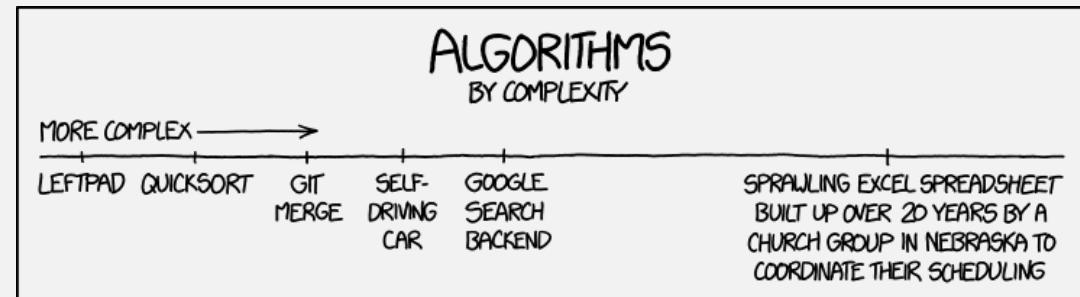


UMB CS622

Time Complexity

Monday, November 8, 2021



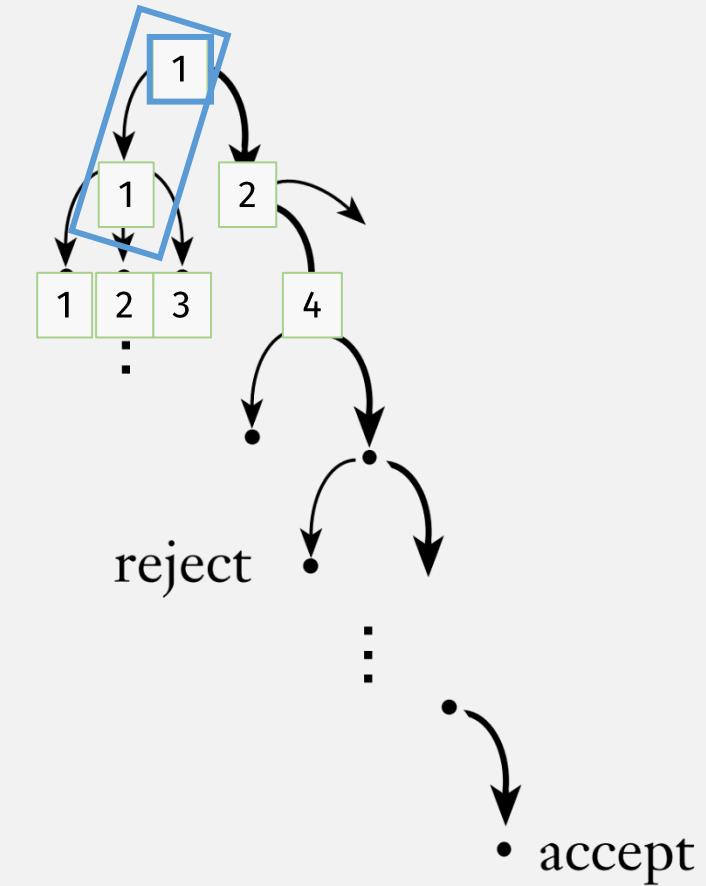
Announcements

- HW7 due Wed 11:59pm EST
- Submit “HW Solution Plans” to Piazza
 - Not at the last minute please
- HW5 grades returned

Flashback: Nondet. TM → Deterministic TM

- To simulate NTM with Det. TM:
 - Number the nodes at each step
 - Deterministically check every tree path, in breadth-first order
 - Root node: 1
 - 1-1

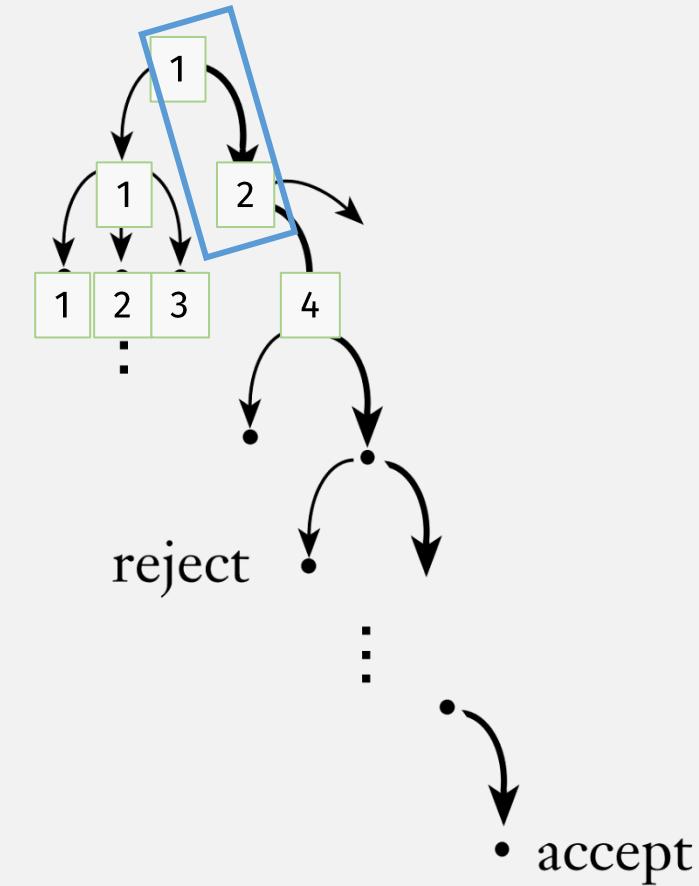
Nondeterministic computation



Flashback: Nondet. TM → Deterministic TM

- To simulate NTM with Det. TM:
 - Number the nodes at each step
 - Deterministically check every tree path, in breadth-first order
 - Root node: 1
 - 1-1
 - 1-2

Nondeterministic computation



Flashback: Nondet. TM → Deterministic TM

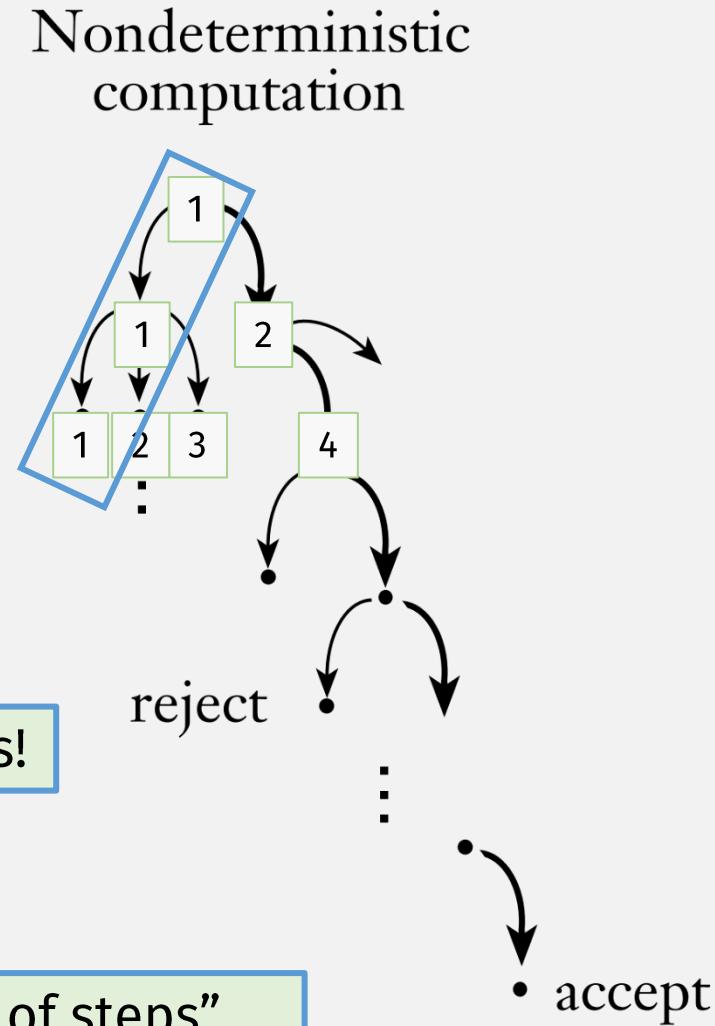
- To simulate NTM with Det. TM:
 - Number the nodes at each step
 - Deterministically check every tree path, in breadth-first order
 - Root node: 1
 - 1-1
 - 1-2
 - 1-1-1

A TM and a NTM are “equivalent” ...

.. but not if we care about the # of steps!

So how inefficient is it?

First, we need a formal way to count “# of steps” ...



A Simpler Example: $A = \{0^k 1^k \mid k \geq 0\}$

M_1 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat if both 0s and 1s remain on the tape:
 3. Scan across the tape, crossing off a single 0 and a single 1.
 4. If 0s still remain after all the 1s have been crossed off, or if 1s still remain after all the 0s have been crossed off, *reject*. Otherwise, if neither 0s nor 1s remain on the tape, *accept*.”

of steps (worst case), n = length of input:

➤ TM Line 1:

- n steps to scan + n steps to return to beginning = $2n$ steps

A Simpler Example: $A = \{0^k 1^k \mid k \geq 0\}$

M_1 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat if both 0s and 1s remain on the tape:
 3. Scan across the tape, crossing off a single 0 and a single 1.
 4. If 0s still remain after all the 1s have been crossed off, or if 1s still remain after all the 0s have been crossed off, *reject*. Otherwise, if neither 0s nor 1s remain on the tape, *accept*.

of steps (worst case), n = length of input:

- TM Line 1:
 - n steps to scan + n steps to return to beginning = $2n$ steps
- Lines 2-3 (loop):
 - steps/iteration (line 3): $n/2$ steps to find “1” + $n/2$ steps to return = n steps
 - # iterations (line 2): Each scan crosses off 2 chars, so at most $n/2$ scans
 - Total = steps/iteration * # iterations = $n(n/2) = n^2/2$ steps

A Simpler Example: $A = \{0^k 1^k \mid k \geq 0\}$

M_1 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat if both 0s and 1s remain on the tape:
3. Scan across the tape, crossing off a single 0 and a single 1.
4. If 0s still remain after all the 1s have been crossed off, or if 1s still remain after all the 0s have been crossed off, *reject*. Otherwise, if neither 0s nor 1s remain on the tape, *accept*.”

$$n^2/2 + 3n$$

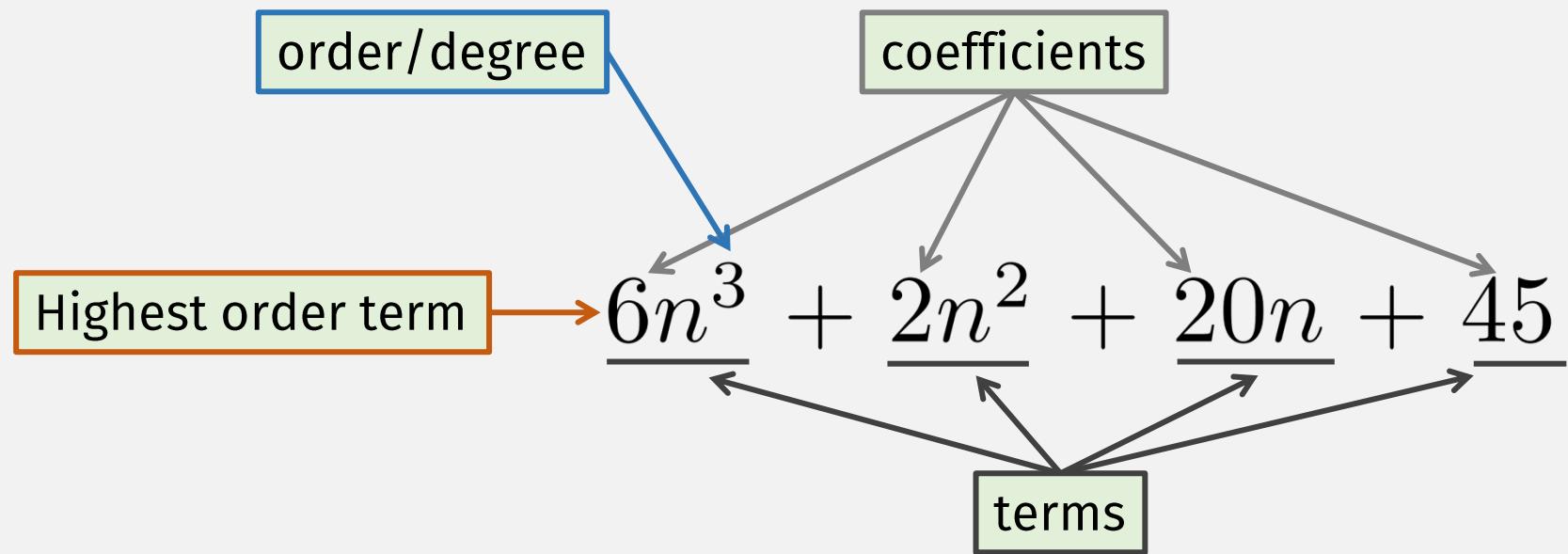
of steps (worst case), n = length of input:

- TM Line 1:
 - n steps to scan + n steps to return to beginning = **$2n$ steps**
- Lines 2-3 (loop):
 - steps/iteration (line 3): $n/2$ steps to find “1” + $n/2$ steps to return = n steps
 - # iterations (line 2): Each scan crosses off 2 chars, so at most $n/2$ scans
 - Total = steps/iteration * # iterations = $n (n/2) = n^2/2$ steps

➤ Line 4:

- **n steps** to scan input one more time
- Total: $2n + n^2/2 + n = n^2/2 + 3n$ steps

Interlude: Polynomials



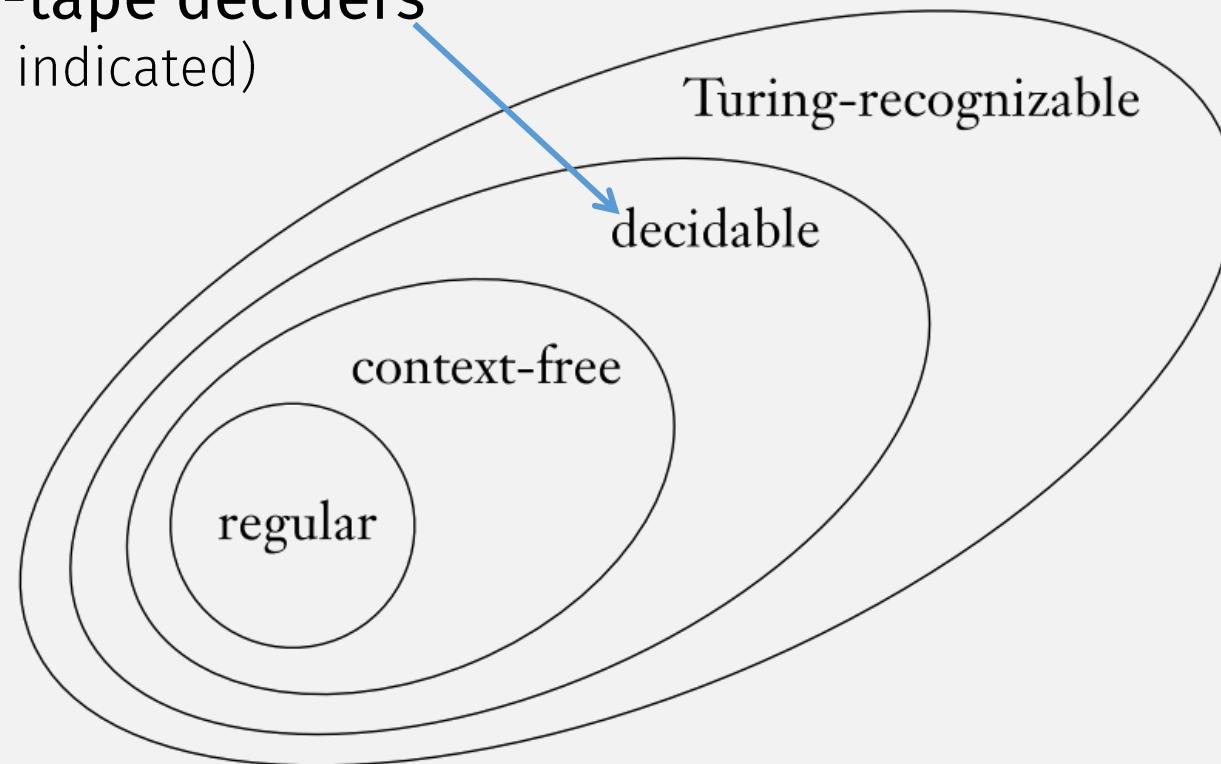
Definition: Time Complexity

Let M be a deterministic Turing machine that halts on all inputs. The ***running time*** or ***time complexity*** of M is the function $f: \mathcal{N} \rightarrow \mathcal{N}$, where $f(n)$ is the maximum number of steps that M uses on any input of length n . If $f(n)$ is the running time of M , we say that M runs in time $f(n)$ and that M is an $f(n)$ time Turing machine. Customarily we use n to represent the length of the input.

i.e., a decider (algorithm)

Where Are We Now?

We are back in here now:
deterministic, single-tape deciders
(unless otherwise indicated)



Definition: Time Complexity

NOTE: n has no units, it's only roughly “length” of the input

n can be:
characters,
states,
nodes, ...

We can use any n that is correlated with the input length

Let M be a deterministic Turing machine that halts on all inputs. The ***running time*** or ***time complexity*** of M is the function $f: \mathcal{N} \rightarrow \mathcal{N}$, where $f(n)$ is the maximum number of steps that M uses on any input of length n . If $f(n)$ is the running time of M ,

say that M runs in time $f(n)$ and that M is an $f(n)$ time Turing machine. Customarily we use n to represent the length of the input.

- Machine M_1 that decides $A = \{0^k 1^k \mid k \geq 0\}$
 - Running time / Time Complexity: $n^2/2+3n$

M_1 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat if both 0s and 1s remain on the tape:
3. Scan across the tape, crossing off a single 0 and a single 1.
4. If 0s still remain after all the 1s have been crossed off, or if 1s still remain after all the 0s have been crossed off, *reject*. Otherwise, if neither 0s nor 1s remain on the tape, *accept*.”

Interlude: Asymptotic Analysis

Total: $n^2 + 3n$

- If $n = 1$
 - $n^2 = 1$
 - $3n = 3$
 - Total = 4
- If $n = 10$
 - $n^2 = 100$
 - $3n = 30$
 - Total = 130
- If $n = 100$
 - $n^2 = 10,000$
 - $3n = 300$
 - Total = 10,300
- If $n = 1,000$
 - $n^2 = 1,000,000$
 - $3n = 3,000$
 - Total = 1,003,000

asymptotic analysis only cares about **large n**

$n^2 + 3n \approx n^2$ as n gets large

Definition: Big-O Notation

Let f and g be functions $f, g: \mathcal{N} \rightarrow \mathcal{R}^+$. Say that $f(n) = O(g(n))$ if positive integers c and n_0 exist such that for every integer $n \geq n_0$,

$$f(n) \leq c g(n).$$

“only care about large n ”

When $f(n) = O(g(n))$, we say that $g(n)$ is an **upper bound** for $f(n)$, or more precisely, that $g(n)$ is an **asymptotic upper bound** for $f(n)$, to emphasize that we are suppressing constant factors.

In other words: Keep only highest order term, drop all coefficients

- Machine M_1 that decides $A = \{0^k 1^k \mid k \geq 0\}$
 - Is an $n^2 + 3n$ time Turing machine
 - Is an $O(n^2)$ time Turing machine
 - Has asymptotic upper bound $O(n^2)$

Definition: Small-*o* Notation (less used)

Let f and g be functions $f, g: \mathcal{N} \rightarrow \mathcal{R}^+$. Say that $f(n) = o(g(n))$ if

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0.$$

In other words, $f(n) = o(g(n))$ means that for any real number $c > 0$, a number n_0 exists, where $f(n) < c g(n)$ for all $n \geq n_0$.

Analogy: Big-*O* : \leq :: small-*o* : $<$

Let f and g be functions $f, g: \mathcal{N} \rightarrow \mathcal{R}^+$. Say that $f(n) = O(g(n))$ if positive integers c and n_0 exist such that for every integer $n \geq n_0$,

$$f(n) \leq c g(n).$$

When $f(n) = O(g(n))$, we say that $g(n)$ is an **upper bound** for $f(n)$, or more precisely, that $g(n)$ is an **asymptotic upper bound** for $f(n)$, to emphasize that we are suppressing constant factors.

Big- O arithmetic

- $O(n^2) + O(n^2)$
 $= O(n^2)$
- $O(n^2) + O(n)$
 $= O(n^2)$
- $2n = O(n)$?
 - TRUE
- $2n = O(n^2)$?
 - TRUE
- $1 = O(n^2)$?
 - TRUE
- $2^n = O(n^2)$?
 - FALSE

Definition: Time Complexity Classes

Let $t: \mathcal{N} \rightarrow \mathcal{R}^+$ be a function. Define the ***time complexity class***, **TIME**($t(n)$), to be the collection of all **languages** that are decidable by an $O(t(n))$ time Turing machine.

Remember: TMs have a time complexity (i.e., a running time),
languages are in a time complexity class

The complexity class of a **language** is determined by the
time complexity (running time) of its deciding **TM**

- Machine M_1 decides language $A = \{0^k 1^k \mid k \geq 0\}$
 - M_1 has time complexity (running time) of $O(n^2)$
 - A is in time complexity class **TIME**(n^2)

A Faster Machine? $A = \{0^k 1^k \mid k \geq 0\}$

Previously:

M_2 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat as long as some 0s and some 1s remain on the tape:
 3. Scan across the tape, checking whether the total number of 0s and 1s remaining is even or odd. If it is odd, *reject*.
 4. Scan again across the tape, crossing off every other 0 starting with the first 0, and then crossing off every other 1 starting with the first 1.
5. If no 0s and no 1s remain on the tape, *accept*. Otherwise, *reject*.”

M_1 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat if both 0s and 1s remain on the tape:
 3. Scan across the tape, crossing off a single 0 and a single 1.
 4. If 0s still remain after all the 1s have been crossed off, or if 1s still remain after all the 0s have been crossed off, *reject*. Otherwise, if neither 0s nor 1s remain on the tape, *accept*.”

A Faster Machine? $A = \{0^k 1^k \mid k \geq 0\}$

M_2 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat as long as some 0s and some 1s remain on the tape:
3. Scan across the tape, checking whether the total number of 0s and 1s remaining is even or odd. If it is odd, *reject*.
4. Scan again across the tape, crossing off every other 0 starting with the first 0, and then crossing off every other 1 starting with the first 1.
5. If no 0s and no 1s remain on the tape, *accept*. Otherwise, *reject*.”

Number of steps (worst case), n = length of input:

➤ Line 1:

- n steps to scan + n steps to return to beginning = $O(n)$ steps

A Faster Machine? $A = \{0^k 1^k \mid k \geq 0\}$

M_2 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat as long as some 0s and some 1s remain on the tape:
 3. Scan across the tape, checking whether the total number of 0s and 1s remaining is even or odd. If it is odd, *reject*.
 4. Scan again across the tape, crossing off every other 0 starting with the first 0, and then crossing off every other 1 starting with the first 1.
5. If no 0s and no 1s remain on the tape, *accept*. Otherwise, *reject*.”

Number of steps (worst case), n = length of input:

- Line 1:
 - n steps to scan + n steps to return to beginning = $O(n)$ steps
- Lines 2-4 (loop):
 - steps/iteration (lines 3-4): a scan takes $O(n)$ steps
 - # iters (line 2): Each iter crosses off half the chars, so at most $O(\log n)$ scans
 - Total: $O(n) * O(\log n) = O(n \log n)$ steps

Interlude: Logarithms

- If $2^x = y \dots$
- ... then $\log_2 y = x$
- $\log_2 n = O(\log n)$
 - “divide and conquer” algorithms = $O(\log n)$
 - E.g., binary search
- **(In computer science, base-2 is the only base!)**

A Faster Machine? $A = \{0^k 1^k \mid k \geq 0\}$

M_2 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat as long as some 0s and some 1s remain on the tape:
3. Scan across the tape, checking whether the total number of 0s and 1s remaining is even or odd. If it is odd, *reject*.
4. Scan again across the tape, crossing off every other 0 starting with the first 0, and then crossing off every other 1 starting with the first 1.
5. If no 0s and no 1s remain on the tape, *accept*. Otherwise, *reject*.”

Number of steps (worst case), n = length of input:

- Line 1:
 - n steps to scan + n steps to return to beginning = $O(n)$ steps
- Lines 2-4 (loop):
 - steps/iteration (lines 3-4): a scan takes $O(n)$ steps
 - # iters (line 2): Each iter crosses off half the chars, so at most $O(\log n)$ scans
 - Total: $O(n) * O(\log n) = O(n \log n)$ steps
- Line 5:
 - $O(n)$ steps to scan input one more time

A Faster Machine? $A = \{0^k 1^k \mid k \geq 0\}$

M_2 = “On input string w :

1. Scan across the tape and *reject* if a 0 is found to the right of a 1.
2. Repeat as long as some 0s and some 1s remain on the tape:
3. Scan across the tape, checking whether the total number of 0s and 1s remaining is even or odd. If it is odd, *reject*.
4. Scan again across the tape, crossing off every other 0 starting with the first 0, and then crossing off every other 1 starting with the first 1.
5. If no 0s and no 1s remain on the tape, *accept*. Otherwise, *reject*.”

$O(n \log n)$

Prev: $n^2/2 + 3n = O(n^2)$

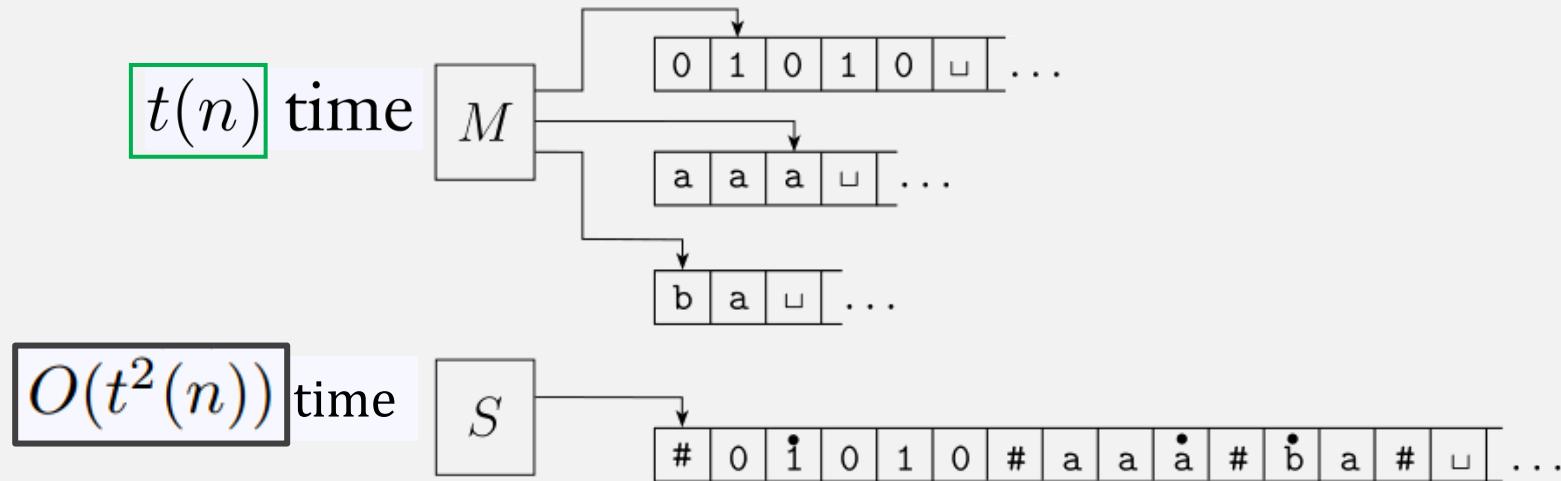
Number of steps (worst case), n = length of input:

- Line 1:
 - n steps to scan + n steps to return to beginning = $O(n)$ steps
- Lines 2-4 (loop):
 - steps/iteration (lines 3-4): a scan takes $O(n)$ steps
 - # iters (line 2): Each iter crosses off half the chars, so at most $O(\log n)$ scans
 - Total: $O(n) * O(\log n) = O(n \log n)$ steps
- Line 5:
 - $O(n)$ steps to scan input one more time
- Total: $O(n) + O(n \log n) + O(n) =$

Terminology: Categories of Bounds

- Exponential time
 - $O(2^{n^c})$, for $c > 0$, or $2^{O(n)}$ (always base 2)
- Polynomial time
 - $O(n^c)$, for $c > 0$
- Quadratic time (special case of polynomial time)
 - $O(n^2)$
- Linear time (special case of polynomial time)
 - $O(n)$
- Log time
 - $O(\log n)$

Multi-tape vs Single-tape TMs: # of Steps



- For single-tape TM to simulate 1 step of multi-tape:
 - Scan to find all “heads” = $O(\text{length of all } M\text{'s tapes})$
 - “Execute” transition at all the heads = $O(\text{length of all } M\text{'s tapes})$
- # single-tape steps to simulate 1 multtape step (worst case)
 - = $O(\text{length of all } M\text{'s tapes})$
 - = $O(t(n))$, If M spends all its steps expanding its tapes
- Total steps (single tape): $O(t(n))$ per step \times $t(n)$ steps =

Flashback: Nondet. TM → Deterministic TM

- Simulate NTM with Det. TM:
 - Number the nodes at each step
 - Deterministically check every tree path, in breadth-first order
 - 1
 - 1-1
 - 1-2
 - 1-1-1

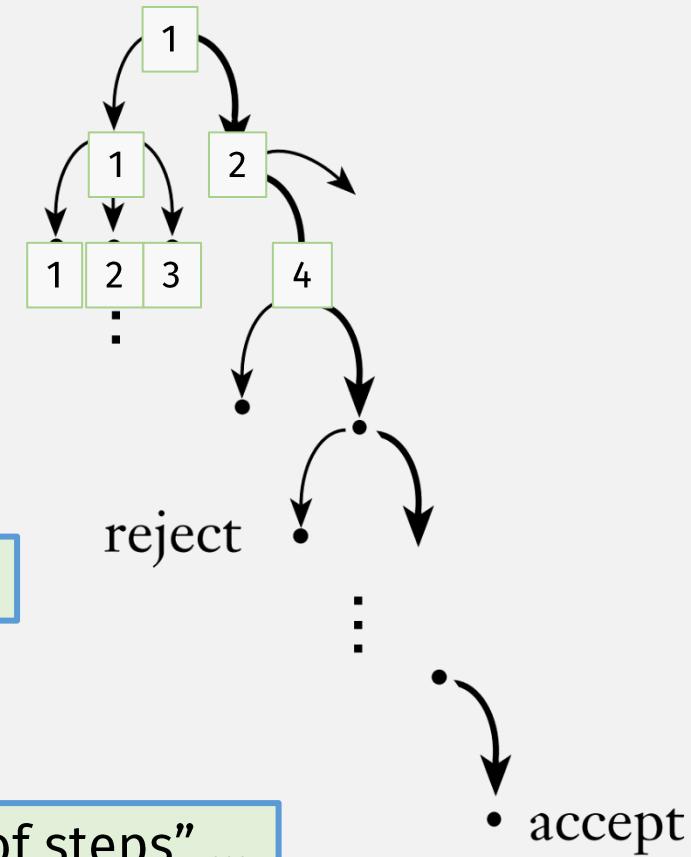
A TM and a NTM are “equivalent” ...

.. but not if we care about the # of steps

How inefficient is it?

First, we need a formal way to count “# of steps” ...

Nondeterministic computation



Flashback: Nondet. TM \rightarrow Deterministic TM

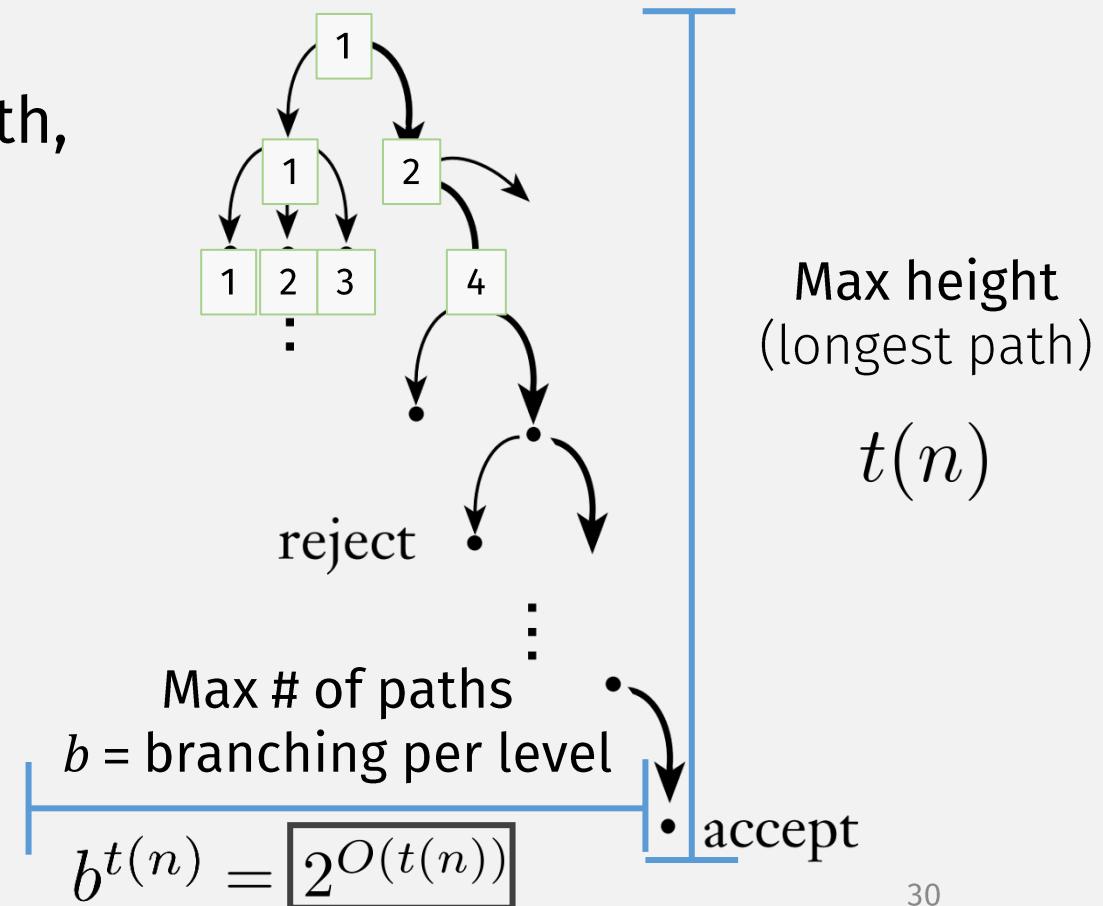
$t(n)$ time $\xrightarrow{} 2^{O(t(n))}$ time

- Simulate NTM with Det. TM:

- Number the nodes at each step
- Deterministically check every tree path, in breadth-first order

- 1
- 1-1
- 1-2
- 1-1-1

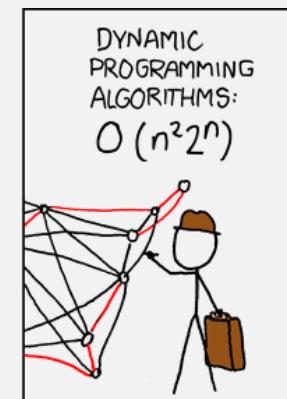
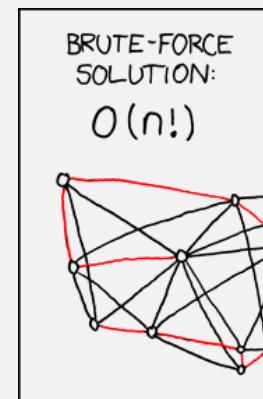
Nondeterministic computation



Summary: TM Variations

- If multi-tape TM: $t(n)$ time
- Then equivalent single-tape TM: $O(t^2(n))$
 - **Quadratically** slower
- If non-deterministic TM: $t(n)$ time
- Then equivalent single-tape TM: $2^{O(t(n))}$
 - **Exponentially** slower

Polynomial Time (P)



The Polynomial Time Complexity Class (**P**)

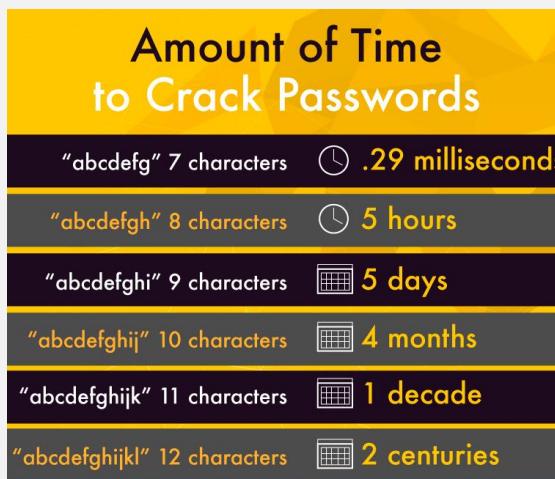
P is the class of languages that are decidable in polynomial time on a deterministic single-tape Turing machine. In other words,

$$P = \bigcup_k \text{TIME}(n^k).$$

- Corresponds to “realistically” solvable problems:
 - Problems in **P** = “solvable” or “tractable”
 - Problems outside **P** = “unsolvable” or “intractable”

“Unsolvable” Problems

- Unsolvable problems (those outside P):
 - usually only have “brute force” solutions
 - i.e., “try all possible inputs”
 - “unsolvable” applies only to large n



Do these problems exist???



3 Problems in P

- A Graph Problem:

$$PATH = \{\langle G, s, t \rangle \mid G \text{ is a directed graph that has a directed path from } s \text{ to } t\}$$

- A Number Problem:

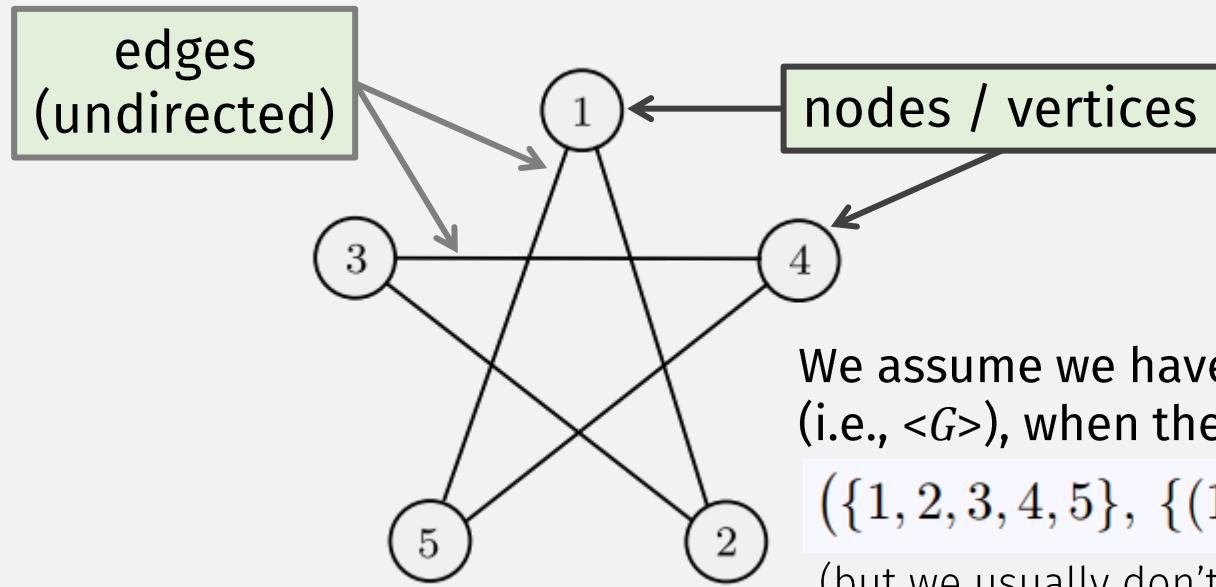
$$RELPRIME = \{\langle x, y \rangle \mid x \text{ and } y \text{ are relatively prime}\}$$

- A CFL Problem:

Every context-free language is a member of P

- To prove that a language is in P ...
 - ... we construct a polynomial time algorithm deciding the language
- (These also have nonpolynomial, i.e., brute force, algorithms)
 - Check all possible ... paths/numbers/strings ...

Interlude: Graphs (see Sipser Chapter 0)



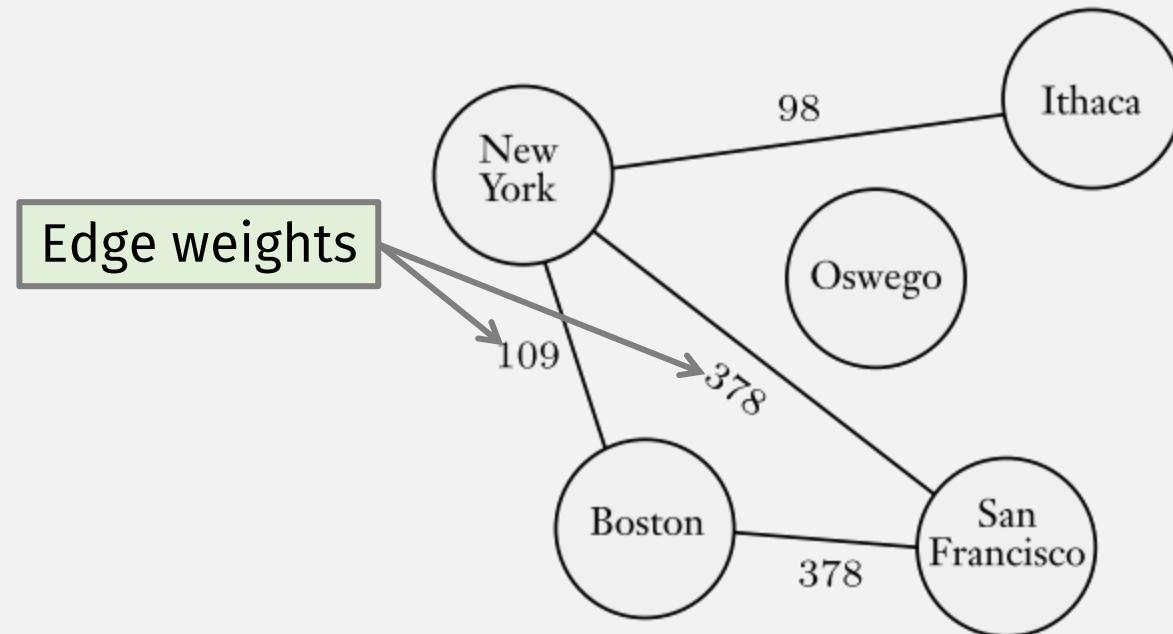
We assume we have **some string encoding of a graph** (i.e., $\langle G \rangle$), when they are args to TMs, e.g.:

$(\{1, 2, 3, 4, 5\}, \{(1, 2), (2, 3), (3, 4), (4, 5), (5, 1)\})$

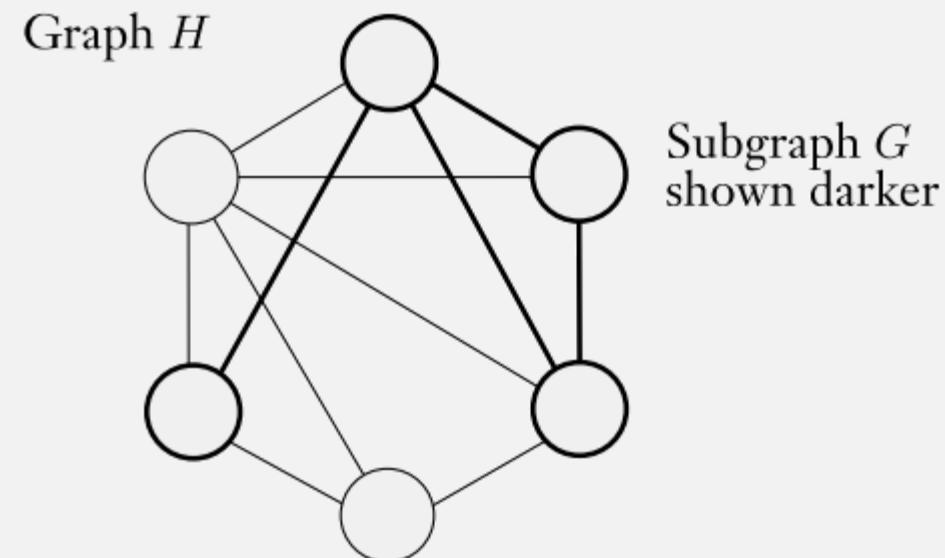
(but we usually don't care about the actual details)

- Edge defined by two nodes (order doesn't matter)
- Formally, a graph = a pair (V, E)
 - Where V = a set of nodes, E = a set of edges

Interlude: Weighted Graphs

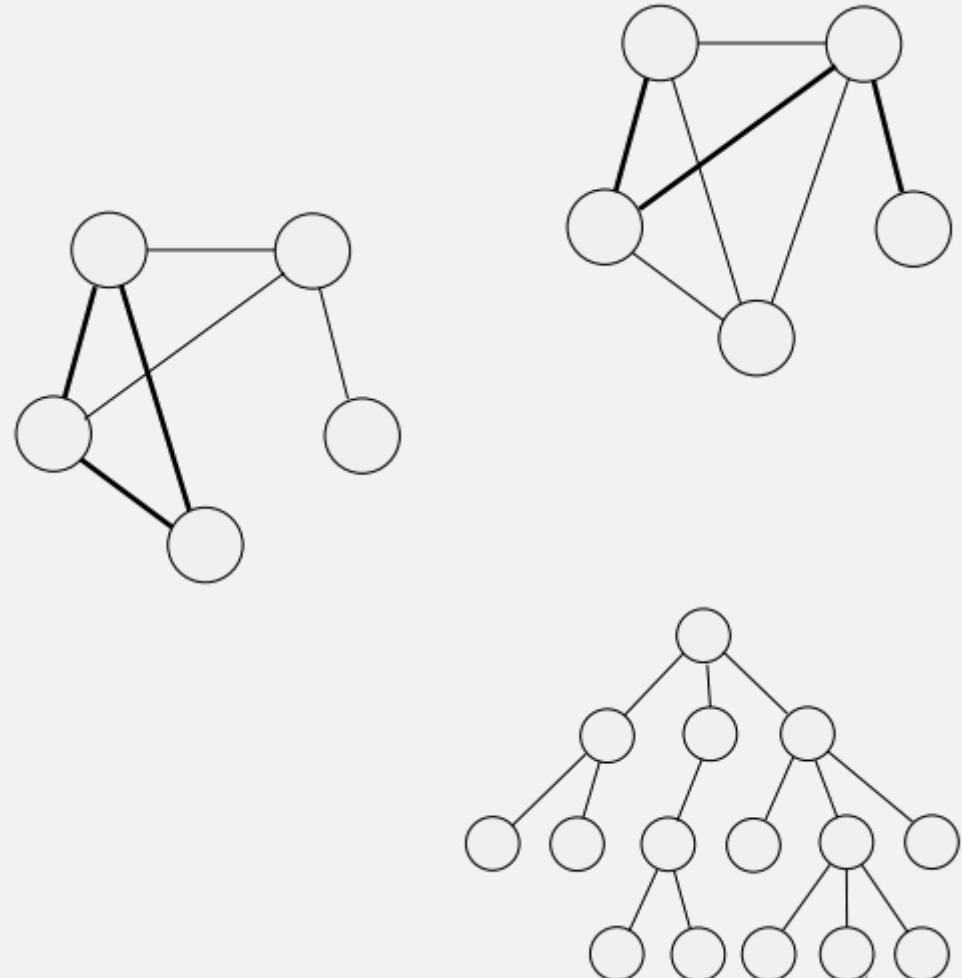


Interlude: Subgraphs

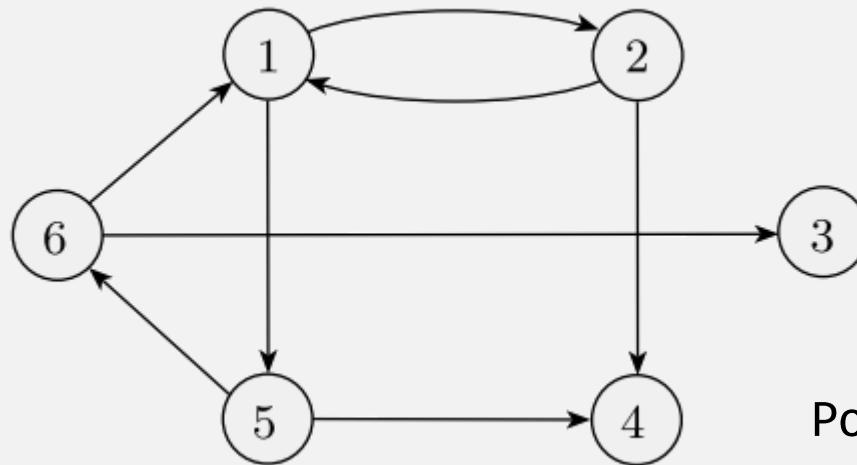


Interlude: Paths and other Graph Things

- Path
 - A sequence of nodes connected by edges
- Cycle
 - A path that starts/ends at the same node
- Connected graph
 - Every two nodes has a path
- Tree
 - A connected graph with no cycles



Interlude: Directed Graphs



Possible string encoding given to TMs:

$(\{1,2,3,4,5,6\}, \{(1,2), (1,5), (2,1), (2,4), (5,4), (5,6), (6,1), (6,3)\})$

- Directed graph = (V, E)
 - V = set of nodes, E = set of edges
- An edge is a pair of nodes (u,v) , **order now matters**
 - u = “from” node, v = “to” node
- “degree” of a node: number of edges connected to the node
 - Nodes in a directed graph have both indegree and outdegree

Each pair of nodes included twice

Interlude: Graph Encodings

$(\{1, 2, 3, 4, 5\}, \{(1, 2), (2, 3), (3, 4), (4, 5), (5, 1)\})$

- For graph algorithms, “length of input” n is usually # of vertices
 - (Not number of chars in the encoding)
- So given graph $G = (V, E)$, $n = |V|$
- Max edges?
 - $= O(|V|^2) = O(n^2)$
- So if a set of graphs (call it lang L) is decided by a TM where
 - # steps of the TM = polynomial in the # of vertices
 - Or # steps of the TM = polynomial in the # of edges
 - Then L is in P

3 Problems in P

- A Graph Problem:

$$PATH = \{\langle G, s, t \rangle \mid G \text{ is a directed graph that has a directed path from } s \text{ to } t\}$$

- A Number Problem:

$$RELPRIME = \{\langle x, y \rangle \mid x \text{ and } y \text{ are relatively prime}\}$$

- A CFL Problem:

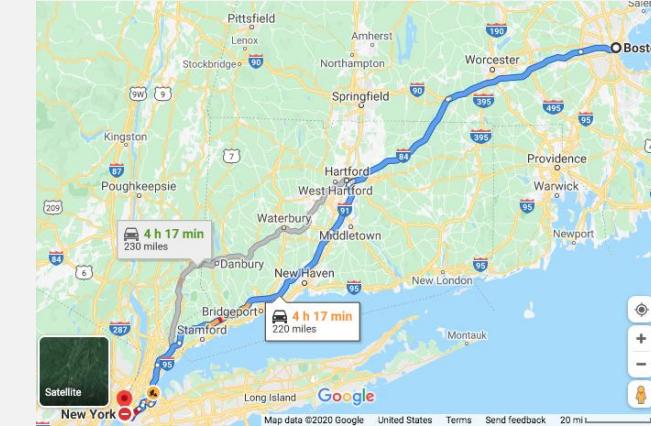
Every context-free language is a member of P

P is the class of languages that are decidable in polynomial time on a deterministic single-tape Turing machine. In other words,

$$\text{P} = \bigcup_k \text{TIME}(n^k).$$

A Graph Theorem: $PATH \in \text{P}$

$PATH = \{\langle G, s, t \rangle \mid G \text{ is a directed graph that has a directed path from } s \text{ to } t\}$



- To prove that a language is in P ...
- ... we must construct a polynomial time algorithm deciding the lang
- A non-polynomial (i.e., "brute force") algorithm:
 - check all possible paths, and see if any connect s to t
 - If $n = \# \text{ vertices}$, then # paths $\approx n^n$

Check-in Quiz 11/8

On gradescope