Advanced Algorithms 9 October 2018

Lecture 4: Approximation Algorithms IV

Lecturer: Mohsen Ghaffari *Scribe: Davin Choo*

1 Randomized approximation schemes

In earlier lectures, we saw PTAS and FPTAS. In this lecture, we study the class of algorithms which extend FPTAS by allowing randomization.

Definition 1 (Fully polynomial randomized approximation algorithm (FPRAS))**.** *For cost metric c, an algorithm $\mathcal{A}$ is a FPRAS if for each fixed $\epsilon > 0$,* $\Pr[|c(\mathcal{A}(I)) - c(OPT(I))| \leq \epsilon \cdot c(OPT(I))] \geq \frac{3}{4}$ *and $\mathcal{A}$ runs in* $\text{poly}(|I|, \frac{1}{\epsilon})$.

A useful inequality that we will use in the proofs below is the Chernoff bound.

Theorem 2 (Chernoff bound)**.** *For independent Bernoulli variables $X_1, \dots, X_n$, let $X = \sum_{i=1}^{n} X_i$. Then,*

$$\Pr[X \geq (1+\epsilon)\mathbb{E}(X)] \leq \exp(\frac{-\epsilon^2 \mathbb{E}(X)}{3}) \quad \text{for } 0 < \epsilon$$
$$\Pr[X \leq (1-\epsilon)\mathbb{E}(X)] \leq \exp(\frac{-\epsilon^2 \mathbb{E}(X)}{2}) \quad \text{for } 0 < \epsilon < 1$$

By union bound, for $0 < \epsilon < 1$, we get $\Pr[|X - \mathbb{E}(X)| \geq \epsilon \mathbb{E}(X)] \leq 2\exp(\frac{-\epsilon^2 \mathbb{E}(X)}{3})$

Remark 1 We usually apply Chernoff bound to show that the probability of bad approximation is low (Pick parameters such that $2\exp(\frac{-\epsilon^2 \mathbb{E}(X)}{3}) \leq \delta$), then negate to get $\Pr[|X - \mathbb{E}(X)| \leq \epsilon \mathbb{E}(X)] \geq 1 - \delta$.

Remark 2 The fraction $\frac{3}{4}$ in the definition of FPRAS is arbitrary. In fact, any fraction $\frac{1}{2} + \alpha$ for $\alpha > 0$ suffices. For any $\delta > 0$, one can invoke $\mathcal{O}(\frac{1}{\delta})$ independent copies of $\mathcal{A}(I)$ then return the median. Then, Chebyshev's inequality tells us that the probability that the median is a correct estimation with probability greater than $\geq 1 - \delta$. This is also sometimes known as *probability amplification*.

2 DNF counting

Definition 3 (Disjunctive Normal Form (DNF))**.** *A formula F on n Boolean variables $x_1, \dots, x_n$ is said to be in DNF:*

- *$F = C_1 \vee \dots \vee C_m$ is a disjuntion of clauses*
- *$\forall i \in \{1, \dots, m\}$, a clause $C_i = l_{i,1} \wedge \dots \wedge l_{i,|C_i|}$ is a conjunction of literals*
- *$\forall i \in \{1, \dots, n\}$, a literal $l_i \in \{x_i, \neg x_i\}$ is either the variable x_i or its negation.*

Let $\alpha : \{1, \dots, n\} \to \{0, 1\}$ be a truth assignment on the n variables. Formula F is said to be satisfiable *if there exists a* satisfying assignment *α such that F evaluates to true under α (i.e. $F[\alpha] = 1$).*

One can see that any clause with both x_i and $\neg x_i$ is trivially false. Since we can remove such clauses in a single scan of F, let us assume that F does not contain such trivial clauses.

Example Let $F = (x_1 \wedge \neg x_2 \wedge \neg x_4) \vee (x_2 \wedge x_3) \vee (\neg x_3 \wedge \neg x_4)$ be a Boolean formula on 4 variables x_1, x_2, x_3, and x_4, where $C_1 = x_1 \wedge \neg x_2 \wedge \neg x_4$, $C_2 = x_2 \wedge x_3$ and $C_3 = \neg x_3 \wedge \neg x_4$. One can draw the truth table and check that there are 9 satisfying assignments to F, one of which is $\alpha(1) = 1, \alpha(2) = \alpha(3) = \alpha(4) = 0$.

Remark Another common normal form for representing Boolean formulas is the *Conjunctive Normal Form* (CNF). Formulas in CNF are disjunctions of conjunctions (as compared to conjunctions of disjunctions in DNF). In particular, one can determine in polynomial time whether a DNF formula is satisfiable but it is $\mathbb{NP}$-complete to determine if a CNF formula is satisfiable.

Suppose F is a DNF Boolean formula. Let $f(F) = |\{\alpha : F[\alpha] = 1\}|$ be the number of satisfying assignments to F. If we let $S_i = \{\alpha : C_i[\alpha] = 1\}$ be the set of satisfying assignments to clause C_i, then we see that $f(F) = |\bigcup_{i=1}^{m} S_i|$. In the above example, $|S_1| = 2$, $|S_2| = 4$, $|S_3| = 4$, and $f(F) = 9$. In the following, we present two failed attempts to compute $f(F)$ and then present Algorithm 1, a FPRAS for DNF counting via sampling.

2.1 Failed attempt 1: Computing $f(F)$ via Principle of Inclusion-Exclusion

By definition of $f(F) = |\bigcup_{i=1}^{m} S_i|$, one may be tempted to apply PIE to expand:

$$|\bigcup_{i=1}^{m} S_i| = \sum_{i=1}^{m} |S_i| - \sum_{i<j} |S_i \cap S_j| + \dots$$

However, there are exponentially many terms and one can show that there exists instances where truncating the sum as a form of approximation can be arbitrarily bad.

2.2 Failed attempt 2: Sampling (wrongly)

Suppose we pick k assignments uniformly at random (u.a.r.). Let X_i be the indicator variable whether the i-th assignment satisfies F, and $X = \sum_{i=1}^{k} X_i$ be the total number of satisfying assignments out of the k sampled assignments. A u.a.r. assignment is satisfying with probability $\frac{f(F)}{2^n}$. By linearity of expectation, $\mathbb{E}(X) = k\frac{f(F)}{2^n}$. Unfortunately, since we only sample $k \in \text{poly}(n, \frac{1}{\epsilon})$ assignments, $\frac{k}{2^n}$ can be exponentially small. That is, this approach will *not* yield a FPRAS for DNF counting.

2.3 A FPRAS for DNF counting via sampling

Consider a m-by-$f(F)$ Boolean matrix M where $M[i,j] = \begin{cases} 1 & \text{if assignment } \alpha_j \text{ satisfies clause } C_i \\ 0 & \text{otherwise} \end{cases}$

Let $|M|$ denote the total number of 1's in M. Since $|S_i| = 2^{n-|C_i|}$, $|M| = \sum_{i=1}^{m} |S_i| = \sum_{i=1}^{m} 2^{n-|C_i|}$. As every column represents a satisfying assignment, there are exactly $f(F)$ "topmost" 1's.

	α_1	α_2	$\dots$	$\alpha_{f(F)}$
C_1	0	1	$\dots$	0
C_2	1	1	$\dots$	1
C_3	0	0	$\dots$	0
$\dots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
C_m	0	1	$\dots$	1

Table 1: Red 1's indicate the ("topmost") smallest index clause C_i satisfied for each assignment α_j

Lemma 4. *Algorithm 1 samples a '1' in the matrix M uniformly at random at each step.*

Proof. Recall that the total number of 1's in M is $|M| = \sum_{i=1}^{m} |S_i| = \sum_{i=1}^{m} 2^{n-|C_i|}$.

$$\begin{aligned}
\Pr[C_i \text{ and } \alpha_j \text{ are chosen}] &= \Pr[C_i \text{ is chosen}] \cdot \Pr[\alpha_j \text{ is chosen} | C_i \text{ is chosen}] \\
&= \frac{2^{n-|C_i|}}{\sum_{i=1}^{m} 2^{n-|C_i|}} \cdot \frac{1}{2^{n-|C_i|}} \\
&= \frac{1}{\sum_{i=1}^{m} 2^{n-|C_i|}} \\
&= \frac{1}{|M|}
\end{aligned}$$

□

Algorithm 1 DNF-COUNT(F, ϵ)

```
X ← 0                                                        ▷ Empirical number of "topmost" 1's sampled
for k = 9m/ε² times do
    C_i ← Sample one of m clauses, where Pr[C_i chosen] = 2^{n-|C_i|}/|M|   ▷ Shorter clauses more likely
    α_j ← Sample one of 2^{n-|C_i|} satisfying assignments of C_i          ▷ Flip coins for x ∉ C_i
    IsTopmost ← True
    for l ∈ {1, ..., i-1} do                                 ▷ Check if α_j is "topmost"
        if C_l[α] = 1 then                                   ▷ Checkable in O(n) time
            IsTopmost ← False
        end if
    end for
    if IsTopmost then
        X ← X + 1
    end if
end for
return |M|·X / k
```

Lemma 5. *In Algorithm 1,* $\Pr[|\frac{|M|\cdot X}{k} - f(F)| \leq \epsilon \cdot f(F)] \geq \frac{3}{4}$.

Proof. Let X_i be the indicator variable whether the i-th sampled assignment is "topmost", where $p = \Pr[X_i = 1]$. By Lemma 4, $p = \Pr[X_i = 1] = \frac{f(F)}{|M|}$. Let $X = \sum_{i=1}^{k} X_i$ be the empirical number of "topmost" 1's. Then, $\mathbb{E}(X) = kp$ by linearity of expectation. By picking $k = \frac{9m}{\epsilon^2}$, Chernoff bound gives:

$$\begin{aligned}
\Pr[|X - kp| \geq \epsilon kp] &\leq 2\exp(-\frac{\epsilon^2 kp}{3}) \\
&= 2\exp(-\frac{3m \cdot f(F)}{|M|}) \quad \text{Since } k = \frac{9m}{\epsilon^2} \text{ and } p = \frac{f(F)}{|M|} \\
&\leq 2\exp(-3) \quad \text{Since } |M| \leq m \cdot f(F) \\
&\leq \frac{1}{8}
\end{aligned}$$

Splitting up the absolute sign, we have: $\Pr[X \geq (1+\epsilon)kp] \leq \frac{1}{8}$ and $\Pr[X \leq (1-\epsilon)kp] \leq \frac{1}{8}$. So,

1. $\Pr[X \geq (1+\epsilon)kp = (1+\epsilon)\frac{k \cdot f(F)}{|M|}] \leq \frac{1}{8}$
2. $\Pr[X \leq (1-\epsilon)kp = (1-\epsilon)\frac{k \cdot f(F)}{|M|}] \leq \frac{1}{8}$

Multiplying both sides by $\frac{|M|}{k}$, union bound gives us:

$$\Pr[|\frac{|M|\cdot X}{k} - f(F)| \geq \epsilon \cdot f(F)] \leq \Pr[X \leq (1-\epsilon)\frac{k \cdot f(F)}{|M|}] + \Pr[X \geq (1+\epsilon)\frac{k \cdot f(F)}{|M|}] \leq \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

Negating, we get:

$$\Pr[|\frac{|M|\cdot X}{k} - f(F)| \leq \epsilon \cdot f(F)] \geq 1 - \frac{1}{4} = \frac{3}{4}$$

□

Lemma 6. *Algorithm 1 runs in poly*$(F, \frac{1}{\epsilon})$ = *poly*$(n, m, \frac{1}{\epsilon})$.

Proof. There are $k \in \mathcal{O}(\frac{m}{\epsilon^2})$ iterations. In each iteration, we spend $\mathcal{O}(m+n)$ sampling C_i and α_j, and $\mathcal{O}(nm)$ for checking if a sampled α_j is "topmost". In total, Algorithm 1 runs in $\mathcal{O}(\frac{m^2 n(m+n)}{\epsilon^2})$ time. □

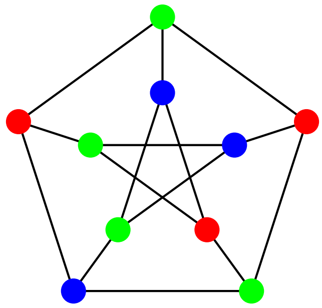
Theorem 7. *Algorithm 1 is a FPRAS for DNF counting.*

Proof. By Lemmas 5 and 6. □

3 Counting graph colourings

Definition 8 (Graph colouring)**.** *Let $G = (V, E)$ be a graph on $|V| = n$ vertices and $|E| = m$ edges. Denote the maximum degree as Δ. Given valid q-colouring of G is an assignment $c : V \to \{1, \ldots, q\}$ such that no adjacent vertices have the same colour. i.e. $(u, v) \in E \Rightarrow c(u) \neq c(v)$.*

Example (3-colouring of the Petersen graph)



For $q \geq \Delta + 1$, one can obtain a valid q-colouring by sequentially colouring a vertex with available colours greedily. In this section, we show a FPRAS for counting the graph colouring $f(G)$ when $q \geq 2\Delta + 1$.

3.1 Sampling a colouring uniformly

When $q \geq 2\Delta + 1$, the Markov chain approach in Algorithm 2 allows us to sample a random colour in $\mathcal{O}(n \log \frac{n}{\epsilon})$ steps.

Algorithm 2 SampleColour($G = (V, E), \epsilon$)

Greedily colour the graph
for $k = \mathcal{O}(n \log \frac{n}{\epsilon})$ times **do**
  Pick a random vertex v uniformly at random from V
  Pick an available colour (different from $N(v)$) uniformly random
  Colour v with new colour ▷ May end up with same colour
end for
return Colouring

Claim 9. *For $q \geq 2\Delta + 1$, the distribution of colourings returned by Algorithm 2 is ϵ-close to a uniform distribution on all valid colourings.*

Proof. Beyond the scope of the course. □

3.2 FPRAS for counting graph colourings for $q \geq 2\Delta + 1$ and $\Delta \geq 2$

Fix an arbitrary ordering of edges in E. For $i = \{1, \ldots, m\}$, let $G_i = (V, E_i)$ be a sequence of graphs such that $E_i = \{e_1, \ldots, e_i\}$ be the first i edges. Define $\Omega_i = \{c : c \text{ is a valid colouring for } G_i\}$ be the set of all proper colourings of G_i, and denote $r_i = \frac{|\Omega_i|}{|\Omega_{i-1}|}$.

One can see that $\Omega_i \subseteq \Omega_{i-1}$ as removal of e_i in G_{i-1} can only increase the number of valid colourings. Furthermore, suppose $e_i = (u, v)$, then $\Omega_{i-1} \setminus \Omega_i = \{c : c(u) = c(v)\}$. Fix the colouring of, say the lower-indexed vertex, u. Then, there are $\geq q - \Delta = 2\Delta + 1 = \Delta + 1$ possible recolourings of v. Hence, $|\Omega_i| \geq (\Delta + 1)|\Omega_{i-1} \setminus \Omega_i| \geq (\Delta + 1)(\Omega_{i-1}| - |\Omega_i|)$. This implies that $r_i = \frac{|\Omega_i|}{|\Omega_{i-1}|} \geq \frac{\Delta+1}{\Delta+2} \geq \frac{3}{4}$ since $\Delta \geq 2$.

Since $f(G) = |\Omega_m| = |\Omega_0| \cdot \frac{|\Omega_1|}{|\Omega_0|} \cdot \frac{|\Omega_m|}{|\Omega_{m-1}|} = |\Omega_0| \cdot \Pi_{i=1}^m r_i = q^m \cdot \Pi_{i=1}^m r_i$, if we can find a good estimate of r_i for each r_i with high probability, then we have a FPRAS for counting the number of valid graph colourings for G.

Lemma 10. *In Algorithm 3, for all $i \in \{1, \ldots, m\}$, $\Pr[|\widehat{r_i} - r_i| \leq \frac{\epsilon}{m} \cdot r_i] \geq \frac{3}{4m}$.*

Proof. Let X_j be the indicator variable whether the i-th sampled colouring for Ω_{i-1} is a valid colouring for Ω_i, where $p = \Pr[X_j = 1]$. From above, we know that $p = \Pr[X_j = 1] = \frac{|\Omega_i|}{|\Omega_{i-1}|} \geq \frac{3}{4}$. Let $X = \sum_{j=1}^k X_j$ be the empirical fraction of colourings that is valid for both Ω_{i-1} and Ω_i, captured by $k \cdot \widehat{r_i}$. Then, $\mathbb{E}(X) = kp$ by linearity of expectation. Picking $k = \frac{128m^3}{\epsilon^2}$, Chernoff bound gives:

$$
\begin{aligned}
\Pr[|X - kp| \geq \tfrac{\epsilon}{2m} kp] &\leq 2\exp\left(-\frac{(\frac{\epsilon}{2m})^2 kp}{3}\right) \\
&= 2\exp\left(-\frac{32mp}{3}\right) && \text{Since } k = \frac{128m^3}{\epsilon^2} \\
&\leq 2\exp(-8m) && \text{Since } p \geq \frac{3}{4} \\
&\leq \frac{1}{4m} && \text{Since } \exp(-x) \leq \frac{1}{x} \text{ for } x > 0
\end{aligned}
$$

Algorithm 3 Colour-Count(G, ϵ)

```
r̂_1, ..., r̂_m ← 0                                      ▷ Estimates for r_i
for i = 1, ..., m do
    for k = 128m^3/ε^2 times do
        c ← Sample colouring of G_{i-1}                 ▷ Using Algorithm 2
        if Adding c is a valid colouring for G_i then
            r̂_i ← r̂_i + 1/k                             ▷ Update empirical count of r_i = |Ω_i|/|Ω_{i-1}|
        end if
    end for
end for
return q^m Π_{i=1}^m r̂_i
```

Dividing by k and negating, we have: $\Pr[|\widehat{r_i} - r_i| \leq \frac{\epsilon}{2m} \cdot r_i] = \Pr[|X - kp| \geq \frac{\epsilon}{2m} kp] \geq 1 - \frac{1}{4m} = \frac{3}{4m}$. □

Lemma 11. *Algorithm 3 runs in poly$(F, \frac{1}{\epsilon})$ = poly$(n, m, \frac{1}{\epsilon})$.*

Proof. There are m r_i's to estimate. Each estimation has $k \in \mathcal{O}(\frac{m^3}{\epsilon^2})$ iterations. In each iteration, we spend $\mathcal{O}(n \log \frac{n}{\epsilon})$ sampling a colouring of G_{i-1} and $\mathcal{O}(n)$ checking if it is a valid colouring for G_i. In total, Algorithm 3 runs in $\mathcal{O}(mk(n \log \frac{n}{\epsilon} + n)) = \mathcal{O}(\frac{m^4 n \log \frac{n}{\epsilon}}{\epsilon^2})$ time. □

Theorem 12. *Algorithm 3 is a FPRAS for counting the number of valid graph colourings for $q \geq 2\Delta + 1$ and $\Delta \geq 2$.*

Proof. By Lemma 11, Algorithm 3 runs in poly$(n, m, \frac{1}{\epsilon})$ time. Since $1 + x \leq e^x$ for all real x, we have $(1 + \frac{\epsilon}{2m})^m \leq e^{\frac{\epsilon}{2}} \leq 1 + \epsilon$. On the other hand, Bernoulli's inequality tells us that $(1 - \frac{\epsilon}{2m})^m \geq 1 - \frac{\epsilon}{2} \geq 1 - \epsilon$. Therefore, via Lemma 10,

$$\Pr[|q^m \Pi_{i=1}^m \widehat{r_i} - f(G)| \leq \epsilon f(G)] = 1 - \Pr[|q^m \Pi_{i=1}^m \widehat{r_i} - f(G)| \geq \epsilon f(G)] \geq (\frac{3}{4m})^m \geq \frac{3}{4}$$

□