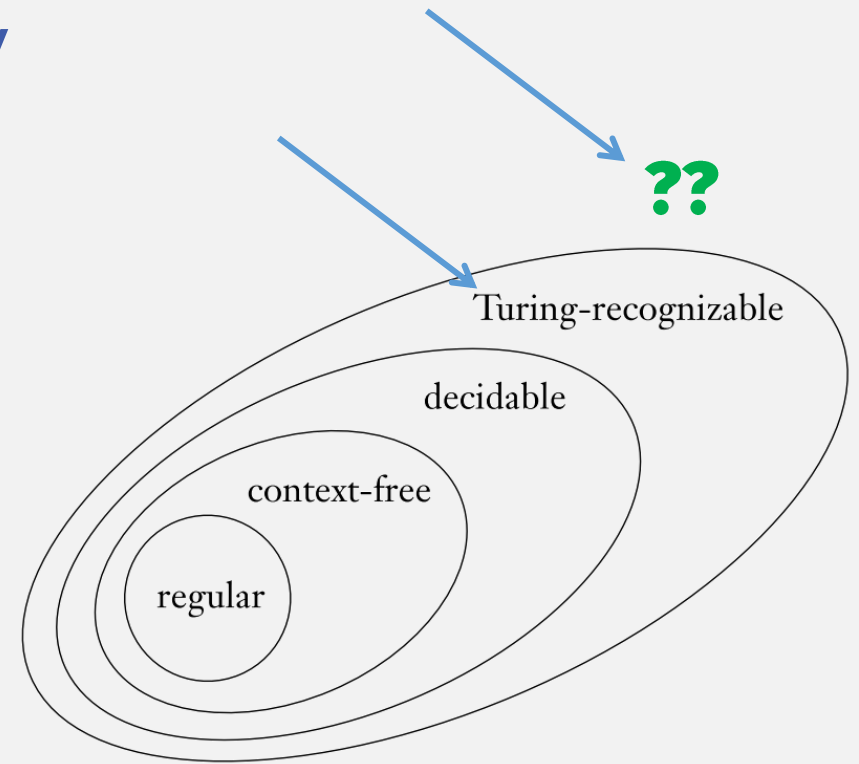


UMB CS 420
Undecidability
Monday, April 22, 2024



Announcements

- HW 9 out
 - due Wednesday 4/22 12pm noon
 - Problems 3 and 4 moved to HW 10

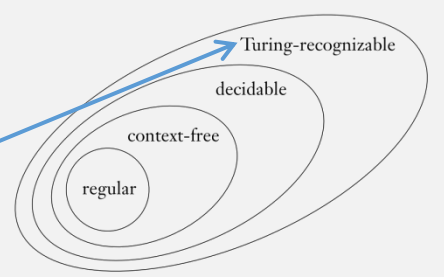
Lecture Participation Question

- Is the Universal Turing Machine (A_{TM}) a decider?

Language: of DFA description + string pairs, i.e., compute whether a DFA accepts a string

Recap: Decidability of Regular and CFLs

- $A_{\text{DFA}} = \{\langle B, w \rangle \mid B \text{ is a DFA that accepts input string } w\}$ Decidable
- $A_{\text{NFA}} = \{\langle B, w \rangle \mid B \text{ is an NFA that accepts input string } w\}$ Decidable
- $A_{\text{REX}} = \{\langle R, w \rangle \mid R \text{ is a regular expression that generates string } w\}$ Decidable
- $E_{\text{DFA}} = \{\langle A \rangle \mid A \text{ is a DFA and } L(A) = \emptyset\}$ Decidable
Compute something about DFA language, from its description
- $EQ_{\text{DFA}} = \{\langle A, B \rangle \mid A \text{ and } B \text{ are DFAs and } L(A) = L(B)\}$ Decidable
- $A_{\text{CFG}} = \{\langle G, w \rangle \mid G \text{ is a CFG that generates string } w\}$ Decidable
- $E_{\text{CFG}} = \{\langle G \rangle \mid G \text{ is a CFG and } L(G) = \emptyset\}$ Decidable
- $EQ_{\text{CFG}} = \{\langle G, H \rangle \mid G \text{ and } H \text{ are CFGs and } L(G) = L(H)\}$ Undecidable?
- $A_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$ Undecidable?
compute whether a TM accepts a string



Thm: A_{TM} is Turing-recognizable

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

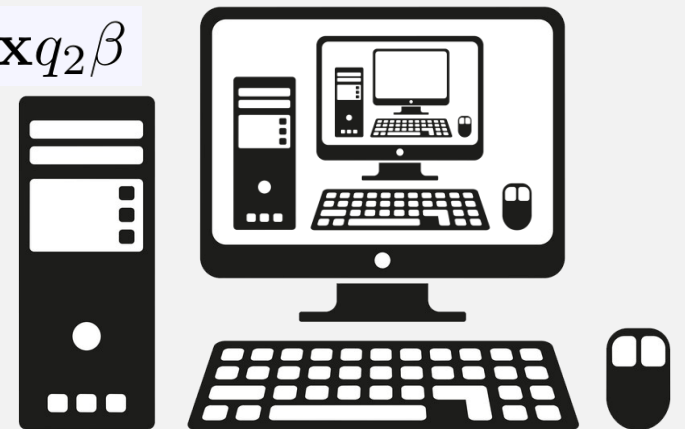
U = “On input $\langle M, w \rangle$, where M is a TM and w is a string:

1. Simulate M on input w . ← Can go into infinite loop, causing U to loop
2. If M ever enters its accept state, *accept*; if M ever enters its reject state, *reject*.”

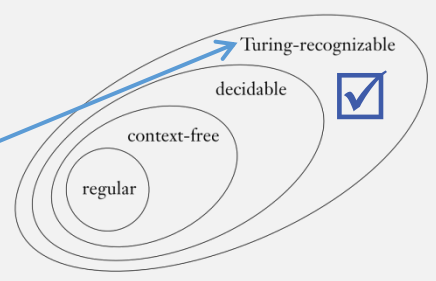
U = Implements TM computation steps $\alpha q_1 a \beta \vdash \alpha x q_2 \beta$

- i.e., “The Universal Turing Machine”
- “Program” simulating other programs (interpreter)
- Problem (Step 1): U loops when M loops

So it's not a decider. Is it recognizer?



Termination argument?



Thm: A_{TM} is Turing-recognizable

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

$U =$ “On input $\langle M, w \rangle$, where M is a TM and w is a string:

1. Simulate M on input w .
2. If M ever enters its accept state, *accept*; if M ever enters its reject state, *reject*.”

“Actual” behavior

“Expected” behavior

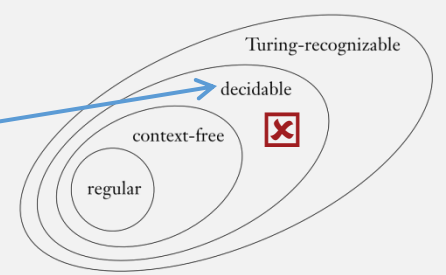
Example Str	M on input w ?	U ?	In A_{TM} lang?
$\langle M_1, 01\#01 \rangle$	Accept	Accept	Yes
$\langle M_1, 00\#11 \rangle$	Reject	Reject	No
$\langle M_{loop}, * \rangle$	Loop!	Loop!	No

Columns must match!

Let:
 - $M_1 =$ “ $w\#w$ ” lang decider
 - $M_{loop} =$ looping TM

Is this right? Yes!

How to prove ... not in here?



Thm: A_{TM} is undecidable

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

- ???

Prove: Spider-Man does not exist ???



In general, proving something not true is different (and harder) than proving it true

In some cases, it's possible, but typically requires new proof techniques!

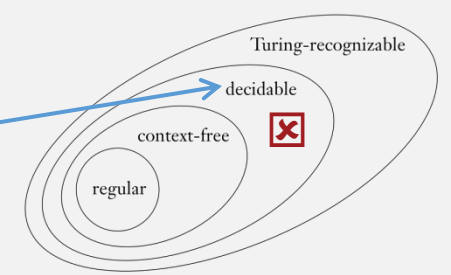
Example (**Regular** Languages)

Prove a language is **regular**:

- Create a DFA

Prove a language is **not regular**:

- Proof by contradiction using **Pumping Lemma**



Not in here?

Thm: A_{TM} is undecidable

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



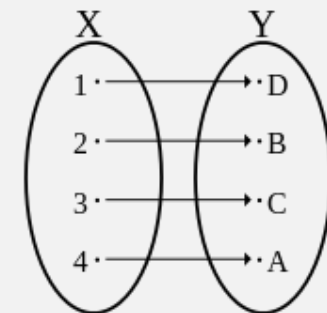
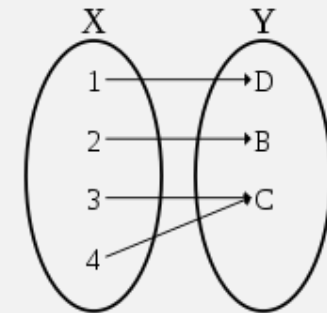
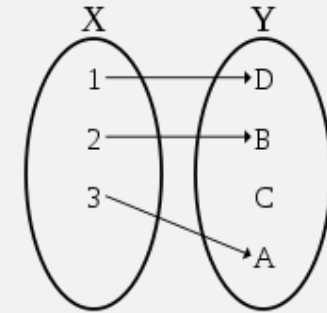
Example (**decidable** languages)
 Prove a language is **decidable**:
 - Create a **decider** TM (with termination argument)

Prove a language is **not decidable**:
 - ????

today

Kinds of Functions (a fn maps DOMAIN \rightarrow RANGE)

- **Injective**, a.k.a., “one-to-one”
 - Every element in DOMAIN has a unique mapping
 - How to remember:
 - Entire DOMAIN is mapped “in” to the RANGE
- **Surjective**, a.k.a., “onto”
 - Every element in RANGE is mapped to
 - How to remember:
 - “Sur” = “over” (eg, survey); DOMAIN is mapped “over” the RANGE
- **Bijective**, a.k.a., “correspondence” or “one-to-one correspondence”
 - Is both injective and surjective
 - Unique pairing of every element in DOMAIN and RANGE



Countability

- A set is “**countable**” if it is:
 - Finite
 - Or, there exists a **bijection** between the set and the natural numbers
 - In this case, the set has the same size as the set of natural numbers
 - This is called “**countably infinite**”

Exercise: Which set is larger?

- The set of:
 - Natural numbers, or
 - Even numbers?
- They are the same size! Both are **countably infinite**
 - Proof: Bijection:

n	$f(n) = 2n$
1	2
2	4
3	6
\vdots	\vdots

Natural numbers Even numbers

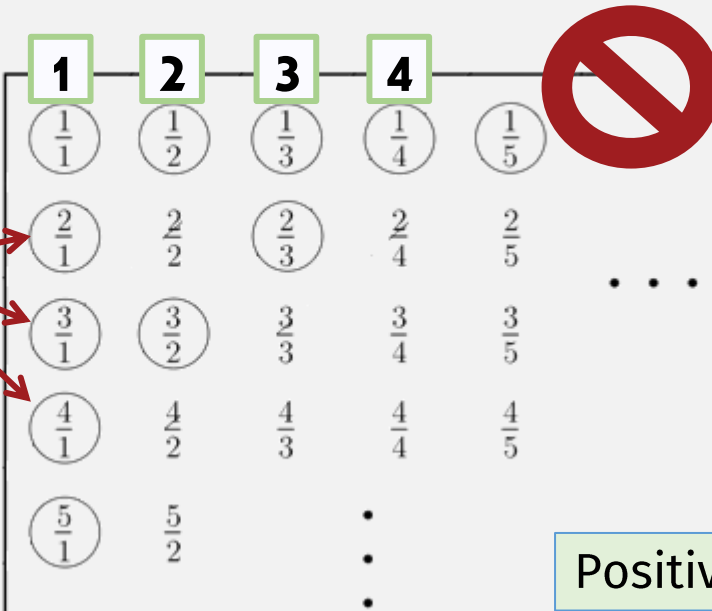
Every natural number maps to a unique even number, and vice versa

Exercise: Which set is larger?

- The set of:
 - Natural numbers \mathcal{N} , or
 - Positive rational numbers? $\mathcal{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathcal{N} \right\}$
- They are the same size! Both are **countably infinite**

A possible mapping of Natural numbers to Positive rationals?

So these don't get mapped to:
(not a **bijection**)



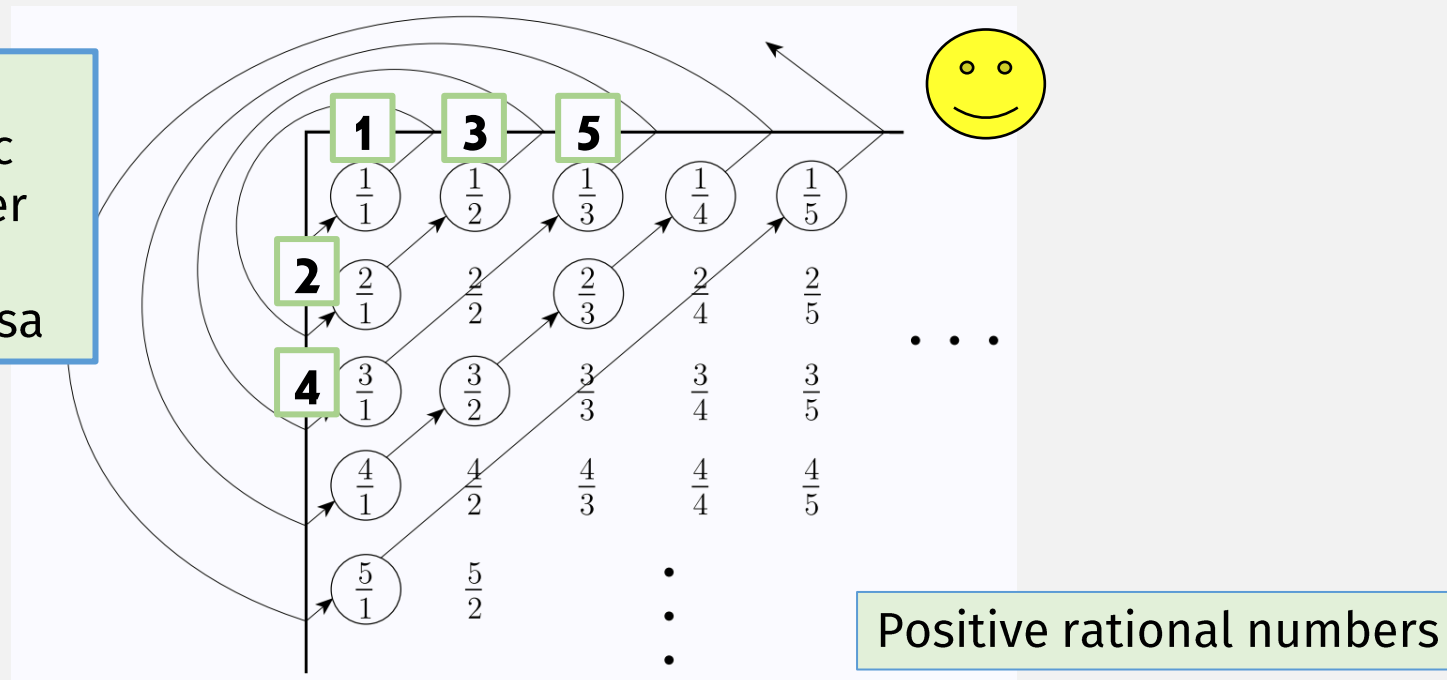
But, each row is infinite

Positive rational numbers

Exercise: Which set is larger?

- The set of:
 - Natural numbers \mathcal{N} , or
 - Positive rational numbers? $\mathcal{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathcal{N} \right\}$
- They are the same size! Both are **countably infinite**

Another mapping:
This is a **bijection** bc
every natural number
maps to a unique
fraction, and vice versa



Exercise: Which set is larger?

- The set of:
 - Natural numbers \mathcal{N} , or
 - Real numbers? \mathcal{R}
- There are more real numbers. It is **uncountably infinite**.

This proof technique is called **diagonalization**

Proof, by contradiction:

- Assume: a bijection between natural and real numbers exists.

- So: every natural num maps to a unique real, and vice versa

But we show that in any given mapping,

- Some real number is not mapped to ...
- E.g., a number that has different digits at each position:

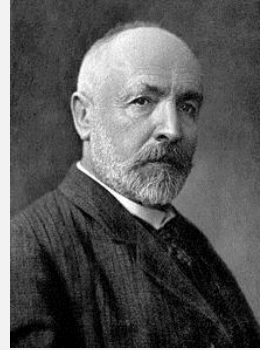
$$x = 0.\overset{\text{different}}{\underset{\text{e.g.}}{\mathbf{4}}}\overset{\text{different}}{\underset{\text{e.g.}}{\mathbf{6}}}\overset{\text{different}}{\underset{\text{e.g.}}{\mathbf{4}}}\overset{\text{different}}{\underset{\text{e.g.}}{\mathbf{1}}}\dots$$

n	$f(n)$
1	3.14159...
2	55.55555...
3	0.12345...
4	0.50000...
\vdots	\vdots

A hypothetical mapping

- This number cannot be in the mapping ...
- ... So we have a **contradiction!**

Georg Cantor



- Invented set theory
- Came up with **countable infinity** (1873)
- **And uncountability:**
 - Also: how to show uncountability with “**diagonalization**” technique



A formative day for Georg Cantor.

Diagonalization with Turing Machines

Diagonal: Result of Giving a TM its own Encoding as Input

All TM Encodings

	$\langle M_1 \rangle$	$\langle M_2 \rangle$	$\langle M_3 \rangle$	$\langle M_4 \rangle$...	$\langle D \rangle$...
M_1	<u>accept</u>	reject	accept	reject		accept	
M_2	accept	<u>accept</u>	accept	accept	...	accept	...
M_3	reject	reject	<u>reject</u>	reject		reject	
M_4	accept	accept	reject	<u>reject</u>		accept	
\vdots			\vdots		\ddots		
D	reject	reject	accept	accept		<u>?</u>	
\vdots							

opposites

All TMs

Try to construct "opposite" TM D

TM D can't exist!

It must both accept and reject!

What should happen here?

Thm: A_{TM} is undecidable

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

Proof by contradiction:

1. Assume A_{TM} is decidable. So there exists a decider H for it:

$$H(\langle M, w \rangle) = \begin{cases} \text{accept} & \text{if } M \text{ accepts } w \\ \text{reject} & \text{if } M \text{ does not accept } w \end{cases}$$

2. Use H in another TM ... the impossible “opposite” machine:

$D =$ “On input $\langle M \rangle$, where M is a TM:

1. Run H on input $\langle M, \langle M \rangle \rangle$.
2. Output the opposite of what H outputs. That is, if H accepts, *reject*; and if H rejects, *accept*.”

From previous slide
(does opposite of
what input TM would
do if given itself)

H computes: M 's result with itself as input

Do the opposite

Thm: A_{TM} is undecidable

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

Proof by contradiction:

This cannot be true

1. Assume A_{TM} is decidable. So there exists a decider H for it:

$$H(\langle M, w \rangle) = \begin{cases} \text{accept} & \text{if } M \text{ accepts } w \\ \text{reject} & \text{if } M \text{ does not accept } w \end{cases}$$

2. Use H in another TM ... the impossible “opposite” machine:

~~$D =$ “On input $\langle M \rangle$, where M is a TM:~~

1. ~~Run H on input $\langle M, \langle M \rangle \rangle$.~~
2. ~~Output the opposite of what H outputs. That is, if H accepts, *reject*; and if H rejects, *accept*.”~~

3. But D does not exist! **Contradiction!** So the assumption is false.

Easier Undecidability Proofs

- We proved $A_{TM} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$ undecidable ...
- ... by contradiction:
 - Use hypothetical A_{TM} decider to create an impossible decider “ D ”!

reduce “ D problem” to A_{TM}

- Step # 1: coming up with “ D ” --- hard!
 - Need to invent **diagonalization**

- Step # 2: **reduce** “ D ” problem to A_{TM} --- easier!

	$\langle M_1 \rangle$	$\langle M_2 \rangle$	$\langle M_3 \rangle$	$\langle M_4 \rangle$	\dots	$\langle D \rangle$
M_1	<u>accept</u>	reject	accept	reject		accept
M_2	accept	<u>accept</u>	accept	accept	\dots	accept
M_3	reject	reject	<u>reject</u>	reject		reject
M_4	accept	accept	reject	<u>reject</u>		accept
\vdots			\vdots		\ddots	
D	reject	reject	accept	accept		<u>?</u>

- From now on: undecidability proofs only need step # 2!
 - And we now have two “impossible” problems to choose from

The Halting Problem

$$HALT_{TM} = \{ \langle M, w \rangle \mid M \text{ is a TM and } M \text{ halts on input } w \}$$

Thm: $HALT_{TM}$ is undecidable

Proof, by contradiction:

reduce (from known undecidable) A_{TM} to $HALT_{TM}$

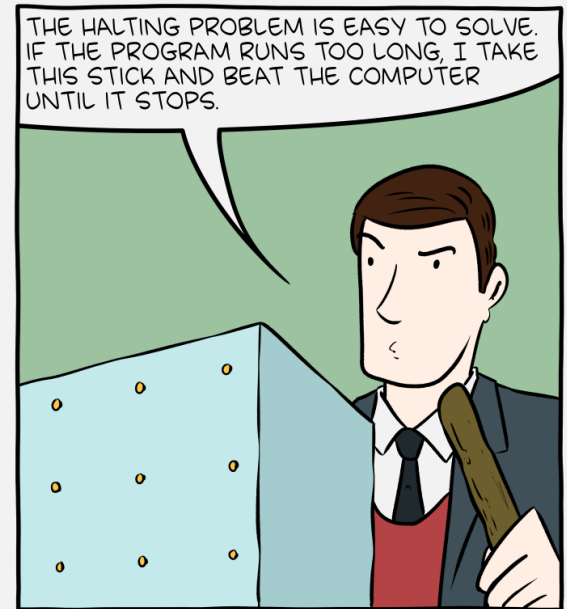
- Assume: $HALT_{TM}$ has decider R ; use it to create decider for A_{TM} :

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

• ...

contradiction

- But A_{TM} is undecidable and has no decider!



What if Alan Turing had been an engineer?

The Halting Problem

$$HALT_{TM} = \{ \langle M, w \rangle \mid M \text{ is a TM and } M \text{ halts on input } w \}$$

Thm: $HALT_{TM}$ is undecidable

Proof, by contradiction: Using our hypothetical $HALT_{TM}$ decider R

- Assume: $HALT_{TM}$ has decider R ; use it to create decider for A_{TM} :

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

$S =$ “On input $\langle M, w \rangle$, an encoding of a TM M and a string w :

1. Run TM R on input $\langle M, w \rangle$.
2. If R rejects, *reject*. ← This means M loops on input w
3. If R accepts, simulate M on w until it halts. ← This step always halts
4. If M has accepted, *accept*; if M has rejected, *reject*.”

Termination argument:

Step 1: R is a decider so always halts

Step 3: M always halts because R said so

Undecidability Proof Technique #1:
Reduce (directly) from A_{TM}
(by creating A_{TM} decider)

The Halting Problem

$$HALT_{TM} = \{ \langle M, w \rangle \mid M \text{ is a TM and } M \text{ halts on input } w \}$$

Thm: $HALT_{TM}$ is undecidable

Proof, by contradiction:

- Assume: $HALT_{TM}$ has *decider* R ; use it to create decider for A_{TM} :

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

~~$S =$ “On input $\langle M, w \rangle$, an encoding of a TM M and a string w :~~

- ~~1. Run TM R on input $\langle M, w \rangle$.~~
- ~~2. If R rejects, *reject*.~~
- ~~3. If R accepts, simulate M on w until it halts.~~
- ~~4. If M has accepted, *accept*; if M has rejected, *reject*.”~~


Now we have three
“impossible” deciders
to choose from

- But A_{TM} is undecidable! I.e., this decider does not exist!
 - So $HALT_{TM}$ is also undecidable!

Interlude: Reducing from $HALT_{TM}$

A practical thought experiment ...
... about compiler optimizations

Your compiler changes your program!

If TRUE then A else B  A

1 + 2 + 3  6

Compiler Optimizations

Optimization - [docs](#)

- `-O0`
 - No optimization, faster compilation time, better for debugging builds.
- `-O2`
- `-O3`
 - Higher level of optimization. Slower compile-time, better for production builds.
- `-Ofast`
 - Enables higher level of optimization than (`-O3`). It enables lots of flags as can be seen [src](#) (`-ffloat-store`, `-ffsast-math`, `-ffinite-math-only`, `-O3 ...`)
- `-finline-functions`
- `-m64`
- `-funroll-loops`
- `-fvectorize`
- `-fprofile-generate`

Types of optimization [\[edit \]](#)

Techniques used in optimization can be broken up among various *scopes* which can affect anything from a single statement to the entire program. Generally speaking, locally scoped techniques are easier to implement than global ones but result in smaller gains. Some examples of scopes include:

Peephole optimizations

These are usually performed late in the compilation process after [machine code](#) has been generated. This form of optimization examines a few adjacent instructions (like "looking through a peephole" at the code) to see whether they can be replaced by a single instruction or a shorter sequence of instructions.^[2] For instance, a multiplication of a value by 2 might be more efficiently executed by [left-shifting](#) the value or by adding the value to itself (this example is also an instance of [strength reduction](#)).

Local optimizations

These only consider information local to a [basic block](#).^[3] Since basic blocks have no control flow, these optimizations need very little analysis, saving time and reducing storage requirements, but this also means that no information is preserved across jumps.

Global optimizations

These are also called "intraprocedural methods" and act on whole functions.^[3] This gives them more information to work with, but often makes expensive computations necessary. Worst case assumptions have to be made when function calls occur or global variables are accessed because little information about them is available.

Loop optimizations

These act on the statements which make up a loop, such as a *for* loop, for example [loop-invariant code motion](#). Loop optimizations can have a significant impact because many programs spend a large percentage of their time inside loops.^[4]

Prescient store optimizations

These allow store operations to occur earlier than would otherwise be permitted in the context of [threads](#) and locks. The process needs some way of knowing ahead of time what value will be stored by the assignment that it should have followed. The purpose of this relaxation is to allow compiler optimization to perform certain kinds of code rearrangement that preserve the semantics of properly synchronized programs.^[5]

Interprocedural, whole-program or link-time optimization

These analyze all of a program's source code. The greater quantity of information extracted means that optimizations can be more effective compared to when they only have access to local information, i.e. within a single function. This kind of optimization can also allow new techniques to be performed. For instance, function [inlining](#), where a call to a function is replaced by a copy of the function body.

Machine code optimization and object code optimizer

These analyze the executable task image of the program after all of an executable machine code has been [linked](#). Some of the techniques that can be applied in a more limited scope, such as macro compression which saves space by collapsing common sequences of instructions, are more effective when the entire executable task image is available for analysis.^[6]

The Optimal Optimizing Compiler

“Full Employment” Theorem

Thm: The Optimal (C++) Optimizing Compiler does not exist

Proof, by contradiction:

Assume: *OPT* is the Perfect Optimizing Compiler

Use it to create $HALT_{TM}$ decider (accepts $\langle M, w \rangle$ if M halts with w , else **rejects**):

$S =$ On input $\langle M, w \rangle$, where M is C++ program and w is string:

- If $OPT(M) == \text{for}(;;)$
 - a) Then **Reject**
 - b) Else **Accept**

In computer science and mathematics, a **full employment theorem** is a term used, often humorously, to refer to a theorem which states that no algorithm can optimally perform a particular task done by some class of professionals. The name arises because such a theorem ensures that there is endless scope to keep discovering new techniques to improve the way at least some specific task is done.

For example, the *full employment theorem for compiler writers* states that there is no such thing as a provably perfect size-optimizing compiler, as such a proof for the compiler would have to **detect non-terminating computations** and reduce them to a one-instruction **infinite loop**. Thus, the existence of a provably perfect size-optimizing compiler would imply a solution to the **halting problem**, which cannot exist. This also implies that there may always be a better compiler since the proof that one has the best compiler cannot exist. Therefore, compiler writers will always be able to speculate that they have something to improve.

Summary: The Limits of Algorithms

- $A_{\text{DFA}} = \{\langle B, w \rangle \mid B \text{ is a DFA that accepts input string } w\}$ Decidable
- $A_{\text{CFG}} = \{\langle G, w \rangle \mid G \text{ is a CFG that generates string } w\}$ Decidable
- $A_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$ **Undecidable**
- $HALT_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ halts on input } w\}$ **Undecidable**

Similar languages

It's straightforward to use hypothetical $HALT_{\text{TM}}$ decider to create A_{TM} decider

Summary: The Limits of Algorithms

- $A_{\text{DFA}} = \{\langle B, w \rangle \mid B \text{ is a DFA that accepts input string } w\}$ Decidable
- $A_{\text{CFG}} = \{\langle G, w \rangle \mid G \text{ is a CFG that generates string } w\}$ Decidable
- $A_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$ **Undecidable**
- $HALT_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ halts on input } w\}$ **Undecidable**
- $E_{\text{DFA}} = \{\langle A \rangle \mid A \text{ is a DFA and } L(A) = \emptyset\}$ Decidable
- $E_{\text{CFG}} = \{\langle G \rangle \mid G \text{ is a CFG and } L(G) = \emptyset\}$ Decidable
- $E_{\text{TM}} = \{\langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset\}$ **Undecidable**

Not as similar languages

next

How can we use a hypothetical E_{TM} decider to create A_{TM} or $HALT_{\text{TM}}$ decider?

Reducibility: Modifying the TM

$$E_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset \}$$

Thm: E_{TM} is undecidable

Proof, by contradiction:

- Assume E_{TM} has *decider* R ; use it to create *decider* for A_{TM} :

$S =$ “On input $\langle M, w \rangle$, an encoding of a TM M and a string w :

First, construct M_1

- Run R on input $\langle M_1 \rangle$ ← Note: M_1 is only used as arg to R ; we never run it!
- If R accepts, *reject* (because it means $\langle M \rangle$ doesn't accept w)
- if R rejects, then *accept* ($\langle M \rangle$ accepts w)

- Idea: Wrap $\langle M \rangle$ in a new TM that can only accept w :

$M_1 =$ “On input x :

1. If $x \neq w$, *reject*. ← Input not w , always reject

Input is w , maybe accept →

2. If $x = w$, run M on input w and *accept* if M does.”

M_1 accepts w if M does

Reducibility: Modifying the TM

$$E_{\text{TM}} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset \}$$

Thm: E_{TM} is undecidable

Proof, by contradiction:

This decider for A_{TM} cannot exist!

- Assume E_{TM} has *decider* R ; use it to create *decider* for A_{TM} :

~~$S =$ “On input $\langle M, w \rangle$, an encoding of a TM M and a string w :~~

~~First, construct M_1~~

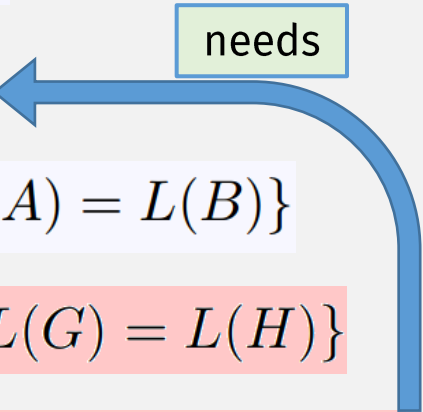
- ~~• Run R on input $\langle M \rangle$~~
- ~~• If R accepts, *reject* (because it means $\langle M \rangle$ doesn't accept w)~~
- ~~• if R rejects, then *accept* ($\langle M \rangle$ accepts w)~~

- Idea: Wrap $\langle M \rangle$ in a new TM that can only accept w :

$M_1 =$ “On input x :

1. If $x \neq w$, *reject*.
2. If $x = w$, run M on input w and *accept* if M does.”

Summary: The Limits of Algorithms

- $A_{\text{DFA}} = \{\langle B, w \rangle \mid B \text{ is a DFA that accepts input string } w\}$ Decidable
- $A_{\text{CFG}} = \{\langle G, w \rangle \mid G \text{ is a CFG that generates string } w\}$ Decidable
- $A_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$ **Undecidable**
- $E_{\text{DFA}} = \{\langle A \rangle \mid A \text{ is a DFA and } L(A) = \emptyset\}$ Decidable
- $E_{\text{CFG}} = \{\langle G \rangle \mid G \text{ is a CFG and } L(G) = \emptyset\}$ Decidable
- $E_{\text{TM}} = \{\langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset\}$ **Undecidable** 
- $EQ_{\text{DFA}} = \{\langle A, B \rangle \mid A \text{ and } B \text{ are DFAs and } L(A) = L(B)\}$ Decidable
- $EQ_{\text{CFG}} = \{\langle G, H \rangle \mid G \text{ and } H \text{ are CFGs and } L(G) = L(H)\}$ **Undecidable**
- $EQ_{\text{TM}} = \{\langle M_1, M_2 \rangle \mid M_1 \text{ and } M_2 \text{ are TMs and } L(M_1) = L(M_2)\}$ **Undecidable**

needs

next

Reduce to something else: EQ_{TM} is undecidable

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

Proof, by contradiction:

- Assume: EQ_{TM} has decider R ; use it to create decider for A_{TM} :

$$E_{TM} = \{\langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset\}$$

$S =$ “On input $\langle M \rangle$, where M is a TM:

1. Run R on input $\langle M, M_1 \rangle$, where M_1 is a TM that rejects all inputs.
2. If R accepts, *accept*; if R rejects, *reject*.”

Reduce to something else: EQ_{TM} is undecidable

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

Proof, by contradiction:

- Assume: EQ_{TM} has decider R ; use it to create decider for E_{TM} :

$$= \{ \langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset \}$$

~~$S =$ “On input $\langle M \rangle$, where M is a TM:~~

- ~~1. Run R on input $\langle M, M_1 \rangle$, where M_1 is a TM that rejects all inputs.~~
- ~~2. If R accepts, *accept*; if R rejects, *reject*.”~~

- But E_{TM} is undecidable!

Summary: Undecidability Proof Techniques

- Proof Technique #1: $A_{TM} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$
 - Use hypothetical decider to implement impossible A_{TM} decider ↓ Reduce
 - Example Proof: $HALT_{TM} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ halts on input } w\}$

- Proof Technique #2:
 - Use hypothetical decider to implement impossible A_{TM} decider
 - But first modify the input M
 - Example Proof: $E_{TM} = \{\langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset\}$ ↓ Reduce

- Proof Technique #3:
 - Use hypothetical decider to implement non- A_{TM} impossible decider
 - Example Proof: $EQ_{TM} = \{\langle M_1, M_2 \rangle \mid M_1 \text{ and } M_2 \text{ are TMs and } L(M_1) = L(M_2)\}$

Can also combine these techniques

Summary: Decidability and Undecidability

- $A_{\text{DFA}} = \{\langle B, w \rangle \mid B \text{ is a DFA that accepts input string } w\}$ Decidable
- $A_{\text{CFG}} = \{\langle G, w \rangle \mid G \text{ is a CFG that generates string } w\}$ Decidable
- $A_{\text{TM}} = \{\langle M, w \rangle \mid M \text{ is a TM and } M \text{ accepts } w\}$ **Undecidable**
- $E_{\text{DFA}} = \{\langle A \rangle \mid A \text{ is a DFA and } L(A) = \emptyset\}$ Decidable
- $E_{\text{CFG}} = \{\langle G \rangle \mid G \text{ is a CFG and } L(G) = \emptyset\}$ Decidable
- $E_{\text{TM}} = \{\langle M \rangle \mid M \text{ is a TM and } L(M) = \emptyset\}$ **Undecidable**
- $EQ_{\text{DFA}} = \{\langle A, B \rangle \mid A \text{ and } B \text{ are DFAs and } L(A) = L(B)\}$ Decidable
- $EQ_{\text{CFG}} = \{\langle G, H \rangle \mid G \text{ and } H \text{ are CFGs and } L(G) = L(H)\}$ **Undecidable**
- $EQ_{\text{TM}} = \{\langle M_1, M_2 \rangle \mid M_1 \text{ and } M_2 \text{ are TMs and } L(M_1) = L(M_2)\}$ **Undecidable**

Also Undecidable ...

next

- $REGULAR_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a regular language} \}$

Thm: $REGULAR_{TM}$ is undecidable

$$REGULAR_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a regular language} \}$$

Proof, by contradiction:

- Assume: $REGULAR_{TM}$ has decider R ; use it to create decider for A_{TM} :

$S =$ “On input $\langle M, w \rangle$, an encoding of a TM M and a string w :

- First, construct M_2 (??)
- Run R on input $\langle M \rangle_2$
- If R accepts, *accept*; if R rejects, *reject*

Want: $L(M_2) =$

- **regular**, if M accepts w
- **nonregular**, if M does not accept w

Thm: $REGULAR_{TM}$ is undecidable (continued)

$REGULAR_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a regular language} \}$

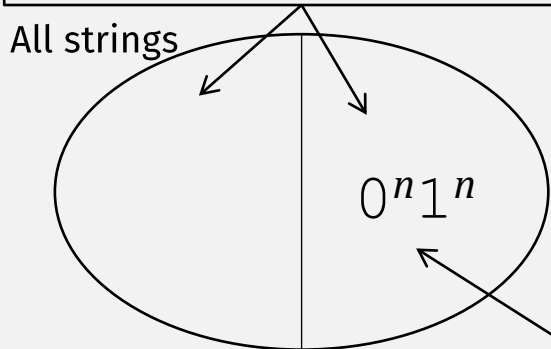
$M_2 =$ “On input x :

1. If x has the form $0^n 1^n$, *accept*.
2. If x does not have this form, run M on input w and *accept* if M accepts w .”

Always accept strings $0^n 1^n$
 $L(M_2) = \text{nonregular}$, so far

If M accepts w ,
accept everything else,
so $L(M_2) = \Sigma^* = \text{regular}$

if M does not accept w , M_2 accepts all strings (**regular lang**)



Want: $L(M_2) =$

- **regular**, if M accepts w
- **nonregular**, if M does not accept w

if M accepts w , M_2 accepts this **nonregular lang**

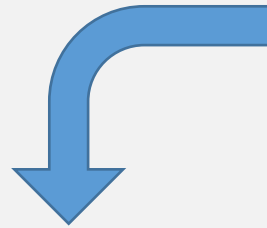
Also Undecidable ...

Seems like no algorithm can compute
anything about
the language of a Turing Machine,
i.e., about the runtime behavior of programs ...

- $REGULAR_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a regular language} \}$
- $CONTEXTFREE_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a CFL} \}$
- $DECIDABLE_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a decidable language} \}$
- $FINITE_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a finite language} \}$

An Algorithm About Program Behavior?

```
main()
{
    printf("hello, world\n");
}
```



Write a program that,
given another program as its argument,
returns TRUE if that argument prints
“Hello, World!”



TRUE

Fermat's Last Theorem
(unknown for ~350 years,
solved in 1990s)

```
main()  
{  
  If  $x^n + y^n = z^n$ , for any integer  $n > 2$   
  printf("hello, world\n");  
}
```

Write a program that,
given another program as its argument,
returns ~~TRUE~~ if that argument prints
"Hello, World!"

?????

Also Undecidable ...

Seems like no algorithm can compute
anything about
the language of a Turing Machine,
i.e., about the runtime behavior of programs ...

- $REGULAR_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a regular language} \}$
- $CONTEXTFREE_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a CFL} \}$
- $DECIDABLE_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a decidable language} \}$
- $FINITE_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and } L(M) \text{ is a finite language} \}$
- ...
- $ANYTHING_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and “... anything ...” about } L(M) \}$

Rice's Theorem

Rice's Theorem: *ANYTHING*_{TM} is Undecidable

*ANYTHING*_{TM} = { $\langle M \rangle$ | M is a TM and ... **anything** ... about $L(M)$ }

- “... **Anything** ...”, more precisely:
 - For any M_1, M_2 ,
 - if $L(M_1) = L(M_2)$
 - then $M_1 \in ANYTHING_{TM} \Leftrightarrow M_2 \in ANYTHING_{TM}$
- Also, “... **Anything** ...” must be “non-trivial”:
 - $ANYTHING_{TM} \neq \{\}$
 - $ANYTHING_{TM} \neq$ set of all TMs

Rice's Theorem: $ANYTHING_{TM}$ is Undecidable

$ANYTHING_{TM} = \{ \langle M \rangle \mid M \text{ is a TM and ... anything ... about } L(M) \}$

Proof by contradiction

- Assume some language satisfying $ANYTHING_{TM}$ has a decider R .
 - Since $ANYTHING_{TM}$ is non-trivial, then there exists $M_{ANY} \in ANYTHING_{TM}$
 - Where R accepts M_{ANY}
- Use R to create decider for A_{TM} :

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

- Create M_w :
 - $M_w =$ on input x :
 - Run M on w
 - If M rejects w : reject x
 - If M accepts w :
 - Run M_{ANY} on x and accept if it accepts, else reject

If M accepts w : $M_w = M_{ANY}$
If M doesn't accept w : M_w accepts nothing

These two cases must be different, (so R can distinguish when M accepts w)

Wait! What if the TM that accepts nothing is in $ANYTHING_{TM}$!

- Run R on M_w
 - If it accepts, then $M_w = M_{ANY}$, so M accepts w , so accept
 - Else reject

Proof still works! Just use the complement of $ANYTHING_{TM}$ instead!

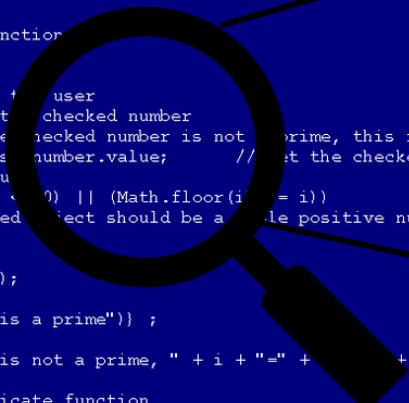
Rice's Theorem Implication

$\{ \langle M \rangle \mid M \text{ is a TM that installs malware} \}$

Undecidable!
(by Rice's Theorem)

```
function check(n)
{ // check if the number n is a prime
  var factor; // if the checked number is not a prime, this is its first factor
  var c;
  factor = 0;
  // try to divide the checked number by all numbers till its square root
  for (c=2; (c <= Math.sqrt(n)); c++)
  {
    if (n%c == 0) // is n divisible by c ?
      { factor = c; break }
  }
  return (factor);
} // end of check function

function communicate()
{ // communicate with the user
  var i; // i is the checked number
  var factor; // if the checked number is not a prime, this is its first factor
  i = document.primeset.number.value; // get the checked number
  // is it a valid input
  if ((isNaN(i)) || (i <= 0) || (Math.floor(i) != i))
  { alert ("The checked object should be a whole positive number") ;
  }
  else
  {
    factor = check (i);
    if (factor == 0)
      { alert (i + " is a prime") ;
    }
    else
      { alert (i + " is not a prime, " + i + "=" + factor + "X" + i/factor) ;
    }
  }
} // end of communicate function
```



Submit in-class participation question

On gradescope