept is based on the idea that the mind is not fixed but plastic, which means that it can change and adapt throughout life in response to environmental and experiential factors. If we understand the intrinsic mechanisms and attributes of our mind. We can apply actuation to change them.

We are currently exploring these principles, for early stage research prototypes targeting our visual perception. In a first example, we are simulating a visual impairment experience using optical see-through glasses [1]. Users can experience both losses of central vision and loss of peripheral vision at different levels. We can raise the awareness regarding We also attempt to transfer eyegaze from one person to another [2].

These are just very early works focusing on visual perception. Compared to our cognition, we understand much more about our vision system, as it’s much easier to test and probe. Yet, in the following years, we plan to extend this approach towards higher level cognitive processes.

References

1 Zhang, Q., Barbareschi, G., Huang, Y., Li, J., Pai, Y. S., Ward, J., and Kunze, K. (2022, October). Seeing our Blind Spots: Smart Glasses-based Simulation to Increase Design Students’ Awareness of Visual Impairment. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (pp. 1-14).

2 Zhang, Q., Huang, Y., Chernyshov, G., Li, J., Pai, Y. S., and Kunze, K. (2022, March). GazeSync: Eye movement transfer using an optical eye tracker and monochrome liquid crystal displays. In 27th International Conference on Intelligent User Interfaces (pp. 54-57).

5.17 Cognitive Augmentation through Sensory Illusion

Katrin Wolf (Berliner Hochschule für Technik, DE)

Imagining future computers will "weave themselves into the fabric of everyday life until they are indistinguishable from it" [1] suggests that future pervasive computers are embedded into traditional materials. Aiming at the creation of a rich and good user experience when interacting with pervasive computers, I see two opportunities. We could either embed other technology that creates a haptic sensation, or we use the concept of sensory illusion that augments our perception without the need for additional technology. Sensory illusion can extend the design space traditional materials properties are offering, e.g., can create the illusion that any surface provides a rich bandwidth of touch feedback when being pressed like a button, such as softness, wetness, and bendability.

In previous work on sensory illusion, I and colleagues applied knowledge from cognitive science to human-computer interaction (a) to better understand the sensory involvement when using certain interface modalities and (b) to explore the possibility of sensory illusion to compensate for shortcomings of interface technologies to extend the user experience. We investigated the tactile and non-visual perception of surface textures, aiming for defining texture patterns that serve as haptic cues for haptic surface texture perception [4]. In another work, we found that sonic cues bias the haptic surface perception by strengthening the haptic perception of the texture pattern if a sonic cue is given in parallel (Wolf and Bennett 2013b). We furthermore investigated how proprioception can support or replace visual feedback on pointing gesture execution [5]. Moreover, we explored the effect of visual distortion in combination with electrotactile feedback on surface illusions [3]. In this investigation, it could be shown that surface illusion can be induced through visual distortion of texture projected on a flat surface as well as through electrotactile stimuli provided when a surface is touched. Finally, I proposed concrete interaction design ideas for illusion-based pervasive interfaces considering any surface to possibly be an interface believing that windows, furniture, and even the floor we walk on will be an interface one day [2].

References

1 Weiser, M. (1991). The Computer for the 21st Century. Scientific american, 265(3), 94-105.

2 Wolf, K. (2017). Augmenting Interface Perception through Sensory Illusion. In Proceedings of the CHI 2017 Workshop on Amplification and Augmentation of Human Perception, May 07, 2017, Denver, CO, USA

3 Wolf, K., & Bäder, T. (2015, April). Illusion of surface changes induced by tactile and visual touch feedback. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (pp. 1355-1360). ACM.

4 Wolf, K., & Bennett, P. D. (2013, April). Haptic cues: texture as a guide for non-visual tangible interaction. In CHI'13 Extended Abstracts on Human Factors in Computing Systems (pp. 1599-1604). ACM.

5 Wolf, K., Müller-Tomfelde, C., Cheng, K., & Wechsung, I. (2012, May). Does proprioception guide back-of-device pointing as well as vision?. In CHI'12 Extended Abstracts on Human Factors in Computing Systems (pp. 1739-1744). ACM.

5.18 Cognitive augmentation is bodily

Florian 'Floyd' Mueller (Exertion Games Lab, Monash University – Melbourne, Australia)

Cognitive augmentation is concerned with the augmentation of cognition, but due to the intertwinedness of the human mind and the body, we argue that augmentation is always bodily. We demonstrate this through a series of research design works around augmented cycling experiences, augmented social cohabitating experiences, and augmented emotion amplification experiences. The results of these works suggest interesting ways forward for augmentation research, in particular how the design of augmentation can highlight experiential aspects, facilitating playful experiences. Ultimately, with our work, we want to enhance our knowledge around the design of augmentation to help people understand who they are, who they want to become, and how to get there.
`http://exertiongameslab.org`

6 Acknowledgements

This work is partially supported by the Cybernetic Being project under JST Moonshot R&D Program (Grant number JPMJMS2013) as well as by the Deutsche Forschungsgemeinschaft (DFG) under the SPP 2199 project Illusionary Surface Interfaces (number 425869442). Florian 'Floyd' Mueller thanks the Australian Research Council, especially DP190102068, DP200102612 and LP210200656. Kai Kunze wants to thank Mai Yamagami and Renate Kunze for their continuous support.

Participants

- Michael Beigl
KIT – Karlsruher Institut für Technologie, DE
- Michael D. Bonfert
Universität Bremen, DE
- Samantha W.T. Chan
MIT – Cambridge, US
- Jiawen Han
Keio University – Yokohama, JP
- Matthias Hoppe
LMU München, DE
- Masahiko Inami
University of Tokyo, JP
- Shoya Ishimaru
TU Kaiserslautern, DE
- Shunichi Kasahara
Sony CSL – Tokyo, JP
- Marion Koelle
OFFIS – Oldenburg, DE
- Thomas Kosch
HU Berlin, DE
- Kai Kunze
Keio University – Yokohama, JP
- Yuichi Kurita
Hiroshima University, JP
- Jie Li
EPAM Systems – Hoofddorp, NL
- Stephan Lukosch
University of Canterbury – Christchurch, NZ
- Paul Lukowicz
DFKI – Kaiserslautern, DE
- Kouta Minamizawa
Keio University – Yokohama, JP
- Qianqian Mu
Aarhus University, DK
- Pat Pataranutaporn
MIT – Cambridge, US
- Rakesh Patibanda
Monash University – Clayton, AU
- Roshan Lalintha Peiris
Rochester Institute of Technology, US
- Enrico Rukzio
Universität Ulm, DE
- Albrecht Schmidt
LMU München, DE
- Valentin Schwind
Frankfurt University of Applied Sciences, DE
- Paul Strohmeier
Max-Planck-Institut für Informatik Saarbrücken, DE
- Steeven Villa
LMU München, DE
- Tobias Wagner
Universität Ulm, DE
- Jamie A. Ward
University of London, GB
- Don Anusha Withanage
The University of Sydney, AU
- Katrin Wolf
Berliner Hochschule für Technik, DE

Remote Participants

- Cindy Hsin-Liu Kao
Cornell University – Ithaca, US
- Zhuying Li
Southeast University – Nanjing, CN
- Pattie Maes
MIT – Cambridge, US
- Florian ‘Floyd’ Mueller
Monash University – Clayton, AU
- Suranga Nanayakkara
National University of Singapore, SG
- Evangelos Niforatos
TU Delft, NL
- Nathalie Overdevest
Monash University – Clayton, AU
- Bektur Ryskeldiev
Mercari – Tokyo, JP
- Aryan Saini
Monash University – Clayton, AU
- Stel Stelarc
Bentley, AU
- Po-Yao (Cosmos) Wang
National Taiwan University – Taipei, TW

Report from Dagstuhl Seminar 22492

Formal Methods and Distributed Computing: Stronger Together

Hagit Attiya*[1], **Constantin Enea***[2], **Sergio Rajsbaum***[3], and **Ana Sokolova***[4]

1 **Technion – Haifa, IL.** hagit@cs.technion.ac.il
2 **Ecole Polytechnique – Palaiseau, FR & CNRS – Palaiseau, FR.** cenea@lix.polytechnique.fr
3 **National Autonomous University of Mexico, MX, on leave at IRIF, Paris, FR.** rajsbaum@im.unam.mx
4 **University of Salzburg, AT.** anas@cs.uni-salzburg.at

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 22492 "Formal Methods and Distributed Computing: Stronger Together", held in December 2022.

Seminar December 4–9, 2022 – http://www.dagstuhl.de/22492

2012 ACM Subject Classification Theory of computation → Distributed algorithms; Theory of computation → Verification by model checking; Software and its engineering → Formal methods

Keywords and phrases automated verification and reasoning, concurrent data structures and transactions, distributed algorithms, large-scale replication

Digital Object Identifier 10.4230/DagRep.12.12.27

1 Executive Summary

Hagit Attiya (Technion – Haifa, IL)
Constantin Enea (Ecole Polytechnique – Palaiseau, FR & CNRS – Palaiseau, FR)
Sergio Rajsbaum (National Autonomous University of Mexico, MX, on leave at IRIF, Paris, FR)
Ana Sokolova (University of Salzburg, AT)

Distributed applications represent nowadays a significant part of our everyday life. To mention just a few examples, our personal data are stored on remote distributed servers, data management relies on remote applications reachable via smartphones or tablets, data-intensive computations are performed on computer clusters, etc. Since distributed applications are increasingly deployed at large scale, they have to be reliable and robust, satisfying stringent correctness criteria. This is the point where a strong interaction of formal methods and of distributed computing becomes a necessity.

The goal of this Dagstuhl Seminar was to achieve a synergy by bringing together researchers working on applying formal methods for concurrent programs and distributed systems, and researchers from distributed computing. Both communities have a deep understanding of distributed computation, but from two different perspectives. Historically, these communities

* Editor / Organizer

have common roots, but since more than two decades they evolved independently. The resulting gap slows down progress in both fields, and limits the applicability of the results obtained in each field, as each one develops its own techniques separately. The seminar addressed several topics that bridge the two research fields, and that have high potential to stimulate the development of the other area:

Concurrent data structures and transactions: Modern multicore architectures enable large performance boosts by executing a number of threads in parallel, which however, poses considerable challenges in maintaining correctness of shared data structures and thread synchronization. These challenges have been addressed using various paradigms like lock-free programming or transactional memory. However, turning these concepts into efficient programming support remains a big challenge, and formal methods may offer new ideas in this direction.

Formal approaches to large-scale replication: Current computing systems are increasingly large-scale distributed systems, for example, distributed databases, distributed ledgers (Blockchains) and key-value stores. At the heart of these systems are fundamental trade-offs between data consistency, availability, and the ability to tolerate failures. A formal approach to studying these issues will provide a common ground for the design, verification, analysis, implementation and use of these systems.

Distributed algorithms for verification: Reasoning about concurrent/distributed software is notoriously difficult due to the inherent non-determinism in its semantics. The different processes in a concurrent program can interleave in many different ways which leads to an enormous number of possible executions. Algorithmic methods are necessary to mitigate the difficulty of reasoning about this huge space of executions, and scalable distributed algorithms may be the answer for the future. These methods can manifest in various forms, e.g., automated testing, deductive verification, model checking, and have led to important results in many timely contexts. Performing verification in a distributed fashion is a particularly promising new direction of research.

The impact of all the areas above on a rigorous development of distributed applications was enhanced by fostering direct interactions between researchers from (automated) formal methods and from distributed computing.

2 Table of Contents

Executive Summary

Experience Reports

Panel discussions

Talk Abstracts

3 Experience Reports

3.1 Seminar Overview

Jennifer Welch (Texas A&M University - College Station, US)

The seminar brought together researchers in distributed computing and in formal verification. It was a successful combination, with the right amount of overlap for appreciating the novelty and understanding the main points of the other community. Participants took part in lively conversations in the Q&A of the presentations, at the coffee breaks, meals, and after dinner. There was an enjoyable excursion to the bustling Christmas market in the nearby picturesque town of St. Wendel.

Monday, Dec. 5

The morning was devoted to two tutorials, to help the participants from the two areas get an overview of the other area.

The first tutorial, by **Victor Vafeiadis**, was entitled "Verification of Concurrent Data Structures." In addition to objects with sequential semantics, he also considered synchronization constructs, which are often bottlenecks in programs and thus are heavily optimized in practice. Victor started by pointing out there there several axes along which to consider the verification problem, in order to determine if the outcome is worth the effort. These include

- At what level of abstraction, between binary code and pseudocode, is the program to be verified?
- What properties are to be verified? These could include memory safety, incomplete functional correctness, linearizability, and liveness. Different properties may need different methods.
- What assumptions should be made about the memory model, memory management, and fairness. It is important that the memory model choice be correct with respect to the level at which the program is being verified.
- How thorough will the verification be? For machine-checked proofs, will all the metatheory be mechanized, or will some properties be axiomatized? Will pragmatic simplifying assumptions be made? Can quantities such as buffer sizes and number of threads be bounded?
- How much human expertise in the approach and tools will be needed? How much information will the user need to supply (e.g., specification, proof outine, invariants)?
- How much human and machine effort (time) will be needed?

Several different verification approaches were outlined. These were simulation of automata and program logic, both of which require a high level of expertise and take weeks; static analysis on a simplified implementation, requiring little-to-medium level of expertise and either takes minutes or fails; model checking of real code which requires little human expertise or time but requires an unpredictable amount of machine time which can be large; and randomized testing, which works on real code, requires little human expertise or effort and a tunable amount of machine effort, but gives no guarantees. State-of-the-art industrial practice is to use a combination of techniques. Barriers to adoption include the human cost and lack of expertise for interactive proofs, and the amount of time and lack of a "progress bar" for model checking.

The rest of the tutorial focused on his research on making model checking practical. To make it faster, one can employ state-space reduction, better algorithms, and reliance on user hints and common code patterns. To estimate the time that will be taken, one can estimate the size of the state space with randomized testing, and then use a "predictable" fast modeling checking algorithm whose time is proportional to the size of the state space. One such approach is his algorithm, TruST, "Truly Stateless C". In order to avoid redundant exploration of the state space and to limit the amount of memory needed, TruST explores alternate executions eagerly, represents executions with graphs, and uses maximal extensions. He went into more detail about the algorithm with a nice small running example.

Pierre Fraignaud gave a two-part tutorial on Distributed Certification.

The first part of the presentation assumed a fixed static network in which one would like to provide some kind of fault-tolerance, such as detecting illegal states, checking correctness of the output of a subroutine, or preventing the launch of a protocol if its prerequisites (such as having a spanning tree) are not satisfied. This problem is formalized as a graph in which each node has as input an id and a label, and we would like to know if the graph satisfies a predicate. For the distributed certification version of the problem, imagine there is a *prover*, a centralized, computationally unlimited, non-trusted orace that assigns certificates to nodes, and a *verifier*, a one-round, synchronous, failure-free distributed algorithm in which nodes exchange certificates with their neighbors to check whether the certificates form a proof that the system satisfies the predicate. Correctness properties are completeness (if the predicate is satisfied, then the verifier accepts at all nodes) and soundness (if the predicate is not satisfied, then the verifier rejects at at least one node). Similar notions in the literature include proof labeling schemes (PLS), locally checkable proofs, and nondeterministic local decisions. The main complexity measure of interest is the size of the certificates. Fortunately, every Turing-decidable predicate has a PLS using certificates of size $O(n^2 + kn)$ bits, where k is the size of the labels. Unfortunately, there exist predicates that require $\tilde{\Omega}(n^2)$ bit certifications, including the properties of not being 3-colorable and having a nontrivial automorphism. A useful primitive that can be certified with $O(\log n)$-bit certificates is a spanning tree: let the certificate be the pair (root id, distance to root). A number of other problems can be certified using $O(\text{polylog}(n))$-bit certificates, including minimum spanning tree, approximate optimization problems, and planar graph. He then presented some more recent results that can be viewed as meta-theorems.

The second part of the tutorial considered analogous ideas in the asynchronous wait-free model of computation. One theorem is that binary consensus is not *checkable* in this model, which can be proved much more easily than showing it is not *computable.* In fact, consensus is checkable with three values (accept, reject, inconclusive). Extensions to k-set agremeent and distributed runtime monitoring were discussed. A general result is that every input-output decidable predicate in an n-process system can be certified with certificates of size at most $\lceil Ack^{-1}(n) + 1 \rceil$ bits using 3-valued decisions; in contrast, $O(1)$-bit certificates do not suffice.

In summary, for the synchronous failure-free model, the theory of distributed certification is mature and evolving rapidly. In contrast, in the asynchronous crash-prone model, there are only partial results that need consolidating. Discussion with the audience during the Q&A indicated there might be some issue with the model in the latter case.

In the afternoon, the "regular" (30-minute) talks began.

Bernd Finkbeiner spoke about hyperproperties in synthesis and verification. A *hyperproperty* is a set of traces; examples include non-interference, secrecy, differential privacy, symmetry, fault detectability over a channel, robustness, partial observation, and dominance. In spite of much work on hyperproperties, there is a gap between general specification logics and practical verification and synthesis algorithms. Bernd presented some research aimed at closing this gap. For example, HyperLT, which is linear-time temporal logic (LTL) plus prenex quantification over traces, is expensive to model-check because of the quantifiers. If we restrict attention to some specific graphs on the Kripke structure, then model-checking can be faster. To model-check HyperLTL, one can reduce the problem to emptiness of Büchi word automata. A recent approach is to apply ideas from game theory; for example, it can be shown that if the existential player has a winning strategy, then the system satisfies the hyperproperty.

The next talk was by **Parosh Aziz Abdula** on checking liveness properties under weak consistency, using TSO as an example. The definition of sequential consistency (SC) is that writes are immediately visible in the memory and reads are from memory; SC is simple and intuitive but expensive to implement. A weaker, and more efficiently-implementable, condition is total-store order (TSO), in which writes are non-atomic while reads are local or from memory. The definition of TSO uses the notion of a store buffer at each process, and data in the store buffer is nondeterministically transferred to memory. There are many existing works on verifying safety properties for TSO but not much on liveness properties. Parosh presented a way to take techniques for proving properties about SC, especially liveness, and extend them to TSO. Assuming some probabilistic fairness of the underlying system, such as that messages eventually leave the store buffers and each message updates memory with the same probability, one can prove that updates occur frequently and buffers tend to be small. Now we can use tools for SC and reduce to reachability analysis.

The rest of the afternoon was devoted to the industry panel, described later.

Tuesday, Dec. 6

The tutorials continued in the morning.

The first tutorial was by **Ahmed Bouajjani** on verification of distributed systems. He considered the situation where clients interact with an application layer, which runs on top of a data storage system, which in turn runs on top of a communication network. The tutorial focused on problems that arise when verifying such systems, especially related to consistency conditions, including isolation levels for transactions. He then presented (1) a formal framework for specifying the conditions, (2) ways to verify an application running under a weak consistency model, and (3) ways to verify that a storage system provides a certain consistency level.

To specify a condition, we define the expected observable behaviors when interacting with memory. The values returned by reads depend on the current set of "visible" actions by each process and the order in which actions are seen by each process. In *strong consistency*, updates are visible without delay and in the same order. If the memory is replicated, a weaker consistency holds, since participants can see different sets of updates and in different orders. Causal consistency, sequential consistency, and snapshot isolation were given as examples.

Challenges in verifying reachability under weak consistency include complex program semantics and reordering of operations, which requires unbounded memory to track dependencies. In many situations, the problem is undecidable; when it is decidable, it can be expensive. "Robustness" is checking whether the semantics of a program is the same under two different consistency models. He surveyed a line of work in this direction, with different papers considering different consistency conditions and different data structures.

In conclusion, decidability and complexity are still open in many cases. We need efficient static analysis algorithms and testing methods. We also need languages for describing different consistency models. And finally, we need to be able to reason systematically about tradeoffs between consistency and performance. During the Q&A session, someone raised the question about analyzing systems that have combinations of different consistency levels. Apparently, there is no formal work in this area.

The last tutorial, by **Petr Kuznetsov**, was on correctness in distributed computing. He focused on correctness conditions that relate the behavior of a concurrent system to that of a sequential system, especially for implementing shared objects in the presence of crash failures. One way of classifying properties is into "safety" (informally, nothing bad ever happens) and "liveness" (informally, eventually something good happens). He compared and contrasted how Lynch formally defined these properties and how Alpern and Schneider did, and their approaches to proving that every property is the intersection of a safety property and a liveness property. He also went over a few techniques for proving safety and liveness properties, including the interesting fact by Padon et al. that liveness can sometimes be reduced to safety. Next he defined the safety properties "linearizability" [1] and "sequential consistency" and discussed some of their pros and cons. A variety of progress properties (from a table in Herlihy and Shavit) were presented. Then he introduced hyperproperties, which are sets of traces.. *There are no formal definitions of safety and liveness for these; such definitions might need a topology on sets of traces.* Then he talked about a particular hyperproperty, "strong linearizability", which is a strengthening of linearizability that preserves probability distributions of the enclosing program.

One way of implementing wait-free linearizable objects is to rely on consensus, which can provide a total order on concurrent, or contending, operations. But this can be expensive or even impossible. An alternate approach to resolve contending operations is with the notion of a lattice, where overlapping operations can return any join of previous and overlapping operations. This definition is suitable for CRDTs (conflict-free replicated data types). Then he talked about using quorum systems.

In conclusion, the philosopher's stone of distributed computing would be a universal machine, allowing any sequential program to be run in a distributed environment with all the good properties. But because of challenges in dealing with real systems, there are numerous lower bounds and impossibility results. Maybe the way around the challenges is to consider weaker properties or friendlier models. One issue raised during the Q&A, is that sequential consistency is not very useful for some data structures that are inherently concurrent, for example, a producer-consumer queue.

The regular 30-minute talkes continued in the afternoon.

Jennifer Welch's talk was on implementing shared objects in the presence of continual churn. She presented a model for crash-prone dynamic systems that allows a limited number of processes to enter and leave during time intervals of a fixed length. The model was inspired by mobile ad hoc networks and may have relevance for permissionless blockchains. Algorithms for implementing a linearizable register and a non-linearizable store-collect object were presented, which make progress even while churn is ongoing. There is a lower bound on the crash resilience, showing that the larger the rate of churn, the smaller must be the fraction of faulty processes.

[1] A later clarification was given: Lynch proved in 1996 that linearizability is a safety property for finite nondeterminism, otherwise, as shown by Guerraoui and Ruppert in 2014, it is not.

Ori Lahav spoke about abstraction for crash-resilient objects. There has been recent interest in *persistent objects*, concurrent objects that can recover from crashes, which are implemented in non-volatile memory (NVM). A natural question is how to define correctness as some variation of linearizability, especially in a way that can be used in verification. His approach is to focus on refinement with respect to another implementation. The new aspect is to include special constructs in the language that allow for intuitive specifications. He focused on the simplest model, persistent sequential consistency and explained how it can be mapped to x86-TSO by adding appropriate fence instructions.

Burcu Kulahcioglu Ozkan's talk was on testing blockchain consensus algorithms, with a focus on consortium/voting-based Byzantine fault-tolerant (BFT) consensus algorithms, inspired by PBFT. The correctness of the consensus algorithm is crucial for the correctness of the overall blockchain. However, many bugs have been found in BFT algorithms, in part because there is a lack of practical testing tools. For instance, software model checking is infeasible due to the possibility of Byzantine faults. Random testing is the current practice for concurrency and network faults. Burcu's talk showed how to apply this idea to blockchain consensus algorithms. First, she explained the challenges to be overcome to make this idea practical. For enumerating executions, she proposed using abstractions of time with protocol rounds. For network partition faults, a small set of random partitions can provide full coverage with high probability. Byzantine faults are modeled by structure-aware and small-scope mutations to protocol messages. This provides equivocation, amnesia, double-voting, losing internal state, and incorrect forward progress. Using this approach, she found a new bug in the Ripple XRP Ledger, namely a violation of termination.

Annette Bieniusa spoke about highly available access control in distributed systems. Assuming the system has POSIX access control policies that specify the types of authorization and the types of access allowed, we need to avoid conflicts without causing internal data leakage and corruption. Her recent work proved the CIA theorem (Confidentiality, Integrity, and Accessibility), analogous to the CAP theorem, showing that one cannot achieve all three properties if there are partitions. In response, she considered weakening the security properties to have a causal order, instead of total order, on access control policies and data. Her work formalized the approach in the modeling system Repliss and tested for counter-examples, which found some corner cases. Her ongoing work is considering decentralized systems with multiple administrators using the idea of forking into parallel universes. She concluded her talk with a discussion of lessons learned.

Gregory Chockler's presentation was about a synchronizer primitive for building correct algorithms under partial synchrony, in the presence of Byzantine failures. He reviewed the motivation for partial synchrony as well as the challenges for designing bug-free consensus algorithms for blockchains. He proposed an abstraction for leader-driven consensus, based in a sequence of views, that cleanly separates safety and liveness properties. He described a case study for proving that PBFT is live using this abstraction. The structure of this liveness proof can be reused for proofs of other protocols, as it establishes properties similar to failure detectors (for crash failures). They are completeness (if a correct process never executes a received command, then it eventually advances to a new view) and eventual accuracy (eventually, in a view of a correct leader and with large enough timeouts, no correct process will advance to a new view). He then explained how to implement the synchronizer using ideas from Bracha's broadcast algorithm, modified to work in bounded space.

Alexey Gotsman presented his work on getting strong consistency and availability under network partitions. He first pointed out that the CAP theorem does not preclude availability in the majority side of a partition. His work looked more deeply into the kind of communication failures that can occur and considers partially synchronous systems with some unidirectional channels that can fail by dropping messages whenever they choose. New results are that consensus (state machine replication) is solvable if and only if at most a minority of processes crash and there is a majority of correct processes that are strongly connected via correct channels. Then he presented a specification and implementation of a synchronizer for this model (cf. immediately previous talk by Gregory Chockler). His last topic was finding the minimal amount of connectivity needed to guarantee availability anywhere at all. For this, he first showed that places where the system is available must be connected, and that to be available more than half the processes must be connected.

Wednesday, Dec. 7

Rupak Majumdar's talk was on random testing with theoretical guarantees. He first pointed out that despite the existence of formal approaches, practitioners usually test their code and it's effective. To explain why, he gave a metaphor using ninjas at a banquet for testing distributed systems where events are partially ordered and we need a schedule for tests. It turns out that many bugs in production software involve just a few events, say d, and thus finding them only requires a d-hitting family of schedules. This concept is related to the "order dimension" of a poset. When the system is running, we can learn an "upgrowing" poset in an on-line fashion. His random testing algorithm PCTCP maintains an on-line chain partition, assigns a random priority to each chain, and executes enabled events from the highest priority chain. At certain random points in the execution, the priority of a chain is decreased. Follow-on work, the Trace Aware HitMC algorithm, exploits knowledge of the algorithm using the notion of communication-closed rounds. A general open question in this area is providing a theoretical explanation for situations when randomized testing behaves well.

Yoram Moses spoke on graph connectivity in distributed algorithms. A new distributed problem called *card consensus*, which cannot be solved under certain circumstances, was used as a running example. A general theorem was presented showing that specifying constraints on actions induces requirements on the knowledge that the processes must have. Using this theorem, he provided an *epistemic* analysis of card consensus, and also discussed the impact of requiring events to occur simultaneously. In general, knowledge can be modeled as a graph, and is the dual of *indistinguishability*, a concept with applications to numerous classical results in distributed computing. Epistemic reasoning about card consensus is the one-dimensional analog of *topological* reasoning for the set consensus and renaming problems. Both epistemic and topological reasoning are useful in analyzing distributed computing. Challenges for future research include finding a clean variant of common knowledge that captures the consensus problem in asynchronous systems, finding topological models that can handle in a natural way timing-related constraints (such as actions that must occur close together), and adapting topological models to handle loss of information with global state.

Serdar Tasiran's presentation was on formal methods for distributed systems at Amazon Simple Storage Services (S3). He is in the automated reasoning group, which has about 20 people, while the whole S3 team has about 1000 developers. Their mandate is to integrate formal methods into software processes across S3. They use a range of methods, starting with lightweight to more powerful. Research is needed on methodology and tooling to enable this transition. Testing is widespread and integrates well into the software processes. Different tools need to be connected and assurance needs to be quantified. They are adapting a "model-first" approach, in which models and code are in the same repositories, and increasingly the models are being written before the code on new projects. An important problem is to ensure that executions are consistent with models, so for instance, while you are watching Netflix, an automaton is being run! S3's new ShardStore storage node pays special attention to crash consistency, has a complex implementation, and is deployed weekly. Model checking and protocol-level deductive proofs are done in TLA+ style, but using Dafny. However, they are not mechanically connected.

Stephan Merz talked about tool support for TLA+. TLA+ is a specification language that has become a set of tools: TLC, Apalache, and TLAPS. Help is also provided by PlusCal and IDES. He used a distributed termination detection algorithm by Safra as a running example. First, he showed how to use TLA to specify the problem. Then one can start expressing correctness properties, including safety and liveness. TLC is an explicit state model checker that works for finite state descriptions, so the number of nodes and maximum number of pending messages need to be fixed. Apalache is a symbolic model checker based on SMT which checks properties of finite prefixes of an execution. Until recently, it only checked safety properties. It relies on constraint solving, not state enumeration. Again, the number of nodes needs to be fixed but it can handle infinitely many messages. The time grows exponentially with the length of the prefix and the number of nodes. However, if you have an inductive invariant, then you only need to check traces of size 0 or 1! TLAPS is a proof assistant; the proof effort here is independent of the size of the instance, although the user has to guide the verification. It uses automatic back-end provers to discharge proof obligations. All the tools share the same input language.

Giuliano Losa's presentation was on formal verification of a classic distributed algorithm using inductive invariants. His example algorithm is a termination detection algorithm of Kumar. He showed a new proof with a simple inductive invariant. He showed how to apply TLA+, Apalache, and Isabelle/HOL. Interestingly, the informal proof of correctness presented is actually wrong! The error is that the algorithm doesn't make sure that all processes are visited. After that problem in the code is fixed, another invariant is proposed. He played around with the tools to help him come up with the invariant, which is fairly simple. Then he did an automatic proof in Isabelle. One lesson from the talk is the advantage of human-machine collaboration to find invariants.

Thursday, Dec. 8

Philipp Woelfel talked about predictable building blocks for randomized shared memory algorithms. He started with a brief history of such algorithms, including the definition of the strong adaptive adversary. Replacing atomic (instantaneous) objects with linearizable objects gives this adversary more power in randomized algorithms. He defined *strong linearizability* (SL), which is sufficient, and necessary, for overcoming this problem. It also has the nice property of being composable. There are general SL constructions using locks or using universal constructions with, for example, compare-and-swap objects. He then summarized more efficient constructions and various positive and negative results. He has work in progress on constructing a strongly linearizable LL/SC object from CAS objects; the motivation is that CAS is available in hardware, but LL/SC is not although it is very useful in designing algorithms. The problem is that a CAS operation may succeed even though an ABA violation occurs, while LL/SC fixes this. One can easily get a strongly linearizable LL/SC from CAS with unbounded tags. The challenge is to bound the tags. He sketched his ideas for bounding the tags. An open question is to find additional efficient strongly linearizable implementations from strong synchronization primitives. Another is how to deal with the oblivious adversary. During the Q&A, the issue of finding complexity lower bounds for SL was raised as well.

Maurice Herlihy's presentation was on correctness conditions for cross-chain transactions. A distributed ledger is an abstraction that can be used for financial transactions and other applications where everyone agrees on the content and is tamper-proof. In the future, there will be many different chains and we want inter-operability. In the past, the classical adversary was considered for solving consensus. But the modern adversary corresponds to real people/governments and makes more sophisticated attacks and brings more complicated problems. Thus we need to rethink correctness. The classical ACID properties for transactions don't work for blockchain models. Instead of a fixed fraction of participants that can behave maliciously while the rest follow the protocol, now we have the notion of "deviating" parties with no bound on the number of deviators. He went through the ACID properties and spoke in more detail about how they need to be adapted for blockchains. (A) Atomicity is impossible; instead, we could use the conditions of liveness (if all parties conform, then all transfers happen) and safety (if some parties deviate, then the conforming parties are "no worse off"). Cross-chain commerce is a cooperative game, where a protocol is a strategy for that game. Parties can form coalitions and the result is the payoff. This area needs formalization! (C) Consistency says that application-specific constraints are respected. This could be restated by saying that conforming to the protocol should be a strong Nash equilibrium in the game. (I) Isolation states that no transaction sees another's intermediate steps. But we don't even want this condition in multiparty swaps, as we may want to use escrow accounts and similar mechanisms. (D) Durability says theat committed transactions survive crashes (a legacy of 20th century technology). In the blockchain world, we want to avoid "censorship" by government, corporations, hackers, etc. In summary, adversarial commerce is here to stay, regardless of specific blockchains, and we need to rethink notions of correctness.

Thomas Wies's talk was on reasoning principles for verifying concurrent search data structures, such as sets and maps. Modularity in proofs used to help us. But now it is the wrong level of abstraction. The challenge is that concurrency and memory safety proofs are intertwined. He gave examples of the problem with a B-link tree algorithm and a hash table algorithm. The proposed solution is to separate the two by finding a common link-technique proof hidden in the specific proofs, abstract it as a template, verify it once, and then use it in different data structure implementations. There are four issues to be dealt with. First, the template must be data-structure-independent. This is done using the concept of edge sets. Second, local operations might have global effects. This is dealt with using keyset resource algebra. Third, global properties must be rendered local. This is done using the inset concept and ideas of flows. Finally, in some implementations (those that are not strongly linearizable), the linearization points depend on the future, or said another way, the linearization points are determined in hindsight. The ability to do this has been added into a separation logic in his current work, supported by a tool called plankton that works well for proof automation. In summary, some modular techniques for reasoning about concurrent search data structures were presented which allow better proof automation.

Azalea Raad spoke about extending Intel-x86 consistency and persistency by formalizing the semantics of Intel-x86 memory types and non-temporal stores. Non-temporal stores write directly to memory, bypassing the cache which avoids cache pollution; they are heavily used in application-level code. One can declare the "type" (cache-ability) of memory with several different options, with the default being write-back (WB), which is used in system-level code. The extended Intel-x86 consistency semantics provide more options, as she explained with some small code examples. It turns out that the behavior of actual programs, especially those that use a mixture of different kinds of accesses, is not consistent with what is written in the manual. The actual, validated, behavior is summarized in an extensive table. The next issue to be considered is persistence. When using non-volatile memory, execution can lag persistence. This behavior has been included in the consistency semantics. She tried to test for post-crash behavior by directly monitoring the memory bus using an expensive piece of hardware, but the results were inconclusive.

Armando Castañeda's presentation covered asynchronous distributed run-time verification and enforcement of linearizability. Efficient linearizable implementations are often complicated and bug-prone, as they use fine-grained synchronization. Formal verification of linearizability is sometimes undecidable or NP-hard. He presented a complementary dynamic approach which is to monitor running executions, ideally in a wait-free manner in order be less intrusive. The key differentiator from most previous work on runtime verification is the emphasis on being wait-free, that is, crash-tolerant in asynchronous systems. He first described his way of modeling the problem, as an interaction between a verifier and an implementation, both of which are concurrent programs. If the (simulated) execution of the implementation satisfies the property of interest (e.g., linearizability), then no verifier process reports an error, otherwise some verifier process reports an error and provides a witness. At first glance, there is a simple proof that linearizability is impossible to verify at run-time for some object. However, looking more closely, one can define a "stretched" version of the original implementation through which the linearizability of the original implementation can be tested indirectly. By-products of this work are simple methodologies to derive fault-tolerant distributed monitors and self-enforced linearizable implementations.

Nathalie Bertrand spoke about a CEGAR approach to parameterized verification of distributed algorithms, where CEGAR stands for counter-example-guided abstraction refinement. Ben-Or's randomized consensus algorithm for Byzantine failures was used as a running example throughout the talk, with randomization replaced by nondeterminism. It is described using three parameters, the number of processes n, the maximum number of faulty processes t, and the actual number of faulty processes f. She presented formal semantics for the algorithm and pointed out that there are two dimensions of infinity. Tools to overcome this difficulty include message abstraction and counting abstraction. Layered threshold automata (LTA) are used for counting abstraction. To get around the problem that parameterized model check of LTAs is undecidable, even just for safety properties, she used predicate abstraction via a guard automaton and CEGAR. The approach has been implemented in a tool PyLTA written in Python and that has been benchmarked with some promising results as well as some inconclusive cases. Future work includes formalizing model extraction from pseudocode and extending the approach to randomized algorithms in order to show, for example, almost-sure temrination of Ben-Or's algorithm.

Rotem Oshman's presentation was on truthful information dissemination in asynchronous networks. She was inspired to choose this topic by Maurice Herlihy's presentation. Suppose the network is composed of *rational* players who try to accomplish a goal but have their own incentives. The algorithm should be "incentive-compatible", meaning that following the algorithm is an equilibrium, with no coalitions: if all players but one follow the algorithm, then that player has no incentive to deviate. Consider the problem of getting all players to output every player's input, where the network graph is known, called the information dissemination problem. Much prior work has focused on "fair" solutions (if many solutions are possible, then choose one uniformly at random). Her work focuses on "good" solutions. First she explained how to set up the information dissemination problem as a Bayesian game and mentioned that her work shows that the graph must be (at least) 2-connected. The main part of the talk was a description of an algorithm for solving the problem in a ring, in which players 0 and 1 learn each other's inputs and commit to a random string, then players learn the other inputs using the random string, and finally all the inputs are revealed and any cheaters are caught. She also outlined the methodology for proving the correctness and mentioned that the algorithm has optimal bit complexity. Future work includes handling coalitions and Byzantine players, as well as relaxing some of the technical assumptions.

Pierre Fraignaud presented a wait-free speedup theorem, which can be used for proving lower bounds. Suppose we could show that there exists a function F such that for every nonnegative t and every problem Π, Π has complexity t if and only if $F(\Pi)$ has complexity $t-1$. Such a theorem would imply that Π has complexity t if and only if $F^{(t)}(\Pi)$ has complexity 0. Then if we can show that the transformed problem $F^{(t)}(\Pi)$ cannot be solved with complexity 0, that would imply that the original problem Π cannot be solved with complexity t. The inspiration includes Linial's round lower bound for 3-coloring in a ring as well as more recent results for the anonymous LOCAL model and for maximal matchings and maximal independent set. The general questions are which models admit speedup theorems and which problems admit speedup theorems for a specific model. Using diagrams to provide intuition, he presented a result for the iterated immediate snapshot model, in which asynchronous crash-prone processes communicate through levels of shared registers that provide operations to write and obtain atomic snapshots. The generic transformation F is instantiated with a specific function called the *closure*, which has a slightly larger set of outputs and requires solving a local task in one round. The (weak) speedup theorem obtained is that if Π is solvable in t rounds, then the closure of Π is solvable in $t-1$ rounds. One application is that the closure of the consensus problem is just the consensus problem, and thus the easily-shown fact that consensus is not solvable in 0 rounds implies that consensus is not solvable at all. Another application is that the closure of ϵ-agreement (approximate agreement where decisions must be within ϵ) is (2ϵ)-agreement, which implies a lower bound of $\lceil \log_2 \frac{1}{\epsilon} \rceil$ for ϵ-agreement. The theorem extends to the use of test-and-set and binary consensus objects. However, the closure of set-agreement is trivial, as it can be solved in 0 rounds (because the theorem is not in both directions). Open questions include identifying which tasks have non-trivial closures, is there an if-and-only-if speedup theorem for asynchronous wait-free computing, and which models allow for the design of useful speedup theorems (e.g., what about t-resilient models).

Sandeep Kulkarni's talk was on using informal methods for distributed computing, in particular, the "war" of consistency violations and self-stabilization. A classic model of distributed computing lets each node see its neighbors' states to decide on its next action, assuming interleaving semantics. Local mutual exclusion provides concurrency in a large system with many processes and small neighborhoods, but it has significant overhead. He want to explore what happens if we don't use local mutual exclusion, and instead allow the resulting consistency violations to occur. In particular, a node can see neighbor states that are stale to different degrees. He then related this behavior to self-stabilization, a form of fault-tolerance that requires a program to eventually reach, and remain in, a set of "legal" states, no matter what state it begins in. The intuition is that if the rate of consistency violations is very high, then the program may not converge; if the rate is low, then the program will (probabilistically) converge, perhaps requiring more steps but without the need for local mutual exclusion or other synchronization mechanisms. He described several case studies (Dijkstra's 3-state token ring, graph coloring, and maximal matching) and made several observations. One is that the benefit of a program transition has exponential distribution. Another is that the cost of a consistency violation transition has exponential distribution. Prior work related to consistency violations includes that on eventually consistent shared memory and that on lattice linearity. In conclusion, systems that tolerate consistency violations have the potential to benefit from the currency; although arbitrary programs may not tolerate such violations, self-stabilizing programs automatically do. Finally the performance of stabilizing programs can be predicted analytically in the presence of consistency violations.

Hagit Attiya spoke about approximately preserving hyper-properties without strong linearizability. When programming with (atomic abstract) objects that are implemented from other objects, a popular consistency condition for the implementation is linearizability, as it preserves trace properties of the program. However, it does not preserve hyper-properties, which include probability distributions. She gave an example of a toy program that terminates with probability at least one-half when it uses atomic registers but that can be made to never terminate if the registers are implemented with the well-known ABD algorithm. Strong linearizability (SL), in which linearization points are prefix-preserved, does preserve probability distributions. She then showed how SL is an example of *strong refinement*, which preserves hyper-properties and is equivalent to the existence of a forward simulation. Unfortunately, numerous objects do not have SL implementations. She then presented a solution to this problem for object implementations that are "tail strongly linearizable" and have "effect-free preambles"; informally this means that there is a prefix of the implementation that does not impact concurrent operations and after which its linearization point is chosen in a prefix-preserving way. The preamble is repeated some number of times and then one of the iterations is chosen at random to be used in the tail. This technique can be applied to several well-known object implementations. She presented a formula showing that the probability of a bad outcome with the transformed object approaches that with atomic objects as the number of preamble iterations increases. Regarding the extra cost incurred, she gave an example of applying the transformation to a specific atomic snapshot implementation for n processes resulting in a wait-free SL implementation taking $O(n^3)$ steps per process, and compared it to a previously known non-wait-free SL implementation that takes $O(n^3)$ steps per process. During the Q&A, the question was asked whether there exist implementations or even objects that cannot be improved in this way.

Friday, Dec. 9

A wrap-up session was held in the morning, during which suggestions for improvement, in case there is a follow-up seminar, were solicited from attendees. The main recommendations were to include more junior researchers, especially PhD students, to have some tools tutorials (e.g., on TLA+), and to put out a call to the distributed computing community requesting open challenges for verification.

4 Panel discussions

4.1 Panel: Applications in Industry

Giuliano Losa (Stellar Development Foundation – San Francisco, US)

The panelists started by each giving a 10 minute talk giving an overview of the use of formal methods at their respective employers. A lively discussion on areas of improvements and missing tools ensued, with emphasis on what academics can do to help.

The panelists were Serdar Tasiran (Principal Scientist, Amazon S3), Akash Lal (Senior Principal Researcher, Microsoft Research), Manuel Bravo (Informal Systems), and Giuliano Losa (Stellar Development Foundation). Cezara Dragoi (Principal Applied Scientist, Amazon) and Bernhard Kragl (Applied Scientist, Amazon) also joined the conversation virtually.

The audience learned that formal methods have made their way in the software development process in many teams at Amazon, Microsoft, and Informal Systems, while the Stellar Development Foundation is investing is more targeted use-cases such as the development of distributed algorithms and smart contracts. Languages and tools such as TLA+, P, Coyote, CBMC, and Ivy are helping software engineers at those companies develop better systems. Moreover, there is heavy investment in new tooling.

The panel identified compositional verification of large systems as an important area in which tooling, best practices, and even scientific experiments identifying the important problems are missing. Academics from the audience highlighted the difficulty of working on large-scale code bases (or even having access to them) in an academic setting where engineering resources are often scarce.

Finally, the panel discussed training software engineers in formal methods. It transpired that typical undergraduate software-engineering classes do not address formal methods, and that the material currently available to train engineers on the job may be insufficient or not accessible to most software engineers because they do not have the prerequisite knowledge.

5 Talk Abstracts

5.1 (Approximately) Preserving Hyper-Properties without Strong Linearizability

Hagit Attiya (Technion – Haifa, IL)

Joint work of Hagit Attiya, Constantin Enea, and Jennifer L. Welch

Atomic shared objects, whose operations take place instantaneously, are a powerful abstraction for designing complex concurrent programs. Since they are not always available, they are typically substituted with software implementations. A prominent condition relating these implementations to their atomic specifications is linearizability, which preserves safety properties of the programs using them. However linearizability does not preserve hyper-properties, which include probabilistic guarantees of randomized programs: an adversary can greatly amplify the probability of a bad outcome, such as nontermination, by manipulating the order of events inside the implementations of the operations. This unwelcome behavior prevents modular reasoning, which is the key benefit provided by the use of linearizable object implementations. A more restrictive property, strong linearizability, does preserve hyper-properties but it is impossible to achieve in many situations.

This paper suggests a novel approach to blunting the adversary's additional power that works even in cases where strong linearizability is not achievable. We show that a wide class of linearizable implementations, including well-known ones for registers and snapshots, can be modified to approach the probabilistic guarantees of randomized programs when using atomic objects. The technical approach is to transform the algorithm of each operation of an existing linearizable implementation by repeating a carefully chosen prefix of the operation several times and then randomly choosing which repetition to use subsequently. We prove that the probability of a bad outcome decreases with the number of repetitions, approaching the probability attained when using atomic objects. The class of implementations to which our transformation applies includes the ABD implementation of a shared register using message-passing, the Afek et al. implementation of an atomic snapshot using single-writer registers, the Vitanyi and Awerbuch implementation of a multi-writer register using single-writer registers, and the Israeli and Li implementation of a multi-reader register using single-reader registers, all of which are widely used in asynchronous crash-prone systems.

5.2 Verification of Liveness Properties on Weakly Consistent Platforms (TSO as an example)

Parosh Aziz Abdulla (Uppsala University, SE)

We present Probabilistic Total Store Ordering (PTSO) – a probabilistic extension of the classical TSO semantics. For a given (finite-state) program, the operational semantics of PTSO induces an infinite-state Markov chain. We resolve the inherent non-determinism due to process schedulings and memory updates according to given probability distributions. We provide comprehensive results showing the decidability of several properties for PTSO. (i) Almost-Sure (Repeated) Reachability: whether a run, starting from a given initial configuration, almost surely visits (resp. almost surely repeatedly visits) a given set of target configurations. (ii) Almost-Never (Repeated) Reachability: whether a run from the initial configuration almost never visits (resp. almost never repeatedly visits) the target. (iii) Approximate Quantitative (Repeated) Reachability: to approximate, up to an arbitrary degree of precision, the measure of runs that start from the initial configuration and (repeatedly) visit the target. (iv) Expected Average Cost: To approximate the expected average run cost from the initial configuration to the target up to an arbitrary degree of precision.

We derive our results through a nontrivial combination of results from the classical theory of (infinite-state) Markov chains, the theories of decisive and eager Markov chains, specific techniques from combinatorics, as well as, decidability and complexity results for the classical (non-probabilistic) TSO semantics. As far as we know, this is the first work considering probabilistic verification of programs running on weak memory models.

5.3 A CEGAR approach to parameterized verification of distributed algorithms

Nathalie Bertrand (INRIA – Rennes, FR)

Joint work of Nathalie Bertrand, Ocan Sankur, Bastien Thomas, Josef Widder

Distributed algorithms are central to many domains such as scientific computing, telecommunications and the blockchain. Even when they aim at performing simple tasks, their behaviour is hard to analyze, due to the presence of faults (crashes, message losses, etc.) and to the asynchrony between the processes. Parameterized verification techniques have been developed to prove correctness of distributed algorithms independently of actual setup, i.e. the number of processes and the potential failures.

In this talk, we present a CEGAR approach to checking safety and liveness properties for fault tolerant distributed algorithms that use threshold conditions, typically on the number of received messages of a given type.

5.4 Highly-available access control in distributed systems

Annette Bieniusa (TU Kaiserslautern, DE)

Highly available distributed systems rely on replication for partition- and fault-tolerance. This results in weaker consistency guarantees for shared data and introduces challenges for the correctness of the application under (temporary) data inconsistencies. In particular regarding application security, it is difficult to determine which inconsistencies can be tolerated and which might lead to security breaches. In this talk, we will introduce state-of-the-art approach to enforcing dynamic access control policies in highly-available systems. As use case, we discuss the interplay of security and consistency in distributed file systems and provide an impossibility result that indicates that confidentiality, integrity and accessibility cannot be achieved together in a highly-available partition-tolerant setting. We further discuss a CRDT-based model, implementing the traditional POSIX access control policy, that guarantees confidentiality and integrity while precluding accessibility only in rare situations while reflecting the users' intention.

5.5 Verification of Distributed Systems

Ahmed Bouajjani (Université Paris Cité, FR)

The talk presents explains the problems to address for the verification of distributed systems, in particular problems related to consistency and isolation levels. We present formal framework for specifying consistency models and isolation levels (in the transactional case). Then, we describe existing approaches and results concerning (1) the verification of application running over weak consistency environments, and (2) the verification that system implementation conform to consistency/isolation levels.

5.6 Asynchronous Distributed Runtime Verification and Enforcement of Linearizability

Armando Castaneda (National Autonomous University of Mexico, MX)

This work studies the problem of distributed runtime verification of linearizability for asynchronous concurrent implementations. It proposes an interactive model for distributed runtime verification and shows that it is impossible to runtime verify this correctness condition for some common sequential objects such as queues, stacks, sets, priority queues, counters and the consensus problem. The impossibility captures informal arguments used in the past that argue distributed runtime verification is impossible. Then, it argues that linearizability can be indirectly verified through a class of implementations. Namely, it shows that (1) linearizability of a class of concurrent implementations can be distributed runtime verified and (2) every implementation can be easily transformed to its counterpart that belongs to the class. From these algorithms, it is easy to derive distributed monitors, as well as concurrent implementations that self-enforce linearizability, namely, these implementations produces outputs that are guaranteed to be linearizable. The same results hold for extensions of linearizability such as set-linearizability and interval-linearizability.

5.7 Synchronizer: a recipe for building correct algorithms under partial synchrony

Gregory Chockler (University of Surrey – Guildford, GB)

Joint work of Manuel Bravo, Gregory V. Chockler, Alexey Gotsman, Alejandro Pastoriza
Main reference Manuel Bravo, Gregory V. Chockler, Alexey Gotsman: "Liveness and Latency of Byzantine State-Machine Replication", in Proc. of the 36th International Symposium on Distributed Computing, DISC 2022, October 25-27, 2022, Augusta, Georgia, USA, LIPIcs, Vol. 246, pp. 12:1–12:19, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022.
URL https://doi.org/10.4230/LIPIcs.DISC.2022.12

Byzantine state-machine replication (SMR) ensures the consistency of replicated state in the presence of malicious replicas and lies at the heart of the modern blockchain technology. Byzantine SMR protocols often guarantee safety under all circumstances and liveness only under synchrony. However, guaranteeing liveness even under this assumption is nontrivial. So far we have lacked systematic ways of incorporating liveness mechanisms into Byzantine SMR protocols, which often led to subtle bugs. To close this gap, we introduce a modular framework to facilitate the design of provably live and efficient Byzantine SMR protocols. Our framework relies on a view abstraction generated by a special SMR synchronizer primitive to drive the agreement on command ordering. We present a simple formal specification of an SMR synchronizer and its bounded-space implementation under partial synchrony. We also apply our specification to prove liveness and analyze the latency of three Byzantine SMR protocols via a uniform methodology. In particular, one of these results yields what we believe is the first rigorous liveness proof for the algorithmic core of the seminal PBFT protocol.

5.8 Hyperproperties in Synthesis and Verification

Bernd Finkbeiner (CISPA – Saarbrücken, DE)

Traditionally, most verification efforts have focused on the satisfaction of trace properties, such as that an assertion is satisfied at a particular program location or that the computation terminates eventually. Many policies from information-flow security, like observational determinism or noninterference, and many other system properties including promptness and knowledge can, however, not be expressed as trace properties, because these properties are hyperproperties, i.e., they relate multiple execution traces. In this talk, I will give an overview on logics for the specification of hyperproperties and on algorithms and tools for verification and synthesis.

5.9 Wait-Free Speedup Theorem

Pierre Fraigniaud (Université de Paris, FR & CNRS, Paris, FR)

Joint work of Pierre Fraigniaud, Ami Paz, Sergio Rajsbaum
Main reference Pierre Fraigniaud, Ami Paz, Sergio Rajsbaum: "A Speedup Theorem for Asynchronous Computation with Applications to Consensus and Approximate Agreement", in Proc. of the PODC '22: ACM Symposium on Principles of Distributed Computing, Salerno, Italy, July 25 – 29, 2022, pp. 460–470, ACM, 2022.
URL https://doi.org/10.1145/3519270.3538422

We study two fundamental problems of distributed computing, consensus and approximate agreement, through a novel approach for proving lower bounds and impossibility results, that we call the asynchronous speedup theorem. For a given n-process task Pi and a given computational model M, we define a new task, called the closure of Pi with respect to M. The asynchronous speedup theorem states that if a task Pi is solvable in $t > 0$ rounds in M, then its closure w.r.t. M is solvable in $t - 1$ rounds in M. We prove this theorem for iterated models, as long as the model allows solo executions. We illustrate the power of our asynchronous speedup theorem by providing a new proof of the wait-free impossibility of consensus using read/write registers, and a new proof of the wait-free impossibility of solving consensus using registers and test&set objects for $n > 2$. The proof is merely by showing that, in each case, the closure of consensus (w.r.t. the corresponding model) is consensus itself. Our main application is the study of the power of additional objects, namely test&set and binary consensus, for wait-free solving approximate agreement faster. By analyzing the closure of approximate agreement w.r.t. each of the two models, we show that while these objects are more powerful than read/write registers from the computability perspective, they are not more powerful as far as helping solving approximate agreement faster is concerned.

5.10 Correctness Conditions for Cross-chain Transactions

Maurice Herlihy (Brown University – Providence, US)

Joint work of Maurice Herlihy, Barbara Liskov, Liuba Shrira

Modern distributed data management systems face a new challenge: how can autonomous, mutually-distrusting parties cooperate safely and effectively? Addressing this challenge brings up questions familiar from classical distributed systems: how to combine multiple steps into a single atomic action, how to recover from failures, and how to synchronize concurrent access to data. Nevertheless, each of these issues requires rethinking when participants are autonomous and potentially adversarial.

5.11 Testing Blockchain Consensus Algorithms

Burcu Kulahcioglu Ozkan (TU Delft, NL)

Byzantine fault-tolerant algorithms promise agreement on a correct value, even if a subset of processes can deviate from the algorithm arbitrarily. While these algorithms provide strong guarantees in theory, in practice, protocol bugs and implementation mistakes may still cause them to go wrong. This talk introduces a simple yet effective method and our recent experience for automatically finding errors in implementations of Byzantine fault-tolerant algorithms through randomized testing. Our approach navigates the space of possible process faults by limiting process faults in an execution to a bounded number of round-based, structure-preserving, small-scope mutations to the protocol messages.

5.12 War of Consistency Violations and Self-Stabilization

Sandeep Kulkarni (Michigan State University – East Lansing, US)

Joint work of Sandeep Kulkarni, Duong Nguyen, Arya Gupta

Consider distributed programs that are designed in a strong memory model (e.g., one that allows a node to read its neighbor atomically). When such program is implemented without synchronization requirements such as local mutual exclusion, they can result in consistency violations (cvfs). Cvfs can cause the program to violate the specification. However, when combined with self-stabilization, we get a war between cvf and stabilization; the fore pushes the program away from the invariant whereas the latter tries to get closer.

We observe the following for various self-stabilizing programs (1) permitting cvfs provides a better performance, (2) benefit of program transition and cost of cvfs follow an exponential distribution, (3) Cvf cost distribution is independent of the number of processes, and (4) Cost of cvf can be computed by sampling the state space.

5.13 Correctness in Distributed Computing

Petr Kuznetsov (Telecom Paris, FR)

We take a walk through the space of correctness criteria that relate concurrent behavior of a distributed system to a sequential specification. In particular, we focus on safety and liveness, linearizability, sequential consistency, progress conditions (from deadlock-freedom to obstruction-freedom). We then discuss the cost of achieving correctness and ways to achieve this kind of correctness, and then highlight several ways to cur the costs by relaxing correctness reqauirements.

5.14 Abstraction for Crash-Resilient Objects

Ori Lahav (Tel Aviv University, IL)

Joint work of Artem Khyzha, Ori Lahav

We study abstraction for crash-resilient concurrent objects using non-volatile memory (NVM). We develop a library-correctness criterion that is sound for ensuring contextual refinement in this setting, thus allowing clients to reason about library behaviors in terms of their abstract specifications, and library developers to verify their implementations against the specifications abstracting away from particular client programs. As a semantic foundation we employ a recent NVM model, called Persistent Sequential Consistency, and extend its language and operational semantics with useful specification constructs. The proposed correctness criterion accounts for NVM-related interactions between client and library code due to explicit persist instructions, and for calling policies enforced by libraries. We illustrate our approach on two implementations and specifications of simple persistent objects with different prototypical durability guarantees. Our results provide the first approach to formal compositional reasoning under NVM.

5.15 Random Testing with Theoretical Guarantees

Rupak Majumdar (MPI-SWS – Kaiserslautern, DE)

We describe an algorithm called PCTCP for random testing of distributed systems. We show that the properties of the algorithm can be understood by looking at online dimension of the underlying partial order of events, which is related to online chain partitioning. We define d-hitting families as a way to organize behaviors. We show that the PCTCP algorithm gives a guarantee on the minimal probability of sampling behaviors from d-hitting families for a fixed d.

5.16 Tool Support for TLA+: TLC, Apalache, and TLAPS

Stephan Merz (INRIA Nancy – Grand Est, FR)

Joint work of Igor Konnov, Markus Kuppe, Stephan Merz
Main reference Igor Konnov, Markus Kuppe, Stephan Merz: "Specification and Verification with the TLA+ Trifecta: TLC, Apalache, and TLAPS", CoRR, Vol. abs/2211.07216, 2022.
URL https://doi.org/10.48550/arXiv.2211.07216

Using an algorithm due to Safra for distributed termination detection as a running example, we present the main tools for verifying specifications written in TLA+. Examining their complementary strengths and weaknesses, we suggest a workflow that supports different types of analysis and that can be adapted to the desired degree of confidence.

Our TLA+ specifications and proofs are available at `https://github.com/tlaplus/Examples/tree/ISoLA2022/specifications/ewd998`.

5.17 On Graph Connectivity in Distributed Algorithms

Yoram Moses (Technion – Haifa, IL)

Knowledge and Topological reasoning capture aspects of local and global information that are essential in a distributed setting. This talk will present a simple problem whose analysis uses graph connectivity in an essential way, and leverage this to discuss how the connection between common knowledge and connectivity has been exploited in the literature. Finally, we present some thoughts regarding the interaction between epistemic analysis and more topological analysis of distributed algorithms and consider future challenges.

5.18 Truthful information Dissemination in Asynchronous Networks

Rotem Oshman (Tel Aviv University, IL)

Joint work of Lior Solodkin, Rotem Oshman
Main reference Lior Solodkin, Rotem Oshman: "Truthful Information Dissemination in General Asynchronous Networks", in Proc. of the 35th International Symposium on Distributed Computing, DISC 2021, October 4-8, 2021, Freiburg, Germany (Virtual Conference), LIPIcs, Vol. 209, pp. 37:1–37:19, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.
URL https://doi.org/10.4230/LIPIcs.DISC.2021.37

We give a protocol for information dissemination in asynchronous networks of rational players, where each player may have its own desires and preferences as to the outcome of the protocol, and players may deviate from the protocol if doing so achieves their goals. We show that under minimalistic assumptions, it is possible to solve the information dissemination problem in a truthful manner, such that no participant has an incentive to deviate from the protocol we design. Our protocol works in any asynchronous network, provided the network graph is at least 2-connected. We complement the protocol with two impossibility results, showing that 2-connectivity is necessary, and also that our protocol achieves optimal bit complexity. As an application, we show that truthful information dissemination can be used to implement a certain class of communication equilibria, which are equilibria that are typically reached by interacting with a trusted third party. Recent work has shown that communication equilibria can be implemented in synchronous networks, or in asynchronous, complete networks; we show that in some useful cases, our protocol yields a lightweight mechanism for implementing communication equilibria in any 2-connected asynchronous network.

5.19 Extending Intel-x86 Consistency and Persistency

Azalea Raad (Imperial College London, GB)

Main reference Azalea Raad, Luc Maranget, Viktor Vafeiadis: "Extending Intel-x86 consistency and persistency: formalising the semantics of Intel-x86 memory types and non-temporal stores", Proc. ACM Program. Lang., Vol. 6(POPL), pp. 1–31, 2022.
URL https://doi.org/10.1145/3498683

Existing semantic formalisations of the Intel-x86 architecture cover only a small fragment of its available features that are relevant for the consistency semantics of multi-threaded programs as well as the persistency semantics of programs interfacing with non-volatile memory.

We extend these formalisations to cover: (1) non-temporal writes, which provide higher performance and are used to ensure that updates are flushed to memory; (2) reads and writes to other Intel-x86 memory types, namely uncacheable, write-combined, and write-through; as well as (3) the interaction between these features. We develop our formal model in both operational and declarative styles, and prove that the two characterisations are equivalent. We have empirically validated our formalisation of the consistency semantics of these additional features and their subtle interactions by extensive testing on different Intel-x86 implementations. Our work on validating the persistency semantics is ongoing.

5.20 Formal Methods for Distributed Systems at Amazon Simple Storage Service (S3)

Serdar Tasiran (Amazon Web Services – New York City, US)

Amazon's Simple Storage Service (S3) is increasingly adopting a "model first" approach, with formal models being first-class artifacts in the software development process. In this approach, we start by model checking or deductively proving that design models of distributed systems provide key properties such as durability or strong consistency. Then, during integration testing, gamma stages, or even in production, we monitor that the implementation code conforms to the design models. This not only detects/prevents implementation errors, but also forces models and implementations to remain in sync, ensuring that the investments in writing and analyzing models continue to pay off. In this talk, I will present several examples of the model-first approach in S3, and challenges of using formal methods at S3 scale.

5.21 Implementing Shared Objects in the Presence of Continual Churn

Jennifer L. Welch (Texas A&M University – College Station, US)

Main reference Hagit Attiya, Sweta Kumari, Archit Somani, Jennifer L. Welch: "Store-collect in the presence of continuous churn with application to snapshots and lattice agreement", Inf. Comput., Vol. 285(Part), p. 104869, 2022.
URL https://doi.org/10.1016/j.ic.2022.104869

I will present a model for dynamic distributed systems that permits a bounded amount of ongoing churn as well as node crashes. In this model I will present an algorithm for simulating a linearizable shared read-write register and a non-linearizable shared store-collect object, as well as a lower bound on the crash-resilience that is possible. The talk is based on the papers.

References

1 Attiya, Chung, Ellen, Kumar, and Welch, "Emulating a Shared Register in a System that Never Stops Changing", IEEE Transactions on Parallel and Distributed Systems, Vol. 30, Issue 3, March 2019 (DOI: 10.1109/TPDS.2018.2867479)

2 Attiya, Kumar, Somani, and Welch, "Store-collect in the presence of continuous churn with application to snapshots and lattice agreement," Information and Computation, Vol. 285, Part B, May 2022 (doi.org/10.1016/j.ic.2022.104869)

5.22 Reasoning Principles for Verifying Concurrent Search Structures

Thomas Wies (New York University, US)

Joint work of Thomas Wies, Siddharth Krishna, Nisarg Patel, Dennis Shasha, Roland Meyer, Sebastian Wolff

Search structures support the fundamental data storage primitives on key-value pairs: insert a pair, delete by key, search by key, and update the value associated with a key. Concurrent search structures are parallel algorithms to speed access to search structures on multicore and distributed servers. These sophisticated algorithms perform fine-grained synchronization between threads, making them notoriously difficult to design correctly. In this talk, I will present a framework for obtaining correctness proofs for concurrent search structures that are modular, reusable, and amenable to automation. The framework takes advantage of recent advances in local reasoning techniques based on concurrent separation logic. I will provide an overview of these techniques and demonstrate there use for verifying realistic search structures such as concurrent B-link trees.

5.23 Predictable Building Blocks for Randomized Shared Memory Algorithms

Philipp Woelfel (University of Calgary, CA)

In this talk I will discuss why linearizability is not a sufficient condition for building blocks used in randomized shared memory algorithms. I will present strong linearizability (introduced by Golab, Higham and Woelfel, 2011), which can be used to preserve the power of the adaptive adversary when replacing atomic operations with implemented ones in randomized algorithms. I will then outline ideas for implementing a wait-free strongly linearizable load-linked/store conditional object from compare-and-swap.

Participants

- Hagit Attiya
Technion – Haifa, IL
- Parosh Aziz Abdulla
Uppsala University, SE
- Nathalie Bertrand
INRIA – Rennes, FR
- Raven Beutner
CISPA – Saarbrücken, DE
- Annette Bieniusa
TU Kaiserslautern, DE
- Ahmed Bouajjani
Université Paris Cité, FR
- Manuel Bravo
Informal Systems – Madrid, ES
- Armando Castaneda
National Autonomous University of Mexico, MX
- Gregory Chockler
University of Surrey – Guildford, GB
- Constantin Enea
Ecole Polytechnique – Palaiseau, FR & CNRS – Palaiseau, FR
- Bernd Finkbeiner
CISPA – Saarbrücken, DE
- Pierre Fraigniaud
Université de Paris, FR & CNRS, Paris, FR
- Alexey Gotsman
IMDEA Software Institute – Madrid, ES
- Maurice Herlihy
Brown University – Providence, US
- Burcu Kulahcioglu Ozkan
TU Delft, NL
- Sandeep Kulkarni
Michigan State University – East Lansing, US
- Petr Kuznetsov
Telecom Paris, FR
- Ori Lahav
Tel Aviv University, IL
- Giuliano Losa
Stellar Development Foundation – San Francisco, US
- Rupak Majumdar
MPI-SWS – Kaiserslautern, DE
- Stephan Merz
INRIA Nancy – Grand Est, FR
- Yoram Moses
Technion – Haifa, IL
- Rotem Oshman
Tel Aviv University, IL
- Azalea Raad
Imperial College London, GB
- Sergio Rajsbaum
National Autonomous University of Mexico, MX, on leave at IRIF, Paris, FR
- Ana Sokolova
University of Salzburg, AT
- Serdar Tasiran
Amazon Web Services – New York City, US
- Viktor Vafeiadis
MPI-SWS – Kaiserslautern, DE
- Jennifer L. Welch
Texas A&M University – College Station, US
- Thomas Wies
New York University, US
- Philipp Woelfel
University of Calgary, CA

Report from Dagstuhl Seminar 22512

Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users II

Ana Aguiar^{*1}, **Onur Altintas**^{*2}, **Falko Dressler**^{*3}, **Gunnar Karlsson**^{*4}, **and Florian Klingler**$^{\dagger 5}$

1 **Universidade do Porto, PT.** anaa@fe.up.pt
2 **Toyota Motor North America – Mountain View, US.** onur@us.toyota-itc.com
3 **TU Berlin, DE.** dressler@ccs-labs.org
4 **KTH Royal Institute of Technology – Stockholm, SE.** gk@kth.se
5 **Universität Paderborn, DE.** klingler@ccs-labs.org

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 21262 "Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users II". Looking back at the last decade, one can observe enormous progress in the domain of vehicular networking. In this growing community, many ongoing activities focus on the design of communication protocols to support safety applications, intelligent navigation, and many others. We shifted the focus from basic networking principles to open challenges in edge computing support and, as a novel aspect, on how to integrate so called vulnerable road users (VRU) into the picture.

Seminar December 18–21, 2022 – http://www.dagstuhl.de/22512

2012 ACM Subject Classification Computing methodologies → Machine learning; Human-centered computing; Networks → Cyber-physical networks; Networks → Network protocols

Keywords and phrases 5G/6G, bicyclists, cooperative driving, edge computing, intelligent transportation systems, pedestrians, tactile internet, V2X, vehicle-to-vehicle communication, vulnerable road users

Digital Object Identifier 10.4230/DagRep.12.12.54

1 Executive Summary

Falko Dressler (TU Berlin, DE)
Ana Aguiar (Universidade do Porto, PT)
Onur Altintas (Toyota Motor North America – Mountain View, US)
Gunnar Karlsson (KTH Royal Institute of Technology – Stockholm, SE)

Looking back at the last decade, one can observe enormous progress in the domain of vehicular networking. In this growing community, many ongoing activities focus on the design of communication protocols to support safety applications, intelligent navigation, and many others. Using the terms Vehicular Ad-hoc Networks (VANETs), Inter-Vehicle Communication (IVC), Car-2-X (C2X), or Vehicle-2-X (V2X), many applications – as interesting as challenging – have been envisioned and (at least) partially realized. Very large projects have been initiated to validate the theoretic work in field tests and protocols are being standardized. With the increasing interest from industry, security and privacy have also become crucial aspects in

[*] Editor / Organizer
[†] Editorial Assistant / Collector

Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users II, *Dagstuhl Reports*, Vol. 12, Issue 12, pp. 54–73
Editors: Ana Aguiar, Onur Altintas, Falko Dressler, Gunnar Karlsson, and Florian Klingler
DAGSTUHL REPORTS Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the stage of protocol design in order to support a smooth and carefully planned roll-out. We are now entering an era that might change the game in road traffic management. Many car makers already supply their recent brands with cellular and WiFi modems, some also adding vehicular WLAN (DSRC, ITS-G5) and C-V2X technologies.

With this latest installment of the "Inter-Vehicular Communication" Dagstuhl Seminar series, we intend to shift the focus from basic networking principles to open challenges in edge computing support and, as a novel aspect, on how to integrate so called vulnerable road users (VRU) into the picture. Edge computing is currently becoming one of the core building blocks of cellular networks, including 5G, and it is necessary to study how to integrate ICT components of moving systems. The trade-offs of computation distribution, system aspects, and the impact on end-to-end latency are still unanswered. Also, vehicular networking and cooperative driving focus almost exclusively on cars but leave out communication and coordination with, for example, pedestrians and bicyclists. And, many of the existing communication solutions for this scenario were designed without having battery constraints in mind. In the meantime, some early research has been initiated on this topic and initial projects report very interesting results on safety features for VRUs. Building upon the great success of the previous Dagstuhl Seminars – as documented, e.g., with results published in widely visible magazine articles [1, 2, 3, 4] – with this follow-up seminar, we aim to again bring together experts from all these fields from both academia and industry.

Seminars in this series focused on general vehicular communication technologies, security and safety impact, cooperative driving concepts and its implications on communication protocol design, and many more. Building upon the online-only seminar in 2021, we now shifted the focus of this seminar from basic networking principles to open challenges in edge computing support and, as a novel aspect, on how to integrate so called vulnerable road users (VRU) into the picture. Edge computing is currently becoming one of the core building blocks of cellular networks, including 5G/6G, and it is necessary to study how to integrate ICT components of moving systems. The trade-offs of computation distribution, system aspects, and the impact on end-to-end latency are still unanswered. Also, vehicular networking and cooperative driving focuses almost exclusively on cars but leaves out communication and coordination with, for ex-ample, pedestrians and bicyclists. For example, many of the existing communication solutions for this scenario were designed without having battery constraints in mind.

The seminar focused intensively on discussions in several working groups. To kick-off these discussions, we invited four keynote talks:

- Vehicles and The Edge: Random thoughts and not so random Perspectives by Jörg Ott (TU Munich, DE)
- Who protects the Unprotected? ITS Services for Vulnerable Road Users by Claudio Casetti (Politecnico di Torino, IT)
- Enabling data spaces: Existing developments and challenges by Gürkan Solmaz (NEC, DE)
- Securing Cooperative Intersection Management by Subjective Trust Networks by Frank Kargl (Ulm University, DE)

We finally organized the following working groups on some of the most challenging issues related to inter-vehicular communication, edge computing, and vulnerable road users:

- Edge computing
- Vulnerable road users
- Vehicle to cloud to vehicle communication
- Sensing and analytics
- Trust

References

1 Falko Dressler, Frank Kargl, Jörg Ott, Ozan K. Tonguz and Lars Wischhof, "Research Chal-lenges in Inter-Vehicular Communication – Lessons of the 2010 Dagstuhl Seminar," IEEE Communications Magazine, vol. 49 (5), pp. 158-164, May 2011.

2 Falko Dressler, Hannes Hartenstein, Onur Altintas and Ozan K. Tonguz, "Inter-Vehicle Communication – Quo Vadis," IEEE Communications Magazine, vol. 52 (6), pp. 170-177, June 2014.

3 Onur Altintas, Suman Banerjee, Falko Dressler and Geert Heijenk, "Executive Summary – Inter-Vehicular Communication Towards Cooperative Driving," Proceedings of Dagstuhl Seminar 18202 on Inter-Vehicular Communication – Towards Cooperative Driving, Schloss Dagstuhl, Germany, May 2018, pp. 31–59.

4 Ana Aguiar, Onur Altintas, Falko Dressler and Gunnar Karlsson, "Executive Summary – Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users," Proceedings of Dagstuhl Seminar 21262 on Inter-Vehicular Communication – From Edge Support to Vulnerable Road Users, vol. 11, Virtual Conference, June 2021, pp. 89–96.

2 Table of Contents

Executive Summary

Overview of Talks

Working groups

Open problems

3 Overview of Talks

3.1 Who protects the Unprotected? ITS Services for Vulnerable Road Users

Claudio Casetti (Polytechnic University of Torino, IT)

In this talk we first defined what is a Vulnerable Road User (VRU) from the point of view of many international standardisation entities. We then discussed four possible approaches for ITS to protect VRUs, highlighting pros and cons. First, the use of smart infrastructure with V2X capability. Then, cooperative perception by vehicles, followed by VRU-awareness messages sent by VRUs themselves. Finally, we discussed the use of edge/cloud support and introduced some open research questions that could be addressed during the rest of the seminar.

3.2 Securing Cooperative Intersection Management by Subjective Trust Networks

Frank Kargl (Universität Ulm, DE)

In this talk, I presented results from recent and newly-starting German and European research projects SecForCARs-SAVE, CONNECT, and ConnRAD where we investigate the role which trust models can play in securing complex cooperative, connected & automated mobility (CCAM).

CCAM systems are highly complex systems-of-systems (SoS) which are composed of many layers of subcomponents. Motivated by incidents like the log4j vulnerability, supply-chain security has recently taken up the challenge to evaluate security in such SoS. In our research, we investigate how to model trust dependencies in such SoS as trust networks or trust graphs in order to allow a quantifiable analysis of the effects that security incidents in one part of the system will have on other parts.

Based on earlier works, we were able to identify Subjective Logic and Subjective Trust Networks as a very useful formalism to model such trust graphs. In the talk, I illustrated this process on the example of Cooperative Intersection Management (CIM) and showed the steps that are needed for a MEC server to establish a trust opinion on the positions that vehicles send as part of their CAM messages.

In the following discussion, we elaborated on different aspects of such trust models, for example, the role of vehicle manufacturers as possible trust brokers as they constantly monitor their vehicle fleet through their backend systems and would be in a very good position to detect intrusions or other incidents that would reduce trust in a specific vehicle.

References

1 `https://www.secforcars.de/`
2 `https://horizon-connect.eu/`

3.3 Vehicles and "The Edge": Random thoughts and not so random Perspectives

Jörg Ott (TU München, DE)

Edge computing has been considered as a promising technology direction to support low-latency applications for end users, by offloading computing-intense and energy-consuming tasks from mobile devices to close by compute resources or by pushing centralized service instances closer to the user. Edge infrastructure should similarly be able to support vehicular applications, for compute offloading or data sharing. But would it need to? And, if so, could it really? In this talk, we explore demands of mobile (vehicular) applications for different latency bounds and see how far those could, in principle, be served by regular data centers. We use Germany as an example and investigate geographic and projected network topology distances from 33M points on German roads to 200+ data centers in 41 locations within the country. We then consider another hypothetical extreme case in which each base station would also serve as an edge server and consider scaling with the number of vehicles obtained from official traffic measurement stations. We finally touch upon the implications, including the need for running, managing and arbitrating all these resources.

3.4 Enabling data spaces: Existing developments and challenges

Gürkan Solmaz (NEC Laboratories Europe – Heidelberg, DE)

This talk presents the existing developments and key technical challenges of data spaces for the future of data ecosystems. Enabling data spaces requires the three layers that are highlighted in the talk: Data connectors/infrastructure, data interoperability, and data value. In the first layer, we consider the existing developments from IDSA, Gaia-X, and FIWARE. These developments target easy and secure data sharing through access and data usage policies, federation of cloud services, and standardized data models and contextualization. The second layer provides the data interoperability to connect and harmonize various data sources through data- and knowledge-driven machine learning. Finally, the third layer focuses on the "value" generation from data by easy and efficient application of advanced data processing functions of prediction, simulation, and optimization.

As a future data space use case, we propose "Green Twin", which aims to minimize energy consumption by creating and utilizing digital twins of entities such as vehicles, buildings, network infrastructure, and people. The talk describes the proof-of-concept project toward the application of Green Twin in the smart campus, building upon the FIWARE open-source ecosystem on the networking infrastructure with 5G and applying machine learning, to improve the efficiency of the energy usage for the buildings and mobility.

4 Working groups

4.1 Sensing and analytics

Ana Aguiar (Universidade do Porto, PT), Khalil Ben Fredj (University of Twente – Enschede, NL), and Gürkan Solmaz (NEC Laboratories Europe – Heidelberg, DE)

Joint work of Carla Fabiana Chiasserini, Geert Heijenk, Klaus David, Jérôme Härri, Onur Altintas, Florian Klingler, Christoph Sommer, Lukas Stratman

The first breakout session initiated with a discussion on the relationship between IoT and intelligent mobility applications, and identification of relevant applications to make discussion more concrete. A plethora of sensing and analytics applications have been considered in research related to mobility such as pedestrian flow detection, trajectory prediction, collision risk detection, user profiling, and so on. Considering vulnerable road users (VRUs), sensing and analytics data services would make use of data such as video data, GPS trajectories, and vehicular sensing.

The sensing may have two types of goals: real-time traffic management and safety applications, or feeding urban planning. Analyzing such data in large geographical domains would bring data communication challenges as well as challenges in the computing and Artificial Intelligence (AI), where specialized algorithms in distributed analytics would be studied. Machine Learning (ML) training and inference problems can be considered for developing vehicle-specific (in-vehicle), local, cooperative and federated models. Particularly, federated and split learning are an on-going research directions that address the distribution of the more computationally expensive phase of ML models: training. These two types of algorithms address privacy constraints by avoiding centralization of the raw data. Other on-going efforts in ML that are relevant to low latency include early-exit models, which process data through the whole pipeline only when necessary, e.g. for increased confidence. This would allow light local processing, moving data out of the mobile device only for some specific (detectable) situations. The distribution of ML model training involves several challenges related to the data itself. Besides addressing the need for i.i.d samples or the bias caused by non-i.i.d samples for most models, some areas currently under-explored are the distributed data pre-processing (e.g. local statistical measures may differ from global ones), and how the annotation of the data in a distributed setting could be achieved.

The trade-offs between different degrees and type of distribution, convergence speed, networking costs and model accuracy are yet in the realm of research. Modelling such trade-offs was identified as a valuable research direction. Distributed computing should make use of cloud and edge resources efficiently, such that the quality-of-service requirements, e.g. latency, from both the communication (network latency) and computation (virtualization and AI latency) angles would be satisfied. One may consider the networking as "in-vehicle" where wired communication would be utilized whereas the information that goes out of the vehicle should be transferred through wireless communication. For the computing side, edge computing may be applied in-vehicle, road-side, or edge data centers that are in the vicinity, materializing the vehicular edge computing paradigm. A brief call of attention was made to the different semantics of the fog, edge and cloud nomenclature in different communities, namely the Internet of Things.
Distributed sensing data collection and analytics are of utmost importance for improved

safety applications for cars, bikes, and pedestrians. For instance, cooperative sensing and perception has the potential to greatly improve the VRU safety in various areas of traffic in and out of cities such as traffic intersections, pedestrian crossings, or blind spots. On the other hand, there is a challenging decision making process between data offloading for improving application performance and keeping critical data in-vehicle for privacy and liability reasons. Especially, liability and accountability concerns are likely to play a determinant role on these scenarios. Other than the safety, there can be various mobility applications. A few of these applications, which can be enabled by distributed sensing and analytics, are listed below.

- Digitalization of the cities: e.g. for improving navigation, transportation and parking services
- Pothole or obstacle detection: Enabling comfort and efficiency.
- Dynamic infrastructure: making dynamic changes to the infrastructure for cost and energy efficiency, e.g. pop-up bike lanes.

The above-mentioned sensing and analytics applications mostly involve dealing with personal and sometimes confidential information, thus privacy-aware system design is a key aspect. The privacy-aware design would have particularly more importance when the new systems evolve from research-level prototypes towards real deployments. Trade-offs between utility and privacy are not well understood, and privacy by design should become the state-of-the-art. Yet, a quantification of the trade-offs would be valuable.

The future of Cooperative-Intelligent Transportation Systems (C-ITS) and the challenges for integration of VRU into the picture was widely discussed. It was consensual that to a large extent the technical problems have been covered in previous research, and the biggest hurdles to actual implementation are of societal, legal, political and economic nature. Nevertheless, several open research directions were identified.

- Improved large scale simulation models and digital twins enabled by high performance computing will enable a better understanding of behaviours. Current simulation solutions are often closed, proprietary, expensive, and of limited access. Lowering the cost for such solutions, e.g. using open source to facilitate evolution, expansion and integration of different simulation environments, enabling this integration on multiple-locations, etc would be of great value to a broad research community. Current status is assessed as very early infancy.
- Sensing and analytics plays a key role to build these models: mobility micro-models for traffic participants, especially VRUs, behavioural models (operational, tactical and strategical) in complex and safety relevant situations, traffic light models. It is of special relevance to study and model unexpected behaviours or behaviour/ intention change, as these are more likely to cause hazardous traffic situations. The metrics to validate such models are also in their infancy.
- Little data exists about accidents and their analysis, and access to existing data is a challenge for collaborative research, e.g. accident analysis databases like GIDA require very strict NDAs. User interface research for interaction with VRUs is another research direction with significant gaps.

Much of this research is strongly inter-disciplinary, and disciplines like transportation, urban planning and human factors need to be involved.

4.2 Edge computing

Falko Dressler (TU Berlin, DE) and Gürkan Solmaz (NEC Laboratories Europe – Heidelberg, DE)

The breakout discussion on edge computing started with the discussion for relevance and applicability of edge computing in terms of the communication and cost of resource usages. There are several questions raised for the applicability of the edge computing:

- Who should manage the edge servers?
- Which computations should happen at the edge?
- What are the communication requirements of mobility applications?

Starting with the first question, there is a question on whether mobile service providers or original equipment manufacturers (OEMs) should deploy and host the edge servers. For instance, mobile service providers may not have interest to realize dense deployments but OEMs can build edge servers on the cars and loads of data can start coming from the cars. For the second question, high-definition map building through video could be considered. Such application would cause high demand in terms of computation cost, bandwidth, and latency due to transfer of videos and the tasks of high-definition map building. There are, however, many more lightweight applications, which intelligent transportation systems applications could benefit from.

As a benefit of edge computing, the edge layer can serve as a buffer, where various computation tasks such as preprocessing could be performed on the edge; and extracted information could be shared with the cloud. Considering the other way around, the edge servers might share data with the vehicle for the internal computation and actuation of the vehicle. However, such critical scenarios will need to be carefully designed to avoid security vulnerabilities. For instance, steering a vehicle by bringing data from edge or cloud might create vulnerabilities to attacks. Furthermore, as vehicles move, they will occasionally be in out-of-coverage areas of the service providers. Thus, basic services and decisions such as steering can stay in the vehicle itself, whereas certain information that could not be collected through in-vehicle sensors may enable a smoother ride.

Edge computing would enable various applications, some of which are discussed during the breakout meeting. These applications include “cooperative” applications that require multiple vehicles as opposed to having a single vehicle behaving independent from other vehicles. At the initial phases of edge computing, a pragmatic approach would be to start with a “minimum viable edge”. The minimum viable edge would have services that are beneficial and easily applicable. For instance, assisting consensus building or creating local dynamic maps based on information collected from various vehicles could be implemented on the edge. For the latter, information can include traffic congestion, obstacles (e.g., potholes), disasters (e.g., water pipe burst), and other unexpected events (e.g., animal nearby). Moreover, applications that are not safety critical such as parking services could be more efficient through the application of edge computing.

One important aspect is the physical placement of the edge computing – if realized on cars. Several possible options exist for the physical placement:

- Intersections, where many vehicles pass by regularly
- Parked vehicles, where the computing platform could be made available
- Charging stations, where vehicles wait for relative long times.

The final part of the breakout discussion focused on the business models for future mobility use cases in terms of edge computing. Recently, vehicle manufacturers have become more like "data" companies as they collect mobility data from the vehicles and users. For the data ecosystem aspects, developments in the fields data contextualization (e.g., FIWARE, smart data models), for understanding data and creating value, as well as data spaces (e.g., Gaia-X, IDSA) for data sharing and exchanges between different parties are becoming highly relevant.

4.3 Trust

Frank Kargl (Universität Ulm, DE) and João P. Vilela (University of Porto, PT & INESC TEC – Porto, PT & CISUC – Coimbra, PT)

Trust can be considered a key aspect of resilient systems, reliably assessing a systems trustworthiness enables informed decisions with respect to, for example, safety- or security-critical functions. This working group discussed trust in cooperative, connected & automated mobility (CCAM) starting with looking at different dimensions of trust. We distinguished a technical, human, and regulatory notion of trust. The technical perspective is based on a notion of functional trust, i.e., the trust that one entity puts in another entity to perform a specific function in a trustworthy way. Alternatively, this can also mean that named other entity can provide certain data accurately. For example, this can refer to another vehicle correctly reporting its position in a CAM message where a receiving node has to rely on this data to predict collision risks. Modeling such trust relationships between components in an automotive system of systems leads to a trust graph that could be modelled using a formal logic to quantify the amount of trust and reason over trust relationships. Subjective logic trust networks are one suitable formalism which was illustrated in the plenary talk of Frank Kargl. With such an approach, a functional model of a system could be augmented with a trust model that allows to reason about technical trust in the system both at design and at run-time. Subjective logic provides the appealing property to allow reasoning under uncertainty with incomplete evidence. Open questions here include how and where to find the initial evidence for trustworthiness to feed the trust model with concrete data. This could come from looking at trust- or node-related trust as distinguished in a survey on misbehavior detection [1]. Furthermore, the structure of trust networks and reasoning approaches and the expected and required levels of trust require further investigation.

Human aspects

A purely technical treatment of trust would deny the fact that such vehicles are meant to transport humans safely and that those humans ultimately also need to trust the technical system to have a comfortable ride. This human aspects focused on the human perception of trust, which is challenging due to the fact that different people reason differently about the trust levels of different entities, be it automated vehicles or other users in the system. This is a highly subjective assessment that depends on many different personal notions of trust. For this, mental models of trust could be devised, as was done previously in privacy research [2], as well as creating user profiles that represent different user perspectives of trust. The generation of such profiles can be helpful to effectively predict user's preferences based on their perspectives on trust. The human aspect of trust perception is not sufficiently explored, in particular not in the context of technical systems and their objective trustworthiness. Open challenges here include being able to convey technical and regulatory trust mechanisms to users in an effective manner, as well as assessing the effect of such functional and regulatory mechanisms on the human perception of trust. Moreover, due to the subjectiveness of trust notions stemming from the distinct risk-perception of users, user profiles may be useful to accurately model an individual's perceptions of trust. However, such profiles must be created while also respecting privacy principles. This can come from federated learning mechanisms to predict users' preferences through user profiles in a privacy-preserving manner [3].

Regulatory aspects

The impact of regulations on trust is another relevant dimension that has an impact on both the technical solutions developed, as well as the human factor of trust perception. On the one hand, regulations define minimum requirements that technical solutions must abide to and thereby guarantee a certain level of safety. This is both a source of trust for us humans, as we assume these regulations to be in place and enforced and thus providing our safety. From a technical perspective, such regulations also provide us with certain assumptions about the trustworthiness of automotive systems and products that we can reflect in our trust models. The challenge here is to translate from regulatory requirements to the trustworthiness reflected, for example, in a Subjective Trust Network. Such a translation is by no means straightforward and defining such a quantification requires additional research. On the other hand, we having precise trust models would allow us to assess trustworthiness and trust requirements for automotive systems in a well-defined way, something that regulations might mandate one day, similar to safety and security analyses are mandated today. With respect to the human aspect, regulations can improve the perception of trust by users, if there is awareness that regulations are strictly enforced and there are visible consequences to institutions that do not comply. A set of challenges arise in this context, namely having auditing mechanisms that are effective in assuring compliance, otherwise a risk assessment may lead to conclusions that the risk and consequences of not being compliant may be worth it. Another challenge lies in the effect of regulations on users' interactions with services. It is known from privacy research that users exchange privacy for small benefits [4]. Additional research is needed to assess if the same holds with respect to trust.

Perceptions of Trust

The discussion group then dived deeper into trust perception. We identified that the perception of trust can be affected by technical solutions and regulatory aspects. With respect to the technical solutions, there are two relevant factors having effect on the trust level achieved:

- The effectiveness of technical solutions in conveying a feeling of trust to the users. For example, are users able to understand what measures a system takes to actually be trustworthy? This requires the ability to translate complex technical solutions into a common language that can be easily communicated to and understood by the average user.
- The usability of the technical solutions. A good technical solution can be easily compromised if it is not usable or affects the level of service that users expectto obtain. The challenge here lies in being able to develop technical solutions that are effective to increase trust, yet without being intrusiveor compromising usability.

A well-known example of a technically sound but not widely adopted trust mechanism is that of Pretty Good Privacy (PGP) to increase security and trust on email communications. It is recognized that PGP is an effective technical solution to increase security and trust of email, but failed from a usability perspective by putting the burden of managing part of the system on users. Users expect technical solutions to be as transparent and automated as possible, although providing visual cues that the system is compliant (e.g., the lock that represents an effective secure HTTP connection when browsing websites). This challenge of devising effective technical solutions that are usable remounts to early days of email, but still holds nowadays. An opposite example is that of trust in avionics. Although the general user/client is not aware of the specifics of regulations that rule the sector, there is a general feeling of trust in such system, instilled by a perception that there are tight regulations and inspections in place, as well as well-defined procedures to adopt in case of incidents. In this case, the user easily accepts tight restrictions and the burden of security controls, in exchange for a more secure system. This is also a natural consequence of the possible impacts of non-compliance: if on the previous example of PGP, the consequence may just be the impersonation of the sender of emails, in this later case it may be a question of life or death. Sousers are likely more willing to accept more complex security- and trust-enhancing mechanisms if their goal is to protect critical assets. Whenever the goal is to protect less tangible or not so critical assets, users expect trust-enhancing mechanisms that are transparent/automated to cause as little disturbance as possible to their operations/usability. Moreover, regulations can have a positive effect on perception of trust, but mostly if there is evidence that such regulations are known to have consequences (e.g.security checks at airports, or auditing of institutions for non-compliance).

Trust in Automated Driving

We then continued to discuss how these insights could be transferred to trust in automated driving compared to trust in today's manually driven vehicles. Regarding the human perception, people tend to desire to at least feel in-control. Therefore, it is important to keep human passengers informed and involved even in a human vehicle. On the other hand, given too much control back to humans, like allowance to make the car speed, might also introduce human error again and might make driving more unsure. So this needs to be investigated and balanced for automated driving. As evidenced by the avionics industry, tight regulations and inspections can instill trust in a system even though the passengers are not in control at all. If trustworthiness of a system can be assessed appropriately, this might even lead to some product certification like, e.g., NCAP or produce some online display of trust status to passengers. If this is helpful or damaging to trust requires investigation. The question should also be if trust assessment is confined to the single vehicle only. As vehicles start to form cooperative systems, the trustworthiness of the overall system should move into focus. Automated driving surely poses many challenges to safety. One involves the fact that it is hardly possible to predict every possible driving situation an automated car might be exposed to beforehand. Therefore, a continuous self-assessment of a vehicle might become a very important feature for autonomous vehicles. First simple self-assessment features are already implemented with today's cars. In such a self-assessment, a sound trust model might be a vital part, as the ultimate question the vehicle has to answer is whether it is still operating trustworthy.

In summary, in order to increase trust in automated driving, we recommend the following steps:

- deepen our understanding how to quantify trust in a complex automotive system-of-systems,
- precisely define what levels of trust are required,
- analyze sources of trust-related information, for example misbehavior detections systems, hardware security mechanisms, or certification,
- investigate how to interface an automated trust management with the driver,
- identify what action to take if insufficient trust is detected, e.g., initiation of a minimum risk maneuver.

References

1 Van der Heijden, R.W., Dietzel, S., Leinmüller, T. and Kargl, F. (2019). Survey on Misbehavior Detection in Cooperative Intelligent Transportation Systems. In IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 779-811, doi: 10.1109/COMST.2018.2873088.

2 Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., and Zhang, J. (2012). Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In Proceedings of the 2012 ACM conference on ubiquitous computing, pages 501–510.

3 Brandão, A., Mendes, R., and Vilela, J. P. (2022). Prediction of mobile app privacy preferences with user profiles via federated learning. In Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy.

4 Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, vol 347, issue 6221, pages 509–515.

4.4 Vulnerable road users

Gunnar Karlsson (KTH Royal Institute of Technology – Stockholm, SE), Khalil Ben Fredj (University of Twente – Enschede, NL), Klaus David (Universität Kassel, DE), and Marie-Christin Hannah Oczko (Paderborn, DE)

Overview and problem formulation

New technologies in motorized vehicles provide active collision avoidance that reduces both the risk of accidents as well as their severity. In addition, drivers and passengers are protected by their cars and trucks. Hence, their safety in traffic is progressing; a vision of zero fatalities might be reached eventually with autonomous vehicles. In contrast to that, traffic safety for vulnerable road users is decreasing as more and more persons switch from transportation by automotive vehicles to walking or to riding bicycles, scooters and other light and unprotected motorized vehicles. The topic of discussion in this breakout groups has targeted how modern computing and communication technologies might be used to improve the safety of vulnerable road users (VRUs). We noted that most developments regarding VRU safety are for heavier vehicles to detect pedestrians crossing roads and bicycles at right turns, and other dangerous situations. Alas, there is little attention to technology support for avoiding collisions among VRUs: bikers with other bikers, bikers with pedestrians who stray into bicycle lanes, and other equally well-known and contentious situations. It is clear by design that developments for autonomous driving have little use for road users who are in direct control of their vehicles, or for pedestrians who do not use a vehicle at all. The system support must hence aim at raising awareness of other road-users and of potential dangers in the infrastructure, such as slippery road conditions, ongoing road works and missing or inconsistent signs and directions.

The group discussed the categorization of vulnerable road users and concluded that it may be broad, considering both pedestrians and runners, people on horseback or in wheelchairs, and users of any two- or three-wheeled vehicle. We decided to focus on three common categories: pedestrians (including runners), bikers (with and without electric motoring), and scooter drivers. As stated, a general problem is lacking awareness of other road users such as bikers approaching pedestrians and faster, overtaking bikers from behind, as well as approaching but visually obstructed bikers. Here the transmission of radio beacons could be used to alert the surrounding traffic. The other type of alert regards the infrastructure where people need warnings of unsafe situations and road conditions, and connectedly how the unsafe spots might be detected, reported and disseminated to others.

Beaconing for increased awareness

The group worked through a design discussion to determine necessary considerations for making a beaconing system for VRUs. We discussed technology design regarding the radio and the protocols for beaconing. Firstly, the group discussed two options of communication: device-to-device communication (ad hoc) and of vehicle-to-infrastructure communication. The second option would use the mobile communication infrastructure (4G and beyond) with positioned users and processing in edge or cloud computing servers. For the D2D option, we discussed needed range and directionality of the transmissions and the position of the receivers, the potential contents of the beacons, the frequency of transmissions, and the use of standard radio technology, such as variants of Bluetooth, WiFi or even ZigBee. We considered a baseline design for a broad use based on radio interfaces in common mobile phones, which could be augmented by external devices (antennae, LIDAR and more), mounted on the vehicle, or on the road user (for instance on the helmet). The second compound of questions relates to the reception and processing of beacons. The contents of beaconing messages could include speed and speed variations, direction and steadiness, as well as type of vehicle and its dimensions. These messages, possibly received from many simultaneously approaching vehicles, should be compiled, prioritized, assessed, and formed into meaningful alerts or warning messages to the road user. This leads to the third topic of discussion: the interaction of the system with the human. We did not have any expert present on human-machine interaction and foresaw that it is a most germane area of research for the system. The alerts must be timely – allowing human reaction time; accurate – not causing false warnings; meaningful – leading to correct actions, and non-distracting – not causing dangerous situations.

Sensing and communication

The other area of safety concerns for VRUs relates to traffic intensity, road conditions, and design and state of the infrastructure (for instance, unsafe solutions, and broken traffic signals). Data for this area might need a centralized collection and compilation of reports from pedestrians and bikers as well as sensing data from the bikes (such as vibrations, and accelerometer measurements indicating sliding, heavy breaking, potholes, or falls). Similar to the beaconing, these messages need to be collected and compiled into meaningful alerts and warnings which have to be locally disseminated to where they have relevance to road users. Aggregated reports could also be provided to road authorities for improving the conditions or expediently removing dangers (for instance by sanding icy patches). For all types of beaconing and reporting of unsafe spots, it is important that the users can remain anonymous and untraceable. Otherwise, a misuse of information endangering individuals might be possible. There are likely other security and privacy aspects that we did not have time to discuss.

Concluding remarks

The two discussion sessions were fruitful and we developed our own understanding of the issues by working through a self-defined scenario for bikers overtaking one another. It opened up suggestions for many additional options such as use of image sensors and radar, and various feedback to the user through smart glasses with displays, tactile signals as well as auditory signals and messages. Several participants in the two sessions were interested in conducting a preliminary study on the feasibility of some of the ideas generated with the hope of defining a larger design study for experimental evaluation. We are grateful for the possibility to meet at Dagstuhl for this engaging discussion around an important problem area.

4.5 Vehicle to cloud to vehicle communication

Michele Segata (University of Trento, IT) and Onur Altintas (Toyota Motor North America – Mountain View, US)

V2V communications and potential applications have been proposed and investigated for more than 20 years. Yet, despite the large effort by both academic and industry research communities, technologies like IEEE 802.11p and C-V2X, as well as the applications that are built on them have not seen widespread deployment. On the other hand, most automakers ship new vehicles with cellular modems to enable, for example, data collection for diagnostic purposes. There is the possibility to exploit such means to potentially realize inter-vehicle communication and applications. Data coming from vehicles to be processed by the car manufacturer is handled by cloud computing facilities, so we refer to this type of inter-vehicle communication as vehicle to cloud to vehicle (V2C2V).

The aim of the breakout session was the analysis of potential benefits and drawbacks of such an approach to inter-vehicle communications. In particular, the breakout group indicated that the first point to be addressed is finding out the set of applications for which V2C2V could actually bring benefits with respect to V2V. One example is data aggregation, where a cloud-supported centralized solution would be easier to implement and more efficient than a fully decentralized one. An additional use case could be the one of cooperative maneuvers in urban scenarios. A centralized approach might ease gathering data and compute the best coordination strategy to be then communicated back to the vehicles, whereas a decentralized V2V solution might incur communication challenges due to the harsh environment. The second point raised by the group is that different OEMs might rely on different mobile operators and, in addition, they might resort to different cloud computing facilities. This opens a problem of interoperability between different car manufacturers.

First of all, in such cases, which operator's resources should handle the communication? Which spectrum should be used? More than technological, this question is mainly answered by agreements and business policies, which can definitely slow down the adoption of such systems.

The group discussed potential solutions to this problem, one being peering agreements between operators. In such case one of the biggest problems would be performance guarantees. As in classical Internet routing, operators might give higher priority to traffic belonging to their customers, but especially for safety applications this might be unacceptable. To encourage operators to cooperate, one solution would be to dedicate a portion of the spectrum for safety-related V2C2V communications which all operators could use for free.

Second, data sharing across different manufacturers is a complex issue. One option could be to agree on minimum amount of safety-related information to be shared so that the safety applications can be deployed without compromise. This clearly requires to define what is the minimum amount of information to be shared that can effectively improve vehicular safety. Here, regulations and standards may play a vital role in determining the minimum necessary set of safety information to be shared among automakers.

Finally, the discussion touched upon issue of deciding who is paying for the service, which relates to the incentives that could be granted by governments. Differently from pure V2V communications such as IEEE 802.11p, the use of cellular technologies does not come for free. A car manufacturer might sell a cellular data plan included in the price of the vehicle, but only for a limited amount of time. If customers have to take over the expenses after this period expires, we incur the risk of them bailing out, with a potentially negatively impacting on safety.

In conclusion, V2C2V might provide benefits to the vehicular domain, but applications and requirements need to be well-defined. In addition, we believe the role of governments and regulators to be fundamental.

5 Open problems

5.1 Connecting Bikes

Ana Aguiar (Universidade do Porto, PT)

Joint work of Pedro Santos, Luís Pinto, José Pintor, Eduardo Soares, Pedro D'Orey, João Mesquita, Miguel Rosa, Luís Almeida
Main reference Luis Ramos Pinto, Pedro M. Santos, Luís Almeida, Ana Aguiar: "Characterization and Modeling of the Bicycle-Antenna System for the 2.4GHz ISM Band", in Proc. of the 2018 IEEE Vehicular Networking Conference, VNC 2018, Taipei, Taiwan, December 5-7, 2018, pp. 1–8, IEEE, 2018.
URL https://doi.org/10.1109/VNC.2018.8628395

Bicycles are a healthy and environmentally friendly transportation mode that is increasingly used for commuting. Connecting bicycles to other vehicles is an enabler for a large variety of services, from safety to infotainment. Conversely useful and comfortable applications could make cycling more attractive. This talk shows demonstrations of motivating applications supported by Bluetooth Low Energy (BLE): a stolen bike detection system [1], and a broadcast walkie-talkie for a platoon [2].

The talk evolves to explore connectivity aspects needed to support such applications, both to connect bicyles to one-another and to the infrastructure. A characterisation 2.4 GHz operating technologies on bike-to-bike links using commodity hardware shows that BLE range exceeds 50m at low packet loss rates [3]. A dependence on relative bike positions was identified. Anechoic chamber measurements with bikes allowed to characterise the dependence of bike-to-anything links on antenna position and bike material [4].

References

1 P. M. Santos, M. Rosa, L. R. Pinto and A. Aguiar, Cooperative Bicycle Localization System via Ad Hoc Bluetooth Networks, IEEE VNC 2020

2 E. Soares, P. Santos, L. Pinto, P. Brandão, R. Prior, A. Aguiar. Demo: VoIP System for Bicycle Platoons IEEE VNC 2018

3 P. Santos, L. Pinto, A. Aguiar, L. Almeida. A Glimpse at Bicycle-to-Bicycle Link Performance in the 2.4GHz ISM Band, IEEE PIMRC 2018

4 P. Santos, L. Pinto, L. Almeida , A. Aguiar. Characterization and Modeling of the Bicycle-Antenna System for the 2.4GHz ISM Band, IEEE VNC 2018

5.2 On Realistic Scenarios for Hazard Perception of Vulnerable Road Users

Jérôme Härri (EURECOM – Biot, FR)

Joint work of Ali Nadar, Mathis Lafon, Jérôme Härri
Main reference Ali Nadar, Mathis Lafon, Jérôme Härri, A RounD-like Roundabout Scenario in CARLA Simulator, in Proceedings of the 4th Symposium on Management of Future Motorway and Urban Traffic Systems (MFTS2022), Dresden, 2022.
URL https://gitlab.eurecom.fr/cats/carla/round-carla

Evaluating contextual hazard of vulnerable road users (VRUs) under realistic driving and sensory contexts are critical to the integration of VRU with legacy and automated vehicles. Over the last decades, various synthetic scenarios have been designed and calibrated for microscopic simulators for SUMO mostly focusing on vehicles. Realistic traffic datasets including VRU such as RounD have been used to extract and learn precise driving and hazard patterns but cannot be modified to evaluate the impact of C-ITS safety applications for VRU in the dataset environment. The driving simulator CARLA has been designed to model robotic and sensory context in highly precise driving environment, which makes it perfectly suitable to model VRU in mixed traffic scenarios. However, most of the studies using CARLA focuses primarily on the modeling or the perception of an ego-vehicle (or a VRU) either isolated or under unrealistic traffic. Considering that realistic traffic interacting with VRU is critical to identifying hazard contexts for VRUs, this talk presents an open-source CARLA [1] scenario reproducing the RounD dataset and discuss its benefit to integrate realistic perception of VRUs.

References

1 RounD scenario for CARLA, `https://gitlab.eurecom.fr/cats/carla/round-carla`

5.3 Joint Communication and Sensing for V2?

Renato Lo Cigno (University of Brescia, IT)

Joint Communication and Sensing (JCS) is a staple of 6G and future Wi-Fi systems. The idea is using SCI (Channel State Information) collected at the PHY layer for MIMO, equalization, and so forth, to *sense* or *sound* the environment. Sensing includes localization, motion recognition and much more. Early works are promising, though not yet definitive, and focus mostly on indoor scenarios. The question is: can this technology be exported to vehicular environments and VRU protection too? Open question that I hope someone can tackle. One further question is if we can also protect the privacy of users against attacks based on EM fingerprinting at the PHY layer that cannot be countered with cryptographic techniques. Also this question has initial positive answers, but more research is needed.

5.4 Reconfigurable Intelligent Surfaces for Edge and Cooperative Driving

Michele Segata (University of Trento, IT)

Joint work of Marios Lestas, Paolo Casari, Taqwa Saeed, Dimitrios Tyrovolas, George Karagiannidis, Christos Liaskos

Reconfigurable Intelligent Surfaces (RIS) are communication devices capable of reflecting impinging wireless signals towards a certain direction, and the reflection angle can differ from the incidence one. These devices can be particularly useful in non-line-of-sight conditions, which are typical for mmWave communications. RIS could find application in vehicular communications to enable around-the-corner communications in the mmWave band, which could be especially beneficial for bandwidth-hungry applications such as cooperative perception or vehicular edge computing. The talk presents such opportunities but also the challenges connected to it, which include huge path loss due to the reflection, RIS scheduling, performance evaluation, and tracking of the users.

5.5 Performance Evaluation of Inter Vehicle Communication (IVC) for Vulnerable Road Users (VRUs)

Christoph Sommer (TU Dresden, DE)

Commoditization of system-level Inter Vehicle Communication (IVC) simulation has beneftitted research greatly. It commonly rests on three pillars: metrics, models, and scenarios. In the past few years, a rich set of all of these has slowly been made available for research on Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) use cases. This meant that high-fidelity experiments were no longer conditioned on the availability of resources for large field operational tests, nor was scale limited to just a few situations or vehicles, as afforded by hardware-in-the-loop type simulation. All of this has propelled research in the area of "traditional" Inter Vehicle Communication (IVC) forward. However, the case could be made that, while fragmented efforts exist [1, 2], a comprehensive set of generalizable metrics, models, and scenarios is still missing for researching Vulnerable Road Users (VRUs) centric use cases. We talk about how such a set might look like with a view towards generalizability and reproducibility of research.

References

1 L. Pinto, P. M. Santos, L. Almeida and A. Aguiar, “Characterization and Modeling of the Bicycle-Antenna System for the 2.4GHz ISM Band,” 2018 IEEE Vehicular Networking Conference (VNC), 2018, pp. 1-8, doi: 10.1109/VNC.2018.8628395.

2 RounD scenario for CARLA, `https://gitlab.eurecom.fr/cats/carla/round-carla`

Participants

- Ana Aguiar
Universidade do Porto, PT
- Onur Altintas
Toyota Motor North America – Mountain View, US
- Khalil Ben Fredj
University of Twente – Enschede, NL
- Claudio Casetti
Polytechnic University of Turin, IT
- Carla-Fabiana Chiasserini
Polytechnic University of Turin, IT
- Klaus David
Universität Kassel, DE
- Falko Dressler
TU Berlin, DE
- Jérôme Härri
EURECOM – Biot, FR
- Geert Heijenk
University of Twente, NL
- Frank Kargl
Universität Ulm, DE
- Gunnar Karlsson
KTH Royal Institute of Technology – Stockholm, SE
- Florian Klingler
Universität Paderborn, DE
- Renato Lo Cigno
University of Brescia, IT
- Marie-Christin Hannah Oczko
Paderborn, DE
- Jörg Ott
TU München, DE
- Michele Segata
University of Trento, IT
- Gürkan Solmaz
NEC Laboratories Europe – Heidelberg, DE
- Christoph Sommer
TU Dresden, DE
- Lukas Stratman
TU Berlin, DE
- João P. Vilela
University of Porto, PT & INESC TEC – Porto, PT & CISUC – Coimbra, PT
- Lars Wolf
TU Braunschweig, DE