CAS CS 535: Complexity Theory

Lecturer: Mark Bun — Fall 2023

Lecture Notes 19:

#P-Completeness, Permanent, Toda's Theorem

Reading.

- Arora-Barak § 17.2-17.4

Last time: Proof of Valiant-Vazirani, Intro to Counting

Definition 1. The complexity class **FP** consists of all functions $f : \{0,1\}^* \to \mathbb{N}$ such that f can be computed in deterministic polynomial time.

Definition 2. A function $f : \{0,1\}^* \to \mathbb{N}$ is in $\#\mathbf{P}$ (pronounced "sharp-P") if there exists a polynomial-time deterministic TM M and a polynomial $p : \mathbb{N} \to \mathbb{N}$ such that

$$f(x) = \#\{u \in \{0,1\}^{p(|x|)} \mid M(x,u) = 1\}.$$

That is, for every x, the function $f(x)$ counts the number of witnesses u such that $M(x,u)$ accepts.

Examples of problems in $\#\mathbf{P}$:

- # spanning trees
- #CYCLE
- #SAT
- #CKTSAT: Given a circuit C, how many satisfying assignments does it have?
- #INDSET

Lots of problems in statistical physics, AI, etc. involve sampling from an implicitly defined probability distribution, i.e, $\mu(x) = \frac{p(x)}{\sum_y p(y)}$ where $p(x)$ is efficiently computable, but the normalization constant (a.k.a., the partition function) $\sum_y p(y) \in \#\mathbf{P}$, and is often hard to compute.

1 Reductions Between Function/Counting Problems

To study the relationships between different counting problems, we'll study two different ways of poly-time reducing between them.

Option 1: Parsimonious reductions. You should think of these as the function-evaluation analogs of Karp reductions (used to define $\mathbf{NP}$-completeness and so forth).

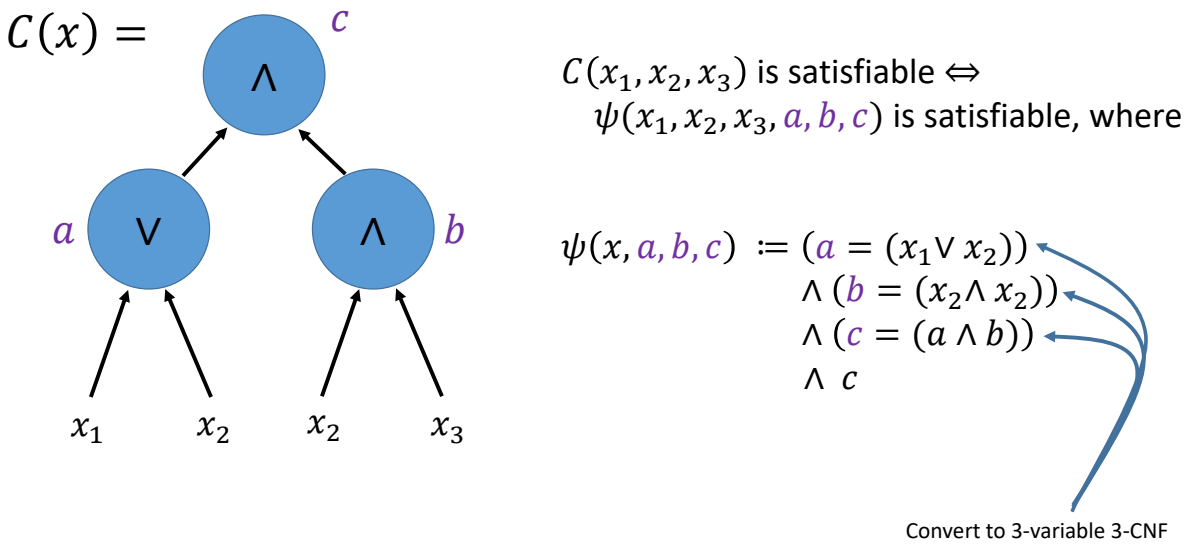
Definition 3. A *parsimonius reduction* R from f to g is a poly-time computable function such that for all $x \in \{0,1\}^*$, we have

$$f(x) = g(R(x)).$$

Interpretation: If computing g is easy, then computing f is also easy.

Example 4. #CKTSAT parsimoniously reduces to #SAT. The reason is that the reduction from the decision version CKTSAT to SAT preserves the number of satisfying assignments.

Proof idea:



Example 5. The Cook-Levin Theorem is parsimonious, so there is a parsimonious reduction from every function $f \in \#\mathbf{P}$ to #SAT.

Theorem 6. #SAT *is* $\#\mathbf{P}$*-complete under parsimonious reductions.*

Option 2: Oracle reductions. These are analogs of Cook reductions. (I.e., language A Cook-reduces to language B if $A \in \mathbf{P}^B$.) Oracle reductions are the more typical way to define $\#\mathbf{P}$-completeness.

We can generalize oracle access to a language to oracle access to a function as follows: We say that TM M has oracle access to function g if it has oracle access to the language $\{\langle x, i\rangle \mid (g(x))_i = 1\}$.

Definition 7. Let $f, g : \{0,1\}^* \to \mathbb{N}$. Then f poly-time oracle reduces to g if $f \in \mathbf{FP}^g$.

Oracle reductions are more powerful than parsimonious reductions. Why should we study them? For one, there are natural examples of counting problems whose complexity is "obviously" equivalent, but which we do not believe are related by efficient parsimonious reductions. For example:

Example 8. $\#(\mathsf{CNF})\mathsf{SAT}$ poly-time reduces to $\#\mathsf{DNFSAT}$. The reduction is as follows.

On input CNF formula $\varphi(x_1, \ldots, x_n)$:

1. Use de Morgan's laws to write $\neg\varphi$ as a DNF ψ.
2. Query $\#\mathsf{DNFSAT}$ on ψ, obtaining a count a.
3. Return $2^n - a$.

Definition 9. A function $g : \{0,1\}^* \to \mathbb{N}$ is $\#\mathbf{P}$-complete if $g \in \#\mathbf{P}$ and every function $f \in \#\mathbf{P}$ poly-time oracle reduces to g.

2 Valiant's Theorem: Permament is $\#\mathrm{P}$-Hard

Let's recall the definition of the determinant of a matrix $A \in \mathbb{R}^{n\times n}$:

$$\det(A) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} A_{i,\sigma(i)}.$$

Here, S_n is the group of permutations over $[n]$, and for a permutation σ,

$$\operatorname{sgn}(\sigma) = \begin{cases} 1 \text{ if } \sigma \text{ has an even number of inversion}(i < j \text{ s.t. } \sigma(i) > \sigma(j)) \\ -1 \text{ otherwise.} \end{cases}$$

It is known that $\det \in \mathbf{FP}$, e.g., by doing Gaussian elimination until A is upper triangular, and then multiplying the diagonal entries.

The matrix *permanent* is defined by removing the sign terms:

$$\operatorname{perm}(A) = \sum_{\sigma\in S_n} \cdot \prod_{i=1}^{n} A_{i,\sigma(i)}.$$

Define the function $0/1\text{-perm}$ to capture the problem of computing the matrix permanent of $\{0,1\}$-valued matrices.

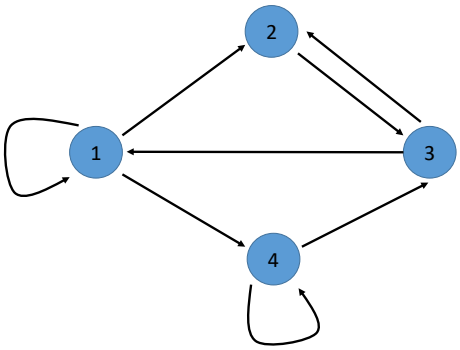
Theorem 10. $0/1\text{-perm}$ *is* $\#\mathbf{P}$*-complete.*

Proof sketch. As usual, there are two things to prove.

$0/1\text{-perm} \in \#\mathbf{P}$**:** Let us introduce a combinatorial view of the matrix permanent, in terms of *cycle covers*.

Definition 11. Let $G = ([n], E)$ be a directed graph, a permutation $\sigma : [n] \to [n]$ is a cycle cover for G if $(i, \sigma(i)) \in E$ for all $i \in [n]$.

Example 12. Consider the following digraph on 4 vertices.



The permutation $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1, 4 \mapsto 4$ is a cycle cover.
The permutation $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 4, 4 \mapsto 1$ is not a cycle cover.

Given a matrix A, let G_A be the digraph whose adjacency matrix is A.

Lemma 13. $\mathrm{perm}(A) = \#$ cycle covers of G_A.

Proof. From the definition,

$$\begin{aligned}
\mathrm{perm}(A) &= \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)} \\
&= \sum_{\sigma \in S_n} \begin{cases} 1 & \text{if } A_{i,\sigma(i)} = 1 \quad \forall i \in [n] \\ 0 & \text{otherwise} \end{cases} \\
&= \sum_{\sigma \in S_n} \begin{cases} 1 & \text{if } (i, \sigma(i)) \in E_A \quad \forall i \in [n] \\ 0 & \text{otherwise} \end{cases} \\
&= \#\text{ cycle covers of } G_A.
\end{aligned}$$

□

Since checking whether a permutation is a cycle cover can be done in poly-time, this implies that counting the number of cycle covers in a graph (and hence, computing the 0/1 permanent) is in $\#\mathbf{P}$.

Note that there's nothing really special about the 0/1 permanent here. This proof generalizes to characterize the integer permanent as summing the products of the weights along all possible weighted cycle covers.

Also, see page 347 in Arora-Barak for a different useful characterization of the permanent as counting perfect matchings in a bipartite graph.

0/1-perm is $\#\mathbf{P}$-hard: This is a quite involved proof involving some very clever gadget trickery. In broad strokes, the structure of the proof goes as follows.

1. Reduce the $\#\mathbf{P}$-complete problem $\#\mathsf{SAT}$ to integer perm. Idea: Map a CNF formula φ to a graph G such that assignments to φ correspond to (structured sets of) cycle covers. Satisfying assignments will all induce covers of weight k (for some parameter k depending on φ), while unsatisfying assignments have weight 0.

2. Reduce integer perm to $\{-1,0,1\}$-perm by replacing weighted edges with sets of parallel paths.

3. Reduce $\{-1,0,1\}$-perm to $\{0,1\}$-perm using modular arithmetic tricks.

□

3 PP and $\oplus$P

Recall that $\#\mathbf{P}$ is the set of functions f such that there is polynomial-time deterministic TM M and a polynomial $p : \mathbb{N} \to \mathbb{N}$ such that for all x,

$$f(x) = \#\{u \in \{0,1\}^{p(|x|)} \mid M(x,u) = 1\}.$$

Any function evaluation problem naturally gives rise to decision problems that correspond to computing individual bits of that function. We can further specialize our attention to the most significant bit and least significant bit of f, which have natural interpretations.

Definition 14. A language $L \in \mathbf{PP}$ if there exists a poly-time deterministic TM M and polynomial p such that

$$x \in L \iff \#\{u \in \{0,1\}^{p(|x|)} \mid M(x,u) = 1\} \geq \frac{1}{2} \cdot 2^{p(n)}.$$

Equivalently: $L \in \mathbf{PP}$ if there is a poly-time probabilistic TM M such that

$$x \in L \iff \Pr[M(x) = 1] \geq \frac{1}{2}.$$

The canonical $\mathbf{PP}$-complete problem is MAJSAT: $\varphi \in \mathsf{MAJSAT} \iff$ the majority of assignments x satisfy φ.

Lemma 15. $\mathbf{P}^{\mathbf{PP}} = \mathbf{P}^{\#\mathbf{P}}$.

Proof. It's enough to show that $\#\mathbf{P} \subseteq \mathbf{FP}^{\mathbf{PP}}$.

To see this, let $f \in \#\mathbf{P}$, so $f(x) = \#\{u \in \{0,1\}^{p(|x|)} \mid M(x,u) = 1\}$ for some machine M. We describe how to compute f in poly-time on a $\mathbf{PP}$-oracle TM as follows. Let $k = p(|x|)$. For each $N \in \{0,1,\ldots,2^k\}$, let M_N be the following TM:

On input $x, \langle b,u\rangle$:

If $b = 0$: Output $M(x,u)$

If $b = 1$: Output 1 iff $u < N$, and 0 otherwise.

Observe that for every x, we have $\#\{\langle b,u\rangle \in \{0,1\}^{k+1} \mid M_N(x,\langle b,u\rangle) = 1\} = f(x) + N$.

Using a $\mathbf{PP}$ oracle, we can thus compare $f(x) + N \stackrel{?}{\geq} \frac{1}{2} \cdot 2^{k+1} = 2^k$ for any N. Now we can binary search over N to find the smallest number N^* for which $f(x) + N^* \geq 2^k$, i.e., $f(x) + N^* = 2^k$, and output the answer $f(x) = 2^k - N^*$. □

Definition 16. A language $L \in \oplus\mathbf{P}$ if there exists a poly-time deterministic TM M and polynomial p such that

$$x \in L \iff \#\{u \in \{0,1\}^{p(|x|)} \mid M(x,u) = 1\} \text{ is odd.}$$

The canonical $\oplus\mathbf{P}$-complete problem is $\oplus$SAT: $\varphi \in \oplus\mathsf{SAT} \iff \#\{x \mid \varphi(x) = 1\}$ is odd.

We do not yet even know whether $\mathbf{NP} \subseteq \mathbf{P}^{\oplus\mathbf{P}}$, but we do know from Valiant-Vazirani that $\mathbf{NP} \subseteq \mathbf{RP}^{\oplus\mathbf{P}}$.

4 Toda's Theorem

Alternation and counting give two ways of generalizing the class $\mathbf{NP}$ which, at first glance, may seem incomparable in power. In 1991, Seinosuke Toda proved the stunning result that counting is, in fact, at least as powerful as alternation.

Theorem 17 (Toda's "First" Theorem)**.** $\mathbf{PH} \subseteq \mathbf{BPP}^{\oplus\mathbf{P}}$.

Theorem 18 (Toda's "Second" Theorem)**.** $\mathbf{BPP}^{\oplus\mathbf{P}} \subseteq \mathbf{P}^{\#\mathbf{P}}$.

Corollary 19 ("Toda's Theorem")**.** $\mathbf{PH} \subseteq \mathbf{P}^{\#\mathbf{P}}$.

I'll sketch a proof of the "First Theorem" following Fortnow's note, "A Simple Proof of Toda's Theorem." The proof makes use of the following three facts:

1. Valiant-Vazirani: There is a poly-time randomized reduction A such that

$$\varphi \in \mathsf{SAT} \implies \Pr[A(\varphi) \in \mathsf{USAT}_Y] \geq \frac{1}{8n} \implies \Pr[A(\varphi) \in \oplus\mathsf{SAT}] \geq \frac{1}{8n}$$
$$\varphi \notin \mathsf{SAT} \implies \Pr[A(\varphi) \in \mathsf{USAT}_N] = 1 \implies \Pr[A(\varphi) \notin \oplus\mathsf{SAT}] = 1.$$

   Corollary 20. $\mathbf{NP} \subseteq \mathbf{RP}^{\oplus\mathsf{SAT}} \subseteq \mathbf{BPP}^{\oplus\mathbf{P}}$. *(Moreover, this proof relativizes.)*

2. $\oplus\mathbf{P}^{\oplus\mathbf{P}} = \oplus\mathbf{P}$. (Maybe we'll prove this Thursday?)

3. If $\mathbf{NP} \subseteq \mathbf{BPP}$, then $\mathbf{PH} \subseteq \mathbf{BPP}$. (This generalizes the inductive proof that $\mathbf{NP} \subseteq \mathbf{P} \implies \mathbf{PH} \subseteq \mathbf{P}$.) Moreover, this proof relativizes.

Using the fact that the Corollary to Fact 1 implies relativizes, we have

$$\mathbf{NP}^{\oplus\mathbf{P}} \subseteq (\mathbf{BPP}^{\oplus\mathbf{P}})^{\oplus\mathbf{P}} = \mathbf{BPP}^{(\oplus\mathbf{P}^{\oplus\mathbf{P}})}.$$

The equality holds because a BPP machine can make its "outer" $\oplus\mathbf{P}$ queries via its "inner" $\oplus\mathbf{P}$ oracle.

Now Fact 2 implies this latter class is just $\mathbf{BPP}^{\oplus\mathbf{P}}$. So using the relativizing version of Fact 3, we get $\mathbf{PH} \subseteq \mathbf{PH}^{\oplus\mathbf{P}} \subseteq \mathbf{BPP}^{\oplus\mathbf{P}}$.