BU CS 332 – Theory of Computation

Lecture 13:

- Decidable Languages
- Universal TM
- Countability and Diagonalization

Reading:

Sipser Ch 4.1, 4.2

Mark Bun

March 8, 2021

A "universal" algorithm for recognizing regular languages

"Given DFA D, string w, does D accept w?"

$A_{\mathrm{DFA}} = \{\langle D, w \rangle \mid \mathrm{DFA}\ D \text{ accepts } w\}$

Theorem: A_{DFA} is decidable

Proof: Define a 3-tape TM M on input $\langle D, w \rangle$:

1. Check if $\langle D, w \rangle$ is a valid encoding (reject if not)
2. Simulate D on w, i.e.,
    - Tape 2: Maintain w and head location of D
    - Tape 3: Maintain state of D, update according to δ
3. Accept if D ends in an accept state, reject otherwise

More Decidable Languages: Emptiness Testing

Theorem: $E_{\mathrm{DFA}} = \{\langle D \rangle \mid D \text{ is a DFA that recognizes } \emptyset\}$ is decidable $= \{\langle D \rangle \mid L(D) = \emptyset\}$

"Given DFA D, does $\underbrace{D \text{ recognize } \emptyset}$?"

$\Updownarrow$

D accepts no strings

Proof: The following TM decides E_{DFA}

Checks if possible to reach an accept state from the start state

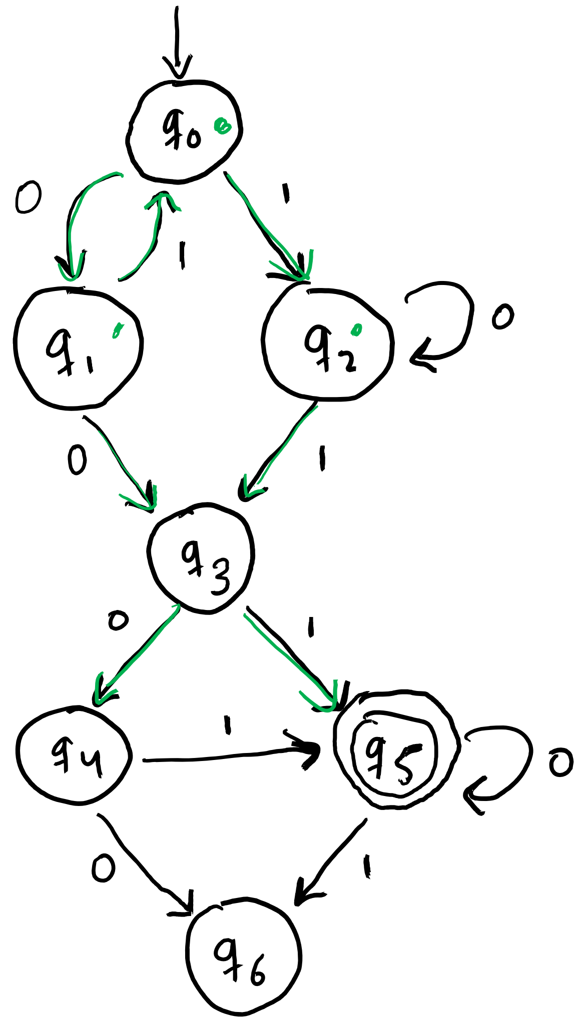
On input $\langle D \rangle$, where D is a DFA with k states:

1. Perform k steps of breadth-first search on state diagram of D to determine if an accept state is reachable from the start state
2. Reject if a DFA accept state is reachable; accept otherwise

E_{DFA} Example

$L(D) \neq \phi$ (e.g. $111 \in L(D)$)

$D =$



Breadth-first search:

- q_0
- q_1
- q_2
- q_3
- q_4
- q_5

Notice q_5 is an accept state

Reject: conclude $L(D) \neq \phi$

New Deciders from Old: Equality Testing

$EQ_{\text{DFA}} = \{\langle D_1, D_2\rangle \mid D_1, D_2 \text{ are DFAs and } L(D_1) = L(D_2)\}$

Theorem: EQ_{DFA} is decidable

Proof: The following TM decides EQ_{DFA}

On input $\langle D_1, D_2\rangle$, where $\langle D_1, D_2\rangle$ are DFAs:

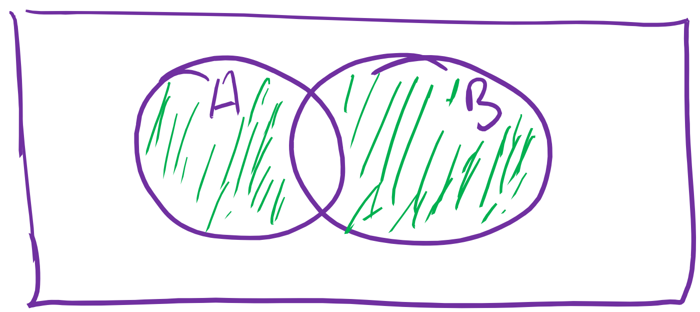
1. Construct a DFA D that recognizes the **symmetric difference** $L(D_1) \triangle L(D_2)$ $= \{ w \mid w$ is in exactly one of $L(D_1)$ or $L(D_2) \}$
2. Run the decider for E_{DFA} on $\langle D\rangle$ and return its output

Analysis (correctness):

- If $\langle D_1, D_2 \rangle \in EQ_{\text{DFA}}$, then $L(D_1) = L(D_2) \Rightarrow L(D_1) \triangle L(D_2) = \emptyset$ $\Rightarrow L(D) = \emptyset \Rightarrow$ accepts ✓
- If $\langle D_1, D_2 \rangle \notin EQ_{\text{DFA}}$, then $L(D_1) \neq L(D_2) \Rightarrow L(D_1) \triangle L(D_2) \neq \emptyset$ $\Rightarrow L(D) \neq \emptyset \Rightarrow$ rejects ✓

Symmetric Difference

$$A \triangle B = \{w \mid w \in A \text{ or } w \in B \text{ but not both}\}$$



$A \setminus B = \{w \mid w \in A \text{ and } w \notin B\}$

$$A \triangle B = (A \setminus B) \cup (B \setminus A)$$
$$= (A \cap \overline{B}) \cup (B \cap \overline{A}) \leftarrow$$
$$= (\overline{\overline{A} \cup B}) \cup (\overline{\overline{B} \cup A})$$

If $A = L(D_1)$, $B = L(D_2)$ $\Rightarrow$ can construct DFA for $L(D_1) \triangle L(D_2)$ using complement, union, and subset constructions

Meta-Computational Languages

$A_{\text{DFA}} = \{\langle D, w \rangle \mid \text{DFA } D \text{ accepts } w\}$

$A_{\text{TM}} = \{\langle M, w \rangle \mid \text{TM } M \text{ accepts } w\}$

$E_{\text{DFA}} = \{\langle D \rangle \mid \text{DFA } D \text{ recognizes the empty language } \emptyset\}$

$E_{\text{TM}} = \{\langle M \rangle \mid \text{TM } M \text{ recognizes the empty language } \emptyset\}$

$EQ_{\text{DFA}} = \{\langle D_1, D_2 \rangle \mid D_1 \text{ and } D_2 \text{ are DFAs}, L(D_1) = L(D_2)\}$

$EQ_{\text{TM}} = \{\langle M_1, M_2 \rangle \mid M_1 \text{ and } M_2 \text{ are TMs}, L(M_1) = L(M_2)\}$

The Universal Turing Machine

$A_{\mathrm{TM}} = \{\langle M, w\rangle \mid M \text{ is a TM that accepts input } w\}$

Theorem: A_{TM} is Turing-recognizable

The following "universal TM" U recognizes A_{TM}

can simulate any other TM

On input $\langle M, w\rangle$:

1. Simulate running M on input w
2. If M accepts, accept. If M rejects, reject.

1) If M accepts w, then U(<M,w>) accepts ✓
2) If M rejects w, then U(<M,w>) rejects ✓
3) If M loops on w, then U(<M,w>) loops
(w ∉ L(M), <M,w> ∉ A_{TM})

Universal TM and A_{TM}

Why is the Universal TM not a decider for A_{TM}?

The following "universal TM" U recognizes A_{TM}

On input $\langle M, w \rangle$: "encoding of a pair consisting of a TM M and string w"

1. Simulate running M on input w
2. If M accepts, accept. If M rejects, reject.

a) It may reject inputs $\langle M, w \rangle$ where M accepts w

b) It may accept inputs $\langle M, w \rangle$ where M rejects w

c) It may loop on inputs $\langle M, w \rangle$ where M loops on w

not a decider

d) It may loop on inputs $\langle M, w \rangle$ where M rejects w

More on the Universal TM

"It is possible to invent a single machine which can be used to compute any computable sequence. If this machine **U** is supplied with a tape on the beginning of which is written the S.D ["standard description"] of some computing machine **M**, then **U** will compute the same sequence as **M**."

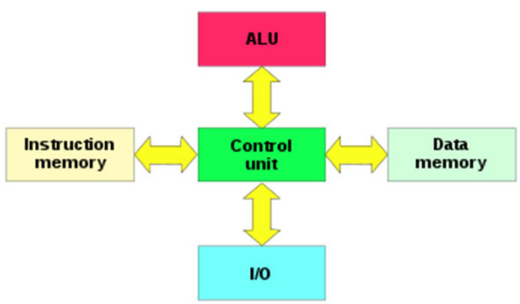
"code of a program M"

- Turing, "On Computable Numbers…" 1936

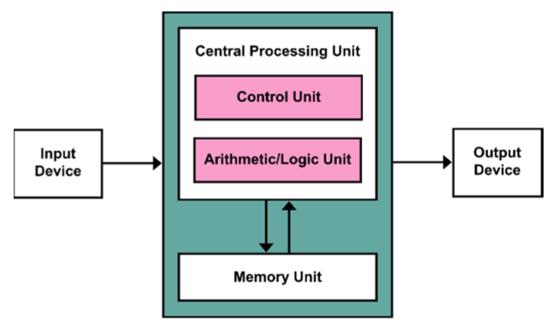
input to M

$U(\langle M, w \rangle)$

- Foreshadowed general-purpose programmable computers
- No need for specialized hardware: Virtual machines as software



Harvard architecture:
Separate instruction and data pathways



von Neumann architecture:
Programs can be treated as data

Undecidability

A_{TM} is Turing-recognizable via Universal TM U

…but it turns out A_{TM} (and $E_{\mathrm{TM}}, EQ_{\mathrm{TM}}$) is **undecidable**

i.e., computers cannot solve these problems no matter how much time they are given

Countability and Diagonalizaiton

What's your intuition?

Which of the following sets is the "biggest"?

POW2 $\subseteq$ E $\subseteq$ $\mathbb{N}$

a) The natural numbers: $\mathbb{N} = \{1, 2, 3, \ldots\}$

b) The even numbers: $E = \{2, 4, 6, \ldots\}$

relabel: 1, 2, 3, ...

c) The positive powers of 2: $POW2 = \{2, 4, 8, 16, \ldots\}$

relabel: 1, 2, 3, 4, ...
($\log_2$)

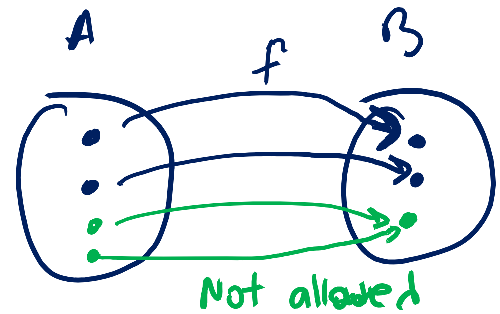
d) They all have the same size ←

Set Theory Review

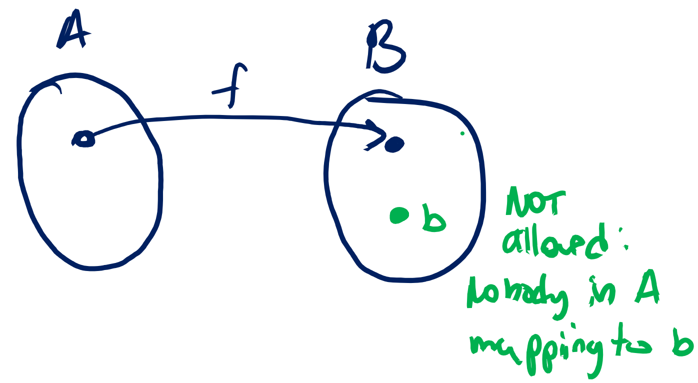
implicit: Every $a \in A$ has some $f(a) \in B$

A function $f: A \to B$ is

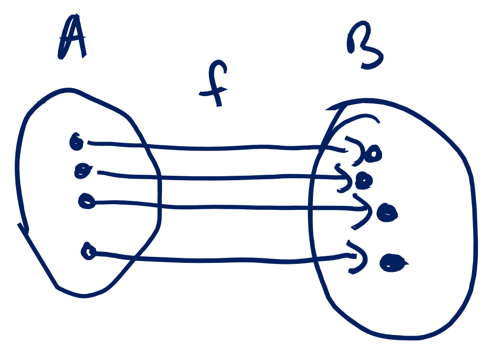
- 1-to-1 (injective) if $f(a) \neq f(a')$ for all $a \neq a'$



- onto (surjective) if for all $b \in B$, there exists $a \in A$ such that $f(a) = b$



- a correspondence (bijective) if it is 1-to-1 and onto, i.e., every $b \in B$ has a unique $a \in A$ with $f(a) = b$



How can we compare sizes of infinite sets?

Definition: Two sets have the same size if there is a bijection between them

$\{x, y, z\}$
$\downarrow \ \downarrow \ \downarrow$
$\{a, b, c\}$

$f(x) = a$
$f(y) = b$
$f(z) = c$

A set S is countable if

- it is a finite set, or
- it has the same size as $\mathbb{N}$, the set of natural numbers

(i.e. $\exists$ a bijection $f: \mathbb{N} \to S$) "countably infinite"

Examples of countable sets

- $\emptyset$
- $\{0,1\}$
- $\{0, 1, 2, \dots, 8675309\}$

finite => countable

- $E = \{2, 4, 6, 8, \dots\}$ $f(i) = 2i$
- $SQUARES = \{1, 4, 9, 16, 25, \dots\}$ $f(i) = i^2$
- $POW2 = \{2, 4, 8, 16, 32, \dots\}$ $f(i) = 2^i$

countably infinite => countable

$$|E| = |SQUARES| = |POW2| = |\mathbb{N}|$$

Construct bijection $f: \mathbb{N} \to \mathbb{N} \times \mathbb{N}$

How to show that $\mathbb{N} \times \mathbb{N}$ is countable?

" $\{(x, y) \mid x \in \mathbb{N},\ y \in \mathbb{N}\}$

$f(1) =$ $(1,1)$	$f(2) =$ $(2,1)$	$f(6) =$ $(3,1)$	$(4,1)$	$(5,1)$	…
$f(3) =$ $(1,2)$	$f(5) =$ $(2,2)$	$(3,2)$	$(4,2)$	$(5,2)$	…
$f(4) =$ $(1,3)$	$(2,3)$	$(3,3)$	$(4,3)$	$(5,3)$	…
$(1,4)$	$(2,4)$	$(3,4)$	$(4,4)$	$(5,4)$	…
$(1,5)$	$(2,5)$	$(3,5)$	$(4,5)$	$(5,5)$	
					⋱

How to argue that a set S is countable

- Describe how to "list" the elements of S, usually in stages:

Ex: Stage 1) List all pairs (x, y) such that $x + y = 2$ (1,1)

Stage 2) List all pairs (x, y) such that $x + y = 3$ (2,1) (1,2)

…

Stage n) List all pairs (x, y) such that $x + y = n + 1$

n'th diagonal

…

- Argue that every element of S appears in the list

Ex: Any $(x, y) \in \mathbb{N} \times \mathbb{N}$ will be listed in stage $x + y - 1$

- Define the bijection $f: \mathbb{N} \to S$ by $f(n) =$ the n'th element in this list (ignoring duplicates if needed)

Subsets of countable sets

If A and B are sets with $A \subseteq B$ (A is a subset of B), which of the following statements are true?

a) If A is countable, then B is countable

b) If B is countable, then A is countable

c) Both are true

d) Neither is true

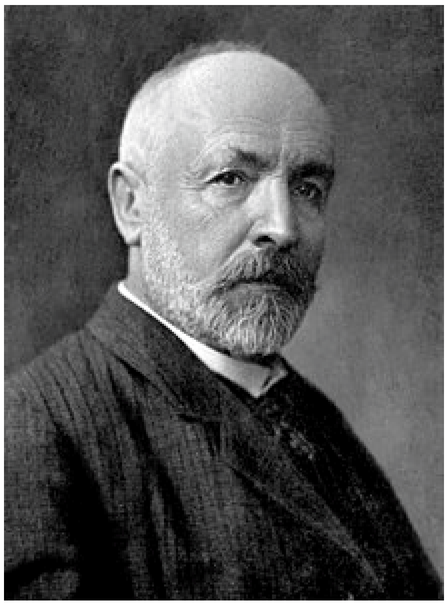
More examples of countable sets

- $\{0,1\}^*$ $\langle \cdot \rangle : \{\text{all TMs}\} \to \{0,1\}^*$
- $\{\langle M \rangle \mid M \text{ is a Turing machine}\}$ $\subseteq \{0,1\}^*$
- $\mathbb{Q} = \{\text{rational numbers}\}$ Use the construction for $\mathbb{N} \times \mathbb{N}$

numerators denominators

So what *isn't* countable?

Cantor’s Diagonalization Method



Georg Cantor 1845-1918

- Invented set theory
- Defined countability, uncountability, cardinal and ordinal numbers, …

Some praise for his work:

“Scientific charlatan…renegade…corruptor of youth”
–L. Kronecker

“Set theory is wrong…utter nonsense…laughable”
–L. Wittgenstein

Uncountability of the reals

Theorem: The real interval $(0, 1)$ is uncountable.

Proof: Assume for the sake of contradiction it were countable, and let $f: \mathbb{N} \to (0,1)$ be a bijection

n	$f(n)$
1	$0 . \boxed{d_1^1} d_2^1\ d_3^1\ d_4^1\ d_5^1 \ldots$
2	$0 . d_1^2 \boxed{d_2^2} d_3^2\ d_4^2\ d_5^2 \ldots$
3	$0 . d_1^3\ d_2^3 \boxed{d_3^3} d_4^3\ d_5^3 \ldots$
4	$0 . d_1^4\ d_2^4\ d_3^4 \boxed{d_4^4} d_5^4 \ldots$
5	$0 . d_1^5\ d_2^5\ d_3^5\ d_4^5\ d_5^5 \ldots$
⋮	⋮

d_j^n = j th digit of $f(n)$

set $b_1 \neq d_1^1$
set $b_2 \neq d_2^2$
$b_3 \neq d_3^3$
⋮

Construct $b \in (0,1)$ which does not appear in this table – contradiction!

b is not f(n) for any n
contradicts f being onto ∴ ✱

$b = 0. b_1 b_2 b_3 \ldots$ where $b_n \neq d_n^n$ (digit n of $f(n)$)

$\forall n \quad b_n \neq d_n^n \Rightarrow b \neq f(n) \Rightarrow$ b does not show up in the table