Notes for Lecture 10

Counting Problems

1 Counting Classes

Definition 1 *R is an* **NP**-relation*, if there is a polynomial time algorithm A such that $(x, y) \in R \Leftrightarrow A(x, y) = 1$ and there is a polynomial p such that $(x, y) \in R \Rightarrow |y| \leq p(|x|)$.*

$\#R$ is the problem that, given x, asks how many y satisfy $(x, y) \in R$.

Definition 2 *#**P** is the class of all problems of the form $\#R$, where R is an **NP**-relation.*

Observe that an **NP**-relation R naturally defines an **NP** language L_R, where $L_R = \{x : x \in R(x, y)\}$, and every **NP** language can be defined in this way. Therefore problems in #**P** can always be seen as the problem of counting the number of witnesses for a given instance of an **NP** problem.

Unlike for decision problems there is no canonical way to define reductions for counting classes. There are two common definitions.

Definition 3 *We say there is a* parsimonious reduction *from $\#A$ to $\#B$ (written $\#A \leq_{par} \#B$) if there is a polynomial time transformation f such that for all x, $|\{y, (x, y) \in A\}| = |\{z : (f(x), z) \in B\}|$.*

Often this definition is a little too restrictive and we use the following definition instead.

Definition 4 *$\#A \leq \#B$ if there is a polynomial time algorithm for $\#A$ given an oracle that solves $\#B$.*

#CIRCUITSAT is the problem where given a circuit, we want to count the number of inputs that make the circuit output 1.

Theorem 1 *#CIRCUITSAT is #**P**-complete under parsimonious reductions.*

Proof: Let $\#R$ be in #**P** and A and p be as in the definition. Given x we want to construct a circuit C such that $|\{z : C(z)\}| = |\{y : |y| \leq p(|x|), A(x, y) = 1\}|$. We then construct $\hat{C}_n$ that on input x, y simulates $A(x, y)$. From earlier arguments we know that this can be done with a circuit with size about the square of the running time of A. Thus $\hat{C}_n$ will have size polynomial in the running time of A and so polynomial in x. Then let $C(y) = \hat{C}(x, y)$. $\square$

Theorem 2 *#3SAT is #**P**-complete.*

PROOF: We show that there is a parsimonious reduction from #CIRCUITSAT to #3-SAT. That is, given a circuit C we construct a Boolean formula φ such that the number of satisfying assignments for φ is equal to the number of inputs for which C outputs 1. Suppose C has inputs $x_1, \ldots, x_n$ and gates $1, \ldots, m$ and φ has inputs $x_1, \ldots, x_n, g_1, \ldots, g_m$, where the g_i represent the output of gate i. Now each gate has two input variables and one output variable. Thus a gate can be complete described by mimicking the output for each of the 4 possible inputs. Thus each gate can be simulated using at most 4 clauses. In this way we have reduced C to a formula φ with $n+m$ variables and $4m$ clauses. So there is a parsimonious reduction from #CIRCUITSAT to #3SAT. $\Box$

Notice that if a counting problem $\#R$ is $\#\mathbf{P}$-complete under parsimonious reductions, then the associated language L_R is **NP**-complete, because $\#3SAT \leq_{par} \#R$ implies $3SAT \leq L_R$. On the other hand, with the less restrictive definition of reducibility, even some counting problems whose decision version is in **P** are #**P**-complete. For example, the problem of counting the number of satisfying assignments for a given 2CNF formula and the problem of counting the number of perfect matchings in a given bipartite graphs are both #**P**-complete.

2 Complexity of counting problems

We will prove the following theorem:

Theorem 3 *For every counting problem $\#A$ in $\#\mathbf{P}$, there is a probabilistic algorithm C that on input x, computes with high probability a value v such that*

$$(1-\epsilon)\#A(x) \leq v \leq (1+\epsilon)\#A(x)$$

in time polynomial in $|x|$ and in $\frac{1}{\epsilon}$, using an oracle for **NP**.

The theorem says that #**P** can be approximate in $\mathbf{BPP}^{\mathbf{NP}}$. We have a remark here that approximating #3SAT is **NP**-hard. Therefore, to compute the value we need at least the power of **NP**, and this theorem states that the power of **NP** and randomization is sufficient.

Another remark concerns the following result.

Theorem 4 (Toda) *For every k, $\Sigma_k \subseteq \mathbf{P}^{\#\mathbf{P}}$.*

This implies that #3SAT is Σ_k-hard for every k, i.e., #3SAT lies outside **PH**, unless the hierarchy collapses. Recall that **BPP** lies inside Σ_2, and hence approximating #3SAT can be done in Σ_3. Therefore, approximating #3SAT cannot be equivalent to computing #3SAT exactly, unless the polynomial hierarchy collapses.[1]

We first make some observations so that we can reduce the proof to an easier one.

- It is enough to prove the theorem for #3SAT.

  If we have an approximation algorithm for #3SAT, we can extend it to any #A in #**P** using the parsimonious reduction from #A to #3SAT.

[1] The above discussion was not very rigorous but it can be correctly formalized.

- It is enough to give a polynomial time $O(1)$-approximation for #3SAT.

  Suppose we have an algorithm C and a constant c such that

  $$\frac{1}{c}\#3\mathrm{SAT}(\varphi) \leq C(\varphi) \leq c\#3\mathrm{SAT}(\varphi).$$

  Given φ, we can construct $\varphi^k = \varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_k$ where each φ_i is a copy of φ constructed using fresh variables. If φ has t satisfying assignments, φ^k has t^k satisfying assignments. Then, giving φ^k to the algorithm we get

  $$\frac{1}{c}t^k \leq C(\varphi^k) \leq ct^k$$
  $$\left(\frac{1}{c}\right)^{1/k} t \leq C(\varphi^k)^{1/k} \leq c^{1/k}t.$$

  If c is a constant and $k = O(\frac{1}{\epsilon})$, $c^{1/k} = 1 + \epsilon$.

- For a formula φ that has $O(1)$ satisfying assignments, $\#3\mathrm{SAT}(\varphi)$ can be found in $\mathbf{P}^{\mathbf{NP}}$.

  This can be done by iteratively asking the oracle the questions of the form: "Are there k assignments satisfying this formula?" Notice that these are **NP** questions, because the algorithm can guess these k assignments and check them.

3 An approximate comparison procedure

Suppose that we had available an approximate comparison procedure `a-comp` with the following properties:

- If $\#3\mathrm{SAT}(\varphi) \geq 2^{k+1}$ then $\mathtt{a} - \mathtt{comp}(\varphi, k) = \mathrm{YES}$ with high probability;
- If $\#3\mathrm{SAT}(\varphi) < 2^k$ then $\mathtt{a} - \mathtt{comp}(\varphi, k) = \mathrm{NO}$ with high probability.

Given `a-comp`, we can construct an algorithm that 2-approximates #3SAT as described in Figure 1.
We need to show that this algorithm approximates #3SAT within a factor of 2. If `a-comp` answers NO from the first time, the algorithm outputs the right answer because it checks for the answer explicitly. Now suppose `a-comp` says YES for all $t = 1, 2, \ldots, i-1$ and says NO for $t = i$. Since $\mathtt{a\text{-}comp}(\varphi, i-1)$ outputs YES, $\#3\mathrm{SAT}(\varphi) \geq 2^{i-1}$, and also since $\mathtt{a\text{-}comp}(\varphi, 2^i)$ outputs NO, $\#3\mathrm{SAT}(\varphi) < 2^{i+1}$. The algorithm outputs $a = 2^i$. Hence,

$$\frac{1}{2}a \leq \#3\mathrm{SAT}(\varphi) < 2 \cdot a$$

and the algorithm outputs the correct answer with in a factor of 2.

Thus, to establish the theorem, it is enough to give a $\mathbf{BPP}^{\mathbf{NP}}$ implementation of the `a-comp`.

```
Input: φ

compute:
   a-comp(φ, 0)
   a-comp(φ, 1)
   a-comp(φ, 2)
   ⋮
   a-comp(φ, n + 1)

if a-comp outputs NO from the first time then
   // The value is either 0 or 1.
   // The answer can be checked by one more query to the NP oracle.
   Query to the oracle and output an exact value.
else
   Suppose that it outputs YES for t = 1, ..., i − 1 and NO for t = i
   Output 2^i
```

Figure 1: How to use `a-comp` to approximate #3SAT.

4 Constructing `a-comp`

The procedure and its analysis is similar to the Valiant-Vazirani reduction: for a given formula φ we pick a hash function h from a pairwise independent family, and look at the number of assignments x that satisfy h and such that $h(x) = \mathbf{0}$.

In the Valiant-Vazirani reduction, we proved that if S is a set of size approximately equal to the size of the range of $h()$, then, with constant probability, exactly one element of S is mapped by $h()$ into $\mathbf{0}$. Now we use a different result, a simplified version of the "Leftover Hash Lemma" proved by Impagliazzo, Levin, and Luby in 1989, that says that if S is sufficiently larger than the range of $h()$ then the number of elements of S mapped into $\mathbf{0}$ is concentrated around its expectation.

Lemma 5 *Let H be a family of pairwise independent hash functions $h : \{0,1\}^n \to \{0,1\}^m$. Let $S \subset 0, 1^n$, $|S| \geq \frac{4 \cdot 2^m}{\epsilon^2}$. Then,*

$$\mathbf{Pr}_{h \in H}\left[\left| |\{a \in S : h(a) = 0\}| - \frac{|S|}{2^m} \right| \geq \epsilon \frac{|S|}{2^m}\right] \leq \frac{1}{4}.$$

From this, `a-comp` can be constructed as in Figure 2.

Notice that the test at the last step can be done with one access to an oracle to **NP**. We will show that the algorithm is in $\mathbf{BPP}^{\mathbf{NP}}$. Let $S \subseteq \{0,1\}^n$ be the set of satisfying assignments for φ. There are 2 cases.

- If $|S| \geq 2^{k+1}$, by Lemma 5 we have:

$$\mathbf{Pr}_{h \in H}\left[\left| |\{a \in S : h(a) = 0\}| - \frac{|S|}{2^m} \right| \leq \frac{1}{4} \cdot \frac{|S|}{2^m}\right] \leq \frac{3}{4},$$

input: φ, k

if $k \leq 5$ **then** check exactly whether $\#3\text{SAT}(\varphi) \geq 2^k$.
if $k \geq 6$,
pick h from a set of pairwise independent hash functions $h : \{0,1\}^n \to \{0,1\}^m$, where $m = k - 5$
answer YES iff there are more then 48 assignments a to φ such that a satisfies φ and $h(a) = 0$.

Figure 2: The approximate algorithm for #3SAT.

(set $\epsilon = \frac{1}{4}$, and $|S| \geq \frac{4 \cdot 2^m}{\epsilon^2} = 64 \cdot 2^m$, because $|S| \geq 2^{k+1} = 2^{m+6}$)

$$\mathbf{Pr}_{h \in H}\left[|\{a \in S : h(a) = 0\}| \leq \frac{3}{4} \cdot \frac{|S|}{2^m}\right] \leq \frac{1}{4},$$
$$\mathbf{Pr}_{h \in H}\left[|\{a \in S : h(a) = 0\}| \geq 48\right] \geq \frac{3}{4},$$

which is the success probability of the algorithm.

- If $|S| < 2^k$:

  Let S' be a superset of S of size 2^k. We have

$$\begin{aligned}
\mathbf{Pr}_{h \in H}[\text{answer YES}] &= \mathbf{Pr}_{h \in H}[|\{a \in S : h(s) = 0\}| \geq 48] \\
&\leq \mathbf{Pr}_{h \in H}[\left|\{a \in S' : h(s) = 0\}\right| \geq 48] \\
&\leq \mathbf{Pr}_{h \in H}\left[\left|\left|\{a \in S' : h(s) = 0\}\right| - \tfrac{|S'|}{2^m}\right| \geq \tfrac{|S'|}{2 \cdot 2^m}\right] \\
&\leq \frac{1}{4}
\end{aligned}$$

  (by Lemma 5 with $\epsilon = 1/2$, $|S'| = 32 \cdot 2^m$.)

Therefore, the algorithm will give the correct answer with probability at least 3/4, which can then be amplified to, say, $1 - 1/4n$ (so that all n invocations of `a-comp` are likely to be correct) by repeating the procedure $O(\log n)$ times and taking the majority answer.

5 The proof of Lemma 5

We finish the lecture by proving Lemma 5.

Proof: We will use Chebyshev's Inequality to bound the failure probability. Let $S = \{a_1, \ldots, a_k\}$, and pick a random $h \in H$. We define random variables $X_1, \ldots, X_k$ as

$$X_i = \begin{cases} 1 & \text{if } h(a_i) = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $|\{a \in S : h(a) = 0\}| = \sum_i X_i$.

We now calculate the expectations. For each i, $\mathbf{Pr}[X_i = 1] = \frac{1}{2^m}$ and $\mathbf{E}[X_i] = \frac{1}{2^m}$. Hence,

$$\mathbf{E}\left[\sum_i X_i\right] = \frac{|S|}{2^m}.$$

Also we calculate the variance

$$\begin{aligned}\mathbf{Var}[X_i] &= \mathbf{E}[X_i^2] - \mathbf{E}[X_i]^2 \\ &\leq \mathbf{E}[X_i^2] \\ &= \mathbf{E}[X_i] = \frac{1}{2^m}.\end{aligned}$$

Because $X_1, \ldots, X_k$ are pairwise independent,

$$\mathbf{Var}\left[\sum_i X_i\right] = \sum_i \mathbf{Var}[X_i] \leq \frac{|S|}{2^m}.$$

Using Chebyshev's Inequality, we get

$$\begin{aligned}\mathbf{Pr}\left[\left| |\{a \in S : h(a) = 0\}| - \frac{|S|}{2^m}\right| \geq \epsilon\frac{|S|}{2^m}\right] &= \mathbf{Pr}\left[\left|\sum_i X_i - \mathbf{E}[\sum_i X_i]\right| \geq \epsilon \mathbf{E}[\sum_i X_i]\right] \\ &\leq \frac{\mathbf{Var}[\sum_i X_i]}{\epsilon^2 \mathbf{E}[\sum_i X_i]^2} \leq \frac{\frac{|S|}{2^m}}{\epsilon^2 \frac{|S|^2}{(2^m)^2}} \\ &= \frac{2^m}{\epsilon^2 |S|} \leq \frac{1}{4}.\end{aligned}$$

□

6 Approximate Sampling

So far we have considered the following question: for an **NP**-relation R, given an input x, what is the size of the set $R_x = \{y : (x, y) \in R\}$? A related question is to be able to sample from the uniform distribution over R_x.

Whenever the relation R is "downward self reducible" (a technical condition that we won't define formally), it is possible to prove that there is a probabilistic algorithm running in time polynomial in $|x|$ and $1/\epsilon$ to approximate within $1 + \epsilon$ the value $|R_x|$ if and only if there is a probabilistic algorithm running in time polynomial in $|x|$ and $1/\epsilon$ that samples a distribution ϵ-close to the uniform distribution over R_x.

We show how the above result applies to 3SAT (the general result uses the same proof idea). For a formula φ, a variable x and a bit b, let us define by $\varphi_{x \leftarrow b}$ the formula obtained by substituting the value b in place of x.[2]

[2]Specifically, $\varphi_{x \leftarrow 1}$ is obtained by removing each occurrence of $\neg x$ from the clauses where it occurs, and removing all the clauses that contain an occurrence of x; the formula $\varphi_{x \leftarrow 0}$ is similarly obtained.

If φ is defined over variables $x_1, \ldots, x_n$, it is easy to see that

$$\#\varphi = \#\varphi_{x\leftarrow 0} + \#\varphi_{x\leftarrow 1}$$

Also, if S is the uniform distribution over satisfying assignments for φ, we note that

$$\mathbf{Pr}_{(x_1,\ldots,x_n)\leftarrow S}[x_1 = b] = \frac{\#\varphi_{x\leftarrow b}}{\#\varphi}$$

Suppose then that we have an efficient sampling algorithm that given φ and ϵ generates a distribution ϵ-close to uniform over the satisfying assignments of φ.

Let us then ran the sampling algorithm with approximation parameter $\epsilon/2n$ and use it to sample about $\tilde{O}(n^2/\epsilon^2)$ assignments. By computing the fraction of such assignments having $x_1 = 0$ and $x_1 = 1$, we get approximate values p_0, p_1, such that $|p_b - \mathbf{Pr}_{(x_1,\ldots,x_n)\leftarrow S}[x_1 = b]| \leq \epsilon/n$. Let b be such that $p_b \geq 1/2$, then $\#\varphi_{x\leftarrow b}/p_b$ is a good approximation, to within a multiplicative factor $(1 + 2\epsilon/n)$ to $\#\varphi$, and we can recurse to compute $\#\varphi_{x\leftarrow b}$ to within a $(1 + 2\epsilon/n)^{n-1}$ factor.

Conversely, suppose we have an approximate counting procedure. Then we can approximately compute $p_b = \frac{\#\varphi_{x\leftarrow b}}{\#\varphi}$, generate a value b for x_1 with probability approximately p_b, and then recurse to generate a random assignment for $\#\varphi_{x\leftarrow b}$.

The same equivalence holds, clearly, for 2SAT and, among other problems, for the problem of counting the number of perfect matchings in a bipartite graph. It is known that it is **NP**-hard to perform approximate counting for 2SAT and this result, with the above reduction, implies that approximate sampling is also hard for 2SAT. The problem of approximately sampling a perfect matching has a probabilistic polynomial solution, and the reduction implies that approximately counting the number of perfect matchings in a graph can also be done in probabilistic polynomial time.

The reduction and the results from last section also imply that 3SAT (and any other **NP** relation) has an approximate sampling algorithm that runs in probabilistic polynomial time with an **NP** oracle. With a careful use of the techniques from last week it is indeed possible to get an *exact* sampling algorithm for 3SAT (and any other **NP** relation) running in probabilistic polynomial time with an **NP** oracle. This is essentially best possible, because the approximate sampling requires randomness by its very definition, and generating satisfying assignments for a 3SAT formula requires at least an **NP** oracle.

7 References

The class $\#\mathbf{P}$ was defined by Valiant [Val79]. An algorithm for approximate counting within the polynomial hierarchy was developed by Stockmeyer [Sto83]. The algorithm presented in these notes is taken from lecture notes by Oded Goldreich. The left-over hash lemma is from [HILL99]. The problem of approximate sampling and its relation to approximate counting is studied in [JVV86].

References

[HILL99] J. Håstad, R. Impagliazzo, L. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999. 7

[JVV86] Mark Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986. 7

[Sto83] L.J. Stockmeyer. The complexity of approximate counting. In *Proceedings of the 15th ACM Symposium on Theory of Computing*, pages 118–126, 1983. 7

[Val79] L.G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979. 7