L
P
U

CSE408
Complexity Classes

Lecture # 40

NP-Completeness

- Poly time algorithm: input size n (in some encoding), worst case running time – $O(n^c)$ for some constant c.
- Three classes of problems
  - P: problems solvable in poly time.
  - NP: problems verifiable in poly time.
  - NPC: problems in NP and as hard as any problem in NP.

NP-Completeness (verifiable)

- Verifiable in poly time: given a certificate of a solution, could verify the certificate is correct in poly time.
- Examples (their definitions come later):
  - Hamiltonian-cycle, given a certificate of a sequence $(v_1, v_2, \ldots, v_n)$, easily verified in poly time.
  - 3-CNF, given a certificate of an assignment 0s, 1s, easily verified in poly time.
  - (so try each instance, and verify it, but 2^n instances)
- Why not defined as “solvable in exponential time?” or “Non Poly time”?

NP-Completeness (why NPC?)

- A problem $p \in$NP, and any other problem $p' \in$NP can be translated as p in poly time.
- So if p can be solved in poly time, then all problems in NP can be solved in poly time.
- All current known NP hard problems have been proved to be NPC.

Relation among P, NP, NPC

- $P \subseteq NP$ (Sure)
- $NPC \subseteq NP$ (sure)
- $P = NP$ (or $P \subset NP$, or $P \neq NP$) ???
- $NPC = NP$ (or $NPC \subset NP$, or $NPC \neq NP$) ???
- $P \neq NP$: one of the deepest, most perplexing open research problems in (theoretical) computer science since 1971.

Arguments about P, NP, NPC

- No poly algorithm found for any NPC problem (even so many NPC problems)
- No proof that a poly algorithm cannot exist for any of NPC problems, (even having tried so long so hard).
- Most theoretical computer scientists believe that NPC is intractable (i.e., hard, and $P \neq NP$).

View of Theoretical Computer Scientists on P, NP, NPC

NP

NPC

P

$P \subset NP$, $NPC \subset NP$, $P \cap NPC = \varnothing$

Why discussion on NPC

- If a problem is proved to be NPC, a good evidence for its intractability (hardness).
- Not waste time on trying to find efficient algorithm for it
- Instead, focus on design approximate algorithm or a solution for a special case of the problem
- Some problems looks very easy on the surface, but in fact, is hard (NPC).

Decision VS. Optimization Problems

- Decision problem: solving the problem by giving an answer “YES” or “NO”
- Optimization problem: solving the problem by finding the optimal solution.
- Examples:
  - SHORTEST-PATH (optimization)
    - Given *G*, *u*,*v*, find a path from *u* to *v* with fewest edges.
  - PATH (decision)
    - Given *G*, *u*,*v*, and *k*, whether exist a path from *u* to *v* consisting of at most *k* edges.

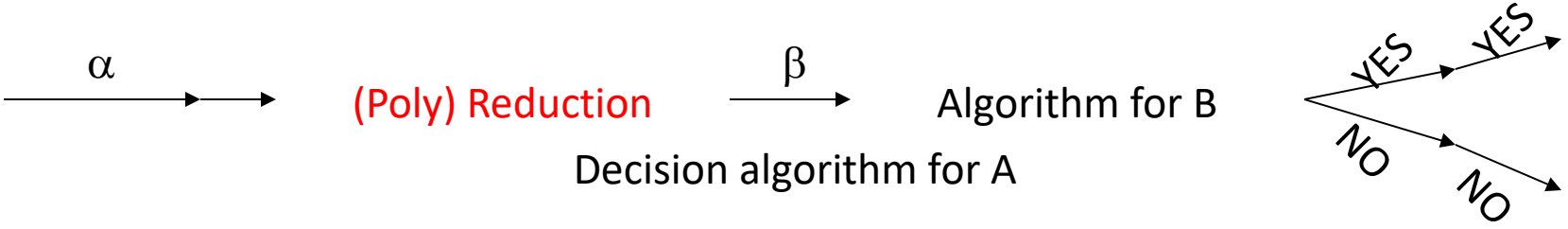
Decision VS. Optimization Problems (Cont.)

- Decision is easier (i.e., no harder) than optimization
- If there is an algorithm for an optimization problem, the algorithm can be used to solve the corresponding decision problem.
  - Example: SHORTEST-PATH for PATH
- If optimization is easy, its corresponding decision is also easy. Or in another way, if provide evidence that decision problem is hard, then the corresponding optimization problem is also hard.
- NPC is confined to decision problem. (also applicable to optimization problem.)
  - Another reason is that: easy to define reduction between decision problems.

(Poly) reduction between decision problems

- Problem (class) and problem instance
- Instance α of decision problem A and instance β of decision problem B
- A reduction from A to B is a transformation with the following properties:
  - The transformation takes poly time
  - The answer is the same (the answer for α is YES if and only if the answer for β is YES).

Implication of (poly) reduction



1. If decision algorithm for B is poly, so does A.
   A is no harder than B (or B is no easier than A)

2. If A is hard (e.g., NPC), so does B.

3. How to prove a problem B to be NPC ??

(at first, prove B is in NP, which is generally easy.)

3.1 find a already proved NPC problem A

3.2 establish an (poly) reduction from A to B

Question: What is and how to prove the first NPC problem?

Circuit-satisfiability problem.

Discussion on Poly time problems

- $\Theta(n^{100})$ vs. $\Theta(2^n)$
  - Reasonable to regard a problem of $\Theta(n^{100})$ as intractable, however, very few practical problem with $\Theta(n^{100})$.
  - Most poly time algorithms require much less.
  - Once a poly time algorithm is found, more efficient algorithm may follow soon.
- Poly time keeps same in many different computation models, e.g., poly class of serial random-access machine ≡ poly class of abstract Turing machine ≡ poly class of parallel computer (#processors grows polynomially with input size)
- Poly time problems have nice closure properties under addition, multiplication and composition.

Encoding impact on complexity

- The problem instance must be represented in a way the program (or machine) can understand.
- General encoding is “binary representation”.
- Different encoding will result in different complexities.
- Example: an algorithm, only input is integer k, running time is $\Theta(k)$.
  - If k is represented in *unary*: a string of k 1s, the running time is $\Theta(k) = \Theta(n)$ on length-n input, poly on n.
  - If k is represented in *binary:* the input length $n = \lfloor \log k \rfloor + 1$, the running time is $\Theta(k) = \Theta(2^n)$, exponential on n.
- Ruling out *unary*, other encoding methods are same.

Examples of encoding and complexity

- Given integer *n*, check whether *n* is a composite.
- Dynamic programming for subset-sum.

Class P Problems

- Let n= the length of binary encoding of a problem (i.e., input size), $T(n)$ is the time to solve it.
- A problem is *poly-time solvable* if $T(n) = O(n^k)$ for some constant k.
- Complexity class **P**=set of problems that are *poly-time solvable*.

Poly Time Verification

- PATH problem: Given <G,*u*,*v*,*k*>, whether exists a path from *u* to *v* with at most *k* edges?
- Moreover, also given a path *p* from *u* to *v*, verify whether the length of *p* is at most *k*?
- Easy or not?

Of course, very easy.

Poly Time Verification, encoding, and language

- Hamiltonian cycles
  - A simple path containing every vertex.
  - HAM-CYCLE={<G>: G is a Hamiltonian graph, i.e. containing Hamiltonian cycle}.
  - Suppose ***n*** is the length of **encoding** of G.
  - HAM-CYCLE can be considered as a **Language** after encoding, i.e. a subset of $\sum^*$ where $\sum=\{0,1\}^*$.
- The naïve algorithm for determining HAM-CYCLE runs in $\Omega(m!)=\Omega(2^m)$ time, where m is the number of vertices, $m \approx n^{1/2}$.
- However, given an ordered sequence of m vertices (called "certificate"), let you verify whether the sequence is a Hamiltonian cycle. Very easy. In $O(n^2)$ time.

Class NP problems

- For a problem *p*, given its certificate, the certificate can be verified in poly time.
- Call this kind of problem an NP one.
- Complement set/class: Co-NP.
  - Given a set S (as a universal) and given a subset A
  - The complement is that S-A.
  - When NP problems are represented as languages (i.e. a set), we can discuss their complement set, i.e., Co-NP.

Relation among P, NP and co-NP={L: L $\in$NP where L=$\overline{\sum^*-L}$}

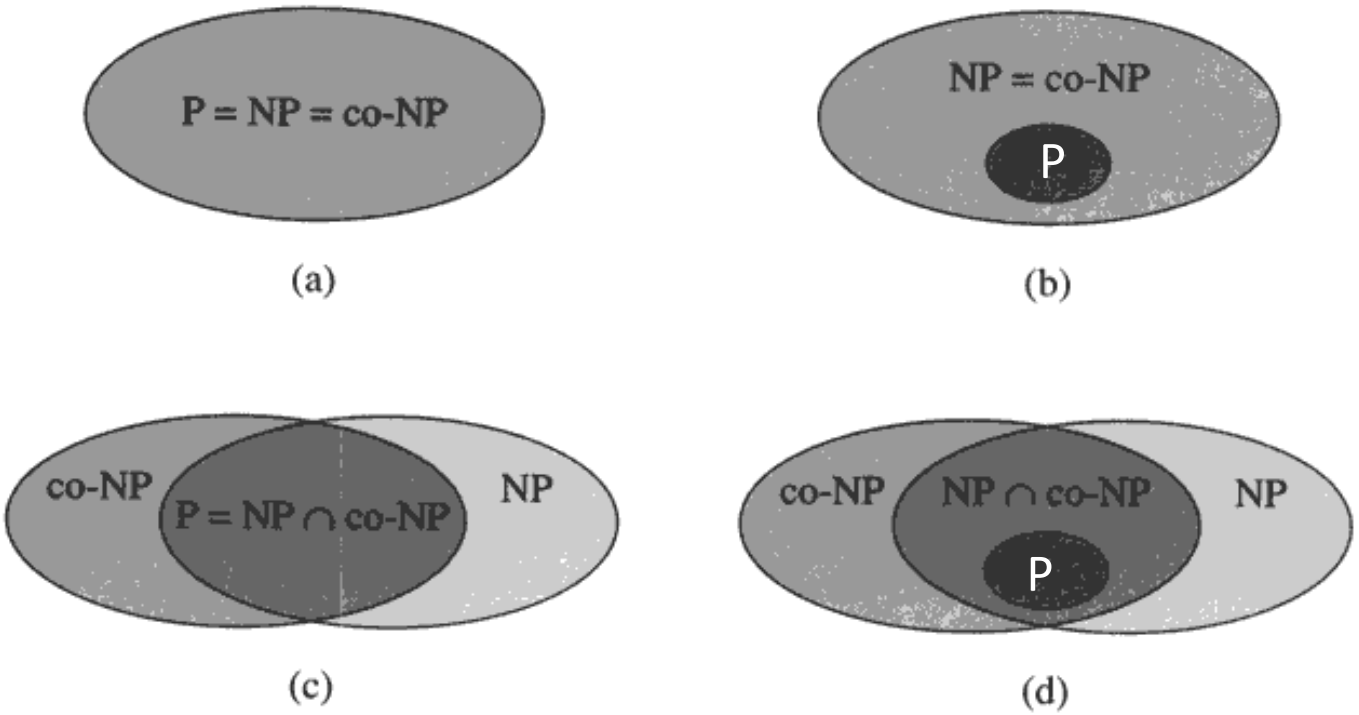


Figure 34.3 Four possibilities for relationships among complexity classes. In each diagram, one region enclosing another indicates a proper-subset relation. **(a)** P = NP = co-NP. Most researchers regard this possibility as the most unlikely. **(b)** If NP is closed under complement, then NP = co-NP, but it need not be the case that P = NP. **(c)** P = NP ∩ co-NP, but NP is not closed under complement. **(d)** NP ≠ co-NP and P ≠ NP ∩ co-NP. Most researchers regard this possibility as the most likely.

NP-completeness and Reducibility

- A (class of) problem P_1 is poly-time reducible to P_2 , written as $P_1 \leq_p P_2$ if there exists a poly-time function f: $P_1 \rightarrow P_2$ such that for any instance of $p_1 \in P_1$, p_1 has "YES" answer if and only if answer to $f(p_1)$ ($\in P_2$) is also "YES".
- *Theorem 34.3*: (page 985)
  - For two problems P_1, P_2, if $P_1 \leq_p P_2$ then $P_2 \in P$ implies $P_1 \in P$.

NP-completeness and Reducibility (cont.)

- A problem p is NP-complete if
  1. $p \in NP$ and
  2. $p' \leq_p p$ for every $p' \in NP$.

  (if p satisfies 2, then p is said NP-hard.)

Theorem 34.4 (page 986)

if any NP-compete problem is poly-time solvable, then P=NP. Or say: if any problem in NP is not poly-time solvable, then no NP-complete problem is poly-time solvable.

First NP-complete problem—Circuit Satisfiability (problem definition)

- Boolean combinational circuit
  - Boolean combinational elements, wired together
  - Each element, inputs and outputs (binary)
  - Limit the number of outputs to 1.
  - Called *logic gates*: NOT gate, AND gate, OR gate.
  - *true table*: giving the outputs for each setting of inputs
  - true assignment: a set of boolean inputs.
  - satisfying assignment: a true assignment causing the output to be 1.
  - A circuit is satisfiable if it has a satisfying assignment.

Circuit Satisfiability Problem: definition

- Circuit satisfying problem: given a boolean combinational circuit composed of AND, OR, and NOT, is it stisfiable?
- CIRCUIT-SAT={<C>: C is a satisfiable boolean circuit}
- Implication: in the area of computer-aided hardware optimization, if a subcircuit always produces 0, then the subcircuit can be replaced by a simpler subcircuit that omits all gates and just output a 0.

Two instances of circuit satisfiability problems

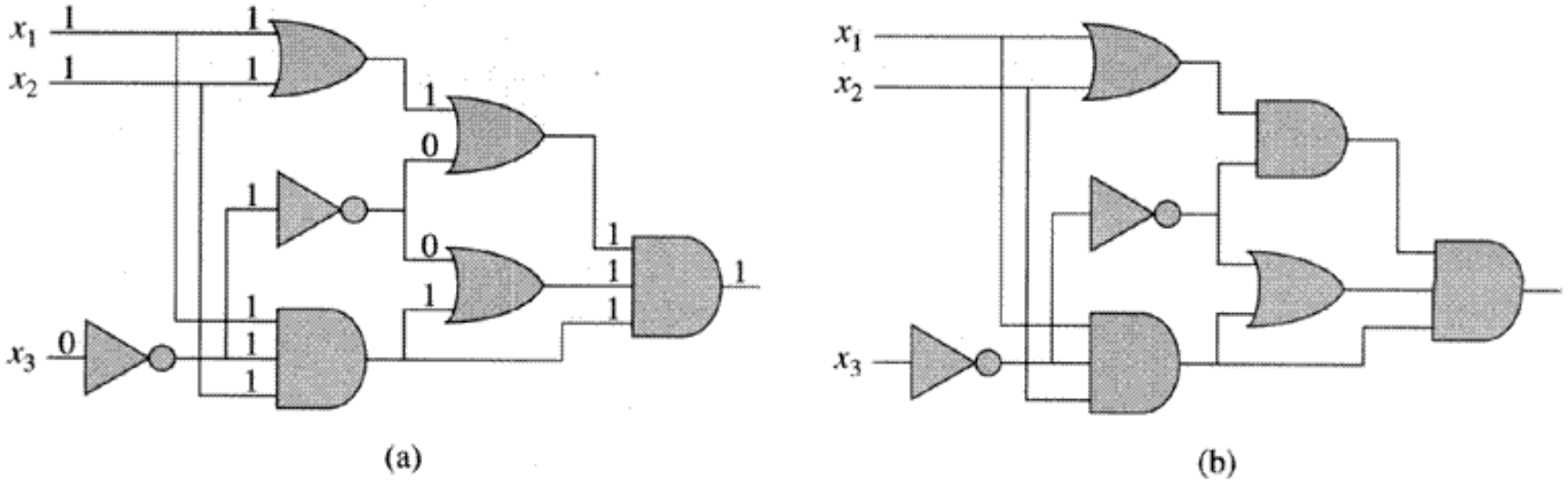


Figure 34.8 Two instances of the circuit-satisfiability problem. **(a)** The assignment $\langle x_1 = 1, x_2 = 1, x_3 = 0 \rangle$ to the inputs of this circuit causes the output of the circuit to be 1. The circuit is therefore satisfiable. **(b)** No assignment to the inputs of this circuit can cause the output of the circuit to be 1. The circuit is therefore unsatisfiable.

Solving circuit-satisfiability problem

- Intuitive solution:
  - for each possible assignment, check whether it generates 1.
  - suppose the number of inputs is k, then the total possible assignments are 2^k. So the running time is $\Omega(2^k)$. When the size of the problem is $\Theta(k)$, then the running time is not poly.

Circuit-satisfiability problem is NP-complete

- *Lemma 34.5*:(page 990)
  - CIRCUIT-SAT belongs to NP.
- Proof: CIRCUIT-SAT is poly-time verifiable.
  - Given (an encoding of) a CIRCUIT-SAT problem C and a certificate, which is an assignment of boolean values to (all) wires in C.
  - The algorithm is constructed as follows: just checks each gates and then the output wire of C:
    - If for every gate, the computed output value matches the value of the output wire given in the certificate and the output of the whole circuit is 1, then the algorithm outputs 1, otherwise 0.
    - The algorithm is executed in poly time (even linear time).
- An alternative certificate: a true assignment to the inputs.

Circuit-satisfiability problem is NP-complete (cont.)

- *Lemma 34.6:* (page 991)
  - CIRCUIT-SAT is NP-hard.
- Proof: Suppose X is *any problem* in NP
  - construct a poly-time algorithm *F* maps every problem instance *x* in X to a circuit C=*f*(*x*) such that the answer to *x* is YES if and only if C$\in$CIRCUIT-SAT (is satisfiable).

Circuit-satisfiability problem is NP-hard (cont.)

- Since $X \in NP$, there is a poly-time algorithm A which verifies X.
- Suppose the input length is n and Let $T(n)$ denote the worst-case running time. Let k be the constant such that $T(n)=O(n^k)$ and the length of the certificate is $O(n^k)$.

Circuit-satisfiability problem is NP-hard (cont.)

- Idea is to represent the computation of A as a sequence of configurations, $c_0, c_1,\ldots,c_i,c_{i+1},\ldots,c_{T(n)}$, each c_i can be broken into
  - (program for A, program counter PC, auxiliary machine state, input x, certificate y, working storage) and
  - c_i is mapped to c_{i+1} by the combinational circuit M implementing the computer hardware.
  - The output of A: 0 or 1– is written to some designated location in working storage. If the algorithm runs for at most $T(n)$ steps, the output appears as one bit in $c_{T(n)}$.
  - Note: $A(x,y)$=1 or 0.

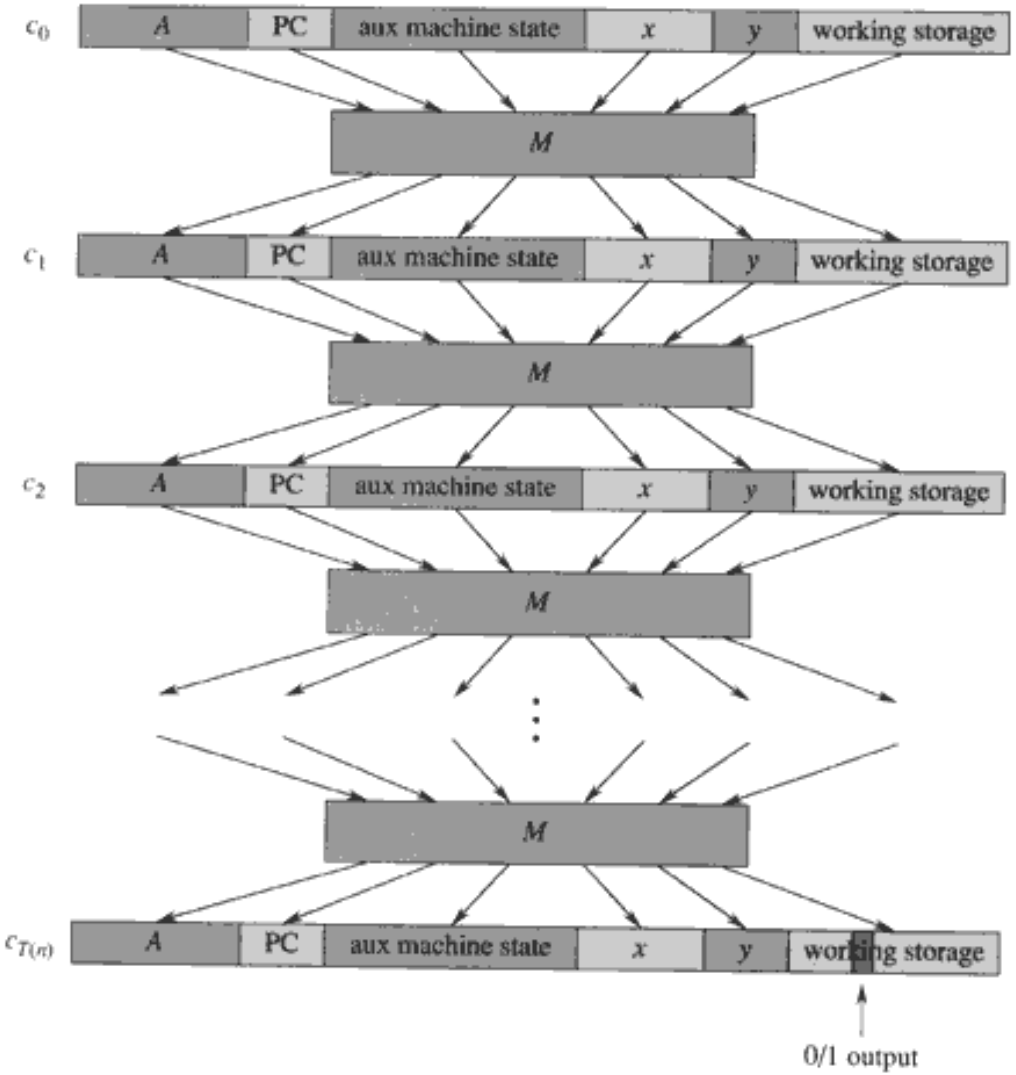


Figure 34.9 The sequence of configurations produced by an algorithm A running on an input x and certificate y. Each configuration represents the state of the computer for one step of the computation and, besides A, x, and y, includes the program counter (PC), auxiliary machine state, and working storage. Except for the certificate y, the initial configuration c_0 is constant. Each configuration is mapped to the next configuration by a boolean combinational circuit M. The output is a distinguished bit in the working storage.

Circuit-satisfiability problem is NP-hard (cont.)

- The reduction algorithm F constructs a single combinational circuit C as follows:
  - Paste together all $T(n)$ copies of the circuit M.
  - The output of the ith circuit, which produces c_i, is directly fed into the input of the $(i+1)$st circuit.
  - All items in the initial configuration, except the bits corresponding to certificate y, are wired directly to their known values.
  - The bits corresponding to y are the inputs to C.
  - All the outputs to the circuit are ignored, except the one bit of $c_{T(n)}$ corresponding to the output of A.

Circuit-satisfiability problem is NP-hard (cont.)

- Two properties remain to be proven:
  - F correctly constructs the reduction, i.e., C is satisfiable if and only if there exists a certificate y, such that $A(x,y)=1$.
    - $\Leftarrow$Suppose there is a certificate y, such that $A(x,y)=1$. Then if we apply the bits of y to the inputs of C, the output of C is the bit of $A(x,y)$, that is $C(y)= A(x,y) =1$, so C is satisfiable.
    - $\Rightarrow$Suppose C is satisfiable, then there is a y such that $C(y)=1$. So, $A(x,y)=1$.
  - F runs in poly time.

Circuit-satisfiability problem is NP-hard (cont.)

- F runs in poly time.
  - Poly space:
    - Size of x is n.
    - Size of A is constant, independent of x.
    - Size of y is $O(n^k)$.
    - Amount of working storage is poly in n since A runs at most $O(n^k)$.
    - M has size poly in length of configuration, which is poly in $O(n^k)$, and hence is poly in n.
    - C consists of at most $O(n^k)$ copies of M, and hence is poly in n.
    - Thus, the C has poly space.
  - The construction of C takes at most $O(n^k)$ steps and each step takes poly time, so F takes poly time to construct C from x.

CIRCUIT-SAT is NP-complete

- In summary
  - CIRCUIT-SAT belongs to NP, verifiable in poly time.
  - CIRCUIT-SAT is NP-hard, every NP problem can be reduced to CIRCUIT-SAT in poly time.
  - Thus CIRCUIT-SAT is NP-complete.

NP-completeness proof basis

- *Lemma 34.8* (page 995)
  - If X is a problem (class) such that $P' \leq_p X$ for some $P' \in$ NPC, then X is NP-hard. Moreover, if $X \in$ NP, then $X \in$ NPC.
- Steps to prove X is NP-complete
  - Prove $X \in$ NP.
    - Given a certificate, the certificate can be verified in poly time.
  - Prove X is NP-hard.
    - Select a known NP-complete P'.
    - Describe a transformation function *f* that maps every instance *x* of P' into an instance *f*(*x*) of X.
    - Prove *f* satisfies that the answer to $x \in P'$ is YES if and only if the answer to $f(x) \in X$ is YES for all instance $x \in P'$.
    - Prove that the algorithm computing *f* runs in poly-time.

NPC proof –Formula Satisfiability (SAT)

- SAT definition
  - n boolean variables: $x_1, x_2,\ldots, x_n$.
  - M boolean connectives: any boolean function with one or two inputs and one output, such as $\wedge,\vee,\neg,\rightarrow,\leftrightarrow,\ldots$ and
  - Parentheses.
- A SAT ϕ is satisfiable if there exists an true assignment which causes ϕ to evaluate to 1.
- SAT=$\{< \phi >: \phi$ is a satifiable boolean formula$\}$.
- The historical honor of the first NP-complete problem ever shown.

SAT is NP-complete

- *Theorem 34.9:* (page 997)
  - SAT is NP-complete.
- Proof:
  - SAT belongs to NP.
    - Given a satisfying assignment, the verifying algorithm replaces each variable with its value and evaluates the formula, in poly time.
  - SAT is NP-hard (show CIRCUIT-SAT$\leq_p$ SAT).

SAT is NP-complete (cont.)

- CIRCUIT-SAT$\leq_p$ SAT, i.e., any instance of circuit satisfiability can be reduced in poly time to an instance of formula satisfiability.
- Intuitive induction:
  - Look at the gate that produces the circuit output.
  - Inductively express each of gate's inputs as formulas.
  - Formula for the circuit is then obtained by writing an expression that applies the gate's function to its input formulas.

- Unfortunately, this is not a poly reduction
  - Shared formula (the gate whose output is fed to 2 or more inputs of other gates) cause the size of generated formula to grow exponentially.

SAT is NP-complete (cont.)

- Correct reduction:
  - For every wire x_i of C, give a variable x_i in the formula.
  - Every gate can be expressed as $x_o \leftrightarrow (x_{i_1} \theta\ x_{i_2} \theta \ldots \theta\ x_{i_l})$
  - The final formula ϕ is the AND of the circuit output variable and conjunction of all clauses describing the operation of each gate. (example Figure 34.10)
- Correctness of the reduction
  - Clearly the reduction can be done in poly time.
  - C is satisfiable if and only if ϕ is satisfiable.
    - If C is satisfiable, then there is a satisfying assignment. This means that each wire of C has a well-defined value and the output of C is 1. Thus the assignment of wire values to variables in ϕ makes each clause in ϕ evaluate to 1. So ϕ is 1.
    - The reverse proof can be done in the same way.

Example of reduction of CIRCUIT-SAT to SAT

$$\phi = x_{10} \wedge (x_{10} \leftrightarrow (x_7 \wedge x_8 \wedge x_9))$$
$$\wedge (x_9 \leftrightarrow (x_6 \vee x_7))$$
$$\wedge (x_8 \leftrightarrow (x_5 \vee x_6))$$
$$\wedge (x_7 \leftrightarrow (x_1 \wedge x_2 \wedge x_4))$$
$$\wedge (x_6 \leftrightarrow \neg x_4))$$
$$\wedge (x_5 \leftrightarrow (x_1 \vee x_2))$$
$$\wedge (x_4 \leftrightarrow \neg x_3)$$

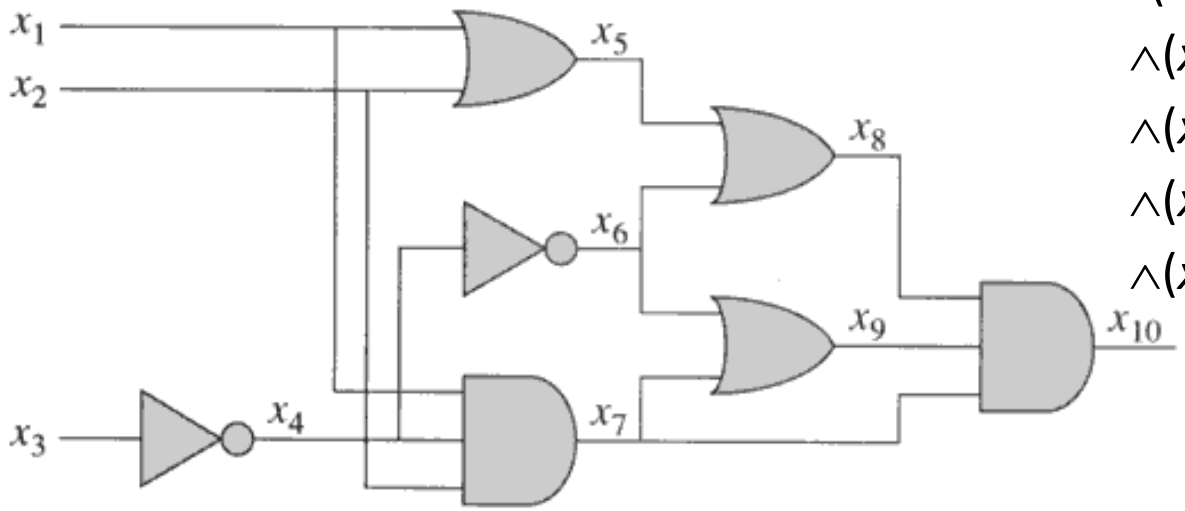


Figure 34.10 Reducing circuit satisfiability to formula satisfiability. The formula produced by the reduction algorithm has a variable for each wire in the circuit.

INCORRECT REDUCTION: $\phi = x_{10} = x_7 \wedge x_8 \wedge x_9 = (x_1 \wedge x_2 \wedge x_4) \wedge (x_5 \vee x_6) \wedge (x_6 \vee x_7)$
$= (x_1 \wedge x_2 \wedge x_4) \wedge ((x_1 \vee x_2) \vee \neg x_4) \wedge (\neg x_4 \vee (x_1 \wedge x_2 \wedge x_4)) = \ldots$

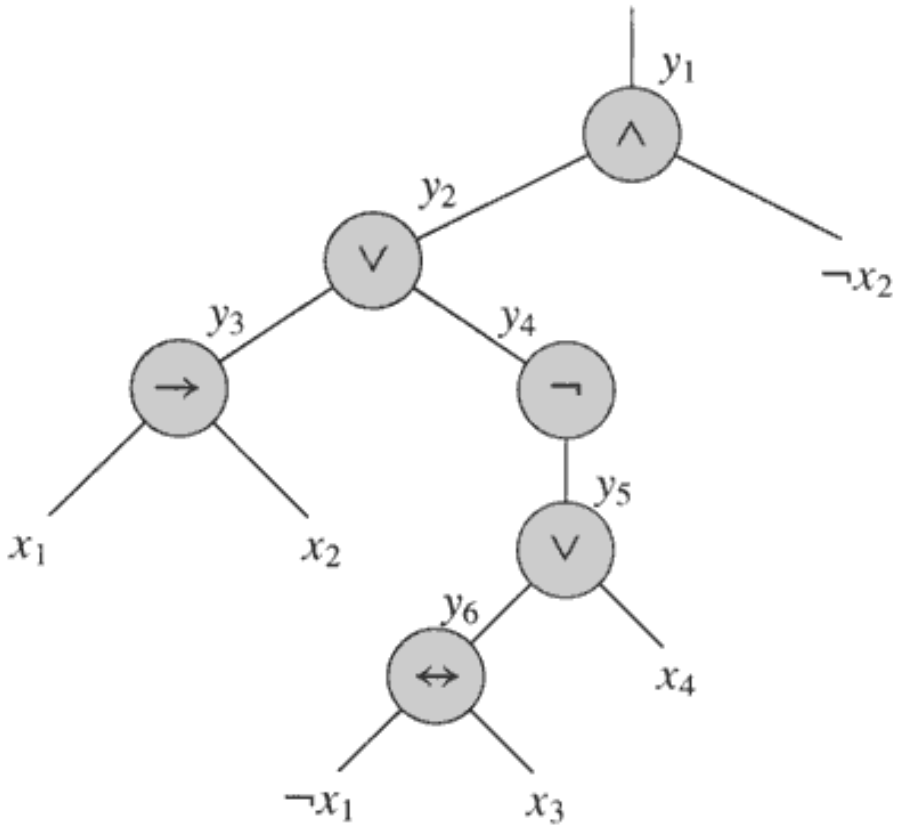
NPC Proof –3-CNF Satisfiability

- 3-CNF definition
  - A *literal* in a boolean formula is an occurrence of a variable or its negation.
  - CNF (Conjunctive Nornal Form) is a boolean formula expressed as AND of clauses, each of which is the OR of one or more literals.
  - 3-CNF is a CNF in which each clause has exactly 3 distinct literals (a literal and its negation are distinct)
- 3-CNF-SAT: whether a given 3-CNF is satiafiable?

3-CNF-SAT is NP-complete

- Proof: 3-CNF-SAT $\in$ NP. Easy.
  - 3-CNF-SAT is NP-hard. (show SAT $\leq_p$ 3-CNF-SAT)
    - Suppose ϕ is any boolean formula, Construct a binary 'parse' tree, with literals as leaves and connectives as internal nodes.
    - Introduce a variable y_i for the output of each internal nodes.
    - Rewrite the formula to ϕ' as the AND of the root variable and a conjunction of clauses describing the operation of each node.
    - The result is that in ϕ', each clause has at most three literals.
    - Change each clause into conjunctive normal form as follows:
      - Construct a true table, (small, at most 8 by 4)
      - Write the disjunctive normal form for all true-table items evaluating to 0
      - Using DeMorgan law to change to CNF.
    - The resulting ϕ'' is in CNF but each clause has 3 or less literals.
    - Change 1 or 2-literal clause into 3-literal clause as follows:
      - If a clause has one literal l, change it to $(l \vee p \vee q) \wedge (l \vee p \vee \neg q) \wedge (l \vee \neg p \vee q) \wedge (l \vee \neg p \vee \neg q)$.
      - If a clause has two literals $(l_1 \vee l_2)$, change it to $(l_1 \vee l_2 \vee p) \wedge (l_1 \vee l_2 \vee \neg p)$.

Binary parse tree for $\phi=((x_1 \rightarrow x_2) \vee \neg((\neg x_1 \leftrightarrow x_3) \vee x_4)) \wedge \neg x_2$



$$\phi' = y_1 \wedge (y_1 \leftrightarrow (y_2 \wedge \neg x_2)) \wedge (y_2 \leftrightarrow (y_3 \vee y_4)) \wedge (y_4 \leftrightarrow \neg y_5) \wedge (y_3 \leftrightarrow (x_1 \rightarrow x_2)) \wedge (y_5 \leftrightarrow (y_6 \vee x_4)) \wedge (y_6 \leftrightarrow (\neg x_1 \leftrightarrow x_3))$$

Figure 34.11 The tree corresponding to the formula $\phi = ((x_1 \rightarrow x_2) \vee \neg((\neg x_1 \leftrightarrow x_3) \vee x_4)) \wedge \neg x_2$.

y_1	y_2	x_2	$(y_1 \leftrightarrow (y_2 \wedge \neg x_2))$
1	1	1	0
1	1	0	1
1	0	1	0
1	0	0	0
0	1	1	1
0	1	0	0
0	0	1	1
0	0	0	1

Disjunctive Normal Form:

$\phi_i' = (y_1 \wedge y_2 \wedge x_2) \vee (y_1 \wedge \neg y_2 \wedge x_2)$
$\vee (y_1 \wedge \neg y_2 \wedge \neg x_2) \vee (\neg y_1 \wedge y_2 \wedge \neg x_2)$

Conjunctive Normal Form:

$\phi_i'' = (\neg y_1 \vee \neg y_2 \vee \neg x_2) \wedge (\neg y_1 \vee y_2 \vee \neg x_2)$
$\wedge (\neg y_1 \vee y_2 \vee x_2) \wedge (y_1 \vee \neg y_2 \vee x_2)$

Figure 34.12 The truth table for the clause $(y_1 \leftrightarrow (y_2 \wedge \neg x_2))$.

3-CNF is NP-complete

- ϕ and reduced 3-CNF are equivalent:
  - From ϕ to ϕ' , keep equivalence.
  - From ϕ' to ϕ'' , keep equivalence.
  - From ϕ'' to final 3-CNF, keep equivalence.
- Reduction is in poly time,
  - From ϕ to ϕ' , introduce at most 1 variable and 1 clause per connective in ϕ.
  - From ϕ' to ϕ'' , introduce at most 8 clauses for each clause in ϕ'.
  - From ϕ'' to final 3-CNF, introduce at most 4 clauses for each clause in ϕ''.

NP-completeness proof structure

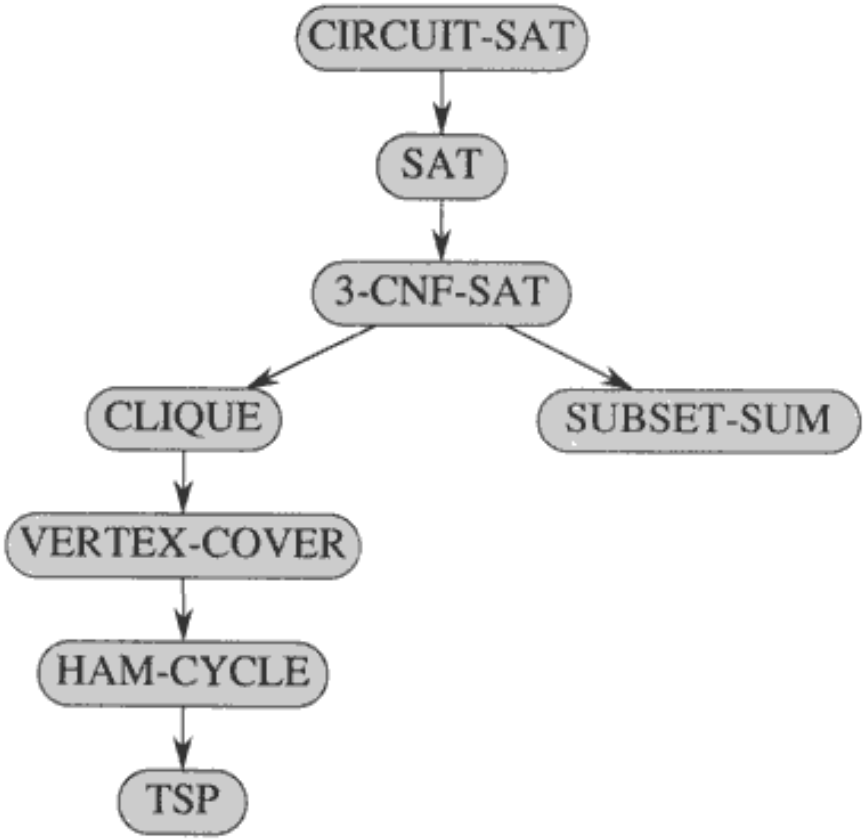


Figure 34.13 The structure of NP-completeness proofs in Sections 34.4 and 34.5. All proofs ultimately follow by reduction from the NP-completeness of CIRCUIT-SAT.

NPC proof -- CLIQUE

- Definition: a clique in an undirected graph G=(V,E) is a subset $V' \subseteq V$ of vertices, each pair of which is connected by an edge in E, i.e., a clique is a complete subgraph of G.
- Size of a clique is the number of vertices in the clique.
- Optimization problem: find the maximum clique.
- Decision problem: whether a clique of given size *k* exists in the graph?
- CLIQUE={<G,*k*>: G is a graph with a clique of size *k*.}
- Intuitive solution: ???

CLIQUE is NP-complete

- *Theorem 34.11*: (page 1003)
  - CLIQUE problem is NP-complete.
- Proof:
  - CLIUEQE $\in$NP: given G=(V,E) and a set V'$\subseteq$V as a certificate for G. The verifying algorithm checks for each pair of $u,v \in$V', whether $\langle u,v \rangle \in$E. time: $O(|V'|^2|E|)$.
  - CLIQUE is NP-hard:
    - show 3-CNF-SAT $\leq_p$CLUQUE.
    - The result is surprising, since from boolean formula to graph.

CLIQUE is NP-complete

- Reduction from 3-CNF-SAT to CLUQUE.
  - Suppose $\phi = C_1 \wedge C_2 \wedge \ldots \wedge C_k$ be a boolean formula in 3-CNF with k clauses.
  - We construct a graph G=(V,E) as follows:
    - For each clause $C_r = (l_1^r \vee l_2^r \vee l_3^r)$, place a triple of v_1^r, v_2^r, v_3^r into V
    - Put the edge between two vertices v_i^r and v_j^s when:
      - $r \neq s$, that is v_i^r and v_j^s are in different triples, and
      - Their corresponding literals are consistent, i.e, l_i^r is not negation of l_j^s .
  - Then ϕ is satisfiable if and only if G has a clique of size k.

$\phi=(x_1\vee\neg x_2\vee\neg x_3)\wedge(\neg x_1\vee x_2\vee x_3)\wedge(x_1\vee x_2\vee x_3)$ and its reduced graph G

$C_1=x_1\vee\neg x_2\vee\neg x_3$

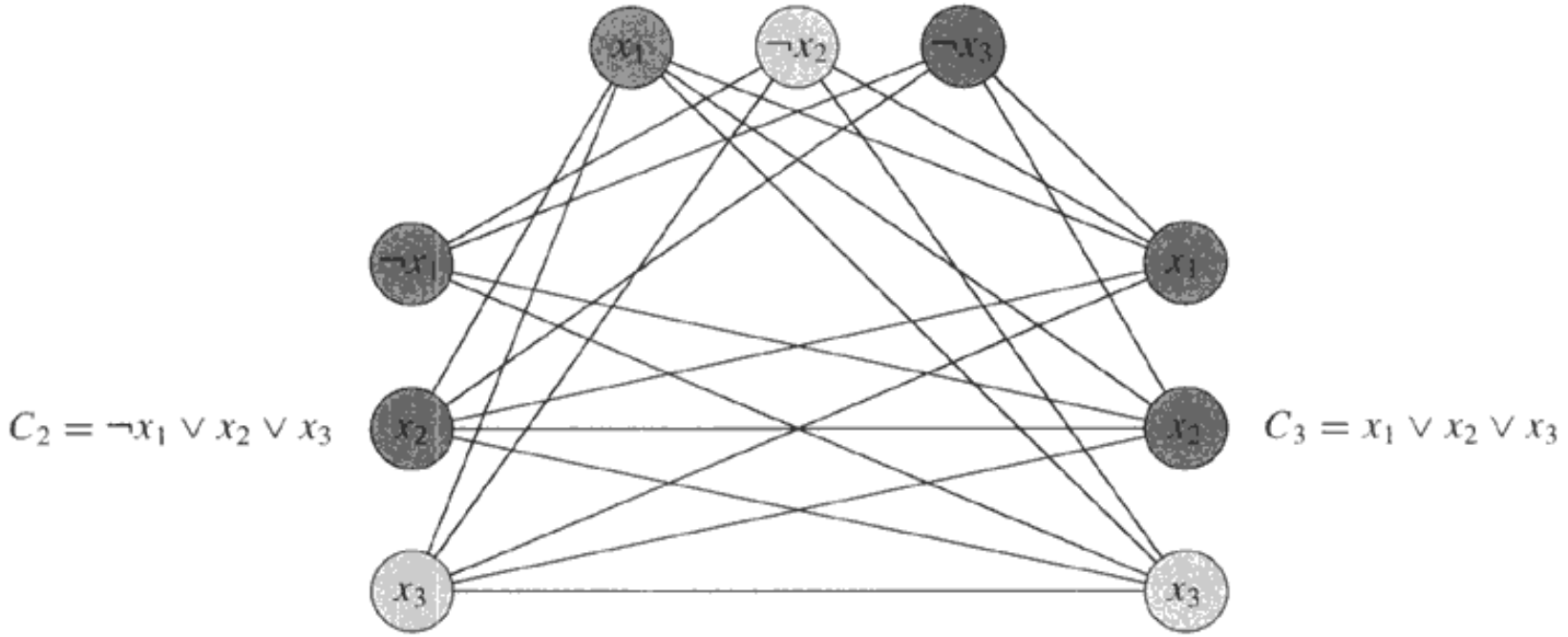


Figure 34.14 The graph G derived from the 3-CNF formula $\phi = C_1 \wedge C_2 \wedge C_3$, where $C_1 = (x_1 \vee \neg x_2 \vee \neg x_3)$, $C_2 = (\neg x_1 \vee x_2 \vee x_3)$, and $C_3 = (x_1 \vee x_2 \vee x_3)$, in reducing 3-CNF-SAT to CLIQUE. A satisfying assignment of the formula has $x_2 = 0$, $x_3 = 1$, and x_1 may be either 0 or 1. This assignment satisfies C_1 with $\neg x_2$, and it satisfies C_2 and C_3 with x_3, corresponding to the clique with lightly shaded vertices.

CLIQUE is NP-complete

- Prove the above reduction is correct:
  - If ϕ is satisfiable, then there exists a satisfying assignment, which makes at least one literal in each clause to evaluate to 1. Pick one this kind of literal in each clause. Then consider the subgraph V' consisting of the corresponding vertex of each such literal. For each pair $v_i^r, v_j^s \in$ V', where $r \neq s$. Since l_i^r, l_j^s are both evaluated to 1, so l_i^r is not negation of l_j^s, thus there is an edge between v_i^r and v_j^s. So V' is a clique of size k.
  - If G has a clique V' of size k, then V' contains exact one vertex from each triple. Assign all the literals corresponding to the vertices in V' to 1, and other literals to 1 or 0, then each clause will be evaluated to 1. So ϕ is satisfiable.
- It is easy to see the reduction is in poly time.
- The reduction of an instance of one problem to a specific instance of the other problem.

Traveling-salesman problem is NPC

- TSP={<G,*c*,*k*>:

  G=(V,E) is a complete graph,

  c is a function from $V \times V \rightarrow Z$,

  $k \in Z$, and G has a traveling salesman tour with cost at most *k*.}

- *Theorem 34.14*: (page 1012)
  - TSP is NP-complete.

TSP is NP-complete

- TSP belongs to NP:
  - Given a certificate of a sequence of vertices in the tour, the verifying algorithm checks whether each vertex appears once, sums up the cost and checks whether at most *k*. in poly time.
- TSP is NP-hard (show HAM-CYCLE$\leq_p$TSP)
  - Given an instance G=(V,E) of HAM-CYCLE, construct a TSP instance <G',c,0) as follows (in poly time):
    - G'=(V,E'), where E'={<*i*,*j*>: *i*,*j*$\in$ V and *i*$\neq$*j*} and
    - Cost function *c* is defined as *c*(*i*,*j*)=0 if (*i*,*j*) $\in$E, 1, otherwise.
  - If G has a hamiltonian cycle *h*, then *h* is also a tour in G' with cost at most 0.
  - If G' has a tour *h*' of cost at most 0, then each edge in *h*' is 0, so each edge belong to E, so *h*' is also a hilmitonian cycle in G.

Subset Sum is NPC

- SUNSET-SUM={<S,*t*>: S is a set of integers and there exists a S'$\subseteq$S such that $t=\sum_{s\in S'} s$.}
- *Theorem 34.15*: (page 1014)
  - SUBSET-SUM is NP-complete.

SUBSET-SUM is NPC

- SUBSET-SUM belongs to NP.
  - Given a certificate S', check whether *t* is sum of S' can be finished in poly time.
- SUBSET-SUM is NP-hard (show 3-CNF-SAT$\leq_p$SUBSET-SUM).

SUBSET-SUM is NPC

- Given a 3-CNF formula $\phi = C_1 \wedge C_2 \wedge \ldots \wedge C_k$ with literals $x_1, x_2, \ldots, x_n$. Construct a SUBSET-SUM instance as follows:
  - Two assumptions: no clause contains both a literal and its negation, and either a literal or its negation appears in at least one clause.
  - The numbers in S are based on 10 and have $n+k$ digits, each digit corresponds to (or is labeled by) a literal or a clause.
  - Target t=1…1||4…4 (n 1's and k 4's)
  - For each literal x_i, create two integers:
    - $v_i = 0\ldots01_{(i)}0\ldots0||0\ldots01_{(l)}0\ldots01_{(w)}0\ldots0$, where x_i appears in $C_l,\ldots,C_w$.
    - $v_i' = 0\ldots01_{(i)}0\ldots0||0\ldots1_{(m)}0\ldots\ldots01_{(p)}0\ldots0$, where $\neg x_i$ appears in $C_m,\ldots,C_p$.
    - Clearly, v_i and v_i' can not be equal in right k digits, moreover all v_i and v_i' in S are distinct.
  - For each clause C_j, create two integers:
    - $s_j = 0\ldots0||0\ldots01_{(j)}0\ldots0$,
    - $s_j' = 0\ldots0||0\ldots02_{(j)}0\ldots0$.
    - all s_j and s_j' are called "slack number". Clearly, all s_j and s_j' in S are distinct.
  - Note: the sum of digits in any one digit position is 2 or 6, so when there is no carries when adding any subset of the above integers.

SUBSET-SUM is NPC

- The above reduction is done in poly time.
- The 3-CNF formula ϕ is satisfiable if and only if there is a subset S' whose sum is *t*.
  - suppose ϕ has a satisfying assignment.
    - Then for i=1,…,n, if x_i=1 in the assignment, then v_i is put in S', otherwise, then v_i' is put in S'.
    - The digits labeled by literals will sum to 1.
    - Moreover, for each digit labeled by a clause C_j and in its three literals, there may be 1, 2, or 3 assignments to be 1. correspondingly, both s_j and s_j' or s_j', or s_j is added to S' to make the sum of the digit to 4.
    - So S' will sum to 1…14…4.
  - Suppose there is a S' which sums to 1…14…4. then S' contains exact one of v_i and v_i' for i=1,…,n. if $v_i \in$ S', then set x_i=1, otherwise, $v_i' \in$ S', then set x_i=0. It can be seen that this assignment makes each clause of ϕ to evaluate to 1. so ϕ is satisfiable.

L
P
U

Thank You !!!