

18.404/6.840 Lecture 11

Last time:

- The Computation History Method for proving undecidability
- The Post Correspondence Problem is undecidable
- Linearly bounded automata, A_{LBA} is decidable
- Configurations, Computation histories
- E_{LBA} and ALL_{CFG} are undecidable

Today: (Sipser §6.1 – §6.2)

- Self-reproducing machines and The Recursion theorem
- Short introduction to mathematical logic

Midterm exam

90 minutes length + 20 minutes for printing/scanning/uploading.

Open book, postings, piazza, notes, and lecture videos, from this year.

Covers through Recursion Theorem presented today.

Will not include section on mathematical logic.

Not permitted: Communication with anyone except course staff, other materials, internet searching.

Not permitted: Providing information about the exam to anyone who hasn't completed it.

Please respect our honor system.

Self-reproduction Paradox

Suppose a Factory makes Cars

- Complexity of Factory > Complexity of Car
(because Factory needs instructions for Car + robots, tools, ...)

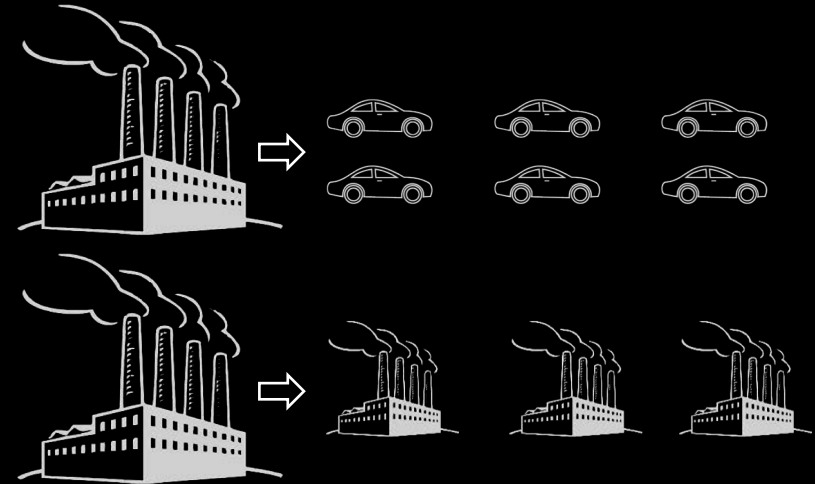
Can a Factory make Factories?

- Complexity of Factory > Complexity of Factory?
- Seems impossible to have a self-reproducing machine

But, living things self-reproduce

How to resolve this paradox?

Self-reproducing machines are possible!



© Source unknown. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <https://ocw.mit.edu/fairuse>.

A Self-Reproducing TM

Theorem: There is a TM *SELF* which (on any input) halts with $\langle SELF \rangle$ on the tape.

Lemma: There is a computable function $q: \Sigma^* \rightarrow \Sigma^*$ such that $q(w) = \langle P_w \rangle$ for every w , where P_w is the TM $P_w = \text{"Print } w \text{ on the tape and halt"}$.

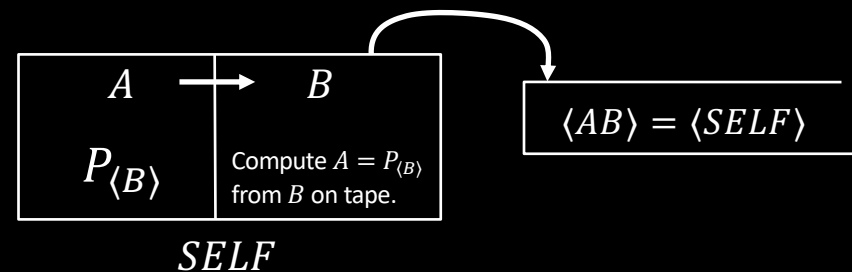
Proof: Straightforward.

Proof of Theorem: *SELF* has two parts, *A* and *B*.

$$A = P_{\langle B \rangle}$$

$$B = P_{\langle A \rangle} ?$$

- B* = "1. Compute $q(\text{tape contents})$ to get *A*.
2. Combine with *B* to get $AB = SELF$.
3. Halt with $\langle SELF \rangle$ on tape."



Can implement in any programming language.

English Implementation

Check-in 11.1

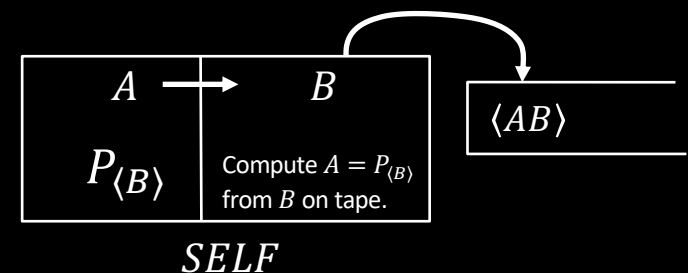
Implementations of the Recursion Theorem have two parts, a Template and an Action. In the TM and English implementations, which is the Action part?

- (a) A and the upper phrase
- (b) A and the lower phrase
- (c) B and the upper phrase
- (d) B and the lower phrase.

Write the following twice, the second time in quotes
“Write the following twice, the second time in quotes”

Write the following twice, the second time in quotes
“Write the following twice, the second time in quotes”

Note on Pset Problem 6: Don't need to worry about quoting.



The Recursion Theorem

A compiler which implements “compute your own description” for a TM.

Theorem: For any TM T there is a TM R where for all w R on input w operates in the same way as T on input $\langle w, R \rangle$.

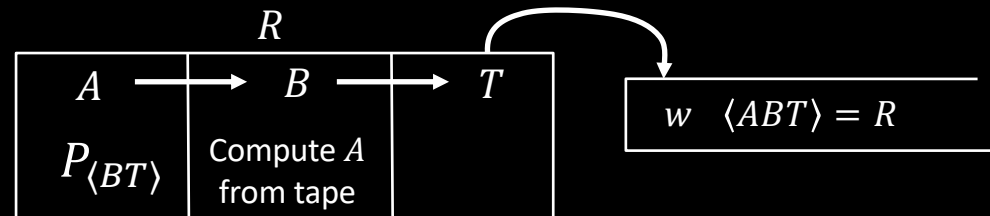
Proof of Theorem: R has three parts: A , B , and T .

T is given

$$A = P_{\langle BT \rangle}$$

- $B =$ “1. Compute q (tape contents after w) to get A .
 2. Combine with BT to get $ABT = R$.
 3. Pass control to T on input $\langle w, R \rangle$.”

Moral: You can use “compute your own description” in describing TMs.



Check-in 11.2

Can we use the Recursion Theorem to design a TM T where $L(T) = \{\langle T \rangle\}$?

- (a) Yes.
- (b) No.

Ex 1: A_{TM} is undecidable - new proof

Theorem: A_{TM} is not decidable

Proof by contradiction: Assume some TM H decides A_{TM} .

Consider the following TM R :

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Use H on input $\langle R, w \rangle$ to determine whether R accepts w .
3. Do the opposite of what H says."

Ex 2: Fixed-point Theorem

Theorem: For any computable function $f: \Sigma^* \rightarrow \Sigma^*$, there is a TM R such that $L(R) = L(S)$ where $f(\langle R \rangle) = \langle S \rangle$.

In other words, consider f to be a program transformation function. Then for some program R , its behavior is unchanged by f .

Proof: Let R be the following TM.

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Compute $f(\langle R \rangle)$ and call the result $\langle S \rangle$.
3. Simulate S on w ."

Ex 3: MIN_{TM} is T-unrecognizable

Defn: M is a minimal TM if $|\langle M' \rangle| < |\langle M \rangle| \rightarrow L(M') \neq L(M)$.

Thus, a minimal TM has the shortest description among all equivalent TMs.

Let $MIN_{TM} = \{\langle M \rangle \mid M \text{ is a minimal TM}\}$.

Theorem: MIN_{TM} is T-unrecognizable.

Proof by contradiction: Assume some TM E enumerates

Consider the following TM R :

$R =$ "On input w

1. Get own description $\langle R \rangle$.
2. Run enumerator E until some TM B appears, where $|\langle R \rangle| < |\langle B \rangle|$.
3. Simulate B on w ."

Thus $L(R) = L(B)$ and $|\langle R \rangle| < |\langle B \rangle|$ so B isn't minimal, but $\langle B \rangle \in L(E)$, contradiction.

Check-in 11.3

Let A be an infinite subset of MIN_{TM} .
Is it possible that A is T-recognizable?

- (a) Yes.
- (b) No.

Other applications

1. Computer viruses.
2. A true but unprovable mathematical statement due to Kurt Gödel:
“This statement is unprovable.”

Intro to Mathematical Logic

Goal: A mathematical study of mathematical reasoning itself.

Formally defines the language of mathematics, mathematical truth, and provability.

Gödel's First Incompleteness Theorem:

In any reasonable formal system, some true statements are not provable.

Proof: We use two properties of formal proofs:

- 1) Soundness: If ϕ has a proof π then ϕ is true.
- 2) Checkability: The language $\{\langle \pi, \phi \rangle \mid \pi \text{ is a proof of statement } \phi\}$ is decidable.

Checkability implies the set of provable statements $\{\langle \phi \rangle \mid \phi \text{ has a proof}\}$ is T-recognizable.

Similarly, if we can always prove $\langle M, w \rangle \in \overline{A_{TM}}$ when it is true, then $\overline{A_{TM}}$ is T-recognizable (false!).

Therefore, some true statements of the form $\langle M, w \rangle \in \overline{A_{TM}}$ are unprovable.

Next, we use the Recursion Theorem to give a specific example of a true but unprovable statement.

A True but Unprovable Statement

Implement Gödel statement “This statement is unprovable.”

Let ϕ_U be the statement $\langle R, 0 \rangle \in \overline{A_{\text{TM}}}$ where R is the following TM:

$R =$ “On any input

1. Obtain $\langle R \rangle$ and use it to obtain ϕ_U .

2. For each possible proof $\pi = \pi_1, \pi_2, \dots$

 Test if π is a proof that ϕ_U is true.

 If yes, then *accept*. Otherwise, continue.”

Theorem: (1) ϕ_U has no proof

(2) ϕ_U is true

Proof:

(1) If ϕ_U has a proof

(2) If ϕ_U is false

ϕ_U

Quick review of today

1. Self-reference and The Recursion Theorem
2. Various applications.
3. Sketch of Gödel's First Incompleteness Theorem in mathematical logic.

MIT OpenCourseWare

<https://ocw.mit.edu>

18.404J / 18.4041J / 6.840J Theory of Computation

Fall 2020

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.