Recall Horn formulas are easy to solve

Recall Horn formulas are easy to solve

How about more general formulas: CNF (conjunction normal form)?

Recall Horn formulas are easy to solve

How about more general formulas: CNF (conjunction normal form)?

Definition

A **CNF formula** is a <u>conjunction</u>, of clauses, where each clause is a <u>disjunction</u> of literals

OR

Recall Horn formulas are easy to solve

How about more general formulas: CNF (conjunction normal form)?

Definition

A **CNF formula** is a conjunction of clauses, where each clause is a disjunction of literals

Example: $(v_1 \lor x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_5 \lor x_6) \land (\bar{x}_4 \lor x_7)$

Recall Horn formulas are easy to solve

How about more general formulas: CNF (conjunction normal form)?

Definition

A **CNF formula** is a conjunction of clauses, where each clause is a disjunction of literals

Example:
$$(v_1 \lor x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_5 \lor x_6) \land (\bar{x}_4 \lor x_7)$$

Definition

A k-CNF is a CNF where each clause contains exactly k literals

The Satisfiability Problem (SAT)

The Satisfiability Problem (SAT)

Instance: A CNF Φ

The Satisfiability Problem (SAT)

Instance: A CNF Φ

Objective: Decide if Φ is satisfiable, i.e., is there an assignment so that

Φ is true?

The Satisfiability Problem (SAT)

Instance: A CNF Φ

Objective: Decide if $\boldsymbol{\Phi}$ is satisfiable, i.e., is there an assignment so that

Φ is true?

The *k*-Satisfiability Problem (*k*-SAT)

Instance: A k-CNF Φ

Objective: Decide if Φ is satisfiable

Theorem

3- $SAT \leq_P Independent Set$

Theorem

3- $SAT \leq_P Independent Set$

Proof.

Mar 3, 2022

Theorem

3- $SAT \leq_P Independent Set$

$$\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$$

Theorem

3- $SAT \leq_P Independent Set$

Proof. First consider an intuition for solving SAT:

• pick one literal from each clause

$$\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$$

Theorem

 $3-SAT <_P Independent Set$

- pick one literal from each clause
- select an assignment that satisfies all selected literals

$$\Phi = (x_1) \lor x_2 \lor x_3) \land (x_2) \lor \bar{x}_3 \lor x_4) \land (x_3) \lor \bar{x}_1 \lor x_5)$$

$$\uparrow_{\mathsf{q}} \qquad \uparrow_{\mathsf{q}} \qquad \uparrow_{\mathsf{q}}$$

Theorem

3- $SAT \leq_P Independent Set$

- pick one literal from each clause
- select an assignment that satisfies all selected literals
- make sure there's no conflict:

$$\Phi = (x_1) \lor x_2 \lor x_3) \land (x_2 \lor (\bar{x}_3) \lor x_4) \land (x_3) \lor \bar{x}_1 \lor x_5)$$

Theorem

3-SAT ≤ $_P$ Independent Set

- pick one literal from each clause
- select an assignment that satisfies all selected literals
- make sure there's no conflict: Don't pick x from one clause and \bar{x} from another

$$\Phi = (x_1 \vee x_2 \vee x_3) \wedge (x_2 \vee \bar{x}_3 \vee x_4) \wedge (x_3 \vee \bar{x}_1 \vee x_5)$$

Theorem

3- $SAT \leq_P Independent Set$

- pick one literal from each clause
- select an assignment that satisfies all selected literals
- make sure there's no conflict: Don't pick x from one clause and \bar{x} from another

$$\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$$

Theorem

3-SAT ≤ $_P$ Independent Set

- pick one literal from each clause
- select an assignment that satisfies all selected literals
- make sure there's no conflict: Don't pick x from one clause and \bar{x} from another

$$\Phi = \bigoplus_{\text{good}}^{\text{bad}} \lor x_2 \lor x_3) \land (x_2 \lor \overline{x_3}) \lor x_4) \land (x_3 \lor \overline{x_1} \lor \overline{x_5})$$

Theorem

3-SAT ≤ $_P$ Independent Set

Proof. First consider an intuition for solving SAT:

- pick one literal from each clause
- select an assignment that satisfies all selected literals
- make sure there's no conflict: Don't pick x from one clause and \bar{x} from another

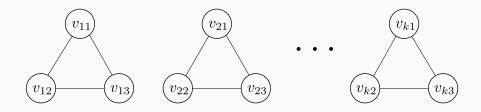
$$\Phi = (\overset{\overset{\text{bad}}{\downarrow}}{\underset{\text{good}}{\downarrow}} \lor x_2 \lor x_3) \land (x_2 \lor \dot{\bar{x}}_3 \lor x_4) \land (\dot{x}_3 \lor \bar{x}_1 \lor x_5)$$

We encode a CNF as a graph, and encode an assignment as independent sets (to keep track of the conflicts)

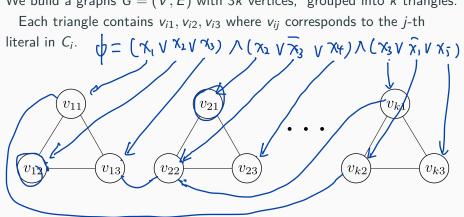


We build a graphs G = (V, E) with 3k vertices,

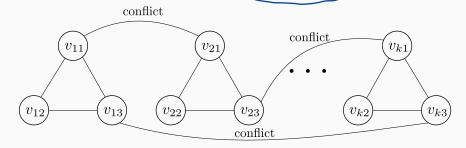
Consider a 3-SAT instance with variables x_1, \ldots, x_n , and clauses C_1, \ldots, C_k We build a graphs G = (V, E) with 3k vertices, grouped into k triangles.



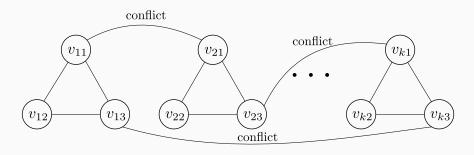
We build a graphs G = (V, E) with 3k vertices, grouped into k triangles. Each triangle contains v_{i1} , v_{i2} , v_{i3} where v_{ii} corresponds to the j-th



We build a graphs G = (V, E) with 3k vertices, grouped into k triangles. Each triangle contains v_{i1}, v_{i2}, v_{i3} where v_{ij} corresponds to the j-th literal in C_i . Add edges for conflicts, i.e., x_j and \bar{x}_j :



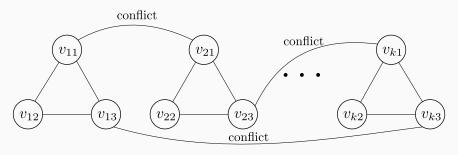
We build a graphs G = (V, E) with 3k vertices, grouped into k triangles. Each triangle contains v_{i1}, v_{i2}, v_{i3} where v_{ij} corresponds to the j-th literal in C_i . Add edges for conflicts, i.e., x_i and \bar{x}_i :



At most one vertex in each triangle can be in an independent set,

We build a graphs G = (V, E) with 3k vertices, grouped into k triangles.

Each triangle contains v_{i1}, v_{i2}, v_{i3} where v_{ij} corresponds to the j-th literal in C_i . Add edges for conflicts, i.e., x_j and \bar{x}_j :



At most one vertex in each triangle can be in an independent set, so the size of an independent set cannot be larger than k

• If there exists a satisfying assignment, there exists a satisfied literal in each clause (triangle).

Mar 3, 2022

 If there exists a satisfying assignment, there exists a satisfied literal in each clause (triangle). Pick such a literal and include it into the independent set

Mar 3, 2022

 If there exists a satisfying assignment, there exists a satisfied literal in each clause (triangle). Pick such a literal and include it into the independent set
 There is no conflicts. It's in fact an independent set

Chunhao Wang

- If there exists a satisfying assignment, there exists a satisfied literal in each clause (triangle). Pick such a literal and include it into the independent set
 There is no conflicts. It's in fact an independent set
- If there exists an independent set S of size k, every triangle contains a vertex from S.

Mar 3, 2022

- If there exists a satisfying assignment, there exists a satisfied literal in each clause (triangle). Pick such a literal and include it into the independent set
 - There is no conflicts. It's in fact an independent set
- If there exists an independent set *S* of size *k*, every triangle contains a vertex from *S*. We can choose an assignment so that all literals (vertices of *S*) are satisfied there's no conflicts

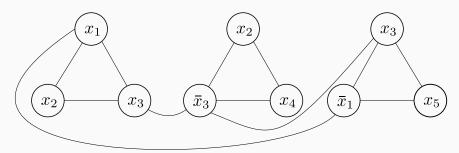
Mar 3, 2022

- If there exists a satisfying assignment, there exists a satisfied literal in each clause (triangle). Pick such a literal and include it into the independent set
 - There is no conflicts. It's in fact an independent set
- If there exists an independent set S of size k, every triangle contains a vertex from S. We can choose an assignment so that all literals (vertices of S) are satisfied — there's no conflicts

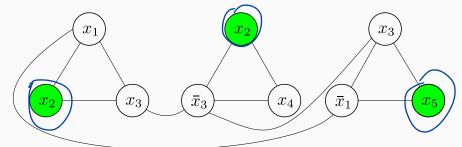
So the 3-CNF has a satisfying assignment if and only if ${\it G}$ has an independent set of size ${\it k}$

Consider
$$\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$$

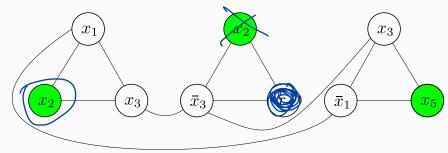
Consider
$$\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$$



Consider
$$\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$$



Consider $\Phi = (x_1 \lor x_2 \lor x_3) \land (x_2 \lor \bar{x}_3 \lor x_4) \land (x_3 \lor \bar{x}_1 \lor x_5)$



Satisfying assignment:
$$x_1 = 0, x_2 = 1, x_3 = 0, x_4 = 0, x_5 = 1$$

NP and Computational Hardness

P, NP, and NP-completeness (Kleinberg-Tardos, Section 8.3, 8.4)

We can encode the input (an instance) of any computational problem as a binary string

We can encode the input (an instance) of any computational problem as a binary string

A decision problem X is the set of strings on which the answer is "yes"

We can encode the input (an instance) of any computational problem as a binary string

A decision problem X is the set of strings on which the answer is "yes"

An algorithm A for a decision problem receives an input string s and

outputs
$$A(s) = \begin{cases} yes \\ no \end{cases}$$

We can encode the input (an instance) of any computational problem as a binary string (χ_{ye}, χ_{h})

A decision problem X is the set of strings on which the answer is "yes"

An algorithm A for a decision problem receives an input string s and

outputs
$$A(s) = \begin{cases} yes \\ no \end{cases}$$

The algorithm A solves X if for all s, A(s) = yes if and only if $s \in X$

We can encode the input (an instance) of any computational problem as a binary string

A decision problem X is the set of strings on which the answer is "yes"

An algorithm A for a decision problem receives an input string s and

outputs
$$A(s) = \begin{cases} yes \\ no \end{cases}$$

The algorithm A solves X if for all s, A(s) = yes if and only if $s \in X$

The algorithm A has **polynomial running time** if there is a polynomial p s.t. for all s, A terminates on s in at most O(p(s)) steps

We can encode the input (an instance) of any computational problem as a binary string

A decision problem X is the set of strings on which the answer is "yes"

An algorithm \boldsymbol{A} for a decision problem receives an input string \boldsymbol{s} and

outputs
$$A(s) = \begin{cases} yes \\ no \end{cases}$$

The algorithm A solves X if for all s, A(s) = yes if and only if $s \in X$

The algorithm A has **polynomial running time** if there is a polynomial p s.t. for all s, A terminates on s in at most O(p(|s|)) steps

Computational class

 \boldsymbol{P} : the class of all problems for which there exists a polynomial-time algorithm

Definition

An algorithm ${\it B}$ is an **efficient certifier** for a problem ${\it X}$ if

Definition

An algorithm B is an efficient certifier for a problem X if

• B is a polynomial-time algorithm that takes two inputs (3,(t)) and

Mar 3, 2022

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and |B(s,t)| = yes

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yes

The string t is called a **certificate**

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

Mar 3, 2022

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

■ 3-SAT:

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

• 3-SAT: certificate: an assignment

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

■ 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x_1} \lor \bar{x}_3 \lor \bar{x}_4)$

Mar 3, 2022

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

■ 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x}_1 \lor \bar{x}_3 \lor \bar{x}_4)$ certificate t: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1$

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

- 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x}_1 \lor \bar{x}_3 \lor \bar{x}_4)$ certificate t: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1$
- Independent set.

Mar 3, 2022

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

- 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x}_1 \lor \bar{x}_3 \lor \bar{x}_4)$ certificate t: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1$
- Independent set. certificate: a set of at least k vertices

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

- 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x}_1 \lor \bar{x}_3 \lor \bar{x}_4)$ certificate t: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1$
- Independent set. certificate: a set of at least k vertices certifier: check if there's no edge joining them

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yes

The string t is called a **certificate**

Example:

- 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x}_1 \lor \bar{x}_3 \lor \bar{x}_4)$ certificate t: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1$
- Independent set. certificate: a set of at least k vertices certifier: check if there's no edge joining them

We can use B to design an algorithm for X: use brute force to find a t.

Definition

An algorithm B is an **efficient certifier** for a problem X if

- B is a polynomial-time algorithm that takes two inputs s, t, and
- there exists a polynomial p s.t. for all s, we have $s \in X$ if and only if there exists a string t s.t. $|t| \le p(|s|)$ and B(s,t) = yesThe string t is called a **certificate**

Example:

- 3-SAT: certificate: an assignment instance s: $(\bar{x}_1 \lor x_2 \lor x_3) \land (x_1 \lor \bar{x}_2 \lor x_3) \land (x_1 \lor x_2 \lor x_4) \land (\bar{x}_1 \lor \bar{x}_3 \lor \bar{x}_4)$ certificate t: $x_1 = 1, x_2 = 1, x_3 = 0, x_4 = 1$
- Independent set. certificate: a set of at least k vertices certifier: check if there's no edge joining them

We can use B to design an algorithm for X: use brute force to find a t. But there might be exponentially many possible t's

Computational class

 $\ensuremath{\mathsf{NP}}$: the class of all problems for which there exists an efficient certifier

Computational class

 $\ensuremath{\mathsf{NP}}$: the class of all problems for which there exists an efficient certifier

It is easy to see: $3-SAT \in \mathbf{NP}$

Independent Set ENP Vertex (nor ENP

Computational class

 $\ensuremath{\mathsf{NP}}$: the class of all problems for which there exists an efficient certifier

It is easy to see: $3-SAT \in \mathbf{NP}$

Lemma

 $\mathbf{P}\subseteq \mathbf{NP}$

Computational class

NP : the class of all problems for which there exists an efficient certifier

It is easy to see: $3-SAT \in \mathbf{NP}$

Lemma

 $\mathsf{P}\subseteq\mathsf{NP}$

Proof.

For any problem in $\bf P$ with algorithm A, we construct a certifier B that just returns A(s) with empty certificate t

Fundamental question in CS: is P = NP?

Fundamental question in CS: is $\mathbf{P}=\mathbf{NP}?$ i.e., does there exist a problem $X\in\mathbf{NP}$ but $X\not\in\mathbf{P}?$

Fundamental question in CS: is $\mathbf{P} = \mathbf{NP}?$ i.e., does there exist a problem $X \in \mathbf{NP}$ but $X \notin \mathbf{P}?$

We don't know the answer, but we try to find the most difficult problems in **NP**:

Fundamental question in CS: is P = NP? i.e., does there exist a problem $X \in NP$ but $X \notin P$?

We don't know the answer, but we try to find the most difficult problems in **NP**:

Definition

A problem X is **NP-complete** if

Mar 3, 2022

Fundamental question in CS: is $\mathbf{P} = \mathbf{NP}$? i.e., does there exist a problem $X \in \mathbf{NP}$ but $X \not\in \mathbf{P}$?

We don't know the answer, but we try to find the most difficult problems in **NP**:

Definition

A problem X is **NP-complete** if

 $\mathbf{X} \in \mathbf{NP}$ and

Fundamental question in CS: is $\mathbf{P}=\mathbf{NP}?$ i.e., does there exist a problem $X\in\mathbf{NP}$ but $X\not\in\mathbf{P}?$

We don't know the answer, but we try to find the most difficult problems in **NP**:

Definition

A problem X is **NP-complete** if

- $\mathbf{X} \in \mathbf{NP}$ and
- for all $Y \in NP$, $Y \leq_P X$

Fundamental question in CS: is $\mathbf{P}=\mathbf{NP}?$ i.e., does there exist a problem $X\in\mathbf{NP}$ but $X\not\in\mathbf{P}?$

We don't know the answer, but we try to find the most difficult problems in **NP**:

Definition

A problem X is **NP-complete** if

- $\mathbf{X} \in \mathbf{NP}$ and
- for all $Y \in \mathbf{NP}$, $Y \leq_P X$

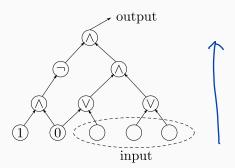
Lemma

If an $\ensuremath{\mathsf{NP}}\xspace$ -complete problem can be solved in polynomial time, then

P = NP

A first $\ensuremath{\textbf{NP}}\xspace\text{-complete}$ problem: Circuit Satisfiability

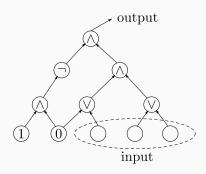
A first NP-complete problem: Circuit Satisfiability



A first NP-complete problem: Circuit Satisfiability

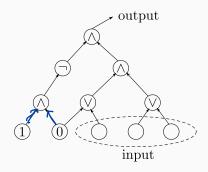
A circuit consists of

inputs



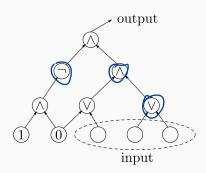
A first NP-complete problem: Circuit Satisfiability

- inputs
- wires



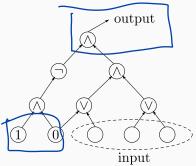
A first **NP**-complete problem: Circuit Satisfiability

- inputs
- wires
- logical gates ∨, ∧, ¬



A first **NP**-complete problem: Circuit Satisfiability

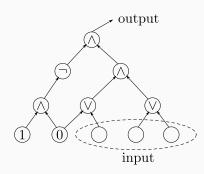
- inputs
- wires
- logical gates ∨, ∧, ¬
- single output



A first **NP**-complete problem: Circuit Satisfiability

A circuit consists of

- inputs
- wires
- logical gates ∨, ∧, ¬
- single output



The Circuit Satisfiability Problem (circuit-SAT)

Instance: A circuit *C*

Objective: Decide if *C* is satisfiable

The Cook-Levin Theorem

Theorem (Cook-Levin)

circuit-SAT is **NP**-complete