The Cook-Levin Theorem

Exposition by William Gasarch—U of MD

Variants of SAT

Definition

1. SAT is the set of all boolean formulas that are satisfiable. That is, $\phi(\vec{x}) \in SAT$ if there exists a vector $\vec{b}$ such that $\phi(\vec{b}) = TRUE$.
2. CNFSAT is the set of all boolean formulas in SAT of the form $C_1 \wedge \cdots \wedge C_m$ where each C_i is an $\vee$ of literals.
3. k-SAT is the set of all boolean formulas in SAT of the form $C_1 \wedge \cdots \wedge C_m$ where each C_i is an $\vee$ of exactly k literals.
4. DNFSAT is the set of all boolean formulas in SAT of the form $C_1 \vee \cdots \vee C_m$ where each C_i is an $\wedge$ of literals.
5. k-DNFSAT is the set of all boolean formulas in SAT of the form $C_1 \vee \cdots \vee C_m$ where each C_i is an $\wedge$ of exactly k literals.

Cook-Levin Theorem

Theorem

CNFSAT is NP*-complete.*

We need to prove two things:

1. $\mathrm{CNFSAT} \in NP$.

$$\mathrm{CNFSAT} = \{\phi : (\exists \vec{y})[\phi(\vec{y}) = T]\}$$

   Formally

$$B = \{(\phi, \vec{y}) : \phi(\vec{y}) = T\}$$

   The satisfying assignment is the witness.

2. For all $X \in \mathrm{NP}$, $X \leq \mathrm{CNFSAT}$. This is the bulk of the proof.

$X \in \mathrm{NP}$ implies $X \le \mathrm{CNFSAT}$

Let $X \in \mathrm{NP}$. We show that $X \le CNFSAT$.
M be a TM and p, q be polynomials such that

$$X = \{x \mid (\exists y)[|y| = p(|x|) \text{ AND } M(x,y) = 1]\}$$

and $M(x,y)$ runs in time $q(|x| + |y|)$.
Let $M = (Q, \Sigma, \delta, q_0, h)$
The machine itself has a tape. Example:

$$\#abba\#ab@ab\#a$$

(Everything to the right that is not seen is a $\#$. Our convention is that you CANNOT go off to the left— from the left most symbol you can't go left.)
If the machine is in state q and the head is looking at (say) the @ sign we denote this by:

$$\#abba\#ab(@, q)a$$

We extend the alphabet and allow symbols $\Sigma \times Q$. The symbol