06051 Abstracts Collection
Kolmogorov Complexity and Applications
— Dagstuhl Seminar —

Marcus Hutter[1], Wolfgang Merkle[2] and Paul M. B. Vitanyi[3]

[1] IDSIA - Lugano, CH
marcus@idsia.ch
[2] Univ. Heidelberg, DE
merkle@math.uni-heidelberg.de
[3] CWI - Amsterdam, NL
paulv@cwi.nl

Abstract. From 29.01.06 to 03.02.06, the Dagstuhl Seminar 06051 "Kolmogorov Complexity and Applications" was held in the International Conference and Research Center (IBFI), Schloss Dagstuhl. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

06051 Executive Summary – Kolmogorov Complexity and Applications

The Kolmogorov complexity of an object is the minimal number of bits required to effectively describe the object. This complexity measure becomes ever more important in the practical setting because it gives the ultimate limits to what is achievable by data compression (a central application area) and in the theoretical setting in an ever more diverse number of areas. Shortest description length is a central theme that is basic to the pursuit of science, and goes back to the principle known as Occam's razor, one version of which amounts to the rule that if presented with a choice between indifferent alternatives, then one ought to select the simplest one. Unconsciously or explicitly, informal applications of this principle in science and mathematics abound.

Kolmogorov complexity (also known as algorithmic information theory) is widely applied in computer science and a plethora of other scientific disciplines. The seminar was meant to be cross-disciplinary and to connect through the common technique of Kolmogorov complexity the areas information theory, individual randomness, algorithmic probability, recursion theory, computational complexity, machine learning and statistics, pattern recognition, data mining, and knowledge discovery. These topics were covered by 38 technical talks; in

Dagstuhl Seminar Proceedings 06051
Kolmogorov Complexity and Applications
http://drops.dagstuhl.de/opus/volltexte/2006/663

addition there were 5 historical talks and a subsequent panel discussion on the early history of Kolmogorov complexity. The seminar was attended by 50 participants, including a large number of leading researchers from the fields listed above. The seminar enabled the participating researchers to assess the state of the art and to inform themselves about new developments, the interdisciplinary character of the seminar helped to forge cohesion in the research community.

In 2005, the field of Kolmogorov complexity is vigorously alive, with new developments consisting of books in the making or just published about

- (i) the "renaissance" of recursion theory focussed on the analysis of individual randomness in terms of Kolmogorov complexity and related formalisms (R. Downey and D. Hirschfeldt, Algorithmic Randomness and Complexity, Springer, to appear);

- (ii) new trends in statistical inference and learning theory, artificial intelligence, based on compression (M. Hutter, Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability, EATCS Monographs, Springer 2004);

- (iii) pattern recognition, clustering and classification based on compression, Kolmogorov's structure function and algorithmic sufficient statistic, and distortion theory MDL and relations between information theory and Kolmogorov complexity (P. Vitanyi, Algorithmic Entropy, Algorithmic Distortion Theory, Springer, in preparation).

There is (iv) the area of the incompressibility method based on Kolmogorov complexity that keeps resolving decade-long (sometimes half-century) open problems in mathematics and computer science. The current trends mentioned above have been very well represented by participants and talks of the seminar.

Further material on the seminar can be found on the external seminar home page maintained by Marcus Hutter, see http://www.idsia.ch/ marcus/dagstuhl/.

Keywords: Executive Summary, Kolmogorov complexity

Joint work of: Hutter, Marcus; Merkle, Wolfgang; Vitanyi, Paul M. B.

See also:
http://www.idsia.ch/ marcus/dagstuhl/

Open Questions in Kolmogorov Complexity and Computational Complexity

Eric Allender (Rutgers Univ. - Piscataway, USA)

In this talk, I consider the question of whether or not there is an efficient reduction from the halting problem to the set of Kolmogorov-random strings.

It is known that there is an efficient nonuniform reduction (that is, a reduction computed by polynomial-size circuits) but the only "efficient" uniform reduction known is the (disjunctive) truth-table that was presented by Kummer – and the running time for that reduction is known to depend on the choice of the universal Turing machine that is used to define Kolmogorov complexity. For disjunctive reductions, it is known that exponential time at least is required. Motivation for this question comes from complexity theory. Derandomization techniques allow us to show that PSPACE is poly-time Turing reducible to the K-random strings, and NEXP is NP-reducible to this same set. No "efficient" uniform reductions to the K-random strings are known for any larger complexity class.

Keywords: Kolmogorov Complexity, Halting Problem, Efficient Reducibility

Time-Bounded Universal Distributions

Luis Antunes (Universidade do Porto, P)

We show that under a reasonable hardness assumptions, the time-bounded Kolmogorov distribution is a universal samplable distribution. Under the same assumption we exactly characterize the worst-case running time of languages that are in average polynomial-time over all P-samplable distributions.

Keywords: Kolmogorov Complexity, Universal Distributions, Average-Case Complexity

Joint work of: Antunes, Luis; Fortnow, Lance

Natural Halting Probabilities, Partial Randomness, and Zeta Functions

Cristian Calude (University of Auckland, NZ)

We introduce the *natural halting probability* and the *natural complexity* of a Turing machine and we relate them to program-size complexity and Chaitin's halting probability. A classification of Turing machines according to their natural (Omega) halting probabilities is proposed: divergent, convergent and tuatara. We prove the existence of universal convergent and tuatara machines. Various results on randomness and partial randomness are proved. For example, we show that the natural halting probability of a universal tuatara machine is c.e. and random. A new type of partial randomness, asymptotic randomness, is introduced. Finally we show that in contrast to classical (algorithmic) randomness—which cannot be characterized in terms of plain complexity—various types of partial randomness admit such characterizations.

Keywords: Natural halting probability, natural complexity

Joint work of: Calude, Cristian S.; Stay, Michael A.

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2006/631

Full Paper:
http://www.cs.auckland.ac.nz/CDMTCS//researchreports/273cris.pdf

See also: C. S. Calude, M. A. Stay. Natural Halting Probabilities, Partial Randomness, and Zeta Functions, *CDMTCS Research Report* 273, 2005.

Complexity Monotone in Conditions and Future Prediction Errors

Alexey Chernov (IDSIA - Lugano, CH)

We bound the future loss when predicting any (computably) stochastic sequence online. Solomonoff finitely bounded the total deviation of his universal predictor M from the true distribution μ by the algorithmic complexity of μ. Here we assume we are at a time $t > 1$ and already observed $x = x_1...x_t$. We bound the future prediction performance on $x_{t+1}x_{t+2}...$ by a new variant of algorithmic complexity of μ given x, plus the complexity of the randomness deficiency of x. The new complexity is monotone in its condition in the sense that this complexity can only decrease if the condition is prolonged. We also briefly discuss potential generalizations to Bayesian model classes and to classification problems.

Keywords: Kolmogorov complexity, posterior bounds, online sequential prediction, Solomonoff prior, monotone conditional complexity, total error, future loss, randomness deficiency

Joint work of: Chernov, Alexey; Hutter, Marcus; Schmidhuber, Jürgen

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/632

Full Paper:
http://www.idsia.ch/idsiareport/IDSIA-16-05.pdf

See also: A.Chernov, M.Hutter. Monotone Conditional Complexity Bounds on Future Prediction Errors. Proc. 16th International Conf. on Algorithmic Learning Theory (ALT-2005), LNCS 3734, Springer, 2005, pp. 414-428.

Randomness and Halting Probabilities

Santiago Figueira (University of Buenos Aires, RA)

We consider the question of randomness of the probability $\Omega_U[X]$ that an optimal Turing machine U halts and outputs a string in a fixed set X. The main results are as follows:

— $\Omega_U[X]$ is random whenever X is Σ_n^0-complete or Π_n^0-complete for some $n \geq 2$.

— However, for $n \geq 2$, $\Omega_U[X]$ is not n-random when X is Σ_n^0 or Π_n^0. Nevertheless, there exists Δ_{n+1}^0 sets such that $\Omega_U[X]$ is n-random.

— There are Δ_2^0 sets X such that $\Omega_U[X]$ is rational. Also, for every $n \geq 1$, there exists a set X which is Δ_{n+1}^0 and Σ_n^0-hard such that $\Omega_U[X]$ is not random.

We also look at the range of Ω_U as an operator. We prove that the set $\{\Omega_U[X] : X \subseteq 2^{<\omega}\}$ is a finite union of closed intervals. It follows that for any optimal machine U and any sufficiently small real r, there is a set $X \subseteq 2^{<\omega}$ recursive in $\emptyset' \oplus r$, such that $\Omega_U[X] = r$.

The same questions are also considered in the context of infinite computations, and lead to similar results.

Keywords: Chaitin's Ω, halting probabilities, randomness, n-randomness, Kolmogorov complexity

Joint work of: Becher, Verónica; Figueira, Santiago; Grigorieff, Serge; Miller, Joseph S.

Inconsistency & Misspecification

Peter Grünwald (CWI - Amsterdam, NL)

We show that Bayesian and MDL inference can be statistically inconsistent under misspecification: for any $a > 0$, there exists a distribution P, a set of distributions (model) M, and a 'reasonable' prior on M such that (a) P is not in M (the model is wrong) (b) There is a distribution $P' \in M$ with KL-divergence $D(P\|P') = a$, yet, if data are i.i.d. according to P, then the Bayesian posterior concentrates on an (ever-changing) set of distributions that all have KL-divergence to P much larger than a. If the posterior is used for classification purposes, it can even perform worse than random guessing.

The result is fundamentally different from existing Bayesian inconsistency results due to Diaconis, Freedman and Barron, in that we can choose the model M to be only countably large; if M were well-specified ('true'), then by Doob's theorem this would immediately imply consistency.

Joint work with John Langford of the Toyota Technological Institute, Chicago.

Keywords: Bayes, MDL, misspecification, consistency, prediction, compression, classificatoin

Joint work of: Grünwald, Peter; Langford, John

Full Paper:
http://www.cwi.nl/ pdg/ftp/bayesinconsistent.ps

Error in Enumerable Sequence Prediction

Nick Hay (University of Auckland, NZ)

We outline a method for quantifying the error of a sequence prediction. With sequence predictions represented by semimeasures $\nu(x)$ we define their error to be $-\log_2 \nu(x)$. We note that enumerable semimeasures are those which model the sequence as the output of a computable system given unknown input. Using this we define the simulation complexity of a computable system C relative to another U giving an *exact* bound on their difference in error. This error in turn gives an exact upper bound on the number of predictions ν gets incorrect.

Keywords: Sequence prediction, Solomonoff induction, enumerable semimeasures

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/633

Extracting Kolmogorov Complexity with Applications to Dimension Zero-One Laws

John M. Hitchcock (University of Wyoming, USA)

We apply recent results on extracting randomness from independent sources to "extract" Kolmogorov complexity. For any $\alpha, \epsilon > 0$, given a string x with $K(x) > \alpha|x|$, we show how to use a constant number of advice bits to efficiently compute another string y, $|y| = \Omega(|x|)$, with $K(y) > (1-\epsilon)|y|$. This result holds for both classical and space-bounded Kolmogorov complexity.

We use the extraction procedure for space-bounded complexity to establish zero-one laws for polynomial-space strong dimension. Our results include:

1. If $Dim_{pspace}(E) > 0$, then $Dim_{pspace}(E/O(1)) = 1$.
2. $Dim(E/O(1)|ESPACE)$ is either 0 or 1.
3. $Dim(E/poly|ESPACE)$ is either 0 or 1.

In other words, from a dimension standpoint and with respect to a small amount of advice, the exponential-time class E is either minimally complex (dimension 0) or maximally complex (dimension 1) within ESPACE.

Keywords: Kolmogorov complexity, resource-bounded dimension, extractors

Joint work of: Fortnow, Lance; Hitchcock, John; Pavan, A.; Vinodchandran, N. V.; Wang, Fengming

On the Convergence of (Non)Universal Semimeasures on Martin-Löf Random Sequences

Marcus Hutter (IDSIA - Lugano, CH)

Solomonoff's central result on induction is that the posterior of a universal semimeasure M converges rapidly and with probability 1 to the true sequence generating posterior μ, if the latter is computable. Hence, M is eligible as a universal sequence predictor in case of unknown μ. Despite some nearby results and proofs in the literature, the stronger result of convergence for all (Martin-Löf) random sequences remained open. Such a convergence result would be particularly interesting and natural, since randomness can be defined in terms of M itself. We show that there are universal semimeasures M which do not converge for all random sequences, i.e. we give a partial negative answer to the open problem. We also provide a positive answer for some non-universal semimeasures. We define the incomputable measure D as a mixture over all computable measures and the enumerable semimeasure W as a mixture over all enumerable nearly-measures. We show that W converges to D and D to μ on all random sequences. The Hellinger distance measuring closeness of two distributions plays a central role.

Keywords: Sequence prediction, Algorithmic Information Theory, universal enumerable semimeasure, mixture distributions, posterior convergence, Martin-Löf randomness

Full Paper:
http://www.idsia.ch/ marcus/ai/mlconvx.htm

Is there a simple theory of prediction?

Shane Legg (IDSIA - Lugano, CH)

Solomonoff's inductive learning model is a powerful, universal and highly elegant theory of sequence prediction. Its critical flaw is that it is incomputable and thus cannot be used in practice. It is sometimes suggested that it may still be useful to help guide the development of very general and powerful theories of prediction which are computable. In this paper it is shown that although powerful algorithms exist, they are necessarily highly complex. This alone makes their theoretical analysis problematic, however it is further shown that beyond a moderate level of complexity the analysis runs into the deeper problem of Gödel incompleteness. This limits the power of mathematics to analyse and study prediction algorithms, and indeed intelligent systems in general.

Points on Computable Curves

Elvira Mayordomo (University of Zaragoza, E)

The "analyst's traveling salesman theorem" of geometric measure theory characterizes those subsets of Euclidean space that are contained in curves of finite length. This result, proven for the plane by Jones (1990) and extended to higher-dimensional Euclidean spaces by Okikiolu (1991), says that a bounded set K is contained in some curve of finite length if and only if a certain "square beta sum", involving the "width of K" in each element of an infinite system of overlapping "tiles" of descending size, is finite. In this paper we characterize those points of Euclidean space that lie on computable curves of finite length. We do this by formulating and proving a computable extension of the analyst's traveling salesman theorem. Our extension, the computable analyst's traveling salesman theorem, says that a point in Euclidean space lies on some computable curve of finite length if and only if it is "permitted" by some computable "Jones constriction". A Jones constriction here is an explicit assignment of a rational cylinder to each of the above-mentioned tiles in such a way that, when the radius of the cylinder corresponding to a tile is used in place of the "width of K" in each tile, the square beta sum is finite. A point is permitted by a Jones constriction if it is contained in the cylinder assigned to each tile containing the point. The main part of our proof is the construction of a computable curve of finite length traversing all the points permitted by a given Jones constriction. Our construction uses the main ideas of Jones's "farthest insertion" construction, but our algorithm for computing the curve must work exclusively with the Jones constriction itself, because it has no direct access to the (typically uncomputable) points permitted by the Jones constriction.

Keywords: Computable analysis, geometric measure theory, analyst's traveling salesman theorem, rectifiable computable curve

Joint work of: Gu, Xiaoyang; Lutz, Jack H.; Mayordomo, Elvira

Full Paper:
http://www.eccc.uni-trier.de/eccc-reports/2005/TR05-157/

See also: Electronic Colloquium on Computational Complexity Report TR05-157

Recent Results in Universal and Non-Universal Induction

Jan Poland (Hokkaido Univ. - Sapporo, J)

We present and relate recent results in prediction based on countable classes of either probability (semi-)distributions or base predictors.

Learning by Bayes, MDL, and stochastic model selection will be considered as instances of the first category. In particular, we will show how analog assertions to Solomonoff's universal induction result can be obtained for MDL and stochastic model selection. The second category is based on prediction with expert advice. We will present a recent construction to define a universal learner in this framework.

Keywords: Bayesian learning, MDL, stochastic model selection, prediction with expert advice, universal learning, Solomonoff induction

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/635

Stability of Kolmogorov Type Properties under Relativization

Andrej Romashchenko (IITP - Moscow, RUS)

Let us have a tuple of strings $x_1, \ldots, x_n$, and some oracle O.

Assume that there is no information in O about x_i's, i.e., $I(x_1, \ldots, x_n; O) \approx 0$.

Can any 'informational' properties of $x_1, \ldots, x_n$ essentially change under relativization conditional to the oracle O? We discuss a formalization of this question and show that some special cases of these general conjecture hold. In particular, we prove that if an oracle has very few mutual information with given strings, then relativization conditional to this oracle cannot help to extract the mutual information between these strings.

Keywords: Kolmogorov complexity, algorithmic information theory, extracting mutual information

Joint work of: Romashchenko, Andrei; Muchnik, Andrei

Forbidden substrings, Kolmogorov complexity and almost periodic sequences

Andrei Rumyantsev (Moscow State University, RUS)

Assume that for some $\alpha < 1$ and for all natural n a set F_n of at most $2^{\alpha n}$ "forbidden" binary strings of length n is fixed.

Then there exists an infinite binary sequence ω that does not have (long) forbidden substrings.

We prove this combinatorial statement by translating it into a statement about Kolmogorov complexity and compare this proof with a combinatorial one based on Laslo Lovasz local lemma. Then we construct an almost periodic sequence with the same property (thus combines the results from Bruno Durand, Leonid Levin, Alexander Shen, Complex tilings and Andrei Muchnik, Alexei Semenov and Maxim Ushakov, Almost periodic sequences).

Both the combinatorial proof and Kolmogorov complexity argument can be generalized to the multidimensional case.

Keywords: Forbidden substrings, almost periodic sequences

Joint work of: Rumyantsev, Andrei; Ushakov, Maxim

See also: STACS, Marseille, 2006

Application of Kolmogorov complexity and universal codes to identity testing and nonparametric testing of serial independence for time series.

Boris Ryabko (Russian Academy of Sc. - Novosibirsk, RUS)

We show that Kolmogorov complexity and such its estimators as universal codes (or data compression methods) can be applied for hypothesis testing in a framework of classical mathematical statistics. The methods for identity testing and nonparametric testing of serial independence for time series are described.

Keywords: Algorithmic complexity, algorithmic information theory, Kolmogorov complexity, universal coding, hypothesis testing

Joint work of: Ryabko, Boris; Astola, Jaakko; Gammerman, Alex

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/636

Learning in Reactive Environments with Arbitrary Dependence

Daniil Ryabko (IDSIA - Lugano, CH)

In reinforcement learning the task for an agent is to attain the best possible asymptotic reward where the true generating environment is unknown but belongs to a known countable family of environments.

This task generalises the sequence prediction problem, in which the environment does not react to the behaviour of the agent.

Solomonoff induction solves the sequence prediction problem for any countable class of measures; however, it is easy to see that such result is impossible for reinforcement learning — not any countable class of environments can be learnt.

We find some sufficient conditions on the class of environments under which an agent exists which attains the best asymptotic reward for any environment in the class. We analyze how tight these conditions are and how they relate to different probabilistic assumptions known in reinforcement learning and related fields, such as Markov Decision Processes and mixing conditions.

Keywords: Reinforcement learning, asymptotic average value, self-optimizing policies, (non) Markov decision processes

Joint work of: Ryabko, Daniil; Hutter, Marcus

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/637

Proof-Based General Search Algorithms and Goedel Machines

Jürgen Schmidhuber (IDSIA - Lugano, CH)

We use Gödel's self-reference trick to build a universal problem solver. A Gödel machine is a computer whose original software includes axioms describing the hardware and the original software (this is possible without circularity) and some formal goal in form of an arbitrary user-defined utility function (e.g., cumulative future expected reward in a sequence of optimization tasks). The original software also includes a proof searcher which uses the axioms to systematically make pairs ("proof", "program") until it finds a proof that a rewrite of the original software through "program" will increase utility. The machine can be designed such that each self-rewrite is necessarily globally optimal in the sense of the utility function, even those rewrites that destroy the proof searcher. We relate the approach to previous asymptotically optimal search algorithms.

Keywords: Self-referential universal problem solvers, provably optimal self-improvements

Full Paper:
http://www.idsia.ch/ juergen/goedelmachine.html

See also:
http://www.idsia.ch/ juergen/gmbib

Multisource aAgorithmic Information Theory

Alexander Shen (IITP - Moscow, RUS)

Multisource information theory is well known in Shannon setting. It studies the possibilities of information transfer through a network with limited capacities. Similar questions could be studied for algorithmic information theory and provide a framework for several known results and interesting questions.

Keywords: Kolmogorov complexity multisource information theory

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/626

Combinatorial proof of Muchnik's theorem

Alexander Shen (IITP - Moscow, RUS)

Original proof of Muchnik's theorem on conditional descriptions can be modified and split into two parts:
1) we construct a graph that allows large online matchings (main part);
2) we use this graph to prove the theorem.
The question about online matching could be interesting in itself.

Keywords: Matching conditional descriptions, Kolmogorov complexity

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/625

Relative n-Randomness for Continuous Measures

Theodore A. Slaman (Univ. California - Berkeley, USA)

We discuss the following question of Reimann.

For which infinite binary sequences X does there exist (a presentation of) a measure μ such that X is n-random for μ?

We show that for each n, for all but countably many X there is such a μ. From the recursion theoretic perspective, the countable set is large, including all the arithmetically definable sets and more.

Keywords: Relative randomness

Self-similar Sets, Dimensions and Kolmogorov Complexity

Ludwig Staiger (Universität Halle-Wittenberg, D)

We present several results on relations between the Kolmogorov complexity of infinite strings and measures of information content (dimensions) known from dimension theory, information theory or fractal geometry.

Special emphasis is laid on bounds on the complexity of strings in self-similar sets in Cantor space. Here a set is called self-similar iff it is the maximum solution of an equation of the form $F = WF$ where W is a set of finite strings of length > 0 and F is the subset of Cantor space defined by this equation.

It turns out that the Hausdorff dimension of these sets can be expressed in terms of the entropy of the language W, and that the Kolmogorov Complexity (or constructive dimension) of infinite strings in F can be estimated via Hausdorff dimension or entropy of languages.

For sets in Cantor space definable by finite automata (regular ω-languages) even tighter bounds of Kolmogorov Complexity via Hausdorff dimension or entropy of languages are obtained.

Here results generalizing P. Martin-Löf's on the bounds for Kolmogorov complexity and oscillation of the complexity in an infinite string are obtained.

Keywords: Infinite strings, Cantor space, Kolmogorov complexity, Hausdorff dimension, packing dimension, entropy of languages

On the difference between Martin-Löf randomness and Schnorr randomness

Sebastiaan Terwijn (TU Wien, A)

In his 1971 book Schnorr introduced a constructive measure that relates to Martin-Löf's notion of constructive measure in much the same way as "computable" relates to "computably enumerable". Schnorr carefully analyzed this notion, and he also characterized it in terms of computable martingales and orders of growth. Schnorr also considered the randomness notion corresponding to computable martingales. In this talk we discuss the differences between these randomness notions from various points of view.

Keywords: Martin-Löf randomnes, Schnorr randomness

Binary Lambda Calculus and Combinatory Logic

John Tromp (CWI - Amsterdam, NL)

We introduce binary representations of both lambda calculus and combinatory logic terms, and demonstrate their simplicity by providing very compact parser-interpreters for these binary languages.

We demonstrate their application to Algorithmic Information Theory with several concrete upper bounds on program-size complexity, including an elegant self-delimiting code for binary strings.

Keywords: Concrete, program size complexity, ambda calculus, combinatory logic, encoding, self-delimiting, binary strings

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/628

On impossibility of sequential algorithmic Forecasting

Vladimir V'yugin (IITP - Moscow, RUS)

The problem of prediction future event given an individual sequence of past events is considered.

Predictions are given in form of real numbers p_n which are computed by some algorithm φ using initial fragments $\omega_1, \ldots, \omega_{n-1}$ of an individual binary sequence $\omega = \omega_1, \omega_2, \ldots$ and can be interpreted as probabilities of the event $\omega_n = 1$ given this fragment.

According to Dawid's *prequential framework* we consider partial forecasting algorithms φ which are defined on all initial fragments of ω and can be undefined outside the given sequence of outcomes.

We show that even for this large class of forecasting algorithms combining outcomes of coin-tossing and transducer algorithm it is possible to efficiently generate with probability close to one sequences for which any partial forecasting algorithm is failed by the method of verifying called *calibration*.

Keywords: Universal forecasting, computable calibration, Dawid's prequential framework, algorithmic randomness, defensive forecasting

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/630

Algorithmic rate-distortion theory

Nikolai K. Vereshchagin (Moscow State University, RUS)

We propose and develop rate-distortion theory in the Kolmogorov complexity setting. This gives the ultimate limits of lossy compression of individual data objects, taking all effective regularities of the data into account.

Keywords: Kolmogorov complexity, rate, distortion, randomness deficiency

Joint work of: Vereshchagin, Nikolai K.; Vitanyi, Paul M.B.

Kolmogorov complexity and Shannon information

Paul M.B. Vitanyi (CWI - Amsterdam, NL)

The paradigms in the computation, manipulation, and transmission of information shift increasingly from being random-variable oriented to individual-outcome oriented.

In theoretical terms, this means a shift from Shannon's Information Theory to Kolmogorov's Complexity Theory.

Complex structural data like sound files or video images are not well modeled by methods that assume they consist of a typical sequence of uncorollated independent outcomes of a simple Bernoulli source or stationary ergodic source. For example, storing or transmitting the first $1,000,000,000,000,000$ bits of $\pi = 3.1415\ldots$ can be done the hard way using Shannon theory, or the easy way using a small program incorporating an approximation algorithm that generates the successive bits π.

Thus, synthetic structural audio compression can reach compression ratios of 10,000 to 1. Such structural data can in principle be modeled by a random source. For example, Tolstoy's "War and Piece" is one element out of the the set of books or meaningful documents that have been produced in the history of mankind, say at most one billion. With every book equally likely, it contains about 30 bits of information. This way, we can in principle assume that there is a relevant random variable that represents the complex structured information objects appropriately. Practically, however, using Shannon theory or general purpose compressors, we estimate the redundancy at about 60%, achieving a compression ratio of about 2 to 1. This redundancy estimate is based on the inherent assumption that the data consists of uncorrolated items and is typical for a simple random source like a Bernoulli or Markov process. Clever or special-purpose compressors analyse structural features of the data to achive better compression. The result is a code that can be viewed as a Shannon-Fano code and then translated back into an appropriate highly skewed distribution.

Viewing this as Shannon theory is begging the question.

Ultimately, one wants identify the unique combination of structural features that characterize the data at hand. One way to do this is by the Kolmogorov complexity approach. We shall show that, for example, the Shannon entropy equals the expected Kolmogorov complexity of the outcomes, but only up to an additive error term that equals the Kolmogorov complexity of the probability mass function involved. In the limit, where we consider an individual object, the Shannon view can be taken as concentrating all probability on the outcome in question.

But then the correspondence breaks down: the Shannon entropy becomes zero, and the error term becomes the complexity of the outcome.Shannon ignores the object itself but considers only the characteristics of the random source of which the object is one of the possible outcomes, while Kolmogorov considers only the object itself to determine the number of bits in the ultimate compressed version irrespective of the manner in which the object arose.

For almost every Shannon theory notion there turns out to be a Kolmogorov complexity theory notion that is equivalent in the sense that the expectation of the latter is closely related to the former. We will review some of these relations.

Joint work of: Grünwald, Peter; Vitanyi, Paul M. B.

Automatic Meaning Discovery Using Google

Paul M.B. Vitanyi (CWI - Amsterdam, NL)

We survey a new area of parameter-free similarity distance measures useful in data-mining, pattern recognition, learning and automatic semantics extraction.

Given a family of distances on a set of objects, a distance is universal up to a certain precision for that family if it minorizes every distance in the family between every two objects in the set, up to the stated precision (we do not require the universal distance to be an element of the family).

We consider similarity distances for two types of objects: literal objects that as such contain all of their meaning, like genomes or books, and names for objects. The latter may have literal embodyments like the first type, but may also be abstract like "red" or "christianity." For the first type we consider a family of computable distance measures corresponding to parameters expressing similarity according to particular features between pairs of literal objects. For the second type we consider similarity distances generated by web users corresponding to particular semantic relations between the (names for) the designated objects.

For both families we give universal similarity distance measures, incorporating all particular distance measures in the family. In the first case the universal distance is based on compression and in the second case it is based on Google page counts related to search terms.

In both cases experiments on a massive scale give evidence of the viability of the approaches.

Keywords: Normalized Compression Distance, Clustering, Clasification, Relative Semantics of Terms, Google, World-Wide-Web, Kolmogorov complexity

Joint work of: Cilibrasi, Rudi; Vitanyi, Paul M.B.

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2006/629

A dual approach to universal prediction

Volodya Vovk (RHUL - London, GB)

The Solomonoff-Levin theory of universal prediction relies on a Bayes-type merging procedure that is applied to a large class of prediction strategies (such as all computable strategies). The desirable properties that we would like universal predictions to satisfy can be roughly classified into theoretical (as many as possible of the standard laws of probability should hold for the predictions and the actual observations) and pragmatic (the decisions based on the predictions should lead to a small loss and/or large gain). Several such properties are known to be satisfied by the Solomonoff-Levin predictions, but it turns out that a less direct approach to universal prediction leads to many more properties of both kinds. A new player is added to the game of prediction: the gambler, who tries to make money on the lack of agreement between the predictions and the actual observations. A large class of potential strategies for the gambler can now be merged into one "universal law of probability", and it can be shown that there is a way of producing predictions that are ideal as far as this universal law is concerned. Such predictions can also be regarded as universal. They can be shown

to satisfy the standard properties of calibration and resolution and to give good predictions in the framework of competitive on-line learning.

Keywords: Universal prediction, defensive forecasting, competitive on-line statistics