GEMS OF TCS

P VS NP

Sasha Golovnev
October 16, 2023
Search Problems
### Definition

A **search problem** is defined by an algorithm $C$ that takes an instance $I$ and a candidate solution $S$, and runs in time polynomial in the length of $I$. We say that $S$ is a solution to $I$ iff $C(S, I) = true$. 
Example

For SAT, $I$ is a Boolean formula, $S$ is an assignment of Boolean constants to its variables. The corresponding algorithm $C$ checks whether $S$ satisfies all clauses of $I$. 
# Class NP

<table>
<thead>
<tr>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A search problem</strong> is defined by an algorithm $C$ that takes an instance $I$ and a candidate solution $S$, and runs in time polynomial in the length of $I$. We say that $S$ is a solution to $I$ iff $C(S, I) = \text{true}$.</td>
</tr>
</tbody>
</table>

**CLASS NP**

**Definition**

A **search problem** is defined by an algorithm $\mathcal{C}$ that takes an instance $I$ and a candidate solution $S$, and runs in time polynomial in the length of $I$. We say that $S$ is a solution to $I$ iff $\mathcal{C}(S, I) = \text{true}$.

**Definition**

**NP** is the class of all search problems.
• NP stands for “non-deterministic polynomial time”: one can guess a solution, and then verify its correctness in polynomial time
• **NP** stands for “non-deterministic polynomial time”: one can guess a solution, and then verify its correctness in polynomial time

• In other words, the class **NP** contains all problems whose solutions can be efficiently verified
Definition

$\mathcal{P}$ is the class of all search problems that can be solved in polynomial time.
Given a complete weighted graph, find a path of minimum total weight (length) visiting each node exactly once.
Given a complete weighted graph and a budget $b$, find a path of total weight (length) $\leq b$ visiting each node exactly once.

length: $6 \leq b$
Given a complete weighted graph and a budget $b$, connect all vertices by $n - 1$ edges of minimum total weight (length)

length: 6
TSP AND MST

MST

Given $n$ cities, connect them by $(n - 1)$ roads of minimal total length
TSP AND MST

MST

Given $n$ cities, connect them by $(n - 1)$ roads of minimal total length

Can be solved efficiently
TSP AND MST

MST
Given \( n \) cities, connect them by \( (n - 1) \) roads of minimal total length
Can be solved efficiently

TSP
Given \( n \) cities, connect them in a path of minimal total length
## TSP AND MST

<table>
<thead>
<tr>
<th>MST</th>
<th>TSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given ( n ) cities, connect them by ((n - 1)) roads of minimal total length</td>
<td>Given ( n ) cities, connect them <em>in a path</em> of minimal total length</td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td>No polynomial algorithm known!</td>
</tr>
</tbody>
</table>
**LONGEST PATH**

<table>
<thead>
<tr>
<th>Longest path</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> A weighted graph, two vertices $s, t$, and a budget $b$.</td>
</tr>
<tr>
<td><strong>Output:</strong> A simple path (containing no repeated vertices) of total length at least $b$.</td>
</tr>
</tbody>
</table>
Example
Example
Example
Shortest path

Find a simple path from s to t of total length at most $b$
<table>
<thead>
<tr>
<th>Shortest path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a simple path from ( s ) to ( t ) of total length at most ( b )</td>
</tr>
</tbody>
</table>

Can be solved efficiently
<table>
<thead>
<tr>
<th>Shortest path</th>
<th>Longest path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a simple path from $s$ to $t$ of total length at most $b$</td>
<td>Find a simple path from $s$ to $t$ of total length at least $b$</td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td>No polynomial algorithm known!</td>
</tr>
<tr>
<td><strong>Shortest path</strong></td>
<td><strong>Longest path</strong></td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Find a simple path from $s$ to $t$ of total length <em>at most</em> $b$</td>
<td>Find a simple path from $s$ to $t$ of total length <em>at least</em> $b$</td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td>No polynomial algorithm known!</td>
</tr>
</tbody>
</table>
Integer Linear Programming Problem

<table>
<thead>
<tr>
<th>Integer linear programming</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> A set of linear inequalities $Ax \leq b$.</td>
</tr>
<tr>
<td><strong>Output:</strong> Integer solution.</td>
</tr>
</tbody>
</table>
Example

\[ x_1 \geq 0.5 \]
\[ -x_1 + 8x_2 \geq 0 \]
\[ -x_1 - 8x_2 \geq -8 \]
Example

\[ x_1 \geq 0.5 \]
\[ -x_1 + 8x_2 \geq 0 \]
\[ -x_1 - 8x_2 \geq -8 \]
Example

\[ x_1 \geq 0.5 \]
\[ -x_1 + 8x_2 \geq 0 \]
\[ -x_1 - 8x_2 \geq -8 \]
Example

\[ x_1 \geq 0.5 \]

\[ -x_1 + 8x_2 \geq 0 \]

\[ -x_1 - 8x_2 \geq -8 \]
Example

\[ x_1 \geq 0.5 \]

\[-x_1 + 8x_2 \geq 0 \]

\[-x_1 - 8x_2 \geq -8 \]
INTEGER LINEAR PROGRAMMING

**LP**

Find a **real** solution of a system of linear inequalities

No polynomial algorithm known!
Integer Linear Programming

**LP**

Find a *real* solution of a system of linear inequalities

Can be solved efficiently

No polynomial algorithm known!
# Integer Linear Programming

<table>
<thead>
<tr>
<th>LP</th>
<th>ILP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a <strong>real</strong> solution of a system of linear inequalities</td>
<td>Find an <strong>integer</strong> solution of a system of linear inequalities</td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td></td>
</tr>
</tbody>
</table>

No polynomial algorithm known!
## Integer Linear Programming

<table>
<thead>
<tr>
<th>LP</th>
<th>ILP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a <strong>real</strong> solution of a system of linear inequalities</td>
<td>Find an <strong>integer</strong> solution of a system of linear inequalities</td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td>No polynomial algorithm known!</td>
</tr>
</tbody>
</table>
### Independent Set Problem

**Input:** A graph and a budget $b$.

**Output:** A subset of vertices of size at least $b$ such that no two of them are adjacent.
Example
INDEPENDENT SETS IN A TREE

A maximum independent set in a tree can be found by a simple greedy algorithm: it is safe to take into a solution all the leaves.
<table>
<thead>
<tr>
<th>Independent set in a tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find an independent set of size at least ( b ) in a given tree</td>
</tr>
<tr>
<td><strong>Independent set in a tree</strong></td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Find an independent set of size at least $b$ in a given <strong>tree</strong></td>
</tr>
</tbody>
</table>

Can be solved efficiently
<table>
<thead>
<tr>
<th>Independent set in a tree</th>
<th>Independent set in a graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find an independent set of size at least $b$ in a given <strong>tree</strong></td>
<td>Find an independent set of size at least $b$ in a given <strong>graph</strong></td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td>No polynomial algorithm known!</td>
</tr>
<tr>
<td>Independent set in a tree</td>
<td>Independent set in a graph</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Find an independent set of size at least $b$ in a given tree</td>
<td>Find an independent set of size at least $b$ in a given graph</td>
</tr>
<tr>
<td>Can be solved efficiently</td>
<td>No polynomial algorithm known!</td>
</tr>
</tbody>
</table>
It turns out that all these hard problems are in a sense a single hard problem: a polynomial time algorithm for any of these problems can be used to solve all of them in polynomial time!
## Class P

Problems whose solution can be **found** efficiently

- MST
- Shortest path
- LP
- IS on trees
Class P
Problems whose solution can be found efficiently

- MST
- Shortest path
- LP
- IS on trees

Class NP
Problems whose solution can be verified efficiently

- TSP
- Longest path
- ILP
- IS on graphs
<table>
<thead>
<tr>
<th>Class P</th>
<th>Class NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problems whose solution can be <strong>found</strong> efficiently</td>
<td>Problems whose solution can be <strong>verified</strong> efficiently</td>
</tr>
<tr>
<td>• MST</td>
<td>• TSP</td>
</tr>
<tr>
<td>• Shortest path</td>
<td>• Longest path</td>
</tr>
<tr>
<td>• LP</td>
<td>• ILP</td>
</tr>
<tr>
<td>• IS on trees</td>
<td>• IS on graphs</td>
</tr>
<tr>
<td>Class P</td>
<td>Class NP</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Problems whose solution can be <strong>found</strong> efficiently</td>
<td>Problems whose solution can be <strong>verified</strong> efficiently</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>• MST</td>
<td>• TSP</td>
</tr>
<tr>
<td>• Shortest path</td>
<td>• Longest path</td>
</tr>
<tr>
<td>• LP</td>
<td>• ILP</td>
</tr>
<tr>
<td>• IS on trees</td>
<td>• IS on graphs</td>
</tr>
</tbody>
</table>
The main open problem in Computer Science

Is $P$ equal to $NP$?
The main open problem in Computer Science

Is $P$ equal to $NP$?

Millenium Prize Problem

Clay Mathematics Institute: $1M prize for solving the problem
• If $P=NP$, then all search problems can be solved in polynomial time.
• If $P=NP$, then all search problems can be solved in polynomial time.

• If $P \neq NP$, then there exist search problems that cannot be solved in polynomial time.
Reductions
INFORMALLY

We say that a search problem $A$ is reduced to a search problem $B$ and write $A \rightarrow B$, if a polynomial time algorithm for $B$ can be used (as a black box) to solve $A$ in polynomial time.
Reduction: $A \rightarrow B$

instance $I$ of $A$
REDUCTION: $A \rightarrow B$

instance $I$ of $A$

Algorithm for $A$

Algorithm for $B$
REDUCTION: $A \rightarrow B$

instance $I$ of $A$

Algorithm for $A$

$f$

Algorithm for $B$
REDUCTION: $A \rightarrow B$

instance $I$ of $A$

Algorithm for $A$

$f$

instance $f(I)$ of $B$

Algorithm for $B$
**REDUCTION:** $A \rightarrow B$

- **Instance** $I$ of $A$

**Algorithm for $A$**

- Function $f$

**Instance $f(I)$ of $B$**

**Algorithm for $B$**

- No solution to $f(I)$
REDUCTION: $A \rightarrow B$

Algorithm for $A$

$f$

instance $I$ of $A$

instance $f(I)$ of $B$

Algorithm for $B$

no solution to $f(I)$

no solution to $I$
Reduction: $A \rightarrow B$

Algorithm for $A$

Instance $I$ of $A$

$\rightarrow$

Algorithm for $B$

Instance $f(I)$ of $B$

No solution to $f(I)$

Solution $S$ to $f(I)$

No solution to $I$
Reduction: $A \rightarrow B$

Instance $I$ of $A$

Algorithm for $A$

Instance $f(I)$ of $B$

Algorithm for $B$

No solution to $f(I)$

Solution $S$ to $f(I)$

No solution to $I$

$h$
**Reduction: A → B**

- **Algorithm for A**
  - Instance $I$ of $A$
  - Function $f$
    - Instance $f(I)$ of $B$
  - Algorithm for B
    - No solution to $f(I)$
    - Solution $S$ to $f(I)$
    - Function $h$
      - Solution $h(S)$ to $I$
**Formally**

<table>
<thead>
<tr>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>We say that a search problem $A$ is reduced to a search problem $B$ and write $A \rightarrow B$, if there exists a polynomial time algorithm $f$ that converts any instance $I$ of $A$ into an instance $f(I)$ of $B$, together with a polynomial time algorithm $h$ that converts any solution $S$ to $f(I)$ back to a solution $h(S)$ to $A$. If there is no solution to $f(I)$, then there is no solution to $I$.</td>
</tr>
</tbody>
</table>
GRAPH OF SEARCH PROBLEMS
NP-COMPLETE PROBLEMS

Definition

A search problem is called \textbf{NP-complete} if all other search problems reduce to it.
A search problem is called **NP-complete** if all other search problems reduce to it.
Do they exist?

It’s not at all immediate that \textbf{NP}-complete problems even exist. We’ll see later that all hard problems that we’ve seen in the previous part are in fact \textbf{NP}-complete!
Two ways of using $A \rightarrow B$:

- if $B$ is easy (can be solved in polynomial time), then so is $A$
- if $A$ is hard (cannot be solved in polynomial time), then so is $B$
Lemma

If $A \rightarrow B$ and $B \rightarrow C$, then $A \rightarrow C$. 
PICTORIALLY

NP
Corollary

If $A \rightarrow B$ and $A$ is NP-complete, then so is $B$. 
Corollary

If $A \rightarrow B$ and $A$ is \textbf{NP}-complete, then so is $B$. 
Corollary

If $A \rightarrow B$ and $A$ is $\textbf{NP}$-complete, then so is $B$. 

\[ \text{NP} \]

\[ \text{A} \rightarrow \text{B} \]
Corollary

If $A \rightarrow B$ and $A$ is $\text{NP}$-complete, then so is $B$. 

Diagram: A directed graph with vertices labeled $A$ and $B$, connected by a directed edge $A \rightarrow B$. The graph is contained within an oval labeled $\text{NP}$. Another oval encompasses the entire diagram.
Corollary

If \( A \rightarrow B \) and \( A \) is \( \text{NP} \)-complete, then so is \( B \).
Corollary

If $A \rightarrow B$ and $A$ is \textbf{NP}-complete, then so is $B$. 
NP-Completeness of SAT
Goal

Show that every search problem reduces to SAT.
Goal

Show that every search problem reduces to SAT.

Instead, we show that any problem reduces to Circuit SAT problem, which, in turn, reduces to SAT.
A circuit is a directed acyclic graph of in-degree at most 2. Nodes of in-degree 0 are called inputs and are marked by Boolean variables and constants. Nodes of in-degree 1 and 2 are called gates: gates of in-degree 1 are labeled with NOT, gates of in-degree 2 are labeled with AND or OR. One of the sinks is marked as output.
<table>
<thead>
<tr>
<th>Circuit-SAT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong>  A circuit.</td>
</tr>
<tr>
<td><strong>Output:</strong> An assignment of Boolean values to the input variables of the circuit that makes the output true.</td>
</tr>
</tbody>
</table>
SAT is a special case of Circuit-SAT as a SAT formula can be represented as a circuit:

**Example:** $(x \lor y \lor z)(y \lor \bar{x})$
To reduce Circuit-SAT to SAT, we need to design a polynomial time algorithm that for a given circuit outputs a SAT formula which is satisfiable, if and only if the circuit is satisfiable.
IDEA

• Introduce a Boolean variable for each gate

• For each gate, write down a few clauses that describe the relationship between this gate and its direct predecessors
NOT GATES

\[(h \lor g)(\bar{h} \lor \bar{g})\]
AND GATES

\[(h_1 \lor \overline{g})(h_2 \lor \overline{g})(\overline{h_1} \lor \overline{h_2} \lor g)\]
OR GATES

\[(\overline{h}_1 \lor g)(\overline{h}_2 \lor g)(h_1 \lor h_2 \lor \overline{g})\]
OUTPUT GATE

\[ g \bigcirc \text{output} \quad (g) \]
The resulting SAT formula is consistent with the initial circuit: in any satisfying assignment of the formula, the value of $g$ is equal to the value of the gate labeled with $g$ in the circuit.
The resulting SAT formula is consistent with the initial circuit: in any satisfying assignment of the formula, the value of $g$ is equal to the value of the gate labeled with $g$ in the circuit.

Therefore, the SAT formula and the circuit are equisatisfiable.
• The resulting SAT formula is consistent with the initial circuit: in any satisfying assignment of the formula, the value of $g$ is equal to the value of the gate labeled with $g$ in the circuit.

• Therefore, the SAT formula and the circuit are equisatisfiable.

• The reduction takes polynomial time.
Goal

Reduce every search problem to Circuit-SAT.
Goal

Reduce every search problem to Circuit-SAT.

• Let $A$ be a search problem
Goal

Reduce every search problem to Circuit-SAT.

- Let $A$ be a search problem
- We know that there exists an algorithm $C$ that takes an instance $I$ of $A$ and a candidate solution $S$ and checks whether $S$ is a solution for $I$ in time polynomial in $|I|$
Goal

Reduce every search problem to Circuit-SAT.

- Let A be a search problem
- We know that there exists an algorithm $C$ that takes an instance $I$ of A and a candidate solution $S$ and checks whether $S$ is a solution for $I$ in time polynomial in $|I|$
- In particular, $|S|$ is polynomial in $|I|$
• Note that a computer is in fact a circuit implemented on a chip
• Note that a computer is in fact a circuit implemented on a chip
• Each step of the algorithm $C(I, S)$ is performed by this computer’s circuit
Turn an Algorithm into a Circuit

• Note that a computer is in fact a circuit implemented on a chip
• Each step of the algorithm $C(I, S)$ is performed by this computer’s circuit
• This gives a circuit of size polynomial in $|I|$ that has input bits for $I$ and $S$ and outputs whether $S$ is a solution for $I$ (a separate circuit for each input length)
To solve an instance $I$ of the problem $A$:

- take a circuit corresponding to $C(I, \cdot)$
To solve an instance $I$ of the problem $A$:

- take a circuit corresponding to $C(I, \cdot)$
- the inputs to this circuit encode candidate solutions
To solve an instance $I$ of the problem $A$:

- take a circuit corresponding to $C(I, \cdot)$
- the inputs to this circuit encode candidate solutions
- use a Circuit-SAT algorithm for this circuit to find a solution (if exists)
SUMMARY

Circuit-SAT
SUMMARY

SAT

Circuit-SAT