Lecture 8: Introduction to Randomized Algorithms II

- Conditional Probability and Derandomization
- p-Biased Coins
- Linear Programming and Randomized Rounding

Reference: Williamson, *Lecture Notes on Approximation Algorithms* 1999

IBM Research Report RC21409, Chapter 5

Can be found at http://www.almaden.ibm.com/cs/people/dpw/

Review of Conditional Probabilities and Expectation

The *Probability of X conditioned on A* is

$$Pr(X \mid A) = \frac{Pr(X \cap A)}{Pr(A)}.$$

Example: Roll two die and let X and Y be the respective number of dots they show. Then

$$\begin{aligned} Pr(X+Y=9 \mid Y \text{ is even}) &= \frac{Pr(X+Y=9 \text{ and } Y \in \{2,4,6\})}{Pr(Y \in \{2,4,6\})} \\ &= \frac{Pr(\,(X,Y) \in \{(5,4),(3,6)\}\,)}{1/2} \\ &= \frac{2/36}{1/2} = \frac{1}{9}. \end{aligned}$$

Some basic formulas are

$$E(X \mid A) = \sum_i i \cdot Pr(X = i \mid A)$$

and

$$E(X) = E(X \mid A) \cdot Pr(A) + E(X \mid \bar{A}) \cdot Pr(\bar{A})$$

where $\bar{A}$ is the event "A does not happen" which has $Pr(\bar{A}) = 1 - Pr(A)$.

<u>Example:</u> For the dice example above

$$E(X + Y \mid Y \text{ is even}) = \sum_i i \cdot Pr(X + Y = i \mid Y \text{ is even})$$

and

$$\begin{aligned} E(X+Y) &= E(X + Y \mid Y \text{ even}) \cdot Pr(Y \text{ even}) \\ &\quad + E(X + Y \mid Y \text{ odd}) \cdot Pr(Y \text{ odd}). \end{aligned}$$

Another Example: $C = x_1 \vee x_2 \vee \bar{x}_3$ and S a "random" truth assignment for the x_i as described before.

$$\begin{aligned} Pr(C \text{ satisfied} \,|\, x_1 = F) &= \frac{Pr(x_1 \vee x_2 \vee \bar{x}_3 \text{ satisfied and } x_1 = F)}{Pr(x_1 = F)} \\ &= \frac{Pr(x_2 \vee \bar{x}_3 \text{ satisfied and } x_1 = F)}{Pr(x_1 = F)} \\ &= \frac{Pr(x_2 \vee \bar{x}_3 \text{ satisfied}) \cdot Pr(x_1 = F)}{Pr(x_1 = F)} \\ &= Pr(x_2 \vee \bar{x}_3 \text{ satisfied}) \\ &= 1 - \left(\frac{1}{2}\right)^2 \\ &= \frac{3}{4}. \end{aligned}$$

But

$$\begin{aligned} Pr(C \text{ satisfied} \,|\, x_1 = T) &= \frac{Pr(x_1 \vee x_2 \vee \bar{x}_3 \text{ satisfied and } x_1 = T)}{Pr(x_1 = T)} \\ &= \frac{Pr(x_1 = T)}{Pr(x_1 = T)} \\ &= 1. \end{aligned}$$

Now let I be an instance of Max Sat, S a "random assignment" and W the weight of this random assignment. Recall that

$$E(W) = \sum_j w_j Pr(C_j \text{ satisfied by } S).$$

The notation $E(W \mid x_1, x_2, \ldots x_k)$ will mean the expected value of W conditioned on the truth assignment of $x_1, x_2, \ldots x_k$ being given. By linearity of expectation we have

$$E(W \mid x_1, x_2, \ldots x_k) = \sum_j w_j Pr(C_j \text{ satisfied by } S \mid x_1, x_2, \ldots x_k)$$

where the probabilities are conditioned on the truth assignment of $x_1, x_2, \ldots x_k$ being given.

Notice that this value can always be calculated in time polynomial in the size of the instance!

Derandomization of the MAX-SAT algorithm

We will now see how to "derandomize" the MAX-SAT algorithm from the last lecture. This will give us a deterministic method for *constructing* a truth assignment S such that the weight associated with S satisfies $W \geq \frac{OPT}{2}$.

The important fact to notice is that, for all $0 < k \leq n$

$$\begin{aligned} E(W \mid x_1, x_2, \ldots x_{k-1}) \;=\; & \frac{1}{2} E(W \mid x_1, x_2, \ldots x_{k-1}, x_k = T) \\ & + \frac{1}{2} E(W \mid x_1, x_2, \ldots x_{k-1}, x_k = F) \end{aligned}$$

DeRandomized MAX-SAT
For $i = 1$ to n do
Calculate $W_T = E(W \mid x_1, x_2, \ldots x_{k-1}, x_k = T)$
Calculate $W_F = E(W \mid x_1, x_2, \ldots x_{k-1}, x_k = F)$
if $W_T \geq W_F$
set $x_k = T$
else set $x_k = F$
Return variable setting S.

Notice that after setting x_k we must have (why?)

$$E(W \mid x_1, x_2, \ldots x_{k-1}) \leq E(W \mid x_1, x_2, \ldots x_k).$$

This implies that

$$\frac{OPT}{2} \leq E(W) \leq E(W \mid x_1, x_2, \ldots x_n)$$

so the truth assignment given by this deterministic algorithm is also a $\frac{1}{2}$ approximation of OPT.

MAX-SAT and p-Biased coins

A p-biased coin is one that, for $0 \le p \le 1$, has

$$Pr(\text{Heads}) = p, \qquad Pr(\text{Tails}) = 1 - p.$$

p-biased random MAX SAT
For $i = 1$ to n do
Flip a p-biased coin.
If "heads", set $x_i = T$.
else set $x_i = F$.
Return variable setting S.

Lemma: If $p \ge \frac{1}{2}$ and C is a clause that is not of the form "single variable negated" $(\bar{x}_i)$ then

$$Pr(C \text{ is satisfied by } S) \ge \min(p, 1 - p^2).$$

Proof: Let l be the number of variables in C.
If $l = 1$ then $C = x_i$ for some i and

$$Pr(C \text{ is satisfied by } S) \;=\; Pr(x_i = T) \;=\; p.$$

If $l > 1$ let n be the number of variables that appear negated in C. Then $l - n$ is the number that do not appear negated and

$$Pr(C \text{ is satisfied by } S) = 1 - p^n(1-p)^{l-n} \ge 1 - p^l \ge 1 - p^2.$$

Let $\phi = \frac{\sqrt{5}-1}{2} \approx 0.618$ be the solution in $[0,1]$ to $p = 1 - p^2$.

Lemma: Given an instance of MAX-SAT in which every length 1 clause is not negated let OPT be the optimal solution and W the solution returned by ϕ-biased random MAX SAT. Then

$$E[W] \geq \phi \cdot OPT.$$

Proof: Set

$$I_j = \begin{cases} 1 & \text{if } C_j \text{ is satisfied by } S \\ 0 & \text{otherwise} \end{cases}.$$

Then

$$\begin{aligned}
E[W] &= E\left[\sum_j w_j I_j\right] \\
&= \sum_j w_j E[I_j] \\
&= \sum_j w_j Pr(C_j \text{ is satisfied by } S) \\
&\geq \sum_j w_j \min(\phi, 1 - \phi^2) \\
&= \phi \sum_j w_j \\
&\geq \phi \cdot OPT.
\end{aligned}$$

The result on the previous page can be improved to work for *all* instances of MAX-SAT, even those that contain length one clauses that are negated variables.

Let I be our instance of MAX-SAT. First suppose that variable x_i appears in some length one negated clause, $C_j = \bar{x}_i$, but does not appear in any length one non-negated clause as $C_{j'} = x_i$. We then introduce a new variable u_i and replace every $\bar{x}_i$ in I by u_i and every instance of x_i in I by $\bar{u}_i$. We can then solve this new instance of MAX-SAT without worrying about the negated length one clause C_j (solutions to the original problem are in one-to-one correspondence to solutions to the new problem).

The only hard case to deal with is when x_i appears both as negated clause $C_{j'} = \bar{x}_i$ and non-negated clause $C_j = x_i$.

We may assume that $w_j \geq w_{j'}$ (why?).

For all i, if $\bar{x}_i$ exists as its own clause let $v_i =$ wt of $\bar{x}_i$. Otherwise $v_i = 0$.

Since a truth assignment can only satisfy *one* of x_i and $\bar{x}_i$ we have

$$OPT \leq \sum_j w_j - \sum_i v_i.$$

But then

$$\begin{aligned}
E[W] &= E\left[\sum_j w_j I_j\right] \\
&= \sum_j w_j Pr(C_j \text{ is satisfied by } S) \\
&\geq \sum_{j, \forall i,\, C_j \neq \bar{x}_i} w_j Pr(C_j \text{ is satisfied by } S) \\
&\geq \sum_{j, \forall i,\, C_j \neq \bar{x}_i} \phi w_j \\
&= \phi \cdot \left(\sum_j w_j - \sum_i v_i\right) \\
&\geq \phi \cdot OPT.
\end{aligned}$$

This gives a 0.618 approximation which is much better than the $\frac{1}{2}$ approximation we started with originally.

MAX-SAT and Randomized Rounding

We start by seeing how to restate the MAX-SAT problem. For every clause C_j let I_j^+ be the set of indices of variables that are not negated in C_j and I_j^- the set of indices of variables that are negated. For example, when $C_j = x_1 \vee \bar{x}_2 \vee x_3 \vee x_4 \vee \bar{x}_5$ then $I_j^+ = \{1, 3, 4\}$ and $I_j^- = \{2, 5\}$. Now construct the problem:

Maximize $\Sigma_j w_j z_j$
subject to conditions
$\forall j,\ \Sigma_{i \in I_j^+} y_i + \Sigma_{i \in I_j^-}(1 - y_i) \geq z_j$
$\forall j,\ z_j \in \{0, 1\}$
$\forall i,\ y_i \in \{0, 1\}$

The idea now is to construct a correspondence between truth assignment S and the values of y_i by setting
$y_i = 1$ if and only if $x_i = T$ and $y_i = 0$ iff $x_i = F$.

Note that with this assignment S satisfies C_j if and only if $\Sigma_{i \in I_j^+} y_i + \Sigma_{i \in I_j^-}(1 - y_i) \geq 1$. For this fixed assignment the objective function $\Sigma_j w_j z_j$ is maximized if we set $z_j = 1$ for every C_j that is satisfied (the unsatisfied C_j have $z_j = 0$). This implies that the objective function $\Sigma_j w_j z_j$ is maximized when the y_i are assigned values corresponding to a MAX-SAT assignment.

Maximize $\Sigma_j w_j z_j$
subject to conditions
$\forall j,\ \Sigma_{i \in I_j^+} y_i + \Sigma_{i \in I_j^-}(1 - y_i) \geq z_j$
$\forall j,\ z_j \in \{0, 1\}$
$\forall i,\ y_i \in \{0, 1\}$

How does rewriting the problem help???

The problem as we wrote it is an *integer linear program* and solving integer linear programs is NP-Hard.

But, if we relax the problem and no longer require the y_i and z_j to be integers we get a regular *linear program*. Linear programs can be solved in polynomial time using off-the-shelf software.

Relaxed LP
Maximize $\Sigma_j w_j z_j$
subject to conditions
$\forall j,\ \Sigma_{i \in I_j^+} y_i + \Sigma_{i \in I_j^-}(1 - y_i) \geq z_j$
$\forall j,\ 0 \leq z_j \leq 1$
$\forall i,\ 0 \leq y_i \leq 1$

We will use the notation (y^*, z^*) to denote the variables in the optimal solution for the relaxed LP. Note that if $z_{LP} = \Sigma_j w_j z_j^*$ is the optimum (maximum) calculated in the relaxed linear program and OPT is the optimal solution for MAX-SAT then

$$z_{LP} \geq OPT.$$

The *Randomized-Rounding* MAX-SAT algorithm is:

Randomized Rounding
Solve the Relaxed Linear Program
Calculate (y^*, z^*)
for $i = 1$ to n do
Flip a y_i^*-biased coin
If Heads set $x_i = T$
else set $x_i = F$.

<u>Lemma:</u> Let W be the weight of the assignment created by *Randomized Rounding*. Then

$$E(W) \geq \left(1 - \frac{1}{e}\right) OPT$$

where $1 - \frac{1}{e} \approx 0.632$.

The proof of the lemma will need two facts.
The first is that if $a_i \geq 0$ then

$$\forall k,\ (a_1 a_2 \ldots a_k)^{\frac{1}{k}} \ \leq \frac{1}{k}\left(a_1 + a_2 + \ldots + a_k\right).$$

The second fact is that if function $f(x)$ is concave on $[l, u]$,
($f(x)$ is concave on $[l, u]$ if $f''(x) \leq 0$ for all $x \in [l, u]$)
$f(l) \geq al + b$, and $f(u) \geq au + b$ then

$$\forall x \in [l, u],\ f(x) \geq ax + b.$$

We will apply this to

$$f(x) = 1 - \left(1 - \frac{x}{k}\right)^k$$

on $[0, 1]$. This $f()$ is concave on $[0, 1]$,

$$f(0) = 0, \qquad f(1) = 1 - \left(1 - \frac{1}{k}\right)^k,$$

so for $x \in \{0, 1\}$, $f(x) \geq \left(1 - \left(1 - \frac{1}{k}\right)^k\right) x$.

This implies that

$$\forall x \in [0, 1],\ f(x) \geq \left(1 - \left(1 - \frac{1}{k}\right)^k\right) x.$$

We start by considering a clause C_j that has only all positive variables, e.g., $C_j = x_1 \vee x_2 \vee \ldots \vee x_k$. Since $I_j^- = \emptyset$ the LP constraint for C_j was $\Sigma_{i=1}^k y_i^* \geq z_j^*$.
From the algorithm $Pr(x_i = T) = y_i^*$ so

$$\begin{aligned}
Pr(C_j \text{ is satisfied}) &= 1 - \prod_{i=1}^{k} (1 - y_i^*) \\
&\geq 1 - \left(\frac{k - \Sigma_{i=1}^k y_i^*}{k} \right)^k \\
&\geq 1 - \left(1 - \frac{z_j^*}{k} \right)^k \\
&\geq \left(1 - \left(1 - \frac{1}{k} \right)^k \right) z_j^*.
\end{aligned}$$

The first inequality comes from the first fact on the previous page, the second from $\Sigma_{i=1}^k y_i^* \geq z_j^*$, and the third from the second fact on the previous page.

Now consider a clause C_j of the form

$$C_j = x_1 \vee x_2 \ldots \vee x_l \vee \bar{x}_{l+1} \ldots \vee \bar{x}_k.$$

The LP constraint for C_j was $\Sigma_{i=1}^{l} y_i^* + \Sigma_{i=l+1}^{k}(1 - y_i^*) \geq z_j^*$. Now

$$\begin{aligned} Pr(C_j \text{ is satisfied}) &= 1 - \prod_{i=1}^{l}(1 - y_i^*) \prod_{i=l+1}^{k} y_i^* \\ &\geq 1 - \left(\frac{l - \Sigma_{i=1}^{l} y_i^* + \Sigma_{i=l+1}^{k} y_i^*}{k}\right)^k \\ &\geq 1 - \left(1 - \frac{z_j^*}{k}\right)^k \\ &\geq \left(1 - \left(1 - \frac{1}{k}\right)^k\right) z_j^*. \end{aligned}$$

The only new aspect of this derivation is in the second equality. It comes from the fact that

$$\begin{aligned} l - \sum_{i=1}^{l} y_i^* + \sum_{i=l+1}^{k} y_i^* &= k - \sum_{i=1}^{l} y_i^* - \sum_{i=l+1}^{k}(1 - y_i^*) \\ &\leq k - z_j^*. \end{aligned}$$

We have actually just proven that if C_j is a clause with l_j variables then

$$Pr(C_j \text{ is satisfied}) \geq \left(1 - \left(1 - \frac{1}{l_j}\right)^{l_j}\right) z_j^*.$$

Thus

$$\begin{aligned}
E(W) &= \sum_j w_j Pr(C_j \text{ is satisfied}) \\
&\geq \sum_j w_j \left(1 - \left(1 - \frac{1}{l_j}\right)^{l_j}\right) z_j^* \\
&\geq \min_j \left(1 - \left(1 - \frac{1}{l_j}\right)^{l_j}\right) \sum_j w_j z_j^* \\
&\geq \left(1 - \frac{1}{e}\right) OPT
\end{aligned}$$

which is what we have been attempting to prove.
We used the facts that $\forall x \geq 0$, $\left(1 - \frac{1}{x}\right)^x \leq e^{-1}$ and that $\Sigma_j w_j z_j^* \geq OPT$.