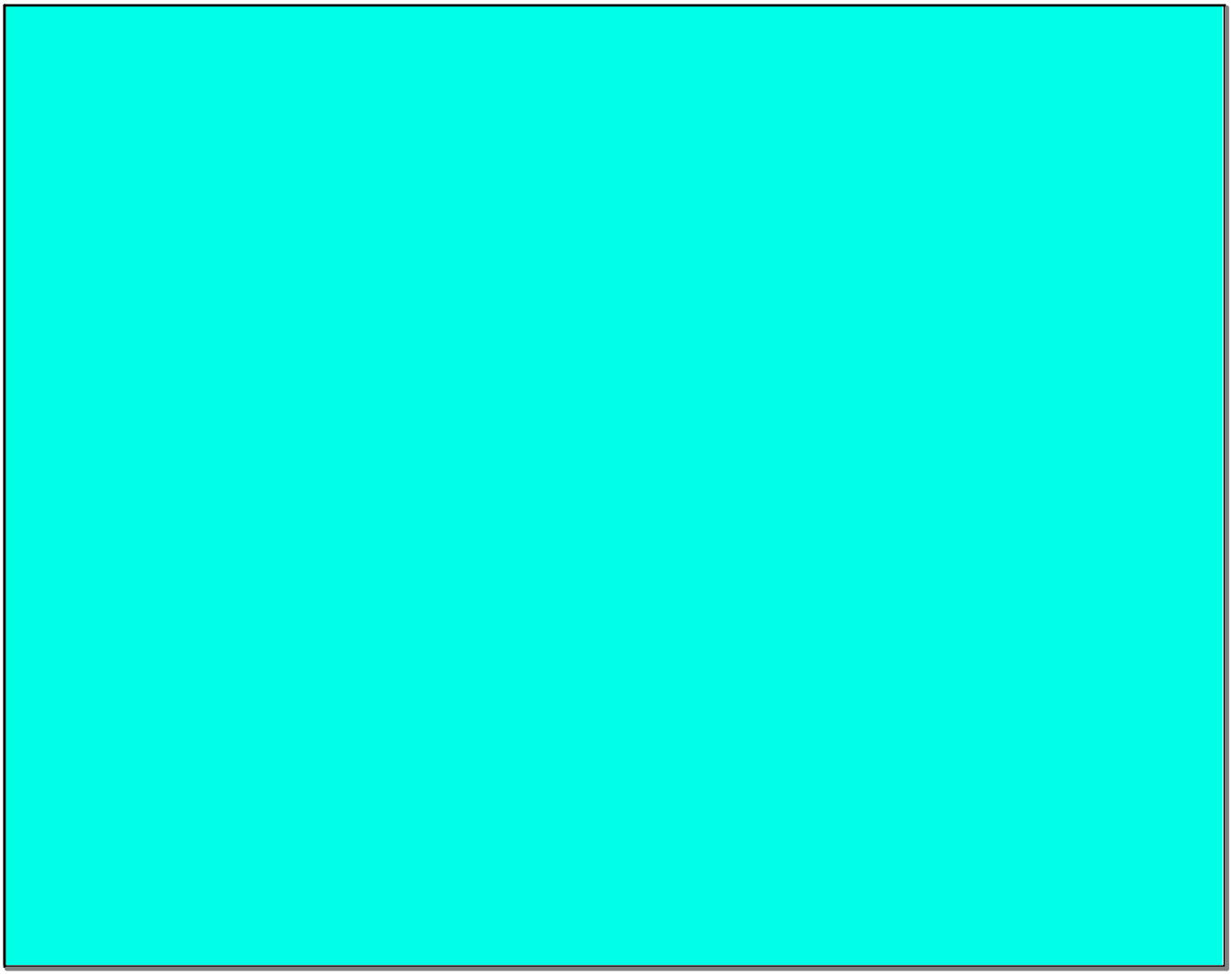


Def: A formula (Boolean) φ is in conjunctive normal form (cnf) if φ has the following syntax:

$\varphi \equiv C_1 \wedge \cdots \wedge C_k$ ($C_1, \ldots, C_k$ are clauses),

each clause is a disjunction

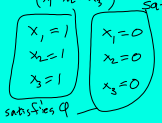
$C \equiv \ell_1 \vee \cdots \vee \ell_m$

of literals ($\ell_1, \ldots, \ell_m$ are literals),

where a literal is either a variable (x, say) or its negation $\bar{x}$ (or $\neg x$).

Ex:

$\varphi \equiv (x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee x_2) \wedge (x_1 \vee x_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_2 \vee x_3)$ satisfiable?



$\varphi = (x \vee y) \wedge (x \vee \bar{y}) \wedge (\bar{x} \vee y) \wedge (\bar{x} \vee \bar{y})$ is not satisfiable.

Def: CNF-SAT is the following decision problem:

Instance: A Boolean fmla φ in cnf.

Question: Is φ satisfiable?

Theorem: CNF-SAT is NP-complete.

(Cook-Levin theorem)

Proof: CNF-SAT $\in$ NP.

Clear (restriction of SAT, which is in NP).

WTS: CNF-SAT is NP-hard.

Fix any any NP language A.

Fix a verifier V that verifies yes-instances of A in p-time.

Fix a polynomial $p(n)$ that bounds the running time of V (n = length of the input, not including the proof string) and the length of the proof string + length of the input (n).

On any string w and proof string y,

$V(\langle w, y \rangle)$ accepts iff y is a valid proof that $w \in A$.

V runs in time $p(|w|)$ and $|y| + |w| \leq p(|w|)$

Need a p-time function f that takes an input string w ($n := |w|$) and outputs a cnf formula φ_w such that $\exists y\, V(\langle x, y \rangle)$ accepts iff φ_w is satisfiable.

Then

$A \leq_p$ CNF-SAT via f.

($w \in A$ iff φ_w is sat).

Construction uses formal aspects of V (standard 1-tape TM).

$V = \langle Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{reject} \rangle$

Technical assumptions (w.l.o.g.):

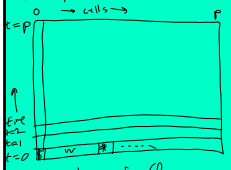
- A symbol $\$ \in \Gamma \setminus \Sigma$ that is initially in cell 0, never written in any other cell, never changed in cell 0, and V does not have a left transition scanning $\$$.

- let $p := p(n)$.
- Initial input is

$\underbrace{\$w\#y\sqcup\cdots}_{p}$

for any $y \in \Gamma$.

$\# \in \Gamma \setminus \Sigma$.



Construction of Q_w:

Variables:

For each state $q \in Q$, $0 \le t \le p$ have variable

$S_{q,t} \equiv$ "V is in state q at time t"

For every $0 \le i \le p$ and $0 \le t \le p$, variable

$h_{i,t} \equiv$ "V's head is scanning cell i at time t"

For every $a \in \Gamma$, $0 \le i \le p$, $0 \le t \le p$, variable

$C_{a,i,t} \equiv$ "a is in cell i at time t".

Clauses of Q_w:

Sanity clauses:

V can't be in 2 states at once. For every $q, r \in Q$ with $q \neq r$, and for all $0 \le t \le p$ add the clause

$$\left(\overline{S_{q,t}} \vee \overline{S_{r,t}}\right)$$

Head can't be in 2 places at once: $\forall\, 0 \le i < j \le p$, $0 \le t \le p$, clause

$$\left(\overline{h_{i,t}} \vee \overline{h_{j,t}}\right)$$

A cell can only hold one symbol at a time. $\forall a, b \in \Gamma$ with $a \neq b$, $\forall 0 \le i \le p$, $\forall 0 \le t \le p$, clause

$$\left(\overline{C_{a,i,t}} \vee \overline{C_{b,i,t}}\right)$$

Initial conditions:

Clauses

$S_{q_0,0}$ (V is in state q_0 at time t=0. 1-literal clause)

$h_{0,0}$ (head scans cell 0 at time 0)

$C_{\$,0,0}$ ($\$$ is in cell 0 initially)

Let $w = w_1 \cdots w_n$ $(w_i \in \Sigma)$

For $1 \le i \le n$, add clause

$C_{w_i,i,0}$ (cell i has w_i for $1 \le i \le n$ initially)

$C_{\#,n+1,0}$ (- - -)

For each i, $n+2 \le i \le p$, clause

$\bigvee_{a \in \Gamma} C_{a,i,0}$ (cells beyond the # have something in them)

Say that V accepts.

(assume wlog, that q_{acc}, q_{rej} are not halting states but if V enters one of them it never leaves.)

Add the clause

$S_{q_{acc},p}$ (V is in state q_{acc} at time p)

Transition clauses:
They all look like this:
"If something holds at time t, then something holds at time $t+1$."
For every $0 \leq t < p$, $0 \leq i \leq p$, $\forall a \in \Gamma$,

$$\left(C_{a,i,t} \wedge \overline{h_{i,t}} \rightarrow C_{a,i,t+1} \right)$$

not a clause, but equiv to:

$$\left(\overline{C_{a,i,t}} \vee h_{i,t} \vee C_{a,i,t+1} \right)$$

Why?

$$(p \wedge q \rightarrow r)$$
$$\equiv \left(\overline{p \wedge q} \vee r \right)$$
$$\equiv \left(\bar{p} \vee \bar{q} \vee r \right)$$

(unscanned cells don't change from time t to $t+1$.)