Seetapun’s Theorem and Kolmogorov Complexity

Patrick Lutz

UCLA
Theorem (Seetapun). For every uncomputable $X$ and set $A \subseteq \mathbb{N}$, either $A$ or $\overline{A} (= \mathbb{N} \setminus A)$ has an infinite subset which does not compute $X$.

Comments.

- Original motivation was reverse math of Ramsey’s theorem
- First explicitly proved by Dzhafarov and Jockusch

Informally: You can’t encode an infinite amount of information into all infinite subsets of both a set and its complement

Question. How much finite information can you encode?
Question. How much finite information can you encode into all infinite subsets of both a set and its complement?

Meta question. How can we measure finite information?
Answer. Use Kolmogorov complexity.

Definition. For a string $\sigma \in 2^{<\omega}$ and set $\mathcal{X} \subseteq \mathcal{P}(\mathbb{N})$, define

$$C(\sigma \mid \mathcal{X}) = \max_{B \in \mathcal{X}} C^B(\sigma).$$

Notation. For $A \subseteq \mathbb{N}$

- $[A]^\omega$ = set of infinite subsets of $A$.
- $\text{Seet}(A) = [A]^\omega \cup [\bar{A}]^\omega$

Question, formal version. Given a string $\sigma$ and set $A \subseteq \mathbb{N}$, how low can $C(\sigma \mid \text{Seet}(A))$ be compared to $C(\sigma)$?
An Example
It is possible to encode “an arbitrary integer larger than \( N \)” (for any \( N \)).

**Definition.** For any string \( \sigma \) and number \( N \), define

\[
C(\sigma \mid \geq N) = \max_{n \geq N} C(\sigma \mid n).
\]

**Proposition.** For any string \( \sigma \) and number \( N \), there is some set \( A \subseteq \mathbb{N} \) such that \( C(\sigma \mid \text{Seet}(A)) \leq C(\sigma \mid \geq N) + O(1) \).

**Proof.**

- \( A \)
  - \[ \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \]

- \( \bar{A} \)
  - \[ \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \]

  \[ \uparrow \]

  \[ N \]
It is possible to encode “an arbitrary integer larger than $N$” (for any $N$).

**Definition.** For any string $\sigma$ and number $N$, define

$$C(\sigma | \geq N) = \max_{n \geq N} C(\sigma | n).$$

**Proposition.** For any string $\sigma$ and number $N$, there is some set $A \subseteq \mathbb{N}$ such that $C(\sigma | \text{Seet}(A)) \leq C(\sigma | \geq N) + O(1)$.

**Theorem (Vereshchagin).** For any string $\sigma$,

$$C^{0'}(\sigma) = \min_{N} C(\sigma | \geq N) \pm O(1).$$

So $C(\sigma | \text{Seet}(A))$ can be as small as $C^{0'}(\sigma)$.

**Question.** Is there any way for all infinite subsets of both $A$ and $\overline{A}$ to lower the complexity of $\sigma$ below $C^{0'}(\sigma)$?

**Answer.** No.
The Main Theorem
Observation. $C(\sigma \mid \text{Seet}(A))$ can be as small as $C^0'(\sigma)$.

Question. Is there any way for all infinite subsets of both $A$ and $\overline{A}$ to lower the complexity of $\sigma$ below $C^0'(\sigma)$?

Answer. No.

Theorem (Harrison-Trainor and L.). For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^0'(\sigma) - O(1).$$

Comment. Standard proofs of Seetapun’s theorem don’t seem to yield anything like this (at least not obviously).
How to Prove It*

*Sort of.
A much easier theorem. For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^X(\sigma) - O(\log |\sigma|)$$

where $X$ is a complete $\Sigma^1_2$ set.

**Proof strategy.** Assume that for all $B \in \text{Seet}(A)$, $C^B(\sigma) < k$ and show that $C^X(\sigma \mid k) \leq k + O(1)$.

**Idea:** Using $X$, enumerate a list of at most $2^k$ strings that “look like” $\sigma$ i.e. a list of at most $2^k$ strings which includes $\sigma$

**Key property of $\sigma$:** There is some set $A$ such that for all $B \in \text{Seet}(A)$, $C^B(\sigma) < k$.

**Claim 1.** At most $2^k$ strings have this property.

**Claim 2.** $X$ can enumerate the set of strings with this property.
A much easier theorem. For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^X(\sigma) - O(\log |\sigma|)$$

where $X$ is a complete $\Sigma^1_2$ set.

**Key property of $\sigma$:** There is some set $A$ such that for all $B \in \text{Seet}(A)$, $C^B(\sigma) < k$.

**Claim 1.** At most $2^k$ strings have this property.

**Proof.** Suppose $\tau_1, \ldots, \tau_n$ all have this property... as witnessed by $A_1, \ldots, A_n$.

Let $B$ be a boolean combination of the $A_i$’s which is infinite. E.g. $B = A_1 \cap \overline{A_2} \cap \overline{A_3} \cap \ldots \cap A_n$.

Then for each $i \leq n$, $C^B(\tau_i) < k$. Impossible if $n > 2^k$. 
A much easier theorem. For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^X(\sigma) - O(\log |\sigma|)$$

where $X$ is a complete $\Sigma^1_2$ set.

**Key property of $\sigma$:** There is some set $A$ such that for all $B \in \text{Seet}(A)$, $C^B(\sigma) < k$.

**Claim 2.** $X$ can enumerate the set of strings with this property.

**Proof.** The property is $\Sigma^1_2$. 
A much easier theorem. For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^X(\sigma) - O(\log |\sigma|)$$

where $X$ is a complete $\Sigma^1_2$ set.

Proof of easier theorem. Assume that for all $B \in \text{Seet}(A)$, $C^B(\sigma) < k$. Identify a property of $\sigma$ which is

- shared by at most $2^k$ other strings
- and which $X$ can recognize.

Theorem (Harrison-Trainor and L.). For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^0'(\sigma) - O(1).$$

Proof idea. Identify a more complicated property of $\sigma$ which is easier to compute.
Theorem (Harrison-Trainor and L.). For all strings \(\sigma\) and sets \(A \subseteq \mathbb{N}\),

\[
C(\sigma \mid \text{Seet}(A)) \geq C^{0'}(\sigma) - O(1).
\]

Proof sketch. Assume that for all \(B \in \text{Seet}(A)\), \(C^B(\sigma) < k\).

Definition. A finite set of strings \(F\) is safe if there is some partition \(A_1, \ldots, A_n\) of \(\mathbb{N}\) such that for all \(i \leq n\) and \(s \subseteq A_i\) finite,

\[
|s| > 1 \implies |\{\tau \mid C^s(\tau) < k\} \cup F| \leq 2^k.
\]

i.e. we can safely assume that (all infinite subsets of) each \(A_i\) will give each \(\tau \in F\) complexity less than \(k\)

Claim 1. No safe set has size larger than \(2^k\).

Claim 2. For any safe set \(F\), \(F \cup \{\sigma\}\) is also safe.

Therefore every maximal safe set contains \(\sigma\).

Claim 3. The set of safe sets is \(0'\)-enumerable.

Therefore \(0'\) can enumerate a maximal safe set.
Theorem (Harrison-Trainor and L.). For all strings \( \sigma \) and sets \( A \subseteq \mathbb{N} \),
\[
C(\sigma \mid \text{Seet}(A)) \geq C^{0'}(\sigma) - O(1).
\]

Proof sketch. Assume that for all \( B \in \text{Seet}(A) \), \( C^B(\sigma) < k \).

Claim 1. No safe set has size larger than \( 2^k \).

Claim 2. For any safe set \( F \), \( F \cup \{\sigma\} \) is also safe.

Claim 3. The set of safe sets is \( 0' \)-enumerable.

The following \( 0' \)-program enumerates a maximal safe set.

Set \( F = \emptyset \)
While true:
\[
\text{Search for } \tau \text{ such that } F \cup \{\tau\} \text{ is safe}
\]
\[
\text{Enumerate } \tau \text{ and set } F = F \cup \{\tau\}
\]

Key point: A maximal safe set has size at most \( 2^k \) and contains \( \sigma \)
A Question
Theorem (Harrison-Trainor and L.). For all strings $\sigma$ and sets $A \subseteq \mathbb{N}$,

$$C(\sigma \mid \text{Seet}(A)) \geq C^0' (\sigma) - O(1).$$

In one sense, this theorem is sharp. But it doesn’t seem to completely capture the following intuition.

Intuition. The only thing you can encode into all infinite subsets of both a set and its complement is “an arbitrary integer larger than $N$” for any single integer $N$.

Question. Fix a set $A \subseteq \mathbb{N}$. Is there a number $N$ such that for all strings $\sigma$,

$$C(\sigma \mid \text{Seet}(A)) \geq C(\sigma \mid \geq N) - O(1)?$$