“Turing recognizable” vs. “Decidable”

Language is **Turing recognizable** if some Turing machine recognizes it

- Also called “recursively enumerable”

Machine that halts on all inputs is a **decider**. A decider that recognizes language $L$ is said to **decide** language $L$.

Language is **Turing decidable**, or just **decidable**, if some Turing machine decides it.

Not all problems can be solved

- Good to know when you might not find an answer
- Get perspective on limits of computation

Decidable problems for regular languages

- Does DFA $D$ accept string $s$?
- Is $L(D)$ of DFA empty?
- Are two DFAs $D_1$ and $D_2$ equivalent?

Specify DFA on input TM, determine control algorithm to run DFA specified on tape.
Arbitrary DFA D accepts string \( w \)

Language: \( A_{\text{DFA}} = \{(D, w) \mid D \text{ is DFA that accepts } w \} \)

Theorem: \( A_{\text{DFA}} \) is decidable

Proof idea:
• Define machine \( M \) that simulates \( D \) on \( w \)
• If simulation ends in an accept, accept; else, reject

Note: control states in \( M \) cannot be states in \( D \)
\( M \) needs to run arbitrary \( D \)

A\(_{\text{DFA}}\) decider Proof Outline

DFA \( D \) described as string: 5-tuple

Use marks on tape to track
• current state in simulated \( D \)
• current symbol read from \( w \)

Implement transition function of \( D \) for current \( D \) state and input \( w \)
• \( D \)'s transition \( \delta \) is different from TM \( M \)'s transition \( \delta \)

Arbitrary DFA D accepts no strings

\( E_{\text{DFA}} = \{D \mid D \text{ is DFA with } L(D) = \emptyset \} \) is decidable language

Proof idea:
• Is there any way to reach accept from start?
• Think of graph-marking

Proof:
• Mark start state of DFA \( D \)
• Repeat until no new states
  • Mark any state that past-marked states transition to
• If an accept state is marked, REJECT; else, accept

Two DFAs are equivalent

\( EQ_{\text{DFA}} = \{(A, B) \mid A \text{ and } B \text{ are DFAs and } L(A) = L(B)\} \) is decidable language

Proof idea:
• Construct new DFA \( C \) from \( A \) and \( B \); \( C \) accepts only strings accepted by either \( A \) or \( B \), but not both
• Check if \( C \)'s language is empty (last slide)
Two DFAs are equivalent

\[ \text{EQ}_{\text{DFA}} = \{(A, B) \mid A \text{ and } B \text{ are DFAs and } L(A) = L(B)\} \] is decidable language

Alternative proposals:
- Proposal 1: For machine A, start at start state and record path from start to all possible accept states; do the same for B; check if bijection between unique A and B paths – might work, but need to confirm how this will deal with self loops
- Proposal 2: Create enumerator for machine A. Loop on \( i = 1 \) to infinite, print out strings in \( L(A) \) with length \( i \) and test if B accepts – theoretically works, but takes infinite time for infinite words – will never halt (this solution does not show decidability)

\[ A_{\text{CFG}} \] is decidable – Proof

For CFG \( G \) and string \( w \), determine if \( G \) generates \( w \)

Idea 1: Simulate \( G \) to go through all derivations
  - May never terminate

Idea 2: Note \(|w| = n\); \( 2n-1 \) steps from CNF rules to each string
  - Produce all words of lengths \( n \)
  - Breadth-first search of finite depth is fixed

\[ B_{\text{CFG}} \] is a decidable language

- For CFG \( G \), determine if there is any terminal string generated by \( G \)
- Mark all variables that generate terminals
- Repeated loop:
  - Mark all variables that have previously-marked variables on its rules right sides
  - If mark \( S \), ACCEPT; otherwise reject

\[ \begin{align*}
S & \rightarrow AB \\
A & \rightarrow An \mid x \\
B & \rightarrow yB \mid d
\end{align*} \]

\[ A_{\text{CFG}} \] is a decidable language

Alternative solution:
- Breadth-first search of rules; block repeat-visits to a branch
- If you reach an all-terminal solution, accept
- If never reach all-terminal solution, reject

Are there a finite number of non-breadth-first traversals? \textbf{Yes}

Rule can lead to \( k \) possible variables: \( S \rightarrow \text{ABACCD...E} \)
Each of \( m \) variables in grammar can have \( p \) possible rules
Given no loops allowed, max # rule evals per variable in first rule is \( pm \)
Max bound on number of steps: \( p(pm)^k \) ... huge but finite
EQ_{CFG} is not a decidable language

• Regular expressions closed under complement and intersection
• CFLs not closed under complement and intersection
• We will prove non-decidable languages later

---

The Halting Problem

Key theorem to theory of computation
Addressing unsolvable problems

Unsolvable: Software verification
• For arbitrary computer program P and precise specification of program's behavior S, determine if P fulfills S

---

Halting Problem specified

A_{TM} = \{(M, w) \mid M \text{ is a TM and } M \text{ accepts } w\}

• If M loops forever on w, our TM for A_{TM} must reject w
• This problem is Turing recognizable, but not decidable!

---

Detour: Cantor diagonalization

Comparing sizes of two infinite sets
• What is larger: set of even positive integers or set of all strings in (0U1)^*?

Diagonalization: two sets have same size if each element of set A can be compared with one element of set B
From CISC 1400: Can you define bijection from set A to set B?
Example pairing

N = \{1, 2, 3, 4, ...\} and E = \{2, 4, 6, 8, ...\}
• N and E have “same size” because there exists a bijection from N to E
• f(x) = 2x

Set is countable if either it is finite or if it has the same size as N.

Q is countable

Let Q = \{m/n: m, n ∈ N\}, positive rational numbers

Follow diagonal, skipping redundant values

Concatenating infinite set of finite lists produces countable list

Take countable steps along diagonal line to reach each number in Q.

Real numbers are uncountable

Real numbers have infinite number of decimal places

Proving uncountability
• Presume we have a list of n real numbers
• Generate new real number x not in current list
  • Pick i\textsuperscript{th} decimal value of x to be different from i\textsuperscript{th} decimal value of element i in list of real numbers
• At end, x will not be in list

| R(1) | 1.532532 |
| R(2) | 0.352144 |
| R(3) | 5.244525 |
| R(4) | 9.327431 |
| R(5) | 5.366324 |
| R(6) | 4.459322 |
| ... | ... |
| x   | 3.646311 |
Uncountability implications

There are uncountably many languages
There are countably many Turing machines
Some languages are not Turing recognizable

“There are countably many Turing machines”
Each TM is captured by finite string $<M> \in \Sigma^*$
• $\Sigma^*$ is countable – add number of strings of length 0, length 1, length 2, ... (like Q)

“There are uncountably many languages”
Represent $L$ as binary sequence
• 1 for each accepted string, 0 for each reject string
• Infinite number of strings – infinite sequence of 0/1s
• Set of possible binary sequences is uncountable (like R)

“Some languages are not Turing decidable”
Set of TMs is countable
Set of Languages is uncountable
Each TM has one language
Some languages not recognized by any TM
Back to the Halting Problem

\[ A_{\text{TM}} = \{ \langle M, w \rangle \mid M \text{ is a TM and accepts } w \} \]

- Proof by contradiction – uses diagonalization

Contradiction

Assume \( A_{\text{TM}} \) is decidable

\( H \) decides \( A_{\text{TM}} \)

- Input \( \langle M, w \rangle \) causes \( H \) to accept if \( M \) accepts \( w \), otherwise \( H \) rejects

Define a TM \( D \) that calls \( H \) on \( \langle M, \langle M \rangle \rangle \), then outputs opposite answer to \( H \)

- \( D \) rejects if \( M \) accepts \( \langle M \rangle \); \( D \) accepts if \( M \) does not accept \( \langle M \rangle \)

Run \( D \) on itself

- \( D(\langle D \rangle) = \text{accept if } D \text{ does not accept } \langle D \rangle; \text{ reject if } D \text{ accepts } \langle D \rangle \)

Contradiction!

Implications

\[ A_{\text{TM}} = \{ (M, w) \mid M \text{ is a TM and } M \text{ accepts } w \} \text{ is not decidable} \]

Some algorithms are decidable

\( A_{\text{TM}} \) is Turing recognizable – just simulate \( M \) on machine
Co-Turing Recognizable

Language is co-Turing recognizable if it is the complement of a Turing-recognizable language

Theorem: Language is decidable if it is Turing-recognizable and co-Turing recognizable

Thus, for any undecidable language L, either L or L’ is not Turing-recognizable
• Is $A_{TM}'$ Turing-recognizable?

Reducibility

If A reduces to B, solution to B will solve A
Example: A: Navigate NYC  B: Reading a map

If A reduces to B
• A is no harder than B
• A could be easier than B

Reduction and decidability

If A is reducable to B and B is decidable
• A is decidable

If A is reducible to B and A is undecidable
• B must be undecidable

HALT$_{TM}$ is undecidable

We can reduce $A_{TM}$ (TM accepts \( w \)) to HALT$_{TM}$ (TM halts on \( w \))
$A_{TM}$ is undecidable, this HALT$_{TM}$ is undecidable

Proof by contradiction:
• Assume HALT$_{TM}$ is decidable – TM R
• Use R to construct TM S to decide $A_{TM}$
• S definition:
  • If R does not indicate halting for \( <M,w> \), reject w
  • If R does indicate halt, simulate M on w