

# 4th international conference in Logic, Computability and Randomness

Abstracts

June 29 to July 3, 2009  
CIRM, Marseille, France

**Eric Allender** (Rutgers University, USA).

*Circuits, Levin's Kt complexity, and some new unconditional lower bounds*

Levin introduced a time-bounded Kolmogorov complexity measure,  $Kt$ , that has proved useful in many different settings. Earlier work by the speaker and others has

- shown that if  $x$  is viewed as the truth table of a function  $f_x$ , then  $Kt(x)$  is polynomially-related to the size of the smallest *oracle* circuit for  $f_x$ , where the oracle comes from the exponential-time complexity class EXP.
- presented a related time-bounded Kolmogorov complexity measure  $KT$ , with the property that, for any oracle  $A$ ,  $KT^A(x)$  is polynomially-related to the circuit size of  $f_x$ , on oracle circuits using oracle  $A$ . (The usual Kolmogorov measures  $K(x)$  and  $C(x)$  are polynomially related to  $KT^H(x)$ , where  $H$  is the halting problem.)
- made use of derandomization techniques to show that the set  $R_{Kt}$  consisting of strings with high  $Kt$  complexity is complete for EXP under reductions computed by polynomial-size *circuits*, although it is *not* complete under the ordinary polynomial-time many-one reductions.

A frustrating open problem has been to prove an *unconditional* lower bound on the set  $R_{Kt}$ . Although  $R_{Kt}$  is complete for EXP under "efficient" reductions, we have no proof that  $R_{Kt}$  is not in P.

This talk will introduce a nondeterministic variant of  $Kt$ , called  $KNt$ , and present some motivation to convince you that it is a fairly natural measure (and indeed is it very closely related to a measure defined earlier by Buhrman, Fortnow, and Laplante).  $KNt(x)$  is polynomially-related to the circuit size of  $f_x$  on oracle circuits with an oracle from NEXP, and  $R_{KNt}$  is complete for NEXP/poly under reductions computed by polynomial-size circuits.

Most interestingly, we present an unconditional lower bound, showing that  $R_{KNt}$  is not in  $NP \cap coNP$ . Although the proof is not difficult, it relies on the theory of interactive proofs and the technology of derandomization.

This is joint work with Michal Koucký, Detlef Ronneburger, and Sambudha Roy. More detail and related work can be found at

<http://ftp.cs.rutgers.edu/pub/allender/pervasive.reach.pdf>.

\* \* \*

**George Barmpalias** (Victoria University of Wellington, New Zealand).  
*Elementary differences between the degrees of unsolvability and degrees of compressibility*

The study of the ‘descriptive’ complexity of strings and streams has naturally lead to the study of relativized complexity (where the Turing machines used have access to external information) in the same way that the theory of computability lead to the theory of relative computation and unsolvability. For example, a set  $A$  was called ‘low for  $K$ ’<sup>1</sup> if the prefix-free complexity relative to  $A$  is the same (modulo a constant) as the unrelativized prefix-free complexity. This means that  $A$  contains no information which could help to achieve a better compression on the binary strings. This notion was studied thoroughly in [4], where it was shown that it coincides with two other notions:  $K$ -triviality and lowness for randomness. A set  $A$  is  $K$ -trivial if its initial segments have minimal prefix-free complexity, i.e. no more (modulo a constant) than the complexity of a trivial sequence like  $0^\infty$ . Moreover,  $A$  is low for random if any random sequence is also random relative to  $A$ . In the following, we will mostly use the name ‘ $K$ -trivial’ to refer to any of its equivalent formulations. Based on the notion of ‘low for  $K$ ’, Nies [4] defined the partial order  $\leq_{LK}$  on the Cantor space: we say that  $A \leq_{LK} B$  for two sets  $A, B$  if the prefix-free complexity relative to  $A$  is at least as much (modulo a constant) as the one relative to  $B$ . In other words,  $B$  can compress at least as well as  $A$ , and in symbols  $K^B(\sigma) \leq K^A(\sigma) + c$  for a constant  $c$  and all strings  $\sigma$ .

This partial ordering defines an equivalence relation on the Cantor space which groups different oracles in a single class provided that they are capable for the same level of compression. These equivalent classes are usually called  $LK$  degrees but we also call them *degrees of compressibility*. We note that two oracles may contain mutually disjoint information (in the sense that they form a minimal pair in the degrees of unsolvability) yet be in the same  $LK$  degree<sup>2</sup>. An apparently weaker partial order is obtained if we only require that the random streams relative to  $B$  (i.e. the streams whose initial segments cannot be compressed using information from  $B$ ) are also random relative to  $A$ . This partial order was also introduced in [4], was denoted by  $\leq_{LR}$  and the induced structure of equivalent classes was called the  $LR$  degrees. Remarkably, Joe Miller (see [5]) has shown that  $\leq_{LR}$  coincides with  $\leq_{LK}$ .

In [2,3] the  $LR$  degrees were studied both locally and globally, and a number of similarities were discovered with the Turing degrees, both with respect to algebraic features of the partially ordered structures and in terms of the methods used to prove them. The applicability of methods from the Turing degrees

<sup>1</sup>this notion was defined by Andrej A. Muchnik during a seminar in 1999.

<sup>2</sup>This follows by the fact that there is a promptly simple set  $A \leq_{LK} \emptyset$  (see [5]) since every promptly simple set computes a minimal pair in the Turing degrees and  $\leq_{LK}$  is an extension of  $\leq_T$ .

to the study of the LR degrees was, to some degree, expected as  $\leq_{LR}$  (and  $\leq_{LK}$ ) is a natural extension of  $\leq_T$  (the Turing reducibility) of the same arithmetical complexity. In the same papers a quite special feature of  $\leq_{LR}$  was discovered, namely the uncountable predecessor property, which provided a dramatic difference with the structure of Turing degrees. On the other hand, this property is not elementary (it is not a first order property) and it does not seem to play an important role in the study of the local structures of the LR/LK degrees, for example the  $\Sigma_1^0$  or the  $\Delta_2^0$  degrees. Here an LR/LK degree is called  $\Sigma_1^0/\Delta_2^0$  if it contains a  $\Sigma_1^0/\Delta_2^0$  set respectively (similar definitions hold for higher arithmetical classes).<sup>3</sup> In this paper we provide the first elementary differences between the local structures of the Turing and the LR/LK degrees. We show the following.

**Theorem 1.** *Let  $X$  be a  $\Delta_2^0$  set which are not K-trivial. Then there exists a c.e. set  $A$  which is not K-trivial such that  $A \leq_{LR} X$ .*

Theorem 1 has interesting consequences.

**Corollary.** *The  $\Delta_2^0$  structure of LR/LK degrees is downward and upward dense. Also, the  $\Delta_2^0$  structures of LR/LK degrees and the Turing degrees are not elementarily equivalent.*

After a modification of the construction behind the proof of Theorem 1 we are able to establish the following.

**Theorem 2.** *Let  $X, Y$  be  $\Delta_2^0$  sets which are not K-trivial. Then there exists a c.e. set  $A$  which is not K-trivial such that  $A \leq_{LR} X$  and  $A \leq_{LR} Y$ .*

Theorem 2 has further consequences, outlined in the following corollary.

**Corollary.** *The  $\Sigma_1^0$  structures of LR/LK degrees and the Turing degrees are not elementarily equivalent. Also, the structure of the LR/LK degrees below the LR/LK degree of the halting problem is not elementarily equivalent to the  $\Sigma_1^0$  and  $\Delta_2^0$  structures of LR/LK degrees.*

Finally, we have the following.

**Corollary.** *Given any finite collection of  $\Delta_2^0$  sets which are not K-trivial, there is an uncountable collection of LR/LK degrees below the LR/LK degrees of all of them. This follows by Theorem 2 in combination with the result in [1] that every  $\Delta_2^0$  set which is not K-trivial LR bounds uncountably many sets and the fact from [4] (also see [6]) that every LR degree is a countable equivalence class.*

[1] George Barmpalias. Relative randomness and cardinality. Submitted preprint, 2007.

[2] George Barmpalias, Andrew E. M. Lewis, and Mariya Soskova. Randomness, Lowness and Degrees. *J. of Symbolic Logic*, 73(2):559–577, 2008.

<sup>3</sup>Recall that  $\Sigma_1^0$  sets or degrees are also called computably enumerable, or c.e. for short.

- [3] George Barmpalias, Andrew E. M. Lewis, and Frank Stephan.  $\Pi_1^0$  classes, LR degrees and Turing degrees. *Ann. Pure Appl. Logic*, 156(1):21–38, 2008.
- [4] André Nies. Lowness properties and randomness. *Adv. Math.*, 197(1):274–305, 2005.
- [5] André Nies. *Computability and Randomness*. Oxford University Press, in preparation, 2009.
- [6] Stephen G. Simpson. Almost everywhere domination and superhighness. *MLQ Math. Log. Q.*, 53(4-5):462–482, 2007.

\* \* \*

**Bruno Bauwens** (Ghent University, Belgium).

*Ideal hypothesis testing and algorithmic information transfer*

Influence testing for two discrete time series of equal length is studied, by defining influence-free semimeasures and influence-free semimeasures associated with enumerable semimeasures. It is shown that the first group of semimeasures has a universal element, whether the second has not. A coding result is proved that characterizes the logarithm of this semimeasure approximately by a variant of conditional prefix-free Kolmogorov complexity, and we describe how influence tests in engineering literature can be considered as approximations of these ideal definitions. We show that these have some nice additive properties.

In statistics a simple hypothesis is defined as a set of logical statements that allow the inference of a unique semimeasure over the set of all a priori possible outcomes of an experiment. Assume that two simple hypothesis have semimeasures  $P_0$  and  $P_1$ . The statistical test  $d(x) = P_1(x)/P_0(x)$  has an optimal power for any significance level [4]. Composite tests are tests that specify a set  $S$  of semimeasures. Multiplicative dominance defines a partial order on the semimeasures, and in some cases the enumerable semimeasures in  $S$  have a maximal element, which is called the the universal element of the hypothesis. The likelihood-ratio test can now define in the same way notions of significance and power.

The hypothesis of a time series  $x$  being influence-free from another time series  $y$  with equal length  $l(x) = l(y) = n$ , is a composite hypothesis. We show that with this hypothesis there corresponds a universal enumerable length conditional semimeasure, whose logarithm is approximately given by:

$$K(x|y \uparrow, n) = \min\{l(p) : \Phi(p, n, y_1 \dots y_i) \downarrow = x_{i+1} \wedge \forall z \in 2^n, \Phi(p, n, z) \downarrow\},$$

where  $\Phi$  is a prefix-free Turing machine and  $2^n$  the set of binary strings of length  $n$ . This definition does not measure the extractable information on the halting problem of  $y$  that is transferred to  $x$ . In some contexts this might be more appropriate. We show that we can extend the set of enumerable influence-free semimeasures to the set of influence-free semimeasures associated with enumerable semimeasures and that this set also measures transferred halting information. However, this set is shown to have no universal element.

Due to the standard coding theorem, ideal influence tests against the universal enumerable semimeasure can now be defined as total algorithmic information transfer (TIT). We also define algorithmic information transfer (IT)

that also measures transferred halting information. IT and TIT can both be considered as analogs of Shannon information transfer [5]. Tests used in engineering practice [3,6] can be seen as approximations of these ideal influence tests. We also show that TIT defines approximate decompositions of algorithmic mutual information and leave the possibility of such a decomposition for IT open. All errors on the above approximate results are of logarithmic order of the algorithmic minimal sufficient statistic of both strings. Generalizations of decomposition results of IT and TIT lead to many open questions.

Identity testing or randomness testing relative to a semimeasure, is testing whether a simple or a composite hypothesis can explain all 'structure' in experimental data. In [2] the existence of such universal independence tests are studied for different computability classes. In [1], some results are shown about the existence of universal identity tests relative to the universal enumerable semimeasure.

Acknowledgements: Supported by a Ph.D grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This talk gives an overview of the theoretical section of my PhD work on theoretical and practical independence and influence testing.

[1] B. Bauwens. Co-enumerable sumtests for the universal distribution. Submitted, 2008.

[2] B. Bauwens and S. Terwijn. Notes on sum-tests and independence tests. Submitted, 2008.

[3] C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 1969.

[4] Neyman J. and Pearson E. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, pages 289–337, 1933.

[5] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2), 2000.

[6] M. Winterhalder, B. Schelter, W. Hesse, K. Schwab, L. Leistritz, R. Bauer, J. Timmer, and H. Witte. Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems. *Signal Processing*, 85:2137–2160, 2005.

\* \* \*

**Cristian S. Calude** (The University of Auckland, New Zealand).  
*Every computably enumerable random real is provably computably enumerable random*

We prove that every c.e. random real is provable in Peano Arithmetic (PA) to be c.e. random, a statement communicated to one author by B. Solovay. A major step in the proof is to show that the theorem stating that “a real is c.e. and random iff it is the halting probability of a universal prefix-free Turing machine” can be proven in PA. The proof, which is simpler than the standard one, can also be used for the original theorem.

Our positive result can be contrasted with the case of computable functions, where not every computable function is provably computable. It is even more interesting to compare our result with the fact that almost all random finite strings are not provably random.

Some proofs in the paper have been obtained using interactively the theorem prover Isabelle.

\* \* \*

**Douglas Cenzer** (University of Florida, USA).  
*Random members of random closed sets*

**(Joint work with Rebecca Weber).** We consider the interaction between reals and sets of reals which are random in different senses.

For example, if two reals (viewed as sets of natural numbers) are relatively random in the usual sense, then their union is not random, since it contains roughly  $3/4$  of the natural numbers, but we show that it is  $\mu$ -random with respect to the  $3/4$ - $1/4$  measure.

If a closed set is random in the usual  $1/3$ - $2/3$  measure, then the leftmost path is not random since it takes the value  $1/2$  of the time, but it is  $\nu$ -random in the  $1/3$ - $2/3$  measure on  $2^{\mathbb{N}}$ .

We show that for any  $p$  with  $1/3 \leq p \leq 2/3$ , any random closed set has a member which is  $\mu_p$  random, that is, random with respect to the  $(p, 1 - p)$  measure on  $2^{\mathbb{N}}$ .

\* \* \*

**Alexey Chernov** (Royal Holloway, University of London, UK).  
*Practical aspects of Levin's neutral measure*

**(Joint work with Yuri Kalnishkan and Vladimir Vovk).** This expository talk will include some of the following topics:

- The definition and history of Leonid Levin's idea of neutral measure, and its further development by Peter Gacs.
- Applications to prediction with expert advice, including:
  - prediction with expert advice for several loss functions simultaneously;
  - prediction with specialist experts' advice;
  - prediction with second-guessing experts' advice.
- Applications to competitive on-line prediction, including competing with function classes of prediction strategies.

\* \* \*

**Adam Day** (Victoria University of Wellington, New Zealand).  
*Increasing the Gap between Descriptive Complexity and Algorithmic Probability*

A well known theorem of Gács shows that the analog of the coding theorem fails for continuous sample spaces. This means that descriptive monotonic complexity ( $K_m$ ) does not coincide within an additive constant with the negative logarithm of algorithmic probability ( $KM$ ). Gács's proof provided a lower bound on the difference between these values. He showed that for infinitely many finite binary strings, this difference was greater than a version of the inverse Ackermann function applied to string length. This talk will explain how this lower bound can be substantially improved. The inverse Ackermann function can be replaced with a function  $O(\log(\log(x)))$ . This shows that in continuous sample spaces, descriptive monotonic complexity and algorithmic probability are very different. The proof of this new lower bound builds on the original work by Gács, it does have a number of new features, in particular, the algorithm at the heart of the proof works on sets of strings as opposed to individual strings.

\* \* \*

**Johanna Franklin** (National University of Singapore, Singapore).  
*Far from weak randomness*

The behavior of reals that are far from random for different randomness notions varies remarkably. I will present some results on reals that are far from random for notions weaker than Martin-Löf randomness.

\* \* \*

**Noam Greenberg** (Victoria University of Wellington, New Zealand).  
*Yet more on strongly jump traceable reals*

I present recent discoveries in the study of strongly jump-traceable sets. In particular, I discuss the characterisation (with Nies and Hirschfeldt) of the strongly jump-traceable c.e. sets using randomness, and the possibilities for going beyond the realm of c.e. sets.

\* \* \*

**Serge Grigorieff** (University of Paris 7, France).  
*Effective Wadge hard sets*

**(Joint work with Verónica Becher).** The classical Wadge theory on totally discontinuous Polish spaces has been extended to  $T_0$  topological spaces with Scott topologies by V. Selivanov. We present an effectivization of Wadge theory in the framework of  $\omega$ -continuous d.c.p.o., i.e. Scott domains. Wadge theory was originally defined with no considerations on computable strategies. The reason being that Martin’s Determinacy gives highly non computable strategies. Though Wadge Duality theorem and the non-trivial part of the Wadge Hardness theorem fail in an effective context, an interesting part of Wadge theory still remains after effectivization. We prove that there is an effectively Wadge hard set for each level of the Arithmetical Hierarchy, and one for each level of the Effective Difference Hierarchy. We also give, for each level, a topological characterization of effectively hard sets, and we illustrate with particular examples. To develop the results of the paper we commit to the space  $P(\mathbb{N})$  of all subsets of  $\mathbb{N}$  with the Scott topology, and then we transfer them to any Scott domain.

\* \* \*

**Mathieu Hoyrup** (LORIA, Nancy, France).

*Layerwise computability*

One of the main applications of algorithmic randomness is that properties that classically hold with probability one can be strengthened in principle, holding at every random point. Classical examples can be found in [5,7,8] for instance. When proving this kind of result the key hypothesis is the *computability* of the random variables involved. However, it is well-known that computability notions are the effective versions of topological ones (the computable functions are precisely the effectively continuous ones, the semi-decidable sets are precisely the effectively open sets, and so on). Hence the computability assumption on random variables is (i) inappropriate in principle, as probability theory is grounded on measure theory and not on topology; (ii) a priori too strong, as in the classical setting only properties as measurability, integrability are required. This leads to the following:

**Problem.** Theorems for random points should hold for “effectively measurable” objects and not only computable ones.

This problem has already been independently investigated in [2,6] where ergodic theorems for random points are proved for different types of “almost everywhere computable” functions. These works are, however, still far from catching the effective version of measurable functions. For instance in Birkhoff’s ergodic theorem, nothing can be said about the mean sojourn time of algorithmically random points in fractal sets having effective constructions, as the *Smith-Volterra-Cantor* (or *fat Cantor*) set (the Smith-Volterra-Cantor set  $A \subseteq [0, 1]$  is homeomorphic to the Cantor set and has Lebesgue measure  $\frac{1}{2}$ ).

In [3], working in the context of computable probability spaces (to which Martin-Löf randomness has been recently extended, see [1,4]), effective versions of measure-theoretic notions were examined and another contribution

of algorithmic randomness to probability theory was developed: the setting of a new framework for computability adapted to the probabilistic context. This was achieved by making a fundamental use of the existence of a universal Martin-Löf test to endow the space with what we call the *Martin-Löf layering*. In this framework, which we call *layerwise computability*, the *layerwise* versions of virtually all computability notions can be naturally defined. The point is that effective measurability notions can be characterized as the layerwise versions of computability notions.

We show how this framework provides a general solution to Problem 1. In particular, we prove general versions of theorems for random points and *effectively measurable* random variables, among which Birkhoff's ergodic theorem. This is a significant improvement of [6,2] as it implies in particular a positive result for the Smith-Volterra-Cantor set. To prove these results we develop tools allowing to adapt the existing techniques (used in the *computable* context) to the *layerwise computable* context. Then, the results for *effectively measurable* objects follow from the characterizations. This strategy is very general and applicable in a wide range of situations.

[1] Gács, P.: Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science* **341** (2005) 91–137

[2] Galatolo, S., Hoyrup, M., Rojas, C.: Effective symbolic dynamics, random points, statistical behavior, complexity and entropy. Submitted (2007) Available on arxiv.

[3] Hoyrup, M., Rojas, C.: An application of Martin-Löf randomness to effective probability theory (2009) CiE '09.

[4] Hoyrup, M., Rojas, C.: Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation* (2009) To appear. Available on arxiv.

[5] Martin-Löf, P.: The definition of random sequences. *Information and Control* **9**(6) (1966) 602–619

[6] Nandakumar, S.: An effective ergodic theorem and some applications. In: *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing*, New York, NY, USA, ACM (2008) 39–44

[7] Vovk, V.G.: The law of the iterated logarithm for random kolmogorov, or chaotic, sequences. *Theory of Probability and Applications* **32** (1987) 413–425

[8] V'yugin, V.V.: Effective convergence in probability and an ergodic theorem for individual random sequences. *SIAM Theory of Probability and Its Applications* **42**(1) (1997) 39–50

\* \* \*

**Pavel Karpovich** (Moscow State University, Russia).

*Monotone complexity of a pair*

There are different version of Kolmogorov complexity: plain ( $KS$  or  $C$ ), prefix ( $KP$  or  $K$ ), decision complexity ( $KR$ ), monotone complexity ( $KM$ ) and others. We extend the notion of monotone complexity to pairs of binary strings in a natural way: programs generates pairs of strings (on two output tapes) in a monotone way, It turns out that monotone complexity of a pair differs significantly from a priori complexity (logarithm of a priori probability) for a simple

reason: the complexity of a pair of strings  $x$  and  $y$  can exceed the sum of the lengths of  $x$  and  $y$  by  $\alpha \log(|x| + |y|)$  (if  $\alpha < 1$ ).

We also show that (quite unexpectedly) the following upper bound is true:  $KR(x, y, z)$  (defined in a natural way) does not exceed  $|x| + |y| + |z| + O(1)$ . (For four strings similar result is not known.)

\* \* \*

**Bjoern Kjos-Hanssen** (University of Hawaii at Manoa, USA).

*Feeble subsets*

Intuitively, one expects that each truly random set of integers has an infinite subset that computes no random set. That is, a *loss of randomness beyond algorithmic repair* can occur when some 1s are turned into 0s.

Let us say that a set is *feeble* if it is infinite and computes no Martin-Löf random set.

Many ML-random sets have feeble subsets: all 2-random sets, and all ML-random sets above  $0'$ .

The main idea in the proofs of these results is a certain use of random variables whose values are closed sets.

We still do not know if each ML-random set has a feeble subset. On the other hand, suppose we replace everywhere the Bernoulli parameter  $1/2$  by a non-computable number  $p$ . In this setting, we can show that each ML-random set or its complement has a feeble subset. The idea of the proof is to adopt the point of view of statistics.

\* \* \*

**Jack Lutz** (University of Iowa, USA).

*Dimension Spectra of Random Fractals*

(Joint work with Xiaoyang Gu, Elvira Mayordomo, and P. Moser). The (constructive Hausdorff) dimension of a point  $x$  in Euclidean space is the *algorithmic information density* of  $x$ . Roughly speaking, this is the least real number  $\dim(x)$  such that  $r \times \dim(x)$  bits suffices to specify  $x$  on a general-purpose computer with arbitrarily high precisions  $2^{-r}$ . The *dimension spectrum* of a set  $X$  in Euclidean space is the subset of  $[0, n]$  consisting of all the dimensions of points in  $X$ .

The dimensions of points have been shown to be geometrically meaningful (Lutz 2003, Hitchcock 2003), and the dimensions of points in self-similar fractals have been completely analyzed (Lutz and Mayordomo 2008). Here we begin the more challenging task of analyzing the dimensions of points in random fractals. We focus on fractals that are randomly selected subfractals of a given self-similar fractal. We formulate the specification of a point in such a subfractal as the outcome of an infinite two-player game between a *selector* that selects the subfractal and a *coder* that selects a point within the subfractal.

Our selectors are algorithmically random with respect to various probability measures, so our selector-coder games are games against nature.

We determine the dimension spectra of a wide class of randomly selected subfractals. We show that each such fractal has a dimension spectrum that is a closed interval whose endpoints can be determined from the parameters of the fractal. In general, the maximum of the spectrum is determined by the degree to which the coder can *reinforce* the randomness in the selector, while the minimum is determined by the degree to which the coder can *cancel* randomness in the selector. This constructive and destructive interference between the players' randomnesses is somewhat subtle, even in the simplest cases. Our proof techniques include van Lambalgen's theorem on independent random sequences, measure preserving transformations, an application of network flow theory, a Kolmogorov complexity lower bound argument, and a nonconstructive proof that this bound is tight.

\* \* \*

**Elvira Mayordomo** (Universidad de Zaragoza, Spain).

*Resource-Bounded Dimension in Computational Learning Theory*

**(Joint work with Maria Lopez-Valdes, and Vinodchandran Variyam).** We study the resource-bounded dimension of learnable classes of concepts by different learning algorithms such as on-line algorithms, query-learning algorithms and probably approximately correct algorithms. For instance, the p-dimension of on-line learnable classes with  $o(2^n)$  mistakes, the pspace-dimension of query learnable classes with  $o(2^n)$  membership queries and the pspace-dimension of PAC-learnable classes is zero. Some bounds on this dimensions for less restrictive learning algorithms are also obtained. We expect that this line of research will provide a method to compare and quantify the performance of different methods of learning.

\* \* \*

**Wolfgang Merkle** (Universität Heidelberg, Germany).

*Separating nonmonotonic notions of randomness*

**(Joint work with Laurent Bienvenu, Thorsten Kräling and Rupert Hölzl).** In the theory of algorithmic randomness, several classes of random sequences are defined via a game-theoretic approach. Perhaps the most well-known is computable randomness, introduced by Schnorr: an infinite binary sequence is computably random if no total computable strategy succeeds on it by betting on bits in order. Computable randomness turns out to be, in some sense, far from Martin-Löf randomness. Muchnik (elaborating on ideas of Kolmogorov

and Loveland) refined Schnorr's model by also allowing non-monotonic strategies, i.e. strategies that do not bet on bits in order. The subsequent "non-monotonic" notion of randomness, now called KL-randomness, has been shown to be quite close to Martin-Löf randomness, but whether these two classes coincide remains a fundamental open question. In order to get a better understanding of non-monotonic randomness, Miller and Nies introduced some interesting intermediate notions, where one only allows non-adaptive strategies, i.e. strategies that can still bet non-monotonically, but such that the sequence of betting positions is known in advance (and computable). These notions were shown by Kastermans and Lempp to differ from Martin-Löf randomness. In this talk, we will continue the study of the Miller-Nies non-monotonic randomness notions, from a Kolmogorov complexity point of view. In particular, we provide a complete classification of these notions by order of strength.

\* \* \*

**Joseph Miller** (University of Madison, USA).  
*Randomness and DNC functions*

We will give a short proof that every DNC function computes an infinite subset of a 1-random set (and every infinite subset of a 1-random computes a DNC function). This result, which is joint work with Noam Greenberg, has nice consequences. For example, Kjos-Hanssen used a version to show that almost every set contains an infinite subset of minimal Turing degree.

\* \* \*

**Keng Meng Ng** (Victoria University of Wellington, New Zealand).  
*Degrees of reals with positive effective packing dimension*

The effective packing dimension of a real is a measure of non-integral dimension arising from packing measures. Having effective packing dimension one is a notion where "category meets measure", since it is a property which are shared by both random and sufficiently generic reals. Greenberg and Downey observed that for c.e. degrees, the Turing degrees containing a real of positive effective packing dimension were exactly the degrees which were not c.e. traceable. In this talk we show that outside of the c.e. degrees, the relationship between dimension and traceability is less clear. In particular we show that this characterization fails in the  $\Delta_2^0$  degrees, and also fails in the hyperimmune-free degrees.

\* \* \*

**Cristobal Rojas** (IML, Marseille, France).

*Randomness on computable probability spaces: a dynamical point of view*

The basic idea of algorithmic randomness is that an individual algorithmic random sequences should satisfy all the “effective” probability laws. There are several different possible definitions, depending on the kind of the considered probability laws and their “degree of effectivity”. A very natural source of probability laws is the ergodic theory of dynamical systems. In this talk we present some recent results relating algorithmic randomness to the statistical properties of the trajectories of points in ergodic dynamical systems. We do this in the framework of general computable probability spaces (the effective version of the spaces where usual ergodic theory takes place). First we extend algorithmic randomness to his setting and develop some useful tools. Then, we introduce a “dynamical” notion of randomness: typicality. Roughly, a point is typical for some ergodic dynamics, if it follows the statistical behavior of the system (given by Birkhoff’s pointwise ergodic theorem) with respect to every bounded continuous function used to follow its trajectory (or equivalently, every computable function). The main result is the following characterization: in any computable probability space, a point is Schnorr random if and only if it is typical for every mixing computable dynamical system.

\* \* \*

**Andrey Romyantsev** (Moscow State University, Russia).

*Forbidden strings and subsequences*

For every  $\alpha < 1$  there exists an infinite binary sequence whose factors of every length  $n$  have complexity at least  $\alpha n + O(1)$  (Levin’s lemma). This is a Kolmogorov complexity version of a combinatorial fact: if for every  $n$  at most  $2^{\alpha n}$  strings are declared “forbidden”, there exists an infinite binary sequence that has no long forbidden factors. Here Kolmogorov complexity can be used as a tool to prove a combinatorial statement (there is a nice simple combinatorial argument, explained in J. Miller’s note, too).

The connection works in both directions: Lovasz Local lemma can be used to prove that there exists an infinite sequence  $\omega$  such that for every finite set  $A$  of indices the string  $\omega(A)$ , the restriction of  $\omega$  onto  $A$ , has complexity at least

$$\alpha \#A - \max_{t \in A} C(A|t) - O(1)$$

This is a generalization of Levin’s lemma: if  $A$  is a contiguous interval, the complexity  $C(A|t)$  is negligible (logarithmic) for every  $t \in A$ .

This statement is quite flexible and can be used as a tool for many combinatorial results (sequence of bounded exponent, quasiperiodic sequences of high complexity, twodimensional sequences with complex rectangles, etc.)

Levin’s lemma can be relativized. However, the “mutually recursive” relativization is not known: we don’t know whether there exist two sequences

$\omega$  and  $\tau$  such that every  $n$ -bit factor of  $\omega$  has complexity  $\alpha n - O(1)$  even with oracle  $\tau$  and vice versa.

Partial negative result: for every two infinite sequences  $\alpha$  and  $\beta$  either  $\alpha$  has a factor of complexity  $O(1)$  if its position in  $\alpha$  is known and  $\beta$  is used as an oracle or  $\beta$  has a factor with similar property.

\* \* \*

**Alexander Shen** (Université de Provence & CNRS, France).  
*Unpublished work of Andrej A. Muchnik*

We will try to survey some results of Andrej Muchnik (24.02.1958 - 18.03.2007) on Kolmogorov complexity (including unpublished ones):

- Mutual information and its extraction.
- Game and probabilistic arguments for Kolmogorov complexity.
- Kolmogorov-style “cryptology”.
- “Muchnik paradox” and on-line randomness.

\* \* \*

**Ludwig Staiger** (Universität Halle-Wittenberg, Germany).  
*On Domains of Universal Machines*

**(Joint work with Cristian S. Calude, André Nies and Frank Stephan).** This paper presents some results of the papers [1,2] on domains of universal machines.

Universal machines  $U$  play a central role in algorithmic information theory. A plain universal machine is used to define the plain complexity  $C$ . For a string  $w$  one lets  $C(w)$  be the length of a shortest string  $p$  such that  $U(p) = w$ . One defines  $H(w)$  in a similar way when  $U$  is a universal prefix-free machine.

The paper addresses the questions which recursively enumerable (r.e.) sets can be the domains of a universal plain [prefix-free] machines, and which sets are [prefix-free] supersets of such domains?

Moreover, as domains of universal prefix-free machines are r.e. prefix codes, it is interesting which information-theoretic properties do such codes have.

In the first part we deal with classical information-theoretic properties of r.e. prefix codes containing domains of universal prefix-free machines, and in the subsequent parts we give characterisations of the domains of universal [prefix-free] machines and their [prefix-free] r.e. supersets. Among them we give combinatorial characterisations based on the number of strings of each length in the set.

We consider the *spectrum function*,  $s_W: \mathbb{N} \rightarrow \mathbb{N}$ , of a set of strings  $W \subseteq \{0, 1\}^*$  defined as

$$\begin{aligned} s_W(\ell) &:= |\{w : w \in W \wedge |w| = \ell\}| \quad \text{and its variant} \\ s_W(\ell, c) &:= |\{w : w \in W \wedge \ell \leq |w| \leq \ell + c\}| \end{aligned}$$

Then the following are true.

**Fact 1.** *Let  $W \subseteq \{0, 1\}^*$  be an r.e. language.*

1. *Then the function  $s_W$  is left computable.*
2. *If  $\sum_{i=0}^{\infty} s_W(i) \cdot 2^{-i} < \infty$  then there is a constant  $d \in \mathbb{N}$  such that  $s_W(0, n) \leq 2^{n-H(n)+d}$ .*

**1. Information-theoretic properties.** We start with a characterisation of r.e. prefix-free supersets of domains of prefix-free universal machines and their relations to “classical” prefix codes. Those supersets will be called *universal computably enumerable (c.e.) prefix codes* (see [2]).

**Theorem 1.** [Embeddability of prefix codes] *An r.e. prefix-free language  $W \subseteq \{0, 1\}^*$  is a universal c.e. prefix code if and only if for every r.e. prefix code  $D$  there is a partial recursive one-to-one function  $\varphi \subseteq: \{0, 1\}^* \rightarrow \{0, 1\}^*$  such that  $\text{dom}(\varphi) \supseteq D$ ,  $\varphi(D) \subseteq W$  and  $\exists k \forall u (u \in \text{dom}(\varphi) \rightarrow |\varphi(u)| \leq |u| + k)$ .*

This yields several consequences.

**Theorem 2.** [A. Nies] *Every universal c.e. prefix code is Turing complete.*

**Corollary 1.** [Maximality] *No universal c.e. prefix code is a maximal prefix code.*

Next we consider a kind of density for codes.

**Theorem 3.** [Maximal entropy of languages] *Let  $W \subseteq \{0, 1\}^*$  be a universal c.e. prefix code. Then its lower entropy is*

$$\underline{H}_W := \liminf_{n \rightarrow \infty} \frac{1}{n} \cdot \log_2 s_W(0, n) = 1$$

There are, however, simple deterministic prefix codes having the same maximal entropy of languages.

**Example.** [W. Kuich] *The language  $\mathbb{L} \subseteq \{0, 1\}^*$  defined by  $\mathbb{L} = 0 \cup 1 \cdot \mathbb{L}^2$  has also  $\underline{H}_{\mathbb{L}} = 1$ .*

**2. Combinatorial properties.** Here we give necessary and sufficient conditions in terms of the spectrum function for r.e. subsets of  $\{0, 1\}^*$  to be domains or supersets of domains of universal machines.

First we strengthen the results of the previous section for prefix codes.

**Theorem 4.** *Let  $W \subseteq \{0, 1\}^*$  be an r.e. prefix code. Then  $W$  is a c.e. universal prefix code if and only if*

$$\exists c, d \forall n (2^{n-H(n)-d} \leq s_W(n, c) \leq 2^{n-H(n)+d}).$$

**Theorem 5.** *An r.e. language  $W \subseteq \{0, 1\}^*$  is the domain of a universal prefix-free machine if and only if*

$$\exists c \forall n (H(\langle n, s_U(n, c) \rangle) \geq n).$$

**Theorem 6.** *There exists a universal c.e. prefix code  $W \subseteq \{0, 1\}^*$  which is not the domain of a universal prefix-free machine.*

Finally, we turn to domains of a universal plain machines.

**Theorem 7.** *An r.e. language  $W \subseteq \{0, 1\}^*$  is the superset of a domain of a plain universal machine if and only if*

$$\exists c \forall n (s_W(n, c) \geq 2^n).$$

**Theorem 8.** *An r.e. language  $W \subseteq X^*$  is the domain of a plain universal machine if and only if*

$$\exists c \forall n (C(s_W(n, c)) \geq n).$$

[1] C.S. Calude, A. Nies, L. Staiger and F. Stephan, Universal Recursively Enumerable Sets of Strings, in: M. Ito, M. Toyama (eds.). *Developments in Language Theory*, Lectures Notes in Comput. Sci. 5257, Springer-Verlag, Berlin, 2008, 170–182.

[2] C.S. Calude and L. Staiger. On universal computably enumerable prefix codes. *Mathematical Structures in Computer Science*, to appear.

<http://www.cs.auckland.ac.nz/CDMTCS//researchreports/312cris.pdf>

\* \* \*

**Kohtaro Tadaki** (Chuo University, Japan).

*A statistical mechanical interpretation of algorithmic information theory*

**Abstract.** We develop a statistical mechanical interpretation of algorithmic information theory by introducing into the theory the notion of thermodynamic quantities at temperature  $T$ , such as partition function  $Z(T)$ , free energy  $F(T)$ , energy  $E(T)$ , statistical mechanical entropy  $S(T)$ , and specific heat  $C(T)$ , which are real functions of a real argument  $T > 0$ . We investigate the properties of these quantities by means of program-size complexity from the point of view of algorithmic randomness. It is then discovered that, in the interpretation, the temperature  $T$  equals to the partial randomness of the values of all these thermodynamic quantities, where the notion of partial randomness is a stronger representation of the compression rate by means of program-size complexity. Furthermore, we show that this situation holds for the temperature  $T$  itself as a thermodynamic quantity. Namely, the computability of the value of partition function  $Z(T)$  gives a sufficient condition

for  $T \in (0, 1)$  to be a fixed point on partial randomness. In addition the computability of each of the thermodynamic quantities  $F(T)$ ,  $E(T)$ , and  $S(T)$  also gives the sufficient condition. Moreover, we show that the computability of  $F(T)$  gives completely different fixed points from the computability of  $Z(T)$  in particular.

*Key words:* algorithmic information theory, algorithmic randomness, partial randomness, statistical mechanics, thermodynamic quantities, temperature, fixed point theorem, Chaitin  $\Omega$  number

## 1. Preliminaries

We first review some basic notation and definitions which will be used in this paper.  $\mathbb{N} = \{0, 1, 2, 3, \dots\}$  is the set of natural numbers, and  $\mathbb{N}^+$  is the set of positive integers.  $\mathbb{R}$  is the set of reals.  $\{0, 1\}^*$  is the set of finite binary strings. For any  $s \in \{0, 1\}^*$ ,  $|s|$  is the *length* of  $s$ . A subset  $S$  of  $\{0, 1\}^*$  is called *prefix-free* if no string in  $S$  is a prefix of another string in  $S$ . For any partial function  $f$ , the domain of definition of  $f$  is denoted by  $\text{dom } f$ . We write “r.e.” instead of “recursively enumerable.” We say that a real  $\alpha$  is *computable* if there exists a computable sequence  $\{a_n\}_{n \in \mathbb{N}}$  of rationals such that  $|\alpha - a_n| < 2^{-n}$  for all  $n \in \mathbb{N}$ . For any  $\alpha \in \mathbb{R}$  and  $n \in \mathbb{N}^+$ , we denote by  $\alpha \upharpoonright n \in \{0, 1\}^*$  the first  $n$  bits of the base-two expansion of  $\alpha - \lfloor \alpha \rfloor$  with infinitely many zeros, where  $\lfloor \alpha \rfloor$  is the greatest integer less than or equal to  $\alpha$ . For example, in the case of  $\alpha = 5/8$ ,  $\alpha \upharpoonright 6 = 101000$ .

In what follows we concisely review some definitions of algorithmic information theory (AIT, for short). For the detail, see [3,4]. A *computer* is a partial recursive function  $C: \{0, 1\}^* \rightarrow \{0, 1\}^*$  such that  $\text{dom } C$  is a prefix-free set. For any computer  $C$  and any  $s \in \{0, 1\}^*$ ,  $H_C(s)$  is defined by  $H_C(s) = \min\{|p| \mid p \in \{0, 1\}^* \ \& \ C(p) = s\}$  (may be  $\infty$ ). A computer  $U$  is said to be *optimal* if for each computer  $C$  there exists  $d \in \mathbb{N}$  such that, for every  $s \in \{0, 1\}^*$ ,  $H_U(s) \leq H_C(s) + d$ . There exists an optimal computer. We choose a particular optimal computer  $U$  as the standard one for use, and define  $H(s)$  as  $H_U(s)$ , which is referred to as the *program-size complexity* of  $s$ .

Let  $T$  be an arbitrary real with  $0 < T \leq 1$ . In the works [6,7], we introduced several notions of the partial randomness of a real by parameterizing the notions of randomness of a real by a real  $T$ , as follows. Let  $\alpha \in \mathbb{R}$ . We say that  $\alpha$  is *weakly Chaitin  $T$ -random* if there exists  $c \in \mathbb{N}$  such that  $Tn - c \leq H(\alpha \upharpoonright n)$  for all  $n \in \mathbb{N}^+$ . On the other hand, we say that  $\alpha$  is  *$T$ -compressible* if  $H(\alpha \upharpoonright n) \leq Tn + o(n)$ . Thus, if  $\alpha$  is weakly Chaitin  $T$ -random and  $T$ -compressible, then  $\lim_{n \rightarrow \infty} H(\alpha \upharpoonright n)/n = T$ , i.e., the *compression rate* of  $\alpha$  equals to  $T$ . Furthermore, we say that  $\alpha$  is *Chaitin  $T$ -random* if  $\lim_{n \rightarrow \infty} H(\alpha \upharpoonright n) - Tn = \infty$ . Obviously, if  $\alpha$  is Chaitin  $T$ -random, then  $\alpha$  is weakly Chaitin  $T$ -random. However, in 2005 Reimann and Stephan [5] showed that, in the case of  $T < 1$ , the converse does not necessarily hold.

## 2. Thermodynamic quantities in AIT

We start a statistical mechanical interpretation of AIT by introducing the notion of thermodynamic quantities into AIT in the following manner.

In statistical mechanics, the partition function  $Z_{\text{sm}}(T)$ , free energy  $F_{\text{sm}}(T)$ , energy  $E_{\text{sm}}(T)$ , entropy  $S_{\text{sm}}(T)$ , and specific heat  $C_{\text{sm}}(T)$  at temperature  $T$  are given as follows (**Eq. 1**)

$$\begin{aligned} Z_{\text{sm}}(T) &= \sum_{x \in X} e^{-\frac{E_x}{k_{\text{B}}T}}, & F_{\text{sm}}(T) &= -k_{\text{B}}T \ln Z_{\text{sm}}(T), \\ E_{\text{sm}}(T) &= \frac{1}{Z_{\text{sm}}(T)} \sum_{x \in X} E_x e^{-\frac{E_x}{k_{\text{B}}T}}, & S_{\text{sm}}(T) &= \frac{E_{\text{sm}}(T) - F_{\text{sm}}(T)}{T}, \\ C_{\text{sm}}(T) &= \frac{d}{dT} E_{\text{sm}}(T), \end{aligned}$$

where  $X$  is a complete set of energy eigenstates of a quantum system and  $E_x$  is the energy of an energy eigenstate  $x$ . The constant  $k_{\text{B}}$  is called the Boltzmann Constant, and the  $\ln$  denotes the natural logarithm.<sup>4</sup>

We introduce the notion of thermodynamic quantities into AIT by performing the following replacements below for the thermodynamic quantities (Eq. 1) in statistical mechanics.

**Replacements:**

1. Replace the complete set  $X$  of energy eigenstates  $x$  by the set  $\text{dom } U$  of all programs  $p$  for  $U$ .
2. Replace the energy  $E_x$  of an energy eigenstate  $x$  by the length  $|p|$  of a program  $p$ .
3. Set the Boltzmann Constant  $k_{\text{B}}$  to  $1/\ln 2$ .

For that purpose, we first choose a particular enumeration  $p_1, p_2, p_3, p_4, \dots$  of the countably infinite set  $\text{dom } U$ .<sup>5</sup> Then, motivated by the formulae (Eq. 1) and taking into account Replacements 1, we introduce the notion of thermodynamic quantities into AIT as follows.

**Definition 2.1** (thermodynamic quantities in AIT, [8]). Let  $T$  be any real with  $T > 0$ .

1. The *partition function*  $Z(T)$  at temperature  $T$  is defined as  $\lim_{k \rightarrow \infty} Z_k(T)$  where

$$Z_k(T) = \sum_{i=1}^k 2^{-\frac{|p_i|}{T}}.$$

2. The *free energy*  $F(T)$  at temperature  $T$  is defined as  $\lim_{k \rightarrow \infty} F_k(T)$  where

$$F_k(T) = -T \log_2 Z_k(T).$$

<sup>4</sup>For the thermodynamic quantities in statistical mechanics, see e.g. Chapter 16 of [1] and Chapter 2 of [10]. To be precise, the partition function is not a thermodynamic quantity but a statistical mechanical quantity.

<sup>5</sup>The enumeration  $\{p_i\}$  can be chosen quite arbitrarily and not necessarily to be recursive. The results of this paper is independent of the choice of  $\{p_i\}$ .

3. The *energy*  $E(T)$  at temperature  $T$  is defined as  $\lim_{k \rightarrow \infty} E_k(T)$  where

$$E_k(T) = \frac{1}{Z_k(T)} \sum_{i=1}^k |p_i| 2^{-\frac{|p_i|}{T}}.$$

4. The *statistical mechanical entropy*  $S(T)$  at temperature  $T$  is defined as  $\lim_{k \rightarrow \infty} S_k(T)$  where

$$S_k(T) = \frac{E_k(T) - F_k(T)}{T}.$$

5. The *specific heat*  $C(T)$  at temperature  $T$  is defined as  $\lim_{k \rightarrow \infty} C_k(T)$  where  $C_k(T) = E'_k(T)$ , the derived function of  $E_k(T)$ .

Note that  $Z(1)$  is precisely a Chaitin  $\Omega$  number, in particular. Then Theorems 2.2 and 2.3 below hold for these thermodynamic quantities in AIT.

**Theorem 2.2** (properties of  $Z(T)$  and  $F(T)$ , [6,7,8]). *Let  $T \in \mathbb{R}$ .*

1. *If  $0 < T \leq 1$  and  $T$  is computable, then each of  $Z(T)$  and  $F(T)$  converges and is weakly Chaitin  $T$ -random and  $T$ -compressible.*
2. *If  $1 < T$ , then  $Z(T)$  and  $F(T)$  diverge to  $\infty$  and  $-\infty$ , respectively.*

**Theorem 2.3** (properties of  $E(T)$ ,  $S(T)$ , and  $C(T)$ , [8]). *Let  $T \in \mathbb{R}$ .*

1. *If  $0 < T < 1$  and  $T$  is computable, then each of  $E(T)$ ,  $S(T)$ , and  $C(T)$  converges and is Chaitin  $T$ -random and  $T$ -compressible.*
2. *If  $1 \leq T$ , then both  $E(T)$  and  $S(T)$  diverge to  $\infty$ . In the case of  $T = 1$ ,  $C(T)$  diverges to  $\infty$ .<sup>6</sup>*

The above two theorems show that if  $T$  is a computable real with  $T \in (0, 1)$  then the temperature  $T$  equals to the partial randomness (and therefore the compression rate) of the values of all the thermodynamic quantities in Definition 2.1. These theorems also show that the values of the thermodynamic quantities: partition function, free energy, energy, and statistical mechanical entropy diverge in the case of  $T > 1$ . This phenomenon might be regarded as some sort of phase transition in statistical mechanics. Note that the weak Chaitin  $T$ -randomness in (i) of Theorems 2.2 is strengthened to the Chaitin  $T$ -randomness in (i) of Theorems 2.3.

### 3. Fixed point theorems on partial randomness

In statistical mechanics or thermodynamics, among all thermodynamic quantities one of the most typical thermodynamic quantities is temperature itself. Inspired by this fact in physics and the observation in the previous section that the temperature  $T$  equals to the partial randomness of the values of the thermodynamic quantities in the statistical mechanical interpretation of AIT, the following question arises naturally:

Can the partial randomness of the temperature  $T$  equal to the temperature  $T$  itself in the statistical mechanical interpretation of AIT?

---

<sup>6</sup>It is still open whether  $C(T)$  diverges or not in the case of  $T > 1$ .

This question is rather self-referential. However, we can answer it affirmatively in the following form.

**Theorem 3.1** (fixed point theorem by partition function, [8]). *For every  $T \in (0, 1)$ , if  $Z(T)$  is computable, then  $T$  is weakly Chaitin  $T$ -random and  $T$ -compressible, and therefore*

$$\lim_{n \rightarrow \infty} \frac{H(T \upharpoonright n)}{n} = T.$$

Theorem 3.1 is just a *fixed point theorem on partial randomness*, where the computability of the value  $Z(T)$  gives a sufficient condition for a real  $T \in (0, 1)$  to be a fixed point on partial randomness. Thus, the above observation that the temperature  $T$  equals to the partial randomness of the values of the thermodynamic quantities in the statistical mechanical interpretation of AIT is further confirmed. In addition, we can show that fixed point theorems of the same form as Theorem 3.1 hold also for the free energy  $F(T)$ , energy  $E(T)$ , and statistical mechanical entropy  $S(T)$ , as follows.<sup>7</sup> Thus we can confirm the above observation much further.

**Theorem 3.2** (fixed point theorem by free energy, [9]). *For every  $T \in (0, 1)$ , if  $F(T)$  is computable then  $T$  is weakly Chaitin  $T$ -random and  $T$ -compressible.*

**Theorem 3.3** (fixed point theorem by energy, [9]). *For every  $T \in (0, 1)$ , if  $E(T)$  is computable then  $T$  is Chaitin  $T$ -random and  $T$ -compressible.*

**Theorem 3.4** (fixed point theorem by statistical mechanical entropy, [9]). *For every  $T \in (0, 1)$ , if  $S(T)$  is computable then  $T$  is Chaitin  $T$ -random and  $T$ -compressible.*

Note that the weak Chaitin  $T$ -randomness of  $T$  in Theorems 3.1 is strengthened to the Chaitin  $T$ -randomness of  $T$  in Theorems 3.3 and 3.4.

We can show the following theorem for the sufficient condition of Theorem 3.1. The exactly same theorem holds for each of  $F(T)$ ,  $E(T)$ , and  $S(T)$  also [9].

**Theorem 3.5** [8]. *The set  $\{T \in (0, 1) \mid Z(T) \text{ is computable}\}$  is dense in  $(0, 1)$ .*

Using the “statistical mechanical” relation  $F(T) = -T \log_2 Z(T)$  we can show Theorem 3.6 below. Thus, the computability of  $F(T)$  gives completely different fixed points from the computability of  $Z(T)$ . This implies that neither the computability of  $Z(T)$  nor the computability of  $F(T)$  is the necessary condition for  $T \in (0, 1)$  to be a fixed point on partial randomness at all.

**Theorem 3.6** [9]. *There does not exist  $T \in (0, 1)$  such that both  $Z(T)$  and  $F(T)$  are computable.*

---

<sup>7</sup>It is still open whether fixed point theorem of the same form as Theorem ?? holds for the specific heat  $C(T)$  or not.

- [1] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. John Wiley & Sons, Inc., Singapore, 1985.
- [2] C. S. Calude and M. A. Stay, "Natural halting probabilities, partial randomness, and zeta functions," *Inform. and Comput.*, vol. 204, pp. 1718–1739, 2006.
- [3] G. J. Chaitin, "A theory of program size formally identical to information theory," *J. Assoc. Comput. Mach.*, vol. 22, pp. 329–340, 1975.
- [4] G. J. Chaitin, *Algorithmic Information Theory*. Cambridge University Press, Cambridge, 1987.
- [5] J. Reimann and F. Stephan, On hierarchies of randomness tests. Proceedings of the 9th Asian Logic Conference, World Scientific Publishing, August 16-19, 2005, Novosibirsk, Russia.
- [6] K. Tadaki, Algorithmic information theory and fractal sets. Proceedings of 1999 Workshop on Information-Based Induction Sciences (IBIS'99), pp. 105–110, August 26-27, 1999, Syuzenji, Shizuoka, Japan. In Japanese.
- [7] K. Tadaki, "A generalization of Chaitin's halting probability  $\Omega$  and halting self-similar sets," *Hokkaido Math. J.*, vol. 31, pp. 219–253, 2002.
- [8] K. Tadaki, A statistical mechanical interpretation of algorithmic information theory. Local Proceedings of Computability in Europe 2008 (CiE 2008), pp. 425–434, June 15-20, 2008, University of Athens, Greece. Extended and Electronic Version Available: <http://arxiv.org/abs/0801.4194v1>
- [9] K. Tadaki, Fixed point theorems on partial randomness. Proceedings of the Symposium on Logical Foundations of Computer Science 2009 (LFCS'09), Lecture Notes in Computer Science, Springer-Verlag, vol. 5407, pp. 422–440, January 3-6, 2009. Extended Version Available: <http://arxiv.org/abs/0903.3433>
- [10] M. Toda, R. Kubo, and N. Saitô, *Statistical Physics I. Equilibrium Statistical Mechanics*, 2nd ed. Springer, Berlin, 1992.

\* \* \*

**Hayato Takahashi** (The University of Tokyo, Japan).

*Algorithmic randomness and monotone complexity on product space*

Let  $\Omega$  be the set of infinite binary sequences with product topology and  $S$  be the set of finite binary strings. Let  $P$  be a computable probability on  $X \times Y$ , where  $X = Y = \Omega$ , and  $P_X$  and  $P_Y$  be its marginal distributions. Let  $\mathbb{R}^P, \mathbb{R}^{P_X}, \mathbb{R}^{P_Y}$  be the set of random sequences (points) with respect to  $P, P_X, P_Y$  in the sense of Martin-Löf [1], respectively. For  $x, y \in S$ , set  $\Delta(x) := \{x\omega \mid \omega \in \Omega\}$ , where  $x\omega$  is the concatenation of  $x$  and  $\omega$ , and  $\Delta(x, y) := \Delta(x) \times \Delta(y)$ . Let  $P(x, y) := P(\Delta(x, y))$ ,  $P_X(x) := P(\Delta(x))$ , and  $P_Y(y) := P(\Delta(y))$  for  $x, y \in S$ . Then  $P_X(x) = P(x, \lambda)$  and  $P_Y(y) = P(\lambda, y)$ , where  $\lambda$  is the empty word and  $\Delta(\lambda) = \Omega$ .

Let

$$P(x|y) := \begin{cases} \frac{P(x,y)}{P_Y(y)}, & \text{if } P_Y(y) > 0 \\ 0, & \text{if } P_Y(y) = 0 \end{cases},$$

and

$$P(x|y^\infty) := \lim_{y \rightarrow y^\infty} P(x|y),$$

for  $y^\infty \in \Omega$  if the right-hand side exists.

In [2], it is shown that if  $y^\infty \in \mathbb{R}^{P_Y}$  then  $P(x|y^\infty)$  exists and  $P(\cdot|y^\infty)$  is a probability measure on  $\Omega$ . If  $P(\cdot|y^\infty)$  is computable relative to  $y^\infty$ , then let

$\mathbb{R}^{P(\cdot|y^\infty), y^\infty}$  be the set of random sequences with respect to  $P(\cdot|y^\infty)$  relative to  $y^\infty$ . Let  $x^\infty, y^\infty \in \Omega$  and  $\mathbb{R}_{y^\infty}^P$  be the section of  $\mathbb{R}^P$  at  $y^\infty$ , i.e.,  $\mathbb{R}_{y^\infty}^P = \{(x^\infty, y^\infty) | (x^\infty, y^\infty) \in \mathbb{R}^P\}$ . In [2], it is shown that  $\mathbb{R}^{P(\cdot|y^\infty), y^\infty} \subset \mathbb{R}_{y^\infty}^P$ , and under uniform computability (see [2]),  $\mathbb{R}^{P(\cdot|y^\infty), y^\infty} = R_{y^\infty}^P$  for  $y^\infty \in \mathbb{R}^{P_Y}$ .

Let  $\mathbb{N}$  and  $\mathbb{Q}^+$  be the set of natural numbers and positive rational numbers, respectively. In the following, for  $A \subset S^2$ , we set  $\tilde{A} = \cup_{a \in A} \Delta(a)$ .

**Proposition 1.** Assume that  $y^\infty \in \mathbb{R}^{P_Y}$  and  $P(\cdot|y^\infty)$  is computable relative to  $y^\infty$ , then  $\mathbb{R}^{P(\cdot|y^\infty), y^\infty} = R_{y^\infty}^P$ .

(Outline of Proof) We prove the proposition by extending  $P(\cdot|y^\infty)$  to a finite measure  $P'$  on  $\Omega^2$ . For simplicity, we assume that  $P(x|y^\infty) > 0$  for all  $x$ . Fix  $y^\infty \in \mathbb{R}_Y^P$ . Since  $P(\cdot|y^\infty)$  is computable relative to  $y^\infty$ , there is a partial computable function  $A : S \times S \times \mathbb{N} \rightarrow \mathbb{Q}^+ \cup \{0\}$  such that  $\forall x, k \exists y \sqsubset y^\infty, |P(x|y^\infty) - A(x, y, k)| < \frac{1}{k}$ , where we write  $x \sqsubset y$  if  $x$  is a prefix of  $y$ . If  $A(x, y, k)$  is defined then  $A(x, y, k) = A(x, z, k)$  for all  $y \sqsubset z$ . Similarly, let  $U^{y^\infty} \subset \mathbb{N} \times S$  be a Martin-Löf test with respect to  $P(\cdot|y^\infty)$  relative to  $y^\infty$ , i.e.,  $U^{y^\infty}$  is a recursively enumerable (r.e.) set relative to  $y^\infty$ , and  $P(\tilde{U}_n^{y^\infty} | y^\infty) < 2^{-n}$  for all  $n$ , where  $U_n^{y^\infty} := \{x | (n, x) \in U^{y^\infty}\}$ . Then there is a partial computable function  $B : \mathbb{N} \times \mathbb{N} \times S \rightarrow S$  such that  $\forall n, U_n^{y^\infty} = \{x | \exists i, y \sqsubset y^\infty, B(i, n, y) = x\}$ . If  $B(i, n, y)$  is defined then  $B(i, n, y) = B(i, n, z)$  for all  $y \sqsubset z$ .

Let  $U_n := \{(x, y) | \exists i, B(i, n, y) = x\}$ . Let  $U'_n$  be a r.e. set such that  $\tilde{U}_n = \tilde{U}'_n$  and  $U'_n$  is prefix-free, i.e., if  $(x, y), (x', y') \in U'_n$  and  $(x, y) \neq (x', y')$  then  $\Delta(x, y) \cap \Delta(x', y') = \emptyset$ . Let  $V_n := \{(x, y, k) | \frac{1}{k} < \frac{1}{2}A(x, y, k), (x, y) \in U'_n\}$ , then  $V_n$  is a r.e. set. From  $V_n$ , we can construct a r.e. set  $W_n \subset S \times S \times \mathbb{N}$  that satisfies (1), (2), (3), (4), and (5):

$$\tilde{W}'_n \subset \tilde{V}'_n \quad (1)$$

where  $W'_n := \{(x, y) | \exists k, (x, y, k) \in W_n\}$  and  $V'_n := \{(x, y) | \exists k, (x, y, k) \in V_n\}$ .

$$(x, y, k), (x, y, k') \in W_n \Rightarrow k = k' \quad (2)$$

$$(x, y, k) \in W_n \Rightarrow \forall z, y \sqsubset z, (x, z, k) \in W_n \quad (3)$$

For all  $z^\infty \in \Omega$  and for any  $A(z^\infty) \subset \{(x, y, k) | y \sqsubset z^\infty, (x, y, k) \in W_n\}$ , if  $A(z^\infty)$  consists of disjoint elements, i.e.,  $(x, y, k), (x', y', k') \in A(z^\infty), (x, y, k) \neq (x', y', k') \Rightarrow x \neq x'$ , then

$$\forall z^\infty \in \Omega, \sum_{(x, y, k) \in A(z^\infty)} A(x, y, k) < 2^{-n+1} \quad (4)$$

$$\tilde{U}_n^{y^\infty} = \tilde{W}'_{n, y^\infty} \quad (5)$$

where  $W'_{n, y^\infty} = \{x | (x, y) \in W'_n, y \sqsubset y^\infty\}$ .

Now let  $P'(x, y) := A(x, y, k)P_Y(y)$  for  $(x, y, k) \in W_n$  and  $P'(x, y) := 0$  for  $(x, y)$  such that  $\Delta(x, y) \cap \tilde{W}'_n = \emptyset$ . Then by (d),  $P'(\tilde{W}'_n) < 2^{-n+1}$ .

Finally let  $W''_n := \{(x, y) \in W'_n | P(x, y) < \frac{3}{2}P'(x, y)\}$ . Then  $W''_n$  is a r.e. set, and  $P(\tilde{W}''_n) < \frac{3}{2}P'(\tilde{W}'_n) < 2^{-n+2}$ . Since  $P(x|y) \rightarrow P(x|y^\infty)$  as  $y \rightarrow y^\infty$  for  $y^\infty \in \mathbb{R}^{P_Y}$  and  $\frac{1}{2}A(x, y, k) < P(x|y^\infty) < \frac{3}{2}A(x, y, k)$  for  $(x, y, k) \in V_n, y \sqsubset y^\infty$ , by (c) and (e), we have

$$\tilde{U}_n^{y^\infty} = \tilde{W}''_{n, y^\infty},$$

where  $W''_{n,y^\infty} = \{x|(x,y) \in W''_n, y \sqsubseteq y^\infty\}$ . We see that  $\limsup_n \tilde{W}''_n \subset (\mathbb{R}^P)^c$  and  $R''_{y^\infty} \subset \mathbb{R}^{P(\cdot|y^\infty),y^\infty}$ . The converse inclusion is shown in [2].  $\square$

Next we introduce monotone complexity on product space. In the following, we use bold-faced symbols such as 1)  $\mathbf{x}^\infty, \mathbf{y}^\infty$  to denote an element of  $\Omega^k$ , 2)  $\mathbf{x}, \mathbf{y}, \mathbf{s}$  to denote an element of  $S^k$  or  $(S \cup \Omega)^k$  (we will specify which space we consider), and 3)  $\boldsymbol{\lambda}$  to denote  $(\lambda, \dots, \lambda) \in S^k$  for  $k \geq 1$ , and  $\Delta(\boldsymbol{\lambda}) = \Omega^k$ . Further, we write  $P(\mathbf{x})$  for  $P(\Delta(\mathbf{x}))$ .

Let  $|s|$  be the length of  $s \in S$ .  $|\lambda| = 0$ , where  $\lambda$  is the empty word, and  $|x^\infty| = \infty$  for  $x^\infty \in \Omega$ . For  $\mathbf{s} = (s^1, \dots, s^k) \in (S \cup \Omega)^k$ , set

$$|\mathbf{s}| := |s^1| + \dots + |s^k|.$$

For  $x^\infty \in \Omega$ , set  $\Delta(x^\infty) := \{x^\infty\}$ , and for  $\mathbf{x} = (x^1, \dots, x^k) \in (S \cup \Omega)^k$ , set  $\Delta(\mathbf{x}) := \Delta(x^1) \times \dots \times \Delta(x^k)$ . For  $\mathbf{x}, \mathbf{y} \in (S \cup \Omega)^k$ , we write  $\mathbf{x} \sqsubseteq \mathbf{y}$  if  $\Delta(\mathbf{x}) \supset \Delta(\mathbf{y})$ .  $\mathbf{x}$  and  $\mathbf{y}$  are called comparable if  $\mathbf{x} \sqsubseteq \mathbf{y}$  or  $\mathbf{y} \sqsubseteq \mathbf{x}$ . Then  $((S \cup \Omega)^k, \sqsubseteq)$  is partial order. Let  $A \subset S^k$ . The least upper bound of  $A$  is denoted by  $\bigvee A$ . The  $\bigvee A$  exists iff  $\bigcap_{\mathbf{x} \in A} \Delta(\mathbf{x}) \neq \emptyset$ . Note that if  $\Delta(\mathbf{x}) \cap \Delta(\mathbf{y}) \neq \emptyset$  then there is  $\mathbf{z}$  such that  $\Delta(\mathbf{x}) \cap \Delta(\mathbf{y}) = \Delta(\mathbf{z})$ . Thus if  $\bigcap_{\mathbf{x} \in A} \Delta(\mathbf{x}) \neq \emptyset$ , there is  $\mathbf{y} \in (S \cup \Omega)^k$  such that  $\bigcap_{\mathbf{x} \in A} \Delta(\mathbf{x}) = \Delta(\mathbf{y})$  and  $\bigvee A = \mathbf{y}$ .

Let  $F \subset S^j \times S^k$  and  $F_{\mathbf{s}} := \{\mathbf{x} | (\mathbf{s}, \mathbf{x}) \in F\}$ . Assume that:

a0)  $\forall \mathbf{s} \in S^j, \boldsymbol{\lambda} \in F_{\mathbf{s}}$ .

a1)  $\forall \mathbf{s} \in S^j, \bigvee_{\mathbf{s}' \sqsubseteq \mathbf{s}} F_{\mathbf{s}'}$  exists, i.e.,  $\bigcap_{\mathbf{x} \in \bigcup_{\mathbf{s}' \sqsubseteq \mathbf{s}} F_{\mathbf{s}'}} \Delta(\mathbf{x}) \neq \emptyset$ .

Set

$$f(\mathbf{s}) := \bigvee_{\mathbf{s}' \sqsubseteq \mathbf{s}, \mathbf{s}' \in S^j} F_{\mathbf{s}'}, \text{ for } \mathbf{s} \in (S \cup \Omega)^j \quad (6)$$

We see that  $f : (S \cup \Omega)^j \rightarrow (S \cup \Omega)^k$  and  $f$  is monotonically increasing, i.e.,  $\mathbf{s}' \sqsubseteq \mathbf{s} \Rightarrow f(\mathbf{s}') \sqsubseteq f(\mathbf{s})$ .

Conversely, let  $f : (S \cup \Omega)^j \rightarrow (S \cup \Omega)^k$  be a monotonically increasing function, and set

$$F := \{(\mathbf{s}, \mathbf{x}) \in S^j \times S^k | \mathbf{x} \sqsubseteq f(\mathbf{s})\}.$$

Then  $\bigvee F_{\mathbf{s}} = f(\mathbf{s})$ ,  $F$  satisfies a0 and a1, and the function defined by  $F$  coincides with  $f$ . If  $F$  is a r.e. set that satisfies a0 and a1, then the function  $f$  defined by (6) is called *computable monotone function*.

The monotone complexity with respect to computable monotone function  $f : (S \cup \Omega)^{k+j} \rightarrow (S \cup \Omega)^k$  is defined as follows:

$$Km_f(\mathbf{x}|\mathbf{y}) := \min\{|\mathbf{p}| | \mathbf{x} \sqsubseteq f(\mathbf{p}, \mathbf{y})\},$$

where  $\mathbf{p} \in (S \cup \Omega)^k$ ,  $\mathbf{y} \in (S \cup \Omega)^j$ , and  $\mathbf{x} \in (S \cup \Omega)^k$ . If there is no  $\mathbf{p}$  such that  $\mathbf{x} \sqsubseteq f(\mathbf{p}, \mathbf{y})$ , then  $Km_f(\mathbf{x}|\mathbf{y}) := \infty$ . In the following, we fix an optimal function  $u$  and let  $Km(\mathbf{x}|\mathbf{y}) := Km_u(\mathbf{x}|\mathbf{y})$  and  $Km(\mathbf{x}) := Km_u(\mathbf{x})$ . By definition, we have

**Proposition 2.** 1) Monotonicity:  $\mathbf{x} \sqsubseteq \mathbf{z} \Rightarrow Km(\mathbf{x}|\mathbf{y}) \leq Km(\mathbf{z}|\mathbf{y})$ , and  $\mathbf{y} \sqsubseteq \mathbf{z} \Rightarrow Km(\mathbf{x}|\mathbf{y}) \geq Km(\mathbf{x}|\mathbf{z})$ .

2) Kraft inequality:  $\forall \mathbf{y}, \sum_{\mathbf{x} \in \mathcal{A}} 2^{-Km(\mathbf{x}|\mathbf{y})} \leq 1$  for prefix-free set  $\mathcal{A} \subset (S \cup \Omega)^k$ , where  $\mathcal{A}$  is called *prefix-free* if  $\Delta(\mathbf{x}) \cap \Delta(\mathbf{y}) = \emptyset$  for  $\mathbf{x}, \mathbf{y} \in \mathcal{A}, \mathbf{x} \neq \mathbf{y}$ .

3) Conditional sub-additivity:  $\exists c \forall \mathbf{x} \in S^k, \mathbf{y} \in S^j, Km(\mathbf{x}, \mathbf{y}) \leq Km(\mathbf{x}|\mathbf{y}) + Km(\mathbf{y}) + c$ .

For  $\mathcal{A} \subset S^k$ , we assume that:

Condition A) if  $\mathbf{x}, \mathbf{y} \in \mathcal{A}$  then,  $\mathbf{x}$  and  $\mathbf{y}$  are comparable or  $\Delta(\mathbf{x}) \cap \Delta(\mathbf{y}) = \emptyset$ .

Condition B) there is a recursive function  $f : S^k \times \mathbb{N} \rightarrow \mathcal{A}$  such that for any  $\mathbf{x} \in S^k$ ,  $\Delta(\mathbf{x}) = \widetilde{f(\mathbf{x}, \mathbb{N})}$  and  $f(\mathbf{x}, \mathbb{N})$  is prefix-free.

**Proposition 3.** *Levin-Schnorr theorem on product space* Let  $P$  be a computable probability on  $\Omega^k$ . Let  $\mathcal{A}(\mathbf{x}^\infty) := \{\mathbf{x} \in S^k \mid \mathbf{x} \in \mathcal{A}, \mathbf{x} \sqsubset \mathbf{x}^\infty\}$ . If a r.e. set  $\mathcal{A} \subset S^k$  satisfies conditions A and B, then  $\mathbf{x}^\infty \in \mathbb{R}^P$  iff  $\sup_{\mathbf{x} \in \mathcal{A}(\mathbf{x}^\infty)} -\log P(\mathbf{x}) - Km(\mathbf{x}) < \infty$ .

[1] P. Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–609, 1966.

[2] H. Takahashi. On a definition of random sequences with respect to conditional probability. *Inform. and Compt.*, 206:1375–1382, 2008.

\* \* \*

## Nikolay Vereshchagin (Moscow State University, Russia). On Algorithmic Sufficient Statistics

Let  $x$  be a binary string. A finite set  $A$  containing  $x$  is called an (algorithmic) sufficient statistic of  $x$  if the sum of Kolmogorov complexity of  $A$  and the log-cardinality of  $A$  is close to Kolmogorov complexity  $C(x)$  of  $x$ :

$$C(A) + \log_2 |A| \approx C(x)$$

More specifically we call  $A$  a  $c$ -sufficient statistic of  $x$  if  $C(A) + \log_2 |A| \leq C(x) + c$ . It is straightforward that  $A$  is a sufficient statistic of  $x$  iff  $C(A|x) \approx 0$  and  $C(x|A) \approx \log |A|$ . A sufficient statistic is called minimal if  $C(A)$  is minimal [1]. Minimal sufficient statistics (MSS) are often considered as containing all useful information from  $x$ , a sort of “denoised  $x$ ” [1,3] We argue here that (1) this notion is not well defined for some  $x$  (although for other  $x$  the notion is well defined) and (2) even for those  $x$  for which the notion of a MSS is well defined not every MSS qualifies for a denoised version of  $x$ . We propose a new definition of a (minimal) sufficient statistic that has better properties.

Why the notion of a MSS is not well defined? It is because there are strings such that the minimal complexity of a  $c$ -sufficient statistic is not a smooth function of  $x$ . For example, fix a large  $n$  and let, say,  $k = n/2$ . Then for all  $n$  there is a string  $x$  of length  $n$  and complexity  $k + O(\log n)$  such that the complexity of minimal  $c$ -sufficient statistic decreases fast, as  $c$  increases from  $O(\log n)$  to  $\sqrt{k}$ : the complexity of minimal  $c$ -sufficient statistic is  $k - c^2 + O(\log n)$ . The reader may consider the values of order  $\sqrt{k}$  as negligible (compared to  $n$ ) and thus think that the complexity of MSS is negligible for this string. However in a similar way one can construct a string  $x$  of length  $n$  and complexity  $k + O(\log n)$  such that the complexity of minimal  $c$ -sufficient statistic is about  $k - c \log c$ , or  $k - 10c$ . All these claims are a corollary of a result from [3].

Thus there are strings for which it is hard to identify the complexity of MSS. There is also another minor point regarding minimal sufficient statistics. Namely, there is a string  $x$  for which the complexity of minimal sufficient

statistic is well defined but there are quite different MSS of  $x$ . Namely, let  $y$  be a string,  $k$  a natural number and  $z$  a string of length  $l(z) = k$  that is random relative to  $y$  (that is,  $C(z|y) \geq k$ ). Consider the string  $x = \langle y, z \rangle$ . Intuitively,  $z$  is a noise in  $x$  and we expect all MSS of  $x$  be similar to  $y$ . However, it turns out that some  $x$  of this form may have a weird MSS.

To ban such weird MSS we propose to strengthen the definition of sufficient statistic as follows. Call  $A$  a strongly sufficient statistic of  $x$  if  $x \in A$ ,  $C(x|A) \approx \log |A|$  and there is a short *total* program  $p$  such that  $p(A) = x$  (note that the equality  $C(A|x) \approx 0$  means that there is a short program  $p$  with  $p(x) = A$  but  $p$  is not necessarily total). Call  $u$  and  $v$  equivalent if there is a short total program  $p$  with  $p(0u) = v$  and  $p(1v) = u$ . One can prove the following

- if  $A$  is a strongly sufficient statistic of  $x$  then  $x$  is equivalent to  $\langle A, z \rangle$  where  $z$  is a  $\log |A|$ -bit index of  $x$  in  $A$  (which is random relative to  $A$ ).
- the Kolmogorov structure function [2,3] of any such  $\langle A, z \rangle$  is identified by that of  $A$  and the length of  $z$ .
- if the notion of MSS is well defined for  $x$ , both  $A, B$  are MSS of  $x$  and both  $A, B$  are strongly sufficient then  $A$  and  $B$  are equivalent.

[1] P. Gács, J. Tromp, P.M.B. Vitányi. Algorithmic statistics, *IEEE Trans. Inform. Th.*, 47:6(2001), 2443–2463.

[2] A.N. Kolmogorov, Talk at the Information Theory Symposium in Tallinn, Estonia, 1974.

[3] N.K. Vereshchagin and P.M.B. Vitányi, Kolmogorov’s structure functions and model selection, *IEEE Trans. Information Theory*, 50:12 (2004) 3265-3290.

\* \* \*

**Paul Vitanyi** (CWI, Amsterdam, The Netherlands).

*Positive and Negative Randomness*

Every finite binary string can be represented by its shortest program. Such a shortest program is incompressible and hence random in the sense of Martin-Lof. If we flip a fair coin  $n$  times, then the resultant string is random of length  $n$ . We can view the string as consisting of its probabilistic cause which is very simple and takes a constant number of bits and the remaining randomness. Kolmogorov in 1974 called such a string ‘positively’ random since the cause is a simple probabilistic process. It turns out that there are incompressible strings that can be divided in a cause (the meaningful information or law) and the remaining randomness. If the meaningful information is large, then we call such a string ‘negatively’ random. Although the string is random, most of it is meaningful information or cause. In extreme cases, almost all of the string is a cause. We discuss the theory and results.

[1] P. Gács, J. Tromp, P.M.B. Vitányi. Algorithmic statistics, *IEEE Trans. Inform. Th.*, 47:6(2001), 2443–2463.

[2] A.N. Kolmogorov. Complexity of Algorithms and Objective Definition of Randomness. A talk at Moscow Math. Soc. meeting 4/16/1974. An abstract available in *Uspekhi Mat. Nauk* 29:4(1974),155

[3] A.Kh. Shen, The concept of  $(\alpha, \beta)$ -stochasticity in the Kolmogorov sense, and its properties, *Soviet Math. Dokl.*, 28:1(1983), 295–299.

[4] N.K. Vereshchagin and P. Vitányi, Kolmogorov’s structure functions and model selection, *IEEE Trans. Inform. Theory*, 50:12(2004), 3265–3290.

[5] P.M.B. Vitányi, Meaningful information, *IEEE Trans. Inform. Th.*, 52:10(2006), 4617–4626.

\* \* \*

**Vladimir V’yugin** (Institute for Information Transmission Problems, Moscow, Russia).

### *Coding-invariant Classification of Infinite Sequences*

In this talk we discuss some coding-invariant classification of infinite sequences [1].

An infinite binary sequence  $\alpha$  is reducible to an infinite binary sequence  $\beta$  if  $\alpha = F(\beta)$  for some computable operation  $F$  (denote this  $\alpha \prec \beta$ ). Two sequences  $\alpha$  and  $\beta$  are information equivalent (denote this  $\alpha \equiv \beta$ ), if  $\alpha \prec \beta$  and  $\beta \prec \alpha$ . Informally speaking, the sequences  $\alpha \equiv \beta$  have the same information content (up to finite descriptions of the coding methods).

Let  $\Omega$  be the set of all infinite binary sequences. A set  $A \subseteq \Omega$  is called *coding invariant* if  $\omega \in A$  and  $\alpha \equiv \omega$  implies  $\alpha \in A$  for any  $\omega$  and  $\alpha$ .

We use the concepts of semicomputable semimeasure (first introduced in [2], see also [1]) and the *a priori* semimeasure  $M$ . For any such semimeasure  $P$  the maximal measure  $\bar{P}$  can be defined such that  $\bar{P} \leq P$ .

Let  $I$  be the Boolean algebra of all coding invariant Borel subsets of  $\Omega$ , and let  $\Upsilon$  be the factor-algebra of  $I$  with respect to the equivalence relation

$$A \sim B \iff \bar{M}((A \setminus B) \cup (B \setminus A)) = 0,$$

where  $A, B \in I$ .

Let  $\mathbf{a} = [A]$  be an element of  $\Upsilon$  defined by a coding invariant set  $A \in I$ . We also define  $P(\mathbf{a}) = \bar{P}(A)$  for any semicomputable semimeasure  $P$ , where  $A \in \mathbf{a}$ .

By ([2], Theorem 3.1) any sequence Martin-Löf random with respect to a computable measure is computable or information equivalent to a sequence random with respect to the uniform measure. Therefore, we can define two natural elements  $\mathbf{r} = [\bar{R}]$  and  $\mathbf{c} = [C]$  of  $\Upsilon$ , where  $R$  be a set of all Martin-Löf sequences random with respect to the uniform measure,  $C$  be a set of all computable sequences. Evidently  $\bar{M}(\mathbf{r}) > 0$  and  $\bar{M}(\mathbf{c}) > 0$  and  $P(\mathbf{r}) = 0$  for each computable measure  $P$ .

The zero element  $\mathbf{0}$  of  $\Upsilon$  consists of all coding invariant subsets of  $\Omega$  of  $\bar{M}$ -measure 0. Then  $\bar{M}(\mathbf{0}) = 0$ . The unit element of  $\Upsilon$  is defined  $\mathbf{1} = [\Omega]$ .

An element  $\mathbf{d}$  of  $\Upsilon$  is called *atom element* if  $\mathbf{d} \neq \mathbf{0}$  and it cannot be represented as a disjoint union  $\mathbf{d} = \mathbf{a} \cup \mathbf{b}$  of two non-zero elements of  $\Upsilon$ , i.e., such that  $\mathbf{a} \cap \mathbf{b} = \emptyset$ ,  $\mathbf{a} \neq \mathbf{0}$  and  $\mathbf{b} \neq \mathbf{0}$ . This can be interpreted as that any set of sequences generating an atom cannot be divided on two nontrivial parts by their information content.

**Theorem 1.** (Levin [1]) The elements  $\mathbf{r}$  and  $\mathbf{c}$  are atom elements of  $\Upsilon$ .

The information-theoretic structure of the algebra  $\Upsilon$  is described by the following theorem.

**Theorem 2.** Let  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots$  be all atom elements of  $\Upsilon$ . The following decomposition of the unit element of the algebra  $\Upsilon$  is valid

$$\mathbf{1} = \bigcup_{i=1}^{\infty} \mathbf{a}_i \cup \mathbf{d},$$

where

- the atom  $\mathbf{a}_1 = \mathbf{r}$  is generated by all Martin-Löf random sequences;
- the atom  $\mathbf{a}_2 = \mathbf{c}$  is generated by all computable sequences;
- the atoms  $\mathbf{a}_3, \mathbf{a}_4, \dots$  are generated by infinite sequences which cannot be Martin-Löf random with respect to any computable measure, and even, they cannot be information equivalent to Martin-Löf random sequences;
- $\mathbf{d}$  is the non-zero infinitely divisible element of  $\Upsilon$  generated by the measure-theoretic complement of all atoms.

By definition the element  $\mathbf{d} = \mathbf{1} \setminus \bigcup_{i=1}^{\infty} \mathbf{a}_i$  is infinitely divisible, i.e., for any non-zero  $\mathbf{x} \subseteq \mathbf{d}$  a decomposition  $\mathbf{x} = \mathbf{x}_1 \cup \mathbf{x}_2$  is valid, where  $\mathbf{x}_1 \cap \mathbf{x}_2 = \mathbf{0}$ ,  $\mathbf{x}_1 \neq \mathbf{0}$  and  $\mathbf{x}_2 \neq \mathbf{0}$ .

[1] Levin L.A., V'yugin V.V. Invariant properties of informational bulks, Springer Lecture Notes on Computer Science, 1977, v.53, p.359-364.

[2] A.K. Zvonkin and L.A. Levin. The complexity of finite objects and the algorithmic concepts of information and randomness, Russ. Math. Surv. 1970, 25, 83-124.

\* \* \*

**Liang Yu** (Nanjing Univeristy, China).  
*The theory of higher randomness*

I shall give a survey about recent progress of the higher randomness theory. The talk will focus on  $\Delta_1^1$ -Kurtz,  $\Delta_1^1$ -,  $\Pi_1^1$ -ML- and  $\Pi_1^1$ -randomness and their lowness properties. Some open questions will be given.

\* \* \*