Mutual Information, the Independence Postulate, and Depth

Christopher P. Porter
University of Florida

Varieties of Algorithmic Information
Heidelberg, Germany
June 16, 2015

Introduction

In “Forbidden Information,” Levin seeks to close a loophole in Gödel’s incompleteness theorem (hereafter, GIT).

As cited by Levin, according to Gödel,

> *[I]t turns out that in the systematic establishment of the axioms of mathematics, new axioms, which do not follow by formal logic from those previously established, again and again become evident. It is not at all excluded by the negative results mentioned earlier [i.e., GIT] that nevertheless every clearly posed mathematical yes-or-no question is solvable in this way. For it is just this becoming evident of more and more new axioms on the basis of the meaning of the primitive notions that a machine cannot imitate.*

Introduction (continued)

Gödel's loophole: Just as new axioms are established on the basis of being evident to the mathematical community, perhaps on the basis of such a process we might eventually come to settle all mathematical questions; in particular, such a process may lead the mathematical community to produce, at least in principle, a complete, consistent extension of Peano arithmetic.

Levin's goal is thus to show that a completion of Peano arithmetic cannot be produced by "non-mechanical means."

Our main goals

In this talk, the three main goals are to:

- sketch Levin's argument that purports to close Gödel's loophole;
- raise some concerns about Levin's argument; and
- discuss the merits of Levin's argument.

Outline

1. Technical background
2. Levin's argument
3. Evaluating Levin's argument
4. Deep objects

1. Technical background

Continuous semi-measures

Continuous semi-measures play a central role in our discussion.

Definition

A *continuous semi-measure* is a function $M : 2^{<\omega} \to [0, 1]$ such that

(i) $M(\varepsilon) = 1$ and

(ii) $M(\sigma) \geq M(\sigma 0) + M(\sigma 1)$ for every $\sigma \in 2^{<\omega}$.

Hereafter, I will refer to continuous semi-measures simply as semi-measures.

Left-c.e. semi-measures

We will restrict our attention to left-c.e. semi-measures.

A semi-measure $M : 2^{<\omega} \to [0,1]$ is *left-c.e.* if, uniformly in $\sigma \in 2^{<\omega}$, $M(\sigma)$ is the limit of a computable, non-decreasing sequence of rationals.

Left-c.e. semi-measures and computation

A *Turing functional* $\Phi : 2^\omega \to 2^\omega$ is induced by a c.e. set S_Φ of pairs of strings (σ, τ) such that if $(\sigma, \tau), (\sigma', \tau') \in S_\Phi$ and $\sigma \preceq \sigma'$, then $\tau \preceq \tau'$ or $\tau' \preceq \tau$.

For $\sigma \in 2^{<\omega}$, we define
$\Phi^{-1}(\sigma) := \{X \in 2^\omega : \exists n\, \exists \sigma' \succeq \sigma\ (X{\upharpoonright}n, \sigma') \in S_\Phi\}$.

Theorem (Levin)

(i) *If Φ is a Turing functional, then λ_Φ, defined by*

$$\lambda_\Phi(\sigma) = \lambda(\Phi^{-1}(\sigma))$$

for every $\sigma \in 2^{<\omega}$, is a left-c.e. semi-measure.

(ii) *For every left c.e. semi-measure M, there is a Turing functional Φ such that $M = \lambda_\Phi$.*

Universal semi-measures

Levin also established the existence of a universal semi-measure.

Theorem

There is a left-c.e. semi-measure $\mathbf{M}$ *such that for every left-c.e. semi-measure* Q*, there is some* $c \in \omega$ *such that for every* $\sigma \in 2^{<\omega}$*,*

$$c \cdot \mathbf{M}(\sigma) \geq Q(\sigma).$$

The measure derived from a semi-measure

If Q is a semi-measure, we can define

$$\overline{Q}(\sigma) := \inf_n \sum_{\tau \succeq \sigma \ \& \ |\tau| = n} Q(\tau).$$

One can verify that $\overline{Q}$ is the largest measure such that $\overline{Q} \leq Q$ (but it is not a probability measure in general).

Proposition

If Q is a left-c.e. semi-measure induced by a Turing functional Φ, then

$$\overline{Q}(\sigma) = \lambda(\{X \in 2^\omega : X \in \Phi^{-1}(\sigma) \ \& \ \Phi(X) \text{ is total}\}).$$

Negligibility

$\overline{\mathbf{M}}$ can be seen as a universal measure (universal with respect to all computable measures, as well as the measures derived from left-c.e. semi-measures).

Definition

$\mathcal{S} \subseteq 2^\omega$ is *negligible* if $\overline{\mathbf{M}}(\mathcal{S}) = 0$.

The intuition behind negligibility

Let $\mathcal{S} \subseteq 2^\omega$.

$\overline{\mathbf{M}}(\mathcal{S}) = 0$ means that the probability of producing some member of $\mathcal{S}$ by means of any Turing functional equipped with any sufficiently random oracle is 0.

To see this, one can show that

$$\overline{\mathbf{M}}(\mathcal{S}) = 0 \text{ if and only if } \lambda\Big(\bigcup_{i \in \omega} \Phi_i^{-1}(\mathcal{S})\Big) = 0.$$

In particular, for each Φ_i, $\lambda(\{X \in \mathrm{MLR} : \Phi_i(X) \in \mathcal{S}\}) = 0$.

Mutual information

The *mutual information* of two strings σ and τ, denoted by $\mathbf{I}(\sigma : \tau)$, is defined by

$$\mathbf{I}(\sigma : \tau) = K(\sigma) + K(\tau) - K(\sigma, \tau)$$

where $K(\sigma, \tau) := K(\langle \sigma, \tau \rangle)$.

Levin extends mutual information to infinite sequences by setting

$$\begin{aligned} \mathbf{I}(X : Y) &= \log \sum_{\sigma, \tau \in 2^{<\omega}} 2^{K(\sigma) - K^X(\sigma) + K(\tau) - K^Y(\tau) - K(\sigma, \tau)} \\ &= \log \sum_{\sigma, \tau \in 2^{<\omega}} 2^{-K^X(\sigma) - K^Y(\tau) + \mathbf{I}(\sigma : \tau)}. \end{aligned}$$

2. Levin's argument

Towards closing Gödel's loophole

Levin arrives at Gödel's loophole by first reviewing Gödel's incompleteness theorem and a strengthening of it.

- First, Gödel originally showed that there is no effective procedure for producing a consistent completion of Peano arithmetic.
- Second, Jockusch and Soare proved that there is no probabilistic algorithm that yields a consistent completion of Peano arithmetic with positive probability.
  - In our terminology, they showed that the collection of consistent completions of Peano arithmetic is *negligible*.

Towards closing Gödel's loophole (continued)

Even though the possibility of producing a consistent completion of Peano arithmetic by any combination of computable and probabilistic means has been essentially ruled out, Levin seeks to rule out the possibility of producing such a completion by other "realistic means."

What these means amount to is unclear, but as a minimum, they include the process by which the mathematical community comes to accept the truth of new axioms.

Let us hereafter refer to this process as *process P*.

The structure of Levin's argument

Levin's argument that purports to close Gödel's loophole has the following two-part structure:

1. The technical core of the argument
2. The philosophical core of the argument

The technical core of Levin's argument

Recall that Chaitin's Ω is defined to be

$$\Omega := \sum_{U(\sigma)\downarrow} 2^{-|\sigma|},$$

where U is a universal, prefix-free machine.

Theorem (Levin)

Let A be a consistent completion of Peano arithmetic. Then $\mathbf{I}(A : \Omega) = \infty$.

The philosophical core of Levin's argument

The main philosophical component of Levin's argument involves what he refers to as the *independence postulate*.

IP: Let X be any mathematically definable sequence and let Y be any physically obtainable sequence. Then $\mathbf{I}(X : Y) < \infty$.

These notions of mathematical definability and physical obtainability are rather unclear; let us bracket this concern for the moment.

Levin's argument for closing Gödel's loophole

1. (Reductio premise) Suppose that a consistent completion of Peano arithmetic X can be produced by process P.
2. (Physical obtainability premise) X is thus physically obtainable.
3. (IP premise) $\mathbf{I}(X : Y) < \infty$ for all mathematically definable sequences Y.
4. (Technical premise) $\mathbf{I}(X : \Omega) = \infty$.
5. (Mathematical definability premise) Ω is mathematically definable.
6. Therefore, no consistent completion of Peano arithmetic can be produced by process P.

3. Evaluating Levin's argument

Premises we won't question

First, the technical premise ($\mathbf{I}(X : \Omega) = \infty$) and the mathematical definability premise (Ω is mathematically definable) are beyond question.

Second, one might question whether we should accept the physical obtainability premise (that a consistent completion of Peano arithmetic produced by process P is thus physically obtainable), but we will not do so here.

Evaluating the independence postulate

This leaves the independence postulate as the remaining premise to consider.

In what follows, I will raise several questions about the status of the IP.

Which sequences are mathematically definable?

This isn't a serious concern.

Levin's argument is still valid if we replace "mathematically definable" with "Δ_2^0 definable" in the statement of the IP.

By using a weaker form of the IP, we can strengthen Levin's argument.

Which sequences are physically obtainable?

Note that the IP does not rule out the possibility that computable sequences are physically obtainable:

For every computable sequence X and any $Y \in 2^\omega$, we have $\mathbf{I}(X : Y) < \infty$.

But which sequences are the physically obtainable ones?

- One possible answer: the physically obtainable sequences = the computable sequences

Which sequences are physically obtainable? (continued)

For those who do not accept this identification, is there a principled reason to hold that no non-computable Δ^0_2 is physically obtainable?

Perhaps more worrisome: How can we account for the non-physical obtainability of non-computable Δ^0_2 sequences, which are obtained as the limit of some computable procedures that can be physically implemented (at least in principle)?

Which definition of mutual information?

Levin's result depends upon a specific definition of mutual information.

Why should we think that this is a reasonable notion of mutual information to use?

Is there a minimal set of conditions for a notion of mutual information that guarantees that the IP holds? Or at least one that allows one to close Gödel's loophole?

Implications for hypercomputation?

A number of proponents of hypercomputation have argued that it is physically possible that the halting set can be computed by some hypercomputational procedure.

The IP rules out this possibility.

Such a strong conclusion ought not come for free.

4. Deep objects

The significance of Levin's argument

In this last part of the talk, I will discuss what I take to be the significance of Levin's argument, which is found in the technical core of the argument.

Levin has identified an instance of a more general phenomenon, namely the behavior of so-called *deep* mathematical objects.

Deep Π^0_1 classes

Let $\mathcal{P}$ be a Π^0_1 class and let T be the canonical co-c.e. tree such that $\mathcal{P} = [T]$. Let $T_n = T \cap 2^n$.

$\mathcal{P}$ is a *deep* Π^0_1 class if there is some computable, non-decreasing, unbounded function $h : \omega \to \omega$ such that

$$\mathbf{M}(T_n) \leq 2^{-h(n)},$$

where $\mathbf{M}(T_n) = \sum_{\sigma \in T_n} \mathbf{M}(\sigma)$.

That is, the probability of producing some initial segment of a path through $\mathcal{P}$ is effectively bounded from above.

Generalizing Levin's technical results

Theorem (Levin/Stephan)

The collection of consistent completions of Peano arithmetic forms a deep Π^0_1 class.

One can verify that every deep Π^0_1 class is negligible, and thus this result strengthens the Jockusch/Soare result referenced earlier.

Generalizing Levin's technical results (continued)

Theorem (Bienvenu, Porter)

For every deep Π^0_1 class $\mathcal{P}$ and every $X \in \mathcal{P}$, we have $\mathbf{I}(X : \Omega) = \infty$.

Note that in light of this theorem, Levin's argument applies to every member of every deep Π^0_1 class, not just consistent completions of Peano arithmetic.

- shift-complex sequences;
- DNC_h functions for sufficiently slow-growing functions h;
- compression functions;
- sequences of finite sets of strings of high Kolmogorov complexity.

The fragility of deep Π^0_1 classes

Theorem (Bienvenu, Porter)

The collection of deep Π^0_1 classes forms a filter in the Medvedev degrees of Π^0_1 classes.

By contrast, we have:

Theorem (Bienvenu, Porter)

Every deep Π^0_1 class is Muchnik equivalent to a negligible Π^0_1 class that is not deep.

Connection to Bennett's logical depth

Bennett defined a sequence X to be logically deep if the complexity of initial segments of X are infinitely often far from the time-bounded complexity of initial segments of X.

Example: the halting set K is logically deep.

Theorem (Bienvenu, Porter)

Every member of a deep Π^0_1 class is logically deep.

The promise of deep classes

There is a considerable amount about deep classes that we don't know:

- exact relationships between different deep classes in the Medvedev degrees?
- relationship between deep classes and other notions of highly structured objects?
- more examples of deep objects?

Despite the inadequacies of Levin's argument, his isolation of the basic properties of deep classes has proven to be very useful, and perhaps will lead to further unification of notions of information in algorithmic randomness.