LEMMA

$SAT \leq_p 3SAT$

PROOF

- We want to come up with a polytime computable function $f : \{0,1\}^* \to \{0,1\}^*$ such that:
  - → for every CNF F, $f(\llcorner F \lrcorner)$ is a 3CNF.
  - → F is satisfiable iff $f(\llcorner F \lrcorner)$ is satisfiable
- The idea can be explained as follows:

CLAUSE, WITH ARBITRARILY MANY LITERALS

$$F = C_1 \wedge C_2 \wedge \dots \wedge C_n$$

$$f(\llcorner F \lrcorner) = C_1^1 \wedge \dots \wedge C_1^{n_1} \wedge C_2^1 \wedge \dots \wedge C_2^{n_2}$$

CLAUSES WITH AT MOST 3 VARIABLES

$$C_j \Leftrightarrow C_j^1 \wedge \dots C_j^{n_1}$$

- Now, let's see how a clause C with arbitrarily many literals can be translated into $C^1 \wedge \dots \wedge C^n$ where C^i only contains three literals.

$$C = l_1 \vee l_2 \vee l_3 \vee l_4 \vee l_5 \vee l_6$$

$$\Big[\underbrace{(\ell_1 \vee \ell_2 \vee z_1)}_{C^1} \wedge \underbrace{(\ell_3 \vee \neg z_1 \vee z_2)}_{C^2} \wedge \underbrace{(\ell_4 \vee \neg z_2 \vee z_3)}_{C^3} \wedge \underbrace{(\ell_5 \vee \ell_6 \vee \neg z_3)}_{C^4}$$

IF z_1 IS TRUE THEN
EITHER ℓ_3 OR z_2

IF z_1 IS FALSE THEN
EITHER ℓ_1 OR ℓ_2

IF z_2 IS TRUE THEN
EITHER ℓ_4 OR z_3

IF z_2 IS FALSE THEN
EITHER ℓ_3 OR $\neg z_1$

$\vdots$

C IS SATISFIABLE

$\Updownarrow$

$C^1 \wedge C^2 \wedge C^3 \wedge C^4$ IS SATISFIABLE

- WE can generalize this idea to the case in which C is a clause with arbitrarily many literals
- The nice thing is that computing $f(\llcorner F \lrcorner)$ from $\llcorner F \lrcorner$ is relatively easy: you just need to substitute any clause C_j in F with $C_j^1 \wedge \ldots \wedge C_j^{n_j}$ where C_j^k contains some of the literals in C_j plus some auxiliary variables.

THEOREM

INDSET IS NP-COMPLETE

PROOF

- We know that

$$\text{INDSET} = \{(G, k) \mid \text{THE UNDIRECTED GRAPH } G \text{ HAS AN INDEPENDENT SET OF SIZE AT LEAST } k\}$$

- We already know that $\boxed{\text{INDSET} \in NP}$, because we know that the certificate can be taken as a subset which has cardinality at least k, and which is independent
- We then want to show that $L \leq_p \text{INDSET}$ where L is any NP-complete problem. It is quite convenient to take $L = 3SAT$.
- We want to construct a polytime computable function f such that

$\Rightarrow f(\llcorner F \lrcorner)$ where F is any CNF is an instance of INDSET

$\Rightarrow$ F is satisfiable iff $f(\llcorner F \lrcorner) \in \text{INDSET}$

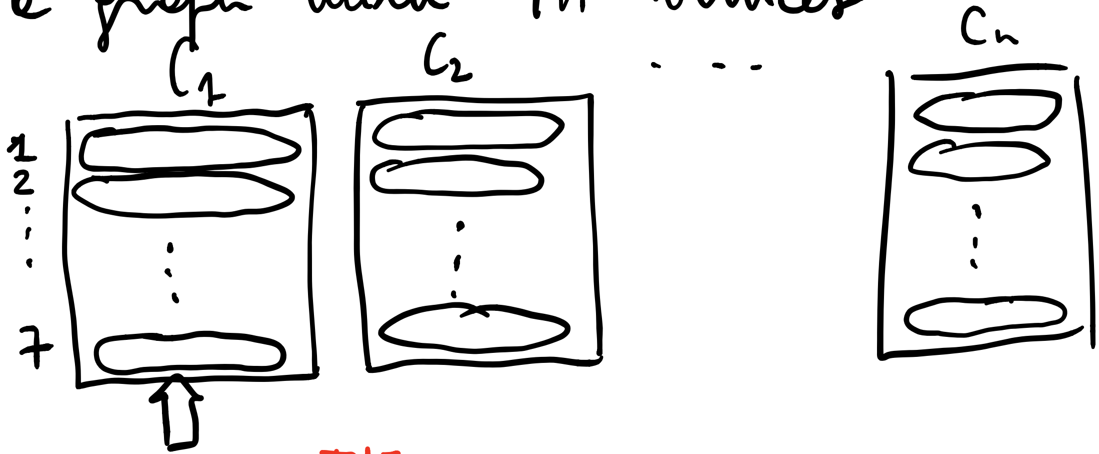
Let's describe f as follows:

$\rightarrow$ Its input is a CNF formula

$$F = C_1 \wedge C_2 \wedge \dots \wedge C_n$$

where C_i is a clause with at most three literals.

→ The idea is to turn F into a graph with $7n$ vertices

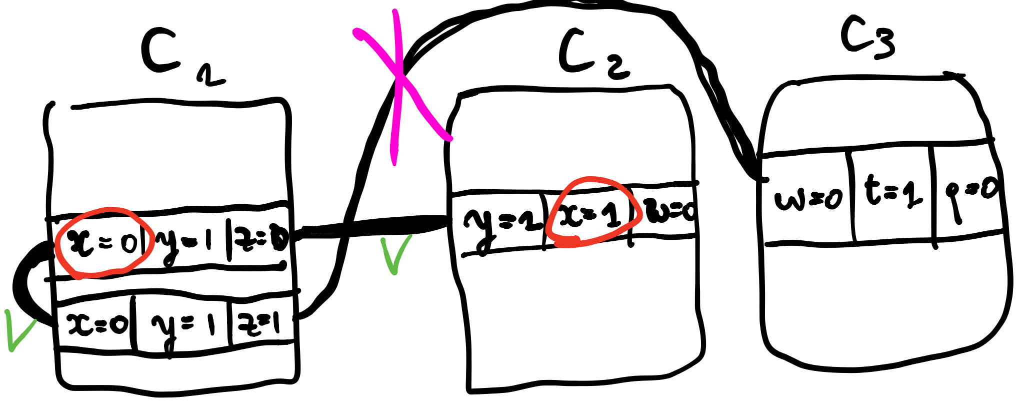


EACH OF THE NODES HERE CORRESPONDS TO AN ASSIGNMENT TO THE VARIABLES OCCURRING IN C_1 WHICH IS CONSISTENT WITH C_1, E.G. IF $C_1 = \neg x \vee \neg y \vee z$ THE ASSIGNMENTS ARE ALL THE ASSIGNMENTS TO x, y, z, EXCEPT THE ONE IN WHIC $x=1$ $y=1$ AND $z=0$

→ The edges of the graph are defined so as to reflect inconsistencies between assignments, that is to say:

→ all assignments in the same clause are inconsistent, of course

→ all assignments in distinct clauses but which mention the same variable inconsistently, are of course incompatible, e.g.



- ==Solving== INDSET on the resulting graph and on the natural number n (i.e. the number of clauses) amounts to solving 3SAT on the CNF we started from, because:
  - → The subset we get in INDSET must have cardinality at least n, meaning that it consists of assignments from <u>each</u> clause
  - → If you have a subset of consistent assignments, this can be easily turned to a proper assignment for <u>all</u> the variables occurring in F.