Time and Space

Definition 1 *Let M be a TM and $x \in \Sigma^*$ be an input. We say that M uses time T on input x if M halts within T steps on x.*

Definition 2 *Let M be a TM with a read-only input tape, and $x \in \Sigma^*$. We say that M uses space S on an input x if M halts on input x and accesses at most S work-tape cells.*

Definition 3 *Let $T : \mathbb{N} \to \mathbb{N}$ be a function. We say that a TM M runs in time T (or space S) if for all $x \in \Sigma^*$, M uses time $T(|X|)$ (or space $S(|X|)$).*

Definition 4 *Let $L \subseteq \Sigma^*$, $T : \mathbb{N} \to \mathbb{N}$ ($S : \mathbb{N} \to \mathbb{N}$). $L \in \mathrm{DTIME}(T)$ ($L \in \mathrm{DSPACE}(S)$) if there exists a TM M deciding L running in time $O(T)$ (space $O(S)$).*

This is a worst case notion, however in certain cases it is more useful to consider a notion of average case complexity. This requires that we have the notion of a distribution over inputs, and while of important practical use, has limited theoretical use due to not entailing any strict bound.

We define the following complexity classes:

$$\mathrm{P} = \bigcup_{k=0}^{\infty} \mathrm{DTIME}(n^k)$$
$$\mathrm{E} = \bigcup_{k=0}^{\infty} \mathrm{DTIME}\left(2^{nk}\right)$$
$$\mathrm{EXP} = \bigcup_{k=0}^{\infty} \mathrm{DTIME}\left(2^{n^k}\right)$$

The extended Church-Turing thesis claims that if a problem can be solved in time T on a "reasonable" model, then it is in $\bigcup_k \mathrm{DTIME}(T^k)$.

Note that we don't particularly care about DTIME(1), because we will always need n steps to read the input. We *do* care about low bounds like DSPACE(1), because it's possible to in certain cases construct TMs which don't require much storage (e.g. for regular languages).

Theorem 1 (Speed-up theorem) *Let $T : \mathbb{N} \to \mathbb{N}$ be a time function. Let $n \in \mathbb{N}$ be a positive integer. If there is a TM M deciding $L \subseteq \Sigma^*$ in T steps, then there is a TM M' deciding L in $T/n + n + 3$ steps.*

To do this, we write each 3 symbols in the input as a single one, expanding Γ to do so. We then expand δ to work with Γ', and with a larger state set we can process and determine $6n$ steps in 6 transitions. (**Consider looking over this again.**)

Theorem 2 (Gap theorem) *Let $f : \mathbb{N} \to \mathbb{N}$ be a computable function. Then there exists a time bound T, space bound S such that $\mathrm{DTIME}(T) = \mathrm{DTIME}(f(T))$ and $\mathrm{DSPACE}(S) = \mathrm{DSPACE}(f(S))$.*

Definition 5 *A time bound $T : \mathbb{N} \to \mathbb{N}$ is time-constructible if there is a TM M that computes $T(n)$ given 1^n as input in time $O(T)$.*

The analogous definition works the same for space-constructibility – we just require that there is a TM computing $S(n)$ from 1^n in space $O(S)$.

Theorem 3 (Time hierarchy) *Let $T, T' : \mathbb{N} \to \mathbb{N}$ such that T' is time-constructible and $T \log T = o(T')$. Then $\mathrm{DTIME}(T) \subset \mathrm{DTIME}(T')$.*

Theorem 4 *Let $S, S' : \mathbb{N} \to \mathbb{N}$ such that S' is space-constructible and $S \log S = o(S')$. Then $\mathrm{DSPACE}(S) \subset \mathrm{DSPACE}(T')$.*

Something something reductions

Theorem 5 (Cook-Levin) *Both* SAT *and SAT*

Polynomial Hierarchies

Note that NP is the set of languages L such that for some polynomial p there is a language R such that

$$L = \{x \in \Sigma^* : \exists w \in \Sigma^* \,.\, |w| \leq p(|x|) \wedge \langle x, w\rangle \in R\}\,.$$

We can in fact extend this to a more general notion of a hierarchy, that with a class $\mathcal{C}$, we define the classes $\exists^{\mathsf{P}}\mathcal{C}$ and $\forall^{\mathsf{P}}\mathcal{C}$ as follows:

$$\exists^{\mathsf{P}}\mathcal{C} := \{\{x \in \Sigma^* : \exists w \in \Sigma^* \,.\, |w| \leq p(x) \wedge \langle x, w\rangle \in R\} : p \in \mathbb{R}[x], R \in \mathcal{C}\}$$
$$\forall^{\mathsf{P}}\mathcal{C} := \{\{x \in \Sigma^* : \forall w \in \Sigma^* \,.\, |w| \leq p(x) \Rightarrow \langle x, w\rangle \in R\} : p \in \mathbb{R}[x], R \in \mathcal{C}\}.$$

Further, we have that

$$\begin{aligned}\mathsf{co}(\exists^{\mathsf{P}}\mathcal{C}) &= \{L^c : L \in \exists^{\mathsf{P}}\mathcal{C}\}\\ &= \{\{x \in \Sigma^* : \forall w \in \Sigma^* \,.\, |w| \leq p(x) \Rightarrow \langle x, w\rangle \notin R\} : p \in \mathbb{R}[x], R \in \mathcal{C}\}\\ &= \{\{x \in \Sigma^* : \forall w \in \Sigma^* \,.\, |w| \leq p(x) \Rightarrow \langle x, w\rangle \in R'\} : p \in \mathbb{R}[x], R' \in \mathsf{co}\mathcal{C}\}\\ &= \forall^{\mathrm{P}}\mathsf{co}\mathcal{C}\end{aligned}$$

allowing us to write, for example, $\mathsf{coNP} = \forall^{\mathsf{P}}\mathsf{coP} = \forall^{\mathsf{P}}\mathsf{P}$.

Randomness

Randomness can often be useful when writing programs, for a few reasons. Firstly, it is useful to break symmetries, because in certain instances it can become impossible to make a choice if all is equal between two objects. Secondly, it can be useful for hiding or obfuscating information. Thirdly, it can be useful for making computation more efficient.

As an example, while it takes deterministically superquadratic time to verify that $AB = C$, one way to randomly verify this is to select a random vector $x \in \{0,1\}^n$, then compute $ABx = A(Bx)$ in quadratic time, and Cx in quadratic time, then if the results are equal we have that $AB = C$ with probability $1/2$. Iterate this k times and we get probability $1 - 1/2^k$.

Alternatively we might want to solve the connectivity problem in an undirected graph. To do this given trying to check if s and t are in the same connected component, take a random walk from s and terminate either after a fixed number of steps to reject, or after the walk reaches t.

A probabilistic turing machine (PTM) is an NTM where there are at most two choices at any stage in the computation. Given a computation $C_0, C_1, \ldots, C_t$ where for all $i \in \{1, \ldots, t-1\}$, $C_i \vdash C_{i+1}$, we say that C_t has probability $1/2^k$ in the computation if there are exactly k probabilistic transitions. We say that a PTM accepts x with probability p in the expected way (sum of probabilities of accepting over possible computations).

Definition 6 *Given PTM M, time function $t : \mathbb{N} \to \mathbb{N}$, error bound $e : \mathbb{N} \to [0,1]$, we say that M decides L in time t with bounded error ε if M always halts in time $t(|x|)$ and $x \in L$ implies that M accepts x with probability $\geq 1 - \varepsilon(|x|)$, and if $x \notin L$ then M accepts x with probability $\leq \varepsilon(|x|)$.*

We say that an algorithm has one-sided error where $x \in L$ means that M accepts with probability $\geq 1 - \varepsilon(|x|)$, and if $x \notin L$ means M accepts with probability 0, and it has zero error if it always outputs the correct answer when it halts, but halts in time $t(|x|)$ with probability $\geq 1 - \varepsilon(|x|)$.

Definition 7

$$\mathsf{BPP} = \{L : \exists k \text{ s.t. } L \text{ is decided by some PTM with error } 1/3 \text{ in } O(n^k)\}$$
$$\mathsf{RP} = \{L : \exists k \text{ s.t. } L \text{ is decided by some PTM with one-sided error } 1/3 \text{ in } O(n^k)\}$$
$$\mathsf{ZPP} = \{L : \exists k \text{ s.t. } L \text{ is decided by some PTM with zero error } 1/3 \text{ in } O(n^k)\}$$

One of the major open questions regarding the ordering is whether $\mathsf{BPP} = \mathsf{P}$, as well as whether $\mathsf{BPP} \subseteq \mathsf{SUBEXP}$ (the set of languages in $O(2^{n^\varepsilon})$ for all $\varepsilon > 0$).

Theorem 6 *Let $p : \Sigma^* \to [0,1]$ be a function. M is a PTM running in polynomial time iff there is a polynomial time bounded deterministic TM N and a constant k such that for any $x \in \Sigma^*$, if M accepts x with probability $p(x)$ then the probability that $N(x, \pi)$ accepts with $\pi \sim U(\{0,1\}^{n^k})$ is $p(x)$.*

The proof by means of constructing N is relatively straightforward. π just decides the entire sequence of transitions, and it is clear how to construct N polynomial to do this.

Theorem 7 *Suppose E does not have circuits of size $2^{\varepsilon n}$ for almost all n with $\varepsilon > 0$. Then $\mathsf{BPP} = \mathsf{P}$.*

Theorem 8 *The following are equivalent:*

- *$L \in \mathsf{RP}$*
- *For all $k \geq 0$ there is a PTM with one-sided error $\leq 1/2^{n^k}$ deciding L.*
- *There is a $k \geq 0$ such that there is a PTM with one-sided error $\leq 1 - 1/n^k$ deciding L.*

Lemma 9 *Let $x_1, \ldots, x_m$ be IID random variables in $[0,1]$. Suppose $\mu = \mathbb{E}[\sum x_i]$. Then $\mathbb{P}(\sum x_i > \mu + \varepsilon) \leq 2e^{-\frac{t^2}{2m}}$.*