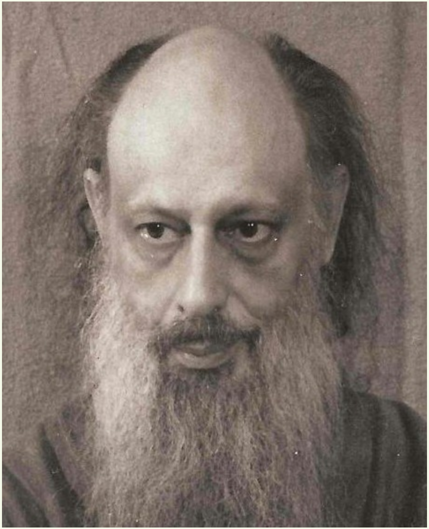
Prediction and Solomonoff

Péter Gács

Boston University

Quantum Foundations worshop, August 2014

- Much of the discussion on the first day of the workshop dealt with the problem of inductive inference in general—quantum physics and cosmology did not seem relevant.
- There is an approach to inductive inference that I felt was ignored, and which can be seen a refinement of Occam's Razor.
- In its generality, this approach is not trying to decide the "true" model to be used for prediction: it is just trying to be (nearly) as good as the best possible predictors that we humans (or computers) can produce.
- Solomonoff achieved (something like) this by chosing a universal prior in a Bayesian framework. It is related to the prior Charlie talked about yesterday, but is not the same.



The probabilities

Turing machine T, one-way binary input tape. One-way output tape.

Experiment: input is an infinite sequence of tosses of an independent unbiased coin. (Monkey at the keyboard.)

$$M_T(x) = \mathbb{P}\{\text{ outputted sequence begins with } x \}.$$

The quantity $M_T(x)$ can be considered the algorithmic probability of the finite sequence x.

Dependence on the choice of T: if T is universal of the type called optimal then this dependence is only minor (Charlie explained this). Fixing such an optimal machine U, write $M(x) = M_U(x)$. This is (the best-known version of) Solomonoff's prior.

The formula

Given a sequence x of experimental results,

$$\frac{M(xy)}{M(x)}$$

assigns a probability to the event that x will be continued by a sequence (or even just a symbol) y.

Attractive: prediction power, combination of some deep principles.

But: incomputable. So in applications, we must deal with the problem of approximating it.

Principle of indifference

In Solomonoff’s theory, Laplace’s principle is revived in the following sense: all descriptions (inputs) of the same length are assigned the same probability.

The prediction theorem

- Solomonoff's theorem restricts consideration to sources $x_1 x_2 \ldots$ with some computable probability distribution $\boldsymbol{P}$.
  Let $\boldsymbol{P}(x)$ = the probability of the set of all infinite sequences starting with x.
- The theorem says that for all $\boldsymbol{P}$, the expression

$$\frac{M(x_1 \ldots x_n b)}{M(x_1 \ldots x_n)}$$

  gets closer and closer to $\frac{\boldsymbol{P}(x_1 \ldots x_n b)}{\boldsymbol{P}(x_1 \ldots x_n)}$ (with very high $\boldsymbol{P}$ probability).
- The proof relies just on the fact that $M(x)$ dominates all computable measures (even all lower semicomputable semimeasures, like itself).

All the usual measures considered by physicists are computable. Here is another example to illustrate the variety.

Example Take a sequence $x_1 x_2 \ldots$ whose even-numbered binary digits are those of π, while its odd-numbered digits are random. Solomonoff's formula will converge to 1/2 on the odd-numbered digits. On the even-numbered digits, it will get closer and closer to 1 if b equals the corresponding digit of π, and to 0 if it does not.