Mutual Information and Weak Lowness Notions

Ian Herbert

University of California, Berkeley

14/2/2013

Mutual Information

Getting on the Same Page

Definition

The *prefix-free Kolmogorov complexity* $K(\sigma)$ of a string $\sigma \in 2^{<\omega}$ is the length of the shortest input to the universal decoding machine, $\mathbb{U}$, that produces σ.

The KC theorem say that for any c.e. set $W \subset 2^{<\omega} \times \mathbb{N}$ (a 'request set') with $\sum_{(\sigma,n)\in W} 2^{-n} < \infty$ ('bounded') there is a machine $\mathcal{M}$ such that for any $(\sigma, n) \in W$ there is a τ with $|\tau| \leq n$ and $\mathcal{M}(\tau) = \sigma$. We can build machines by just asking nicely and not being greedy.

Lowness Notions

Definition

A real A is *K-trivial* if for all n, $K(A \upharpoonright n) \leq^+ K(n)$.

For any real A we can relativize (prefix-free) Kolmogorov complexity to A, K^A, by allowing $\mathbb{U}$ to access A as an oracle.

Definition

A real A is *low for K* if for all σ, $K(\sigma) \leq^+ K^A(\sigma)$.

Mutual Information

We can think of the difference $K(\sigma) - K^A(\sigma)$ as A's information about the string σ. A low for K real has some constant bound on its information about strings.

Definition (Levin)

The *mutual information* of reals A, B is

$$I(A : B) = \log \sum_{\sigma,\tau} 2^{K(\sigma)-K^A(\sigma)+K(\tau)-K^B(\tau)-K(\sigma,\tau)},$$

where $K(\sigma,\tau)$ is the complexity of the pair (σ,τ).
The *simplified mutual information* of reals A, B is

$$I^s(A : B) = \log \sum_{\sigma} 2^{K(\sigma)-K^A(\sigma)-K^B(\sigma)}.$$

It is open whether $I(A : B) = I^s(A : B)$ for all A, B.

Some Nice Properties

- $I(A : B) =^+ I(B : A)$
- If $A \geq_T C$, then $I(A : B) \geq^+ I(C : B)$
- If A is MLR relative to B, then $I(A : B) < \infty$.
- If A is low for K, then $I(A : B) < \infty$ for every B.
- If $I(A : B) < \infty$ for every B, then A is low for K.

Self-Information

We can also examine the quantity $I(A : A)$, the *self-information* of A.

Definition

A real A has *finite self-information* if $I(A : A) < \infty$.

Some Nice Properties

- If $A \geq_T B$ and A has finite self-information, then so does B.
- If A is low for K, then A has finite self-information.
- If $A \geq_T 0'$, then A does not have finite self-information.
- If A is MLR, then A does not have finite self-information.
- If A has finite self-information, then A is jump traceable and so A is GL_1 ($A' \leq_T A \oplus \emptyset'$).

Finite Self-Information

So, having finite self-information is a lowness notion that contains the K-trivials. Is it another characterization of being K-trivial?

Theorem (Hirschfeldt, Weber)

There is a c.e real that has finite self-information and is not low for K.

So we can ask about where these reals occur, how they behave, etc. We know there are none above $\emptyset'$ and some below $\emptyset'$. What about incomparable to $\emptyset'$?

Theorem (H.)

There is a perfect Π^0_1 class of reals with finite self-information. Therefore they are not all Δ^0_2.

A bound on information

The proofs of these theorems depend on the following lemma of Hirschfeldt and Weber.

Lemma

There is a Δ^0_2 function $f : 2^{<\omega} \to \mathbb{N}$ with $\forall i \forall^\infty \forall s [f_s(\sigma) > i]$ such that $\sum_{\sigma,\tau} 2^{f(\sigma)+f(\tau)-K(\sigma,\tau)} < \infty$

Now if we have an A such that for all σ, $K(\sigma) - K^A(\sigma) \leq^+ f(\sigma)$, then A has finite self-information. We just need to build a perfect set of reals that obey this bound on their information. This brings us to a much broader topic.

Weak Lowness Notions

Definition

For a function $g : \mathbb{N} \to \mathbb{N}$, a real A is *K-trivial up to g* if for all n, $K(A \upharpoonright n) \leq^+ K(n) + g(n)$.
For a function $f : 2^{<\omega} \to \mathbb{N}$, a real A is *low for K up to f* if for all σ, $K(\sigma) \leq^+ K^A(\sigma) + f(\sigma)$.

We denote the set of reals that are K-trivial up to g, $\mathcal{KT}(g)$, and the set of reals low for K up to f, $\mathcal{LK}(f)$.
We usually think of functions that are *orders*, that is, unbounded and nondecreasing, but in a very fluid way. All we really need is that they have *finite-to-one* approximations, i.e. $\forall i \forall^\infty \sigma \forall s f_s(\sigma) > i$. This does restrict us to Δ^0_2 functions, but

Theorem (Baartse, Barmpalias)

There is a Δ^0_3 order g such that $\mathcal{KT}(g)$ is exactly the set of K-trivials.

How much weaker are these notions?

- $\mathcal{LK}(0) = \mathcal{KT}(0)$ (Nies, 2005). There is a computable order f such that $\mathcal{LK}(f) \neq \mathcal{KT}(g)$ for any Δ^0_2 order g.
- $\mathcal{LK}(0)$ is closed downwards under $\leq_T$. So is $\mathcal{LK}(f)$ for any f. For any Δ^0_2 order g, for any real B, there is a real in $\mathcal{KT}(g)$ computing B.
- $\mathcal{LK}(0)$ is closed under effective join (Downey, Hirschfeldt, Nies, Stephan, 2003). For any Δ^0_2 order f, $\mathcal{LK}(f)$ is not (unless it is all of $2^{<\omega}$).
- $\mathcal{LK}(0)$ has only countably many elements, and they are all Δ^0_2(Chaitin, 1976). For any Δ^0_2 order f, $\mathcal{LK}(f)$ contains a perfect set.

This last one is the theorem we want.

The Proof

To prove the theorem, we build a computable tree, T, all of whose paths satisfy $K(\sigma) \leq^+ K^A(\sigma) + f(\sigma)$. To ensure this inequality holds, we build a KC set, L, alongside T and enumerate requests for short descriptions of strings when we see them get short descriptions relative to some path through T.

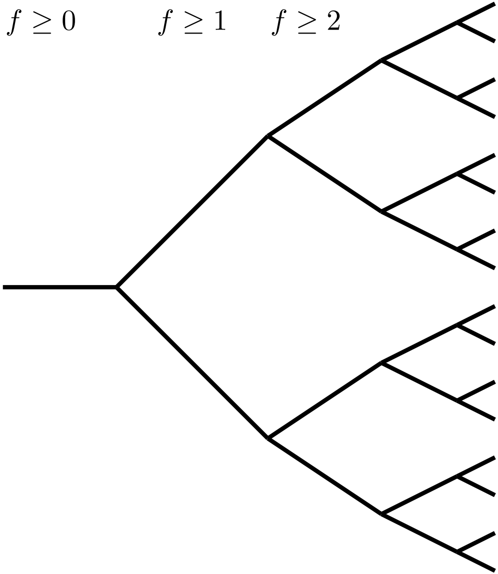
The problem, of course, is that as the number of branches through T increases, the same mass might be used to give short descriptions of different σ over and over again. We need to know we can keep up while keeping the mass in L bounded.

The Proof

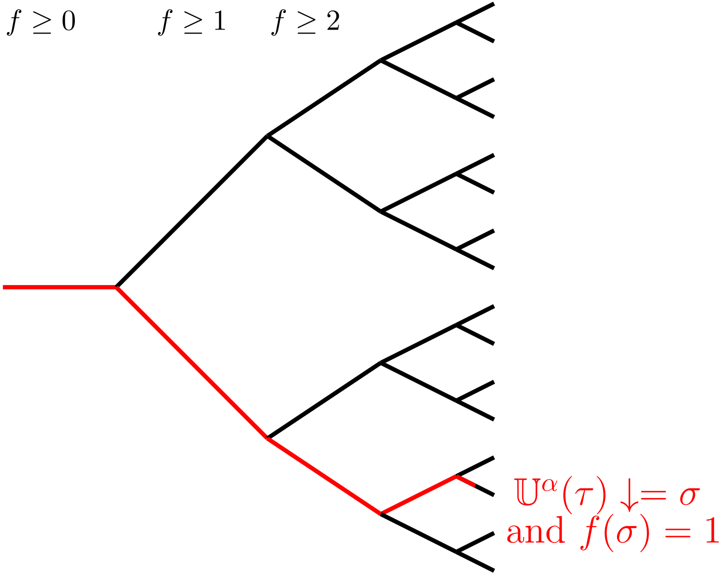
We use the fact that we are only paying up to a factor of f. We know f has a finite-to-one approximation, so for any i, only finitely many σ ever take value i. We contrive to ensure that all descriptions of σ with $f(\sigma) = i$ converge on T below the level where T branches for the $i + 1$st time.

If it looks like this is failing for a given i, keep the path above the $i + 1$st branching level with the most mass (and all identical extensions from nodes at that level), kill all other paths, and move the branching level up (picture to come).

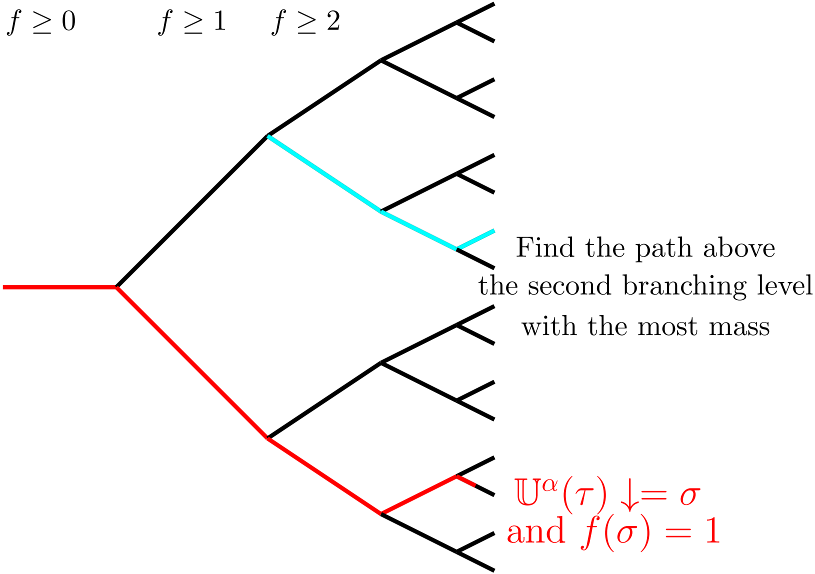
Picture



Picture

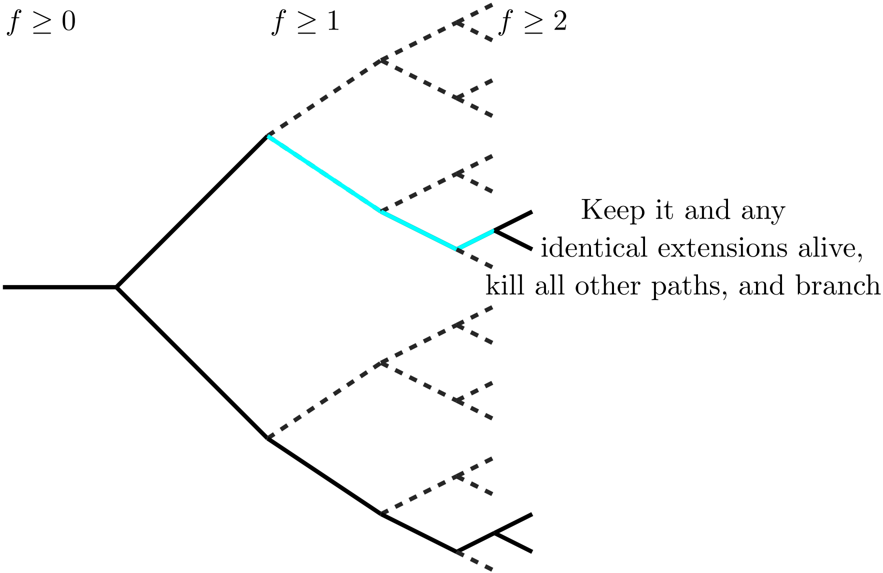

$f \geq 0$
$f \geq 1$
$f \geq 2$
$\mathbb{U}^{\alpha}(\tau) \downarrow = \sigma$
and $f(\sigma) = 1$

Picture

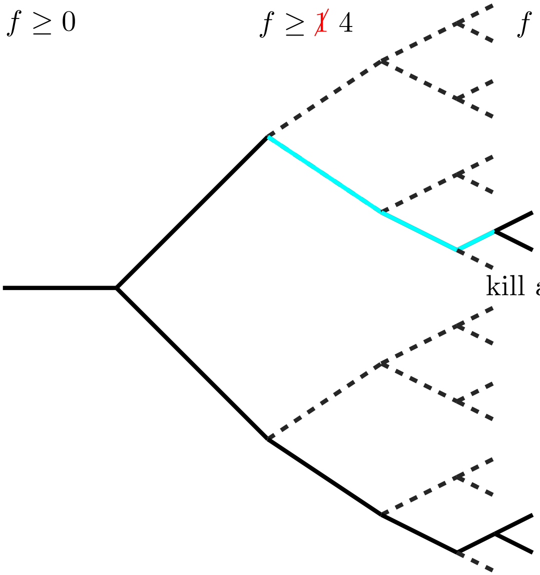


$f \geq 0$
$f \geq 1$
$f \geq 2$
Find the path above
the second branching level
with the most mass
$\mathbb{U}^{\alpha}(\tau)\downarrow = \sigma$
and $f(\sigma) = 1$

Picture


$f \geq 0$
$f \geq 1$
$f \geq 2$
Keep it and any
identical extensions alive,
kill all other paths, and branch

Picture



Keep it and any identical extensions alive, kill all other paths, and branch

We need to make sure the branching levels eventually settle. This is where having a finite-to-one approximation is important. Only finitely many σ ever are in a position to threaten the ith branching level, and each of these can only injure it finitely often.

Next, we need to show the mass we put into L is bounded. We can consider separately the mass paid into L for the living subtree of T and for the nodes killed during the construction. For the second part, we use the following lemma.

Lemma

For any injury to the ith branching level in the construction, the amount that has been paid into L on the paths above that branching level (those kept and those killed) is no more than 2^{-c_i-1} times the mass, m, that converges on the path chosen during this injury.

Now we can bound the mass that is wasted by bounding the mass that is lost by each injury. Each injury fixes some mass to a lower level in the tree and we have a rough bound (1) on how much mass can converge on any level. We keep track of the mass as it trickles down.

This finishes the proof that every path through our perfect tree is in $\mathcal{LK}(f)$.

Now, using the Hirschfeldt-Weber function f_{HW} we get:

Theorem

There is a perfect Π^0_1 set of reals with finite self-information. Moreover, for any real A there are reals B_0, B_1 with finite self-information such that $A \leq_T B_0 \oplus B_1$.

Some Other Applications

Definition

The *effective Hausdorff dimension* of a real S is $\dim(S) = \liminf_{n\to\infty} \frac{K(S\restriction n)}{n}$.

The *effective packing dimension* of a real S is $\mathrm{Dim}(S) = \limsup_{n\to\infty} \frac{K(S\restriction n)}{n}$.

Definition

A real A is *low for effective Hausdorff dimension* if for every real S, $\dim(S) = \dim^A(S)$.

A real A is *low for effective packing dimension* if for every real S, $\mathrm{Dim}(S) = \mathrm{Dim}^A(S)$.

Note that if a real A is in $\mathcal{LK}(\log|\sigma|)$ then A is low for both notions of dimension.

Theorem (H.; Lempp, Miller, Ng, Turetsky, Weber)

There is a perfect, non-ideal Π^0_1 set of reals that are low for dimension.

More Fun Applications

It is clear that the function $K(\sigma)$ has a finite-to-one approximation. Then so does $\lfloor \epsilon K(\sigma) \rfloor$ for any rational ϵ. So we can build a perfect non-ideal Π^0_1 set of reals that are in $\mathcal{LK}(\lfloor \epsilon K(\sigma) \rfloor)$, i.e., that satisfy

$$\lfloor (1-\epsilon) K(\sigma) \rfloor \leq^+ K^A(\sigma).$$

It is not hard to build a tree that works for countably many orders (only let the first i-many injure the first i-many branching levels), so we can also build a perfect non-ideal Π^0_1 set of reals that are in $\mathcal{LK}(\lfloor \epsilon K(\sigma) \rfloor)$ for *every* ϵ. These reals are arbitrarily close to being low for K, but it's still not good enough.

Extending to all orders

We can even extend our construction to handle all Δ^0_2 orders simultaneously, at the cost of some complexity. We'll want to use the same strategy as for countably many orders, but we don't know which Δ^0_2 functions even *are* orders. So, we guess.

We add more branching levels to our tree, which we use to guess whether of not $\phi_{e,s}$ is a finite-to-one approximation. The subtree generated by making all the right guesses at these branching nodes will have its infinite paths in $\mathcal{LK}(f)$ for every Δ^0_2 order f.

Extending to all orders

Some branches will be injured infinitely often, but if they are it will be because they are wrong, and so we can ignore them. The verification that the masses are finite for the different L_e's goes more or less the same as before.

Theorem (H.)

There is a perfect set $\mathcal{P}$ of reals such that for every Δ^0_2 order f $\mathcal{P} \subseteq \mathcal{LK}(f)$.

Thanks!