A Hard Input Distribution

Chris Calabro

September 7, 2004

An input distribution is a map $D : \Sigma^* \rightarrow \mathbb{R}$ such that $\forall x \in \Sigma^* \ D(x) \geq 0$ and $\forall n \in \omega \ \sum_{x \in \Sigma^n} D(x) = 1$. Let $L$ be any NP-complete language, say SAT. We show, assuming $P \neq \text{NP}$,

$$P \neq \text{NP},$$

that there is a single input distribution $D$ on which no deterministic algorithm solving $L$ runs in polytime.

Let $\mathcal{A} = \{A_1, A_2, \ldots\}$ be the set of deterministic algorithms solving $L$. From (1), we know that

$$\forall A \in \mathcal{A}, n, k \in \omega \ \exists x \in \Sigma^*, |x| \geq n \ \text{time}(A(x)) \geq |x|^k.$$  

(2)

Let $f : \omega \rightarrow \omega^2$ be a bijection, and let $f_1, f_2$ be the components of $f$ so that $\forall n \in \omega \ f(n) = (f_1(n), f_2(n))$. Define recursively the sequence $x_0, x_1, \ldots \in \Sigma^*$ so that $x_0 = \epsilon$ and $\forall n \geq 1, x_n$ is the lexicographically smallest $x \in \Sigma^*$ such that $|x| > |x_{n-1}|$ and time($A_{f_1(n)}(x)$) > $|x|^n$. That such an $x$ exists is assured by (2).

Then $x_1, x_2, \ldots$ forms a sequence of strings of strictly increasing size and such that $x_n$ is hard for algorithm $A_{f_1(n)}$. But $\forall i \in \omega \ |f_1^{-1}(i)| = \infty$, and so for each $A \in \mathcal{A}$, an infinite number of the $x_i$’s are hard for $A$.

We can now choose our distribution $D$ so that $D(x)$ is 1 if $x \in \{x_1, x_2, \ldots\}$, 0 if $x \notin \{x_1, x_2, \ldots\}$ and $|x| \in \{|x_1|, |x_2|, \ldots\}$, and $D(x)$ may be defined as an arbitrary distribution on input sizes other than those in $\{|x_1|, |x_2|, \ldots\}$. Notice that $D$ is probably not sampleable by any algorithm whatsoever.