An Open Problem Regarding the Convergence of Universal A Priori Probability

Marcus Hutter[*]

IDSIA, Galleria 2, CH-6928 Manno-Lugano, Switzerland

marcus@idsia.ch http://www.idsia.ch/~marcus

10 June 2003

Abstract

Is the textbook result that Solomonoff's universal posterior converges to the true posterior for all Martin-Löf random sequences true?

Universal induction. Induction problems can be phrased as sequence prediction tasks. This is, for instance, obvious for time series prediction, but also includes classification tasks. Having observed data x_t at times $t<n$, the task is to predict the t-th symbol x_t from sequence $x=x_1...x_{t-1}$. The key concept to attack general induction problems is Occam's razor and to a less extent Epicurus' principle of multiple explanations. The former/latter may be interpreted as to keep the simplest/all theories consistent with the observations $x_1...x_{t-1}$ and to use these theories to predict x_t. Solomonoff [Sol64, Sol78] formalized and combined both principles in his universal prior $M(x)$ which assigns high/low probability to simple/complex environments, hence implementing Occam and Epicurus. $M(x)$ is defined as the probability that a universal Turing machine U outputs a string starting with x, when provided with fair coin flips on the input tape.

Posterior convergence. Solomonoff's [Sol78] central result is that if the probability $\mu(x_t|x_1...x_{t-1})$ of observing x_t at time t, given past observations $x_1...x_{t-1}$ is a computable function, then the universal posterior $M_t := M(x_t|x_1...x_{t-1})$ converges (rapidly!) *with μ-probability 1* (w.p.1) for $t\to\infty$ to the true posterior $\mu_t := \mu(x_t|x_1...x_{t-1})$, hence M represents a universal predictor in case of unknown μ. Convergence of M_t to μ_t w.p.1 tells us that M_t is close to μ_t for sufficiently large t for "almost all" sequences $x_{1:\infty}$ (we abbreviate $x_{1:n} := x_1...x_n$). It says nothing about whether convergence is true for any *particular* sequence (of measure 0).

Martin-Löf randomness is the standard notion to capture convergence for individual sequences and is closely related to Solomonoff's universal prior. Levin gave a characterization equivalent to Martin-Löf's (M.L.) original definition [Lev73]:

A sequence $x_{1:\infty}$ is μ-random (in the sense of M.L.) iff *there is a constant c such that $M(x_{1:n})\leq c\cdot\mu(x_{1:n})$ for all n.*

One can show that a μ-random sequence $x_{1:\infty}$ passes *all* thinkable effective randomness tests, e.g. the law of large numbers, the law of the iterated logarithm, etc. In particular, the set of all μ-random sequences has μ-measure 1.

[*]A prize of 128 Euro for a solution of this problem is offered.

Open problem. An interesting open question is whether M_t converges to μ_t (in difference or ratio) individually for all Martin-Löf random sequences. Clearly, Solomonoff's result shows that convergence may at most fail for a set of sequences with μ-measure zero. A convergence result for μ-random sequences is particularly interesting and natural in this context, since μ-randomness can be defined in terms of M itself (see above).

Proof attempts. Attempts to convert the convergence results w.p.1 to effective μ-randomness tests fail, since M_t is not lower semi-computable. In [LV97, Th.5.2.2] and [VL00, Th.10] the following Theorem is stated:

"*Let μ be a positive recursive measure. If the length of y is fixed and the length of x grows to infinity, then $M(y|x)/\mu(y|x)\to 1$ with μ-probability one. The infinite sequences ω with prefixes x satisfying the displayed asymptotics are precisely ['$\Rightarrow$'* and *'$\Leftarrow$'] the μ-random sequences.*"

While convergence w.p.1 is correct if appropriately interpreted,[1] the proof that convergence holds for μ-random sequences is incomplete: "$M(x_{1:n})\leq c\cdot\mu(x_{1:n})\forall n\Rightarrow \lim_{n\to\infty}M(x_{1:n})/\mu(x_{1:n})$ *exists*" has been used, but not proven, and may indeed be wrong. Vovk [Vov87] shows that for two finitely computable semi-measures μ and ρ, and $x_{1:\infty}$ being μ- *and* ρ-random that $\rho_t/\mu_t\to 1$. If M were recursive, then this would imply $M_t/\mu_t\to 1$ for every μ-random sequence $x_{1:\infty}$, since *every* sequence is M-random. Since M is *not* recursive Vovk's theorem cannot be applied and it is not obvious how to generalize it. So the question of individual convergence remains open.

Conclusions. Contrary to what was believed before, the question of posterior convergence $M_t/\mu_t\to 1$ (also $M_t\to\mu_t$) for all μ-random sequences is still open. In [Hut03] we introduce a new flexible notion of randomness which contains Martin-Löf randomness as a special case. This notion is used to show that standard proof attempts of $M_t/\mu_t\xrightarrow{M.L.}1$ based on so called dominance only must fail, indicating that this problem may be a hard one.

References.

[Hut03] M. Hutter. On the existence and convergence of computable universal priors. Technical Report IDSIA-05-03, 2003. http://arxiv.org/abs/cs.LG/0305052.

[Lev73] L. A. Levin. On the notion of a random sequence. *Soviet Math. Dokl.*, 14(5):1413–1416, 1973.

[LV97] M. Li and P. M. B. Vitányi. *An introduction to Kolmogorov complexity and its applications*. Springer, 2nd edition, 1997.

[Sol64] R. J. Solomonoff. A formal theory of inductive inference: Part 1 and 2. *Inform. Control*, 7:1–22, 224–254, 1964.

[Sol78] R. J. Solomonoff. Complexity-based induction systems: comparisons and convergence theorems. *IEEE Trans. Inform. Theory*, IT-24:422–432, 1978.

[VL00] P. M. B. Vitányi and M. Li. Minimum description length induction, Bayesianism, and Kolmogorov complexity. *IEEE Trans. on Inform. Theory*, 46(2):446–464, 2000.

[Vov87] V. G. Vovk. On a randomness criterion. *Soviet Mathematics Doklady*, 35(3):656–660, 1987.

[1] The formulation of the Theorem is quite misleading in general: First, for off-sequence y convergence w.p.1 does not hold (xy must be demanded to be a prefix of $x_{1:\infty}$). Second, the proof of '$\Leftarrow$' has loopholes (see main text). Last, '$\Rightarrow$' is given without proof and is probably wrong. Also the assertion in [LV97, Th.5.2.1] that S_t converges to zero faster than $1/t$ cannot be made, since S_t may not decrease monotonically.