Kolmogorov Complexity and Chaitin's Incompleteness

Harry Sha

Fall 2026

Today

Today we're going to study another proof of the incompleteness theorem via Kolmogorov Complexity.

The proof builds on our previous machinery for computability and arithmetization.

Incompleteness v1

We constructed a formula φ_M such that φ_M is true iff M halts.

F :

- $\langle F \rangle =$ `get_source()`
- For each string $\pi \in \Sigma^*$
  - If π is a proof of φ_F, loop.
  - If π is a proof of $\neg\varphi_F$, halt.

The paradox

The paradox encoded in this proof is that F is a TM that does the following

"I halt if and only if there's a proof I don't halt"

Combined with soundness and completeness assumptions, this is the paradox

"I halt if and only if I don't halt"

Incompleteness V2

We constructed a formula α_x such that M_x accepts implies α_x has a proof, and M_x rejects implies $\neg\alpha_x$ has a proof. Then we considered the following function.

R :

- $\langle R \rangle =$ `get_source()`
- For each string $\pi \in \Sigma^*$
  - If π is a proof of $\alpha_{\langle R \rangle}$, reject.
  - If π is a proof of $\neg\alpha_{\langle R \rangle}$, accept.

The paradox

I'm gonna speak in first person as if I were the algorithm.

If there's a proof that I accept, then I reject, so there's a proof that I reject.

Thus, the paradox is

"There's a proof that I accept if and only if there's a proof that I reject"

(and there's a proof of exactly one of these by consistency and completeness)

Gödel's Original Paradox

"I have a proof if and only if I don't have a proof"

Today's Paradox

"The smallest positive integer not definable in under sixty letters" (a phrase with fifty-seven letters)

This is called the Berry paradox.

We're going to derive the incompleteness theorem again by showing that given a set of properties we want our formal system to satisfy, we can encode this paradox.

Section 1

Kolmogorov Complexity

Information and Compression

We want to define a measure of complexity on strings.

Intuitively, a string is more complex if it cannot be compressed very much.

A description of a string x is a string $\langle M, w \rangle$ where M when run on w outputs x.

Kolmogorov Complexity

The Kolmogorov Complexity of a string x is then the length of a shortest description of x.

- The Kolmogorov Complexity of x is denoted $K(x)$
- The shortest description of x is denoted $d(x)$. If there are multiple, take the first and lexigraphical order.

Facts about Kolmogorov Complexity

- $K(x) \leq |x| + c$ for some constant c.
- $K(xx) \leq K(x) + c$ for some constant c.
- We could have defined the K in terms of Python programs (or any other programming language) instead of TMs. The definitions are off by only an additive constant.

Compressibility

Let x be a string. Say x is c-compressible if $K(x) \leq |x| - c$.

Say x is incompressible if $K(x) \geq |x|$.

Facts about Compressibility

- Incompressible strings of every length exist.
- For each fixed length, a fraction of at least $1 - 2^{1-c}$ of the strings are not c-compressible.

$K(x)$ is not computable

Section 2

Chaitin's Incompleteness

Incompleteness V1

Let Σ be a set of sentences in the language of arithmetic, and $\mathcal{N}$ be the standard model.

1. **Completeness**. For all sentences A, $\Sigma \vdash A$ or $\Sigma \vdash \neg A$
2. **Soundness**. $\mathcal{N} \models \Sigma$
3. **Recognizable axioms**. $\Sigma \in \mathbf{TR}$

A formula that encodes K

Let $x \in \{0,1\}^*$ and $L \in \mathbb{N}$. Let γ be a formula with two free variables such that

$$\mathcal{N} \models \gamma(\overline{\ulcorner x \urcorner}, \overline{L}) \iff K(x) \geq L$$

Let's convince ourselves informally that we can construct such a formula.

Proof

$H(\langle c \rangle)$:

- for $\pi \in \{0,1\}^*$ in SSO:
  - If π is a proof of $\gamma(\overline{\ulcorner x \urcorner}, \overline{c})$, for some x, output x.

Here $c \in \mathbb{N}$. The input to the TM is a string representation of c over the alphabet $\{0,1\}$.

Proof, cont.

Claim. There is a constant $L \in \mathbb{N}$, such that for all strings x, $\Sigma \nvdash \gamma(\overline{\ulcorner x \urcorner}, \overline{L})$.

Let L be a constant to be chosen later. By contradiction, suppose the claim is false. Then $H(L)$ halts and outputs some string x, such that $\Sigma \vdash \gamma(\overline{\ulcorner x \urcorner}, \overline{L})$.

Proof, cont.

Then, by the claim, and completeness, we have that for all strings x, $\Sigma \vdash \neg\gamma(\overline{\ulcorner x \urcorner}, \overline{L})$.