Introduction to Analysis of Algorithms	The Cook–Levin Construction
CS 4820 Spring 2020	Monday, November 30, 2020

Here is an outline of the Cook–Levin construction that shows that SAT is NP-hard.

Given an arbitrary nondeterministic polynomial-time TM $M = (Q, \Sigma, \Gamma, \vdash, \sqcup, \delta, s, t, r)$ and string $x \in \Sigma^*$, we wish to construct a Boolean formula φ that is satisfiable iff M accepts x. This construction reduces the set $L(M) \in \text{NP}$ to SAT. Recall that the type of the transition function for nondeterministic TMs is $\delta : Q \times \Gamma \to \mathcal{P}(Q \times \Gamma \times \{L, R\})$, where $\mathcal{P}(-)$ denotes the powerset operator. Intuitively, if $(q, b, D) \in \delta(p, a)$, then when in state p reading symbol a, it can print b, move its head in direction D, and enter state q. The set $\delta(p, a)$ determines the set of possible next moves of the machine.

Suppose M runs in time $N = n^k$. Thus all paths in the computation tree of M on inputs of length $n \geq 2$ are of length at most n^k. Our formula will use the following Boolean variables with their intuitive meanings:

- P^a_{ij}, $0 \leq i, j \leq N$, $a \in \Gamma$.
  "The symbol occupying tape cell j at time i is a."
- Q^q_{ij}, $0 \leq i, j \leq N$, $q \in Q$.
  "The machine is in state q at time i scanning tape cell j."

We need to write down constraints in the form of Boolean formulas that describe an accepting computation of M on input x. There will be an accepting computation iff there is a truth assignment that satisfies the conjunction of all the constraints.

First we include clauses that ensure that for each time i, $0 \leq i \leq N$, the values of the variables P^a_{ij} and Q^q_{ij} specify a unique configuration of the machine; that is, there is exactly one symbol on each tape cell j at time i, and the machine is scanning exactly one tape cell j in exactly one state $q \in Q$ at time i.

- "There is exactly one symbol on each tape cell j at time i."

$$\bigwedge_{j=0}^{N} \bigvee_{a \in \Gamma} (P^a_{ij} \wedge \bigwedge_{b \in \Gamma,\ b \neq a} \neg P^b_{ij})$$

  for $0 \leq i \leq N$. This says that for all j, there exists $a \in \Gamma$ such that a occupies tape cell j, and no other symbol besides a occupies tape cell j.

- "The machine is scanning exactly one tape cell j in exactly one state $q \in Q$ at time i."

$$\bigvee_{j=0}^{N} (\bigvee_{q \in Q} (Q^q_{ij} \wedge \bigwedge_{p \in Q,\ p \neq q} \neg Q^p_{ij}) \wedge \bigwedge_{k \neq j} \bigwedge_{q \in Q} \neg Q^q_{ij})$$

  for $0 \leq i \leq N$. This says that there exists j and $q \in Q$ such that the machine is scanning cell j in state q and no other state, and for all cells $k \neq j$, the machine is not scanning cell k in any state.

Now we include clauses that say that the machine starts correctly on input x, runs correctly, and accepts. Suppose $x = x_1 x_2 \cdots x_n$, $x_j \in \Sigma$.

- "The machine starts correctly on input x."

$$Q^s_{00} \wedge P^{\vdash}_{00} \wedge \bigwedge_{j=1}^{n} P^{x_j}_{0j} \wedge \bigwedge_{j=n+1}^{N} P^{\sqcup}_{0j}$$

  This says that the machine starts in the start state s scanning the left endmarker and that the tape initially contains the input string $x = x_1, \ldots, x_n$ to the right of the endmarker and padded out to distance N by blanks $\sqcup$. Thus the values of P^a_{0j} and Q^q_{0j} specify the correct start configuration of M on input x.

- "The machine accepts."

$$\bigvee_{j=0}^{N} Q_{Nj}^t$$

This just says that at time N, the machine is in its accept state scanning some tape cell.

The final clauses ensure that the configuration at time $i+1$ follows by the transition rules of the machine from the configuration at time i. This means that the correct symbol is printed on the cell that the machine is scanning at time i, the head moves in the proper direction, and the machine enters the correct next state. Moreover, all other symbols on the tape are preserved from time i to time $i+1$.

- "The machine runs correctly."

$$P_{ij}^a \wedge Q_{ij}^p \Rightarrow \bigvee_{(q,b,L)\in\delta(p,a)} (P_{i+1,j}^b \wedge Q_{i+1,j-1}^q) \vee \bigvee_{(q,b,R)\in\delta(p,a)} (P_{i+1,j}^b \wedge Q_{i+1,j+1}^q)$$

for all $0 \le i \le N-1$, $0 \le j \le N$, $a \in \Gamma$, and $p \in Q$. This says that whenever the machine is scanning tape cell j in state p and the current symbol occupying cell j is a, then in the next step, the contents of cell j are updated correctly, the head moves in the proper direction, and the machine enters the correct next state according to some possible next move of the machine as given by $\delta(p,a)$. The disjunction on the right-hand side is over all possible nondeterministic choices that the machine could make (recall that the type of δ for nondeterministic machines is $\delta : Q \times \Gamma \to \mathcal{P}(Q \times \Gamma \times \{L,R\})$).

- "The symbol on tape cell j does not change from time i to $i+1$ if the machine is not scanning cell j at time i."

$$(P_{ij}^a \wedge \bigwedge_{q\in Q} \neg Q_{ij}^q) \Rightarrow P_{i+1,j}^a$$

for all $0 \le i \le N-1$, $0 \le j \le N$, and $a \in \Gamma$. This says that if the symbol on tape cell j is a at time i, and if the machine is not scanning tape cell j at time i, then the symbol on tape cell j is still a at time $i+1$.

The conjunction of all these clauses is our formula φ. If there is an accepting computation of M on input x, then setting the values of P_{ij}^a and Q_{ij}^q according to the tape contents and state of the finite control at time i and cell j gives a truth assignment satisfying φ. Conversely, a satisfying assignment to φ has exactly one P_{ij}^a true for each i,j and exactly one Q_{ij}^q true for each i, and this determines an accepting computation of M on input x since all constraints are satisfied.

The size of φ is quadratic in the running time of M (that is, if M runs in time n^k, then $|\varphi|$ is $O(n^{2k})$), and φ can be produced in quadratic time from the description of M and x.