Week 10 at a glance

For Monday, Definition 7.1 (page 276).

For Wednesday, Definition 7.7 (page 279).

For Friday: skim through examples in Chapter 7.

We will be learning and practicing to:

- Know, select and apply appropriate computing knowledge and problem-solving techniques. Reason about computation and systems.
  - Use mapping reduction to deduce the complexity of a language by comparing to the complexity of another.
    - **Use appropriate reduction (e.g. mapping, Turing, polynomial-time) to deduce the complexity of a language by comparing to the complexity of another.**
    - **Use polynomial-time reduction to prove NP-completeness**
  - Classify the computational complexity of a set of strings by determining whether it is decidable or undecidable and recognizable or unrecognizable.
    - **Distinguish between computability and complexity**
    - **Articulate motivating questions of complexity**
    - **Define NP-completeness**
    - **Give examples of PTIME-decidable, NPTIME-decidable, and NP-complete problems**
  - Describe several variants of Turing machines and informally explain why they are equally expressive.
    - **Define nondeterministic Turing machines**
    - **Use high-level descriptions to define and trace machines (Turing machines and enumerators)**

TODO:

Student Evaluations of Teaching forms: Evaluations are open for completion anytime BEFORE 8AM on Saturday, December 7. Access your SETs from the Evaluations site

`https://academicaffairs.ucsd.edu/Modules/Evals`

You will separately evaluate each of your listed instructors for each enrolled course.

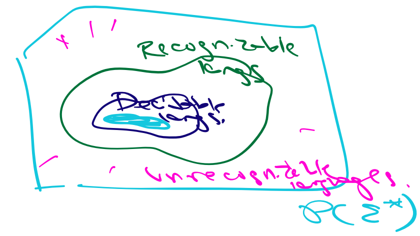
Homework 6 submitted via Gradescope (https://www.gradescope.com/), due Tuesday 12/3/2024

Summary from Week 9

Two models of computation are called **equally expressive** when every language recognizable with the first model is recognizable with the second, and vice versa.

To prove the existence of a Turing machine that decides / recognizes some language, it's enough to construct an example using any of the equally expressive models.

But: some of the **performance** properties of these models are not equivalent.



Monday: Church-Turing Thesis and Complexity

In practice, computers (and Turing machines) don't have infinite tape, and we can't afford to wait unboundedly long for an answer. "Decidable" isn't good enough - we want "Efficiently decidable".

For a given algorithm working on a given input, how long do we need to wait for an answer? How does the running time depend on the input in the worst-case? average-case? We expect to have to spend more time on computations with larger inputs.

A language is **recognizable** if there is a TM that accepts all and only strings in this language.

A language is **decidable** if there is a TM that accepts all strings in this language and rejects all strings not in the language.

A language is **efficiently decidable** if language is decidable and - - - -

A function is **computable** if there is a TM that for each string has a halting computation and the contents of tape at the end is function value.

A function is **efficiently computable** if function is computable and - - - -

Definition (Sipser 7.1): For M a deterministic decider, its **running time** is the function $f : \mathbb{N} \to \mathbb{N}$ given by

$$f(n) = \text{max number of steps } M \text{ takes before halting, over all inputs of length } n$$

Definition (Sipser 7.7): For each function $t(n)$, the **time complexity class** $TIME(t(n))$, is defined by

$$TIME(t(n)) = \{L \mid L \text{ is decidable by a Turing machine with running time in } O(t(n))\}$$

$t(n) = 1 \; \forall n$

Big O
asymptotic rate of growth
- drop constants
- drop lower order terms

An example of an element of $TIME(1)$ is $\emptyset$ witnessed by

An example of an element of $TIME(n)$ is $\{w \in \Sigma^* \mid |w| = 0 \bmod 3\}$ or any regular language

Note: $TIME(1) \subseteq TIME(n) \subseteq TIME(n^2)$

i.e. each language solvable by some TM that runs in constant time is solvable by some TM that runs in (at most) linear time

AND each language solvable by some TM that runs in (at most) linear time is solvable by some TM that runs in (at most) quadratic time

Definition (Sipser 7.12) : P is the class of languages that are decidable in polynomial time on a deterministic 1-tape Turing machine

$$P = \bigcup_k TIME(n^k)$$

Collection of languages that have efficient algorithms to decide membership.

Theorem (Sipser 7.8): Let $t(n)$ be a function with $t(n) \geq n$. Then every $t(n)$ time deterministic multitape Turing machine has an equivalent $O(t^2(n))$ time deterministic 1-tape Turing machine.

P is robust against many changes to the TM model.

What about with respect to nondeterministism?

Definitions (Sipser 7.1, 7.7, 7.12): For M a deterministic decider, its **running time** is the function $f : \mathbb{N} \to \mathbb{N}$ given by

$$f(n) = \text{max number of steps } M \text{ takes before halting, over all inputs of length } n$$

For each function $t(n)$, the **time complexity class** $TIME(t(n))$, is defined by

$$TIME(t(n)) = \{L \mid L \text{ is decidable by a Turing machine with running time in } O(t(n))\}$$

P is the class of languages that are decidable in polynomial time on a deterministic 1-tape Turing machine

$$P = \bigcup_k TIME(n^k)$$

Definition (Sipser 7.9): For N a nodeterministic decider. The **running time** of N is the function $f : \mathbb{N} \to \mathbb{N}$ given by

$$f(n) = \text{max number of steps } N \text{ takes on any branch before halting, over all inputs of length } n$$

Definition (Sipser 7.21): For each function $t(n)$, the **nondeterministic time complexity class** $NTIME(t(n))$, is defined by

$$NTIME(t(n)) = \{L \mid L \text{ is decidable by a nondeterministic Turing machine with running time in } O(t(n))\}$$

$$NP = \bigcup_k NTIME(n^k)$$

NP is the class of languages that are each decidable efficiently by nondeterministic Turing machines.

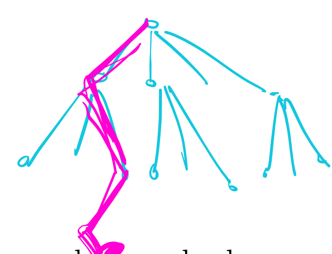
True ~~or **False**~~: $TIME(n^2) \subseteq NTIME(n^2)$

If L has a quadratic time 1-tape deterministic TM that decides it, can use the same algorithm to witness quadratic nondeterministic solution.

True or **False**: $NTIME(n^2) \subseteq TIME(n^2)$

?

Every problem in NP is decidable with an exponential-time algorithm



Nondeterministic approach: guess a possible solution, verify that it works.

Brute-force (worst-case exponential time) approach: iterate over all possible solutions, for each one, check if it works.

Wednesday: P and NP

Examples in P

Can't use nondeterminism; Can use multiple tapes; Often need to be "more clever" than naïve / brute force approach

string encoding

$$PATH = \{\langle G, s, t\rangle \mid G \text{ is digraph with } n \text{ nodes there is path from s to t}\}$$

Use breadth first search to show in P

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

Use Euclidean Algorithm to show in P

$$L(G) = \{w \mid w \text{ is generated by } G\}$$

(where G is a context-free grammar). Use dynamic programming to show in P.

Examples in NP

"Verifiable" i.e. NP, Can be decided by a nondeterministic TM in polynomial time, best known deterministic solution may be brute-force, solution can be verified by a deterministic TM in polynomial time.

$HAMPATH = \{\langle G, s, t\rangle \mid G$ is digraph with n nodes, there is path from s to t that goes through every node exa

$$VERTEX-COVER = \{\langle G, k\rangle \mid G \text{ is an undirected graph with } n \text{ nodes that has a } k\text{-node vertex cover}\}$$

$$CLIQUE = \{\langle G, k\rangle \mid G \text{ is an undirected graph with } n \text{ nodes that has a } k\text{-clique}\}$$

$$SAT = \{\langle X\rangle \mid X \text{ is a satisfiable Boolean formula with } n \text{ variables}\}$$

Assignment
x_1 T
x_2 F
x_3 T

literal $x_1, \ldots, x_n$ $\neg x_1, \ldots, \neg x_n$

$(x_1 \wedge x_2 \wedge x_3) \vee (\neg x_1 \wedge x_2 \wedge \neg x_3) \vee (\quad) \vee (\quad)$

Problems in P	**Problems in NP**
(Membership in any) regular language	Any problem in P
(Membership in any) context-free language	
A_{DFA}	SAT
E_{DFA}	$CLIQUE$
EQ_{DFA}	$VERTEX-COVER$
$PATH$	$HAMPATH$
$RELPRIME$	$\ldots$
$\ldots$	

Notice: $NP \subseteq \{L \mid L \text{ is decidable}\}$ so $A_{TM} \notin NP$

Million-dollar question: Is $P = NP$?

build clever algorithm to solve one of these "hardest problems"

One approach to trying to answer it is to look for *hardest* problems in NP and then (1) if we can show that there are efficient algorithms for them, then we can get efficient algorithms for all problems in NP so $P = NP$, or (2) these problems might be good candidates for showing that there are problems in NP for which there are no efficient algorithms.

prove a lower bound

Before: $A \leq_m B$ means there's a computable function such that $\forall x \in \Sigma^* \ (x \in A$ iff $f(x) \in B)$

Definition (Sipser 7.29) Language A is **polynomial-time mapping reducible** to language B, written $A \leq_P B$, means there is a polynomial-time computable function $f : \Sigma^* \to \Sigma^*$ such that for every $x \in \Sigma^*$

$$x \in A \quad \text{iff} \quad f(x) \in B.$$

The function f is called the polynomial time reduction of A to B.

Theorem (Sipser 7.31): If $A \leq_P B$ and $B \in P$ then $A \in P$.

Proof:

COPY PASTE THEOREM 5.28

Pf: Let A and B be languages
Assume $A \leq_P B$.
By def, there is f efficiently computable function that is witnessed by a TM with polynomial-time running time
So that for each string $x \in \Sigma^*$, $x \in A$ iff $f(x) \in B$.

WTS if B is polynomial time decidable so is A.

Assume B is polynomial time decidable
By definition, there is a TM M_B Turing machine, decider, $L(M_B) = B$ running time $O(n^k)$ for some k.
WTS A is polynomial time decision

Define M_A = " On input x.
1 Compute $f(x)$ polynomial many steps b/c f is efficiently computable
2. Run M_B on $f(x)$ takes polynomial time
3. If accepts, accept; if rejects, reject "

To show M_A decides A, consider arbitrary string x
Case 1: Assume $x \in A$. By def of $f(x)$, $f(x) \in B$. Running M_A on x, step 1 takes finitely long to compute $f(x)$ and M_B on $f(x)$ halts and accepts so M_A accepts x in step 3 ✓

Case 2: Assume $x \notin A$. By def of $f(x)$, $f(x) \notin B$. Running M_A on x, step 1 takes finitely long to compute $f(x)$ and M_B on $f(x)$ halts and rejects so M_A rejects x in step 3 . . .

AND running time of M_A is polynomial + polynomial + c which is polynomial.

Definition (Sipser 7.34; based in Stephen Cook and Leonid Levin's work in the 1970s): A language B is **NP-complete** means (1) B is in NP **and** (2) every language A in NP is polynomial time reducible to B.

$A \leq_P B$

Theorem (Sipser 7.35): If B is NP-complete and $B \in P$ then $P = NP$.

Proof:

Assume B is NP-complete.

Since $P \subseteq NP$, WTS $NP \subseteq P$.

Let A be any language in NP, WTS $A \in P$.

We have $A \leq_P B$ so by Theorem 7.31,

Since $B \in P$, so is A

Friday: NP-Completeness

Garey
Johnson

CSE 202.

NP-Complete Problems

3SAT: A literal is a Boolean variable (e.g. x) or a negated Boolean variable (e.g. $\bar{x}$). A Boolean formula is a **3cnf-formula** if it is a Boolean formula in conjunctive normal form (a conjunction of disjunctive clauses of literals) and each clause has three literals.

$$3SAT = \{\langle \phi \rangle \mid \phi \text{ is a satisfiable 3cnf-formula}\}$$

Example string in $3SAT$

$$\langle (x \underset{\text{OR}}{\vee} \bar{y} \vee \bar{z}) \overset{\text{AND}}{\wedge} (\bar{x} \underset{\text{OR}}{\vee} y \vee z) \overset{\text{AND}}{\wedge} (x \underset{\text{OR}}{\vee} y \vee z) \rangle$$

witness assignment

x	T
y	T
z	F

Example string not in $3SAT$

$$\langle (x \vee y \vee z) \wedge (x \vee y \vee \bar{z}) \wedge (x \vee \bar{y} \vee z) \wedge (x \vee \bar{y} \vee \bar{z}) \wedge (\bar{x} \vee y \vee z) \wedge (\bar{x} \vee y \vee \bar{z}) \wedge (\bar{x} \vee \bar{y} \vee z) \wedge (\bar{x} \vee \bar{y} \vee \bar{z}) \rangle$$

Cook-Levin Theorem: $3SAT$ is NP-complete.

in NP : efficiently verifiable
"hardest among all NP" - Ch 7.

Are there other NP-complete problems? To prove that X is NP-complete

- *From scratch*: prove X is in NP and that all NP problems are polynomial-time reducible to X.
- *Using reduction*: prove X is in NP and that a known-to-be NP-complete problem is polynomial-time reducible to X.

When B_1 NP-complete
a proof that B_2 is in NP and $B_1 \leq_p B_2$
gives that B_2 is NP-complete as well.

Why enough? $\leq_m$ and $\leq_p$ are transitive!

i.e. If $A \leq_m B$ and $B \leq_m C$, then $A \leq_m C$

and If $A \leq_p B$ and $B \leq_p C$ then $A \leq_p C$

① X is in NP

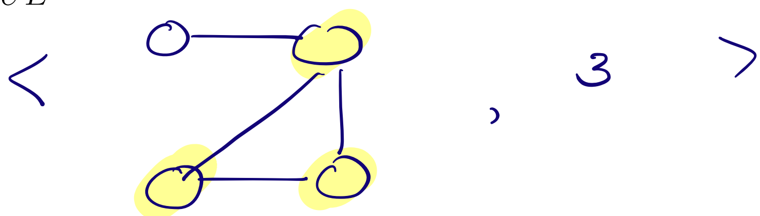
and ② 3-SAT $\leq_p$ X

gives for all A in NP $A \leq_p X$
because Cook-Levin gives $A \leq_p$ 3-SAT.

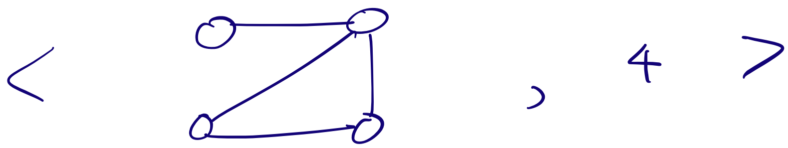
CLIQUE: A k**-clique** in an undirected graph is a maximally connected subgraph with k nodes.

$$CLIQUE = \{\langle G, k\rangle \mid G \text{ is an undirected graph with a } k\text{-clique}\}$$

Example string in $CLIQUE$



Example string not in $CLIQUE$



Theorem (Sipser 7.32):

$$3SAT \leq_P CLIQUE$$

$f: \Sigma^* \to \Sigma^*$

$f(x) = \begin{cases} \langle \ldots, 4 \rangle & x \neq \langle \varphi, k\rangle \\ \langle \ldots, k \rangle & x = \langle \varphi, k \rangle \end{cases}$

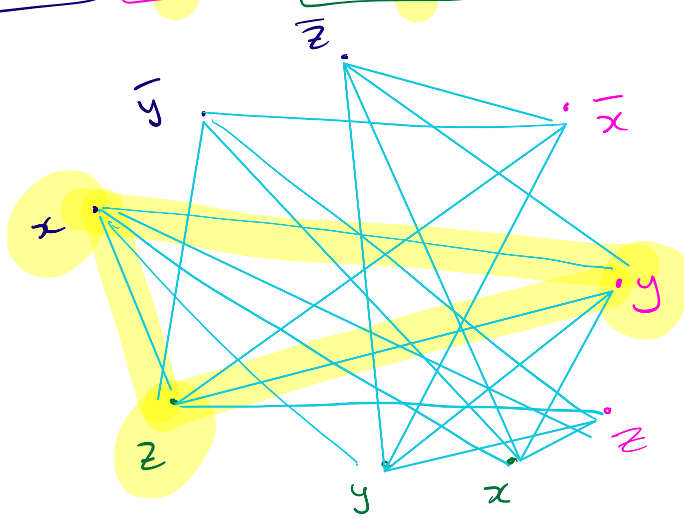
Given a Boolean formula in conjunctive normal form with k clauses and three literals per clause, we will map it to a graph so that the graph has a clique if the original formula is satisfiable and the graph does not have a clique if the original formula is not satisfiable.

The graph has $3k$ vertices (one for each literal in each clause) and an edge between all vertices except

- vertices for two literals in the same clause
- vertices for literals that are negations of one another

Example: $(x \vee \bar{y} \vee \bar{z}) \wedge (\bar{x} \vee y \vee z) \wedge (x \vee y \vee z)$

Satisfiable $x = T$ $y = T$ $z = F$



Model of Computation	Class of Languages
Deterministic finite automata: formal definition, how to design for a given language, how to describe language of a machine? **Nondeterministic finite automata**: formal definition, how to design for a given language, how to describe language of a machine? **Regular expressions**: formal definition, how to design for a given language, how to describe language of expression? *Also*: converting between different models.	**Class of regular languages**: what are the closure properties of this class? which languages are not in the class? using **pumping lemma** to prove nonregularity.
Push-down automata: formal definition, how to design for a given language, how to describe language of a machine? **Context-free grammars**: formal definition, how to design for a given language, how to describe language of a grammar?	**Class of context-free languages**: what are the closure properties of this class? which languages are not in the class?
Turing machines that always halt in polynomial time Nondeterministic Turing machines that always halt in polynomial time	P NP
Deciders (Turing machines that always halt): formal definition, how to design for a given language, how to describe language of a machine?	**Class of decidable languages**: what are the closure properties of this class? which languages are not in the class? using diagonalization and mapping reduction to show undecidability
Turing machines formal definition, how to design for a given language, how to describe language of a machine?	**Class of recognizable languages**: what are the closure properties of this class? which languages are not in the class? using closure and mapping reduction to show unrecognizability

Given a language, prove it is regular

Strategy 1: construct DFA recognizing the language and prove it works.

Strategy 2: construct NFA recognizing the language and prove it works.

Strategy 3: construct regular expression recognizing the language and prove it works.

"Prove it works" means …

Example: $L = \{w \in \{0,1\}^* \mid w \text{ has odd number of 1s or starts with } 0\}$

Using NFA

Using regular expressions

Example: Select all and only the options that result in a true statement: "To show a language A is not regular, we can…"

a. Show A is finite

b. Show there is a CFG generating A

c. Show A has no pumping length

d. Show A is undecidable

Example: What is the language generated by the CFG with rules

$$\begin{aligned} S &\to aSb \mid bY \mid Ya \\ Y &\to bY \mid Ya \mid \varepsilon \end{aligned}$$

Example: Prove that the language $T = \{\langle M \rangle \mid M \text{ is a Turing machine and } L(M) \text{ is infinite}\}$ is undecidable.

Example: Prove that the class of decidable languages is closed under concatenation.


Im
POSSIBLE