Section 5.4: Randomized rounding

In Section 5.3 we saw that biasing the prob. of setting each variable true resulted in a better approx. guarantee.
The approximation ratio can be further improved by allowing a different bias for each variable
We will develop an LP-formulation of the problem

For each clause, C_j, we define:
P_j : the set of indices of variables that occur positively in C_j
N_j : —————— " —————————— negatively —"—
Then, C_j can be written as

$$\bigvee_{i \in P_j} x_i \vee \bigvee_{i \in N_j} \overline{x_i}$$

If $y_i = 0$ corresponds to x_i being false and $y_i = 1$ corresponds to y_i being true, then C_j is true, if

$$\sum_{i \in P_j} y_i + \sum_{i \in N_j} (1 - y_i) \geq 1$$

This leads to the following IP-formulation:

$$\max \sum_{j=1}^{m} z_j w_j$$

subject to

$$\sum_{i \in P_j} y_i + \sum_{i \in N_j} (1 - y_i) \geq z_j \,, \qquad 1 \leq j \leq m$$

$$y_i \in \{0,1\}, \qquad 1 \leq i \leq n$$

$$z_j \in \{0,1\}, \qquad 1 \leq j \leq m$$

Clearly, $Z^*_{LP} \geq Z^*_{IP} = OPT$

value of opt. sol. to LP-relax.

value of opt. sol. to IP

value of opt. sol. to corresponding MAXSAT problem

RandRounding (ϕ)

$(\vec{y^*}, \vec{z^*}) \leftarrow$ opt. sol. to LP-relax. corresponding to ϕ
For $i \leftarrow 1$ to n
Set x_i true with prob. y_i^*

The approx. ratio of RandRounding is at least $1 - \frac{1}{e} \approx 0.632$.
For proving this, we will use the following two facts:

Fact 5.8 (Arithmetic-geometric mean inequality):

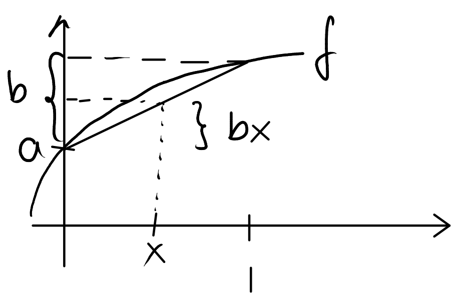
For any $a_1, a_2, \dots, a_k \geq 0$,

$$\left(\prod_{i=1}^{k} a_i\right)^{1/k} \leq \frac{1}{k}\sum_{i=1}^{k} a_i$$

A function f is concave on an interval I, if $f''(x) \leq 0$ for any $x \in I$.

Fact 5.9:

$$\left.\begin{array}{l} f \text{ is concave on } [0,1] \\ f(0) = a, \quad f(1) = a+b \end{array}\right\} \Rightarrow f(x) \geq a+bx, \text{ for } x \in [0,1]$$



$\boxed{\text{Theorem 5.10: RandRounding is a } (1-\tfrac{1}{e})\text{-approx. alg}}$

Proof:

For each clause C_j, the prob. that C_j is not satisfied is

by Fact 5.8

$$\prod_{i\in P_j}(1-y_i^*)\prod_{i\in N_j} y_i^* \leq \left(\frac{1}{\ell_j}\left(\sum_{i\in P_j}(1-y_i^*)+\sum_{i\in N_j} y_i^*\right)\right)^{\ell_j}$$

$$= \left(\frac{1}{\ell_j}\left(|P_j| - \sum_{i\in P_j} y_i^* + |N_j| - \sum_{i\in N_j}(1-y_i^*)\right)\right)^{\ell_j}$$

$$= \left(1 - \frac{1}{\ell_j}\left(\sum_{i\in P_j} y_i^* + \sum_{i\in N_j}(1-y_i^*)\right)\right)^{\ell_j}$$

$$\leq \left(1-\frac{z_j^*}{\ell_j}\right)^{\ell_j}$$

Thus, the prob. that C_j is satisfied is at least

$$f(z_j^*) \equiv 1 - \left(1-\frac{z_j^*}{\ell_j}\right)^{\ell_j},$$

where f is concave on $[0,1]$:

$$f'(z_j^*) = -\ell_j\left(1-\frac{z_j^*}{\ell_j}\right)^{\ell_j-1}\cdot\left(-\frac{1}{\ell_j}\right) = \left(1-\frac{z_j^*}{\ell_j}\right)^{\ell_j-1}$$

$$f''(z_j^*) = (\ell_j-1)\left(1-\frac{z_j^*}{\ell_j}\right)^{\ell_j-2}\cdot\left(-\frac{1}{\ell_j}\right) = \underbrace{\left(\frac{1}{\ell_j}-1\right)}_{\leq 0}\underbrace{\left(1-\frac{z_j^*}{\ell_j}\right)^{\ell_j-2}}_{\geq 0}$$

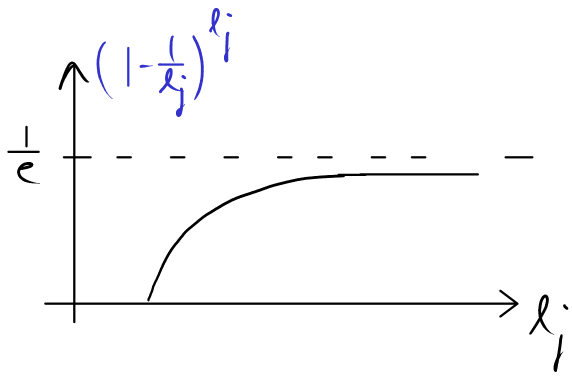
$$\leq 0$$

Thus, the prob. that C_j is satisfied is at least

$$1 - \left(1 - \frac{z_j^*}{\ell_j}\right)^{\ell_j} \overset{\text{by Fact 5.9}}{\geq} 0 + \left(\left(1 - \left(1 - \frac{1}{\ell_j}\right)^{\ell_j}\right) - 0\right) \cdot z_j^*$$
$$= \left(1 - \left(1 - \frac{1}{\ell_j}\right)^{\ell_j}\right) \cdot z_j^*$$

Hence,

$$E[\text{RandRounding}] \geq \sum_{j=1}^{m} \left(1 - \left(1 - \frac{1}{\ell_j}\right)^{\ell_j}\right) z_j^* w_j$$
$$\geq \left(1 - \frac{1}{e}\right) \cdot \underbrace{\sum_{j=1}^{m} z_j^* w_j}_{= Z_{LP}^* \geq OPT}$$



$\square$

Note that RandRounding can be derandomized exactly like Rand and Randp.

Section 5.5 : Choosing the better of two solutions

Combining the alg.s of Sections 5.1 and 5.4 gives a better expected performance than using any one of them separately.

This is because they have different worst-case inputs:

Rand satisfies clause C_j with prob. $p_R = 1-(\frac{1}{2})^{-\ell_j}$.

RandRounding sat. clause C_j with prob. $p_{RR} = (1-(1-\frac{1}{\ell_j})^{\ell_j}) z_j^*$.

p_R increases with ℓ_j and p_{RR} decreases with ℓ_j

BestOfTwo(ϕ)

$\vec{x}_R \leftarrow$ Rand(ϕ)
$\vec{x}_{RR} \leftarrow$ RandRounding(ϕ)
If $w(\phi, \vec{x}_R) \geq w(\phi, \vec{x}_{RR})$
    Return $\vec{x}_R$
Else
    Return $\vec{x}_{RR}$

Note that BestOfTwo is derandomized by using the derandomized versions of Rand and RandRounding.

Theorem 5.11: BestOfTwo is a 3/4-approx. alg.

Proof:

$$E[\text{BestOfTwo}(\phi)] = E[\max\{\text{Rand}(\phi), \text{RandRounding}(\phi)\}]$$

$$\geq E\left[\tfrac{1}{2}\text{Rand}(\phi) + \tfrac{1}{2}\text{RandRounding}(\phi)\right]$$

$$= \tfrac{1}{2}E[\text{Rand}(\phi)] + \tfrac{1}{2}E[\text{RandRounding}(\phi)], \text{ by lin. of exp.}$$

$$\geq \tfrac{1}{2}\sum_{j=1}^{m}(1-2^{-l_j})w_j + \tfrac{1}{2}\sum_{j=1}^{m}\left(1-\left(1-\tfrac{1}{l_j}\right)^{l_j}\right)z_j^*w_j$$

$$\geq \sum_{j=1}^{m} z_j^* w_j \cdot \underbrace{\tfrac{1}{2}\left(1-2^{-l_j} + 1-\left(1-\tfrac{1}{l_j}\right)^{l_j}\right)}_{=\rho_j}, \text{ since } z_j^* \leq 1.$$

For $l_j = 1$, $\rho_j = \frac{1}{2}\left(1-\frac{1}{2}+1-0\right) = \frac{3}{4}$

For $l_j = 2$, $\rho_j = \frac{1}{2}\left(1-\frac{1}{4}+1-\left(1-\frac{1}{2}\right)^2\right) = \frac{3}{4}$

For $l_j \geq 3$, $\rho_j \geq \frac{1}{2}\left(1-\frac{1}{8}+1-\frac{1}{e}\right) > \frac{3}{4}$

Hence,

$$E[\text{BestOfTwo}] \geq \sum_{j=1}^{m} z_j^* w_j \cdot \frac{3}{4} \geq \frac{3}{4}\cdot OPT \qquad \square$$

Section 5.6: Nonlinear randomized rounding

RandRounding$_f(\phi)$

$(\vec{y^*}, \vec{z^*}) \leftarrow$ opt. sol. to LP-relax. corresponding to ϕ
For $i \leftarrow 1$ to n
    Set x_i true with prob. $f(y_i^*)$

Theorem 5.12

RandRounding$_f$ is a 3/4-approx. alg., if $1-4^{-x} \leq f(x) \leq 4^{x-1}$

<u>Proof</u>:
Prob. that C_j is <u>not</u> satisfied:

$$\prod_{i \in P_j} (1 - f(y_i^*)) \prod_{i \in N_j} f(y_i^*) \leq \prod_{i \in P_j} 4^{-y_i^*} \prod_{i \in N_j} 4^{y_i^* - 1}$$

$$= 4^{-\left(\sum_{i \in P_j} y_i^* + \sum_{i \in N_j} 1 - y_i^*\right)}$$

$$\leq 4^{-z_j^*}$$

Prob. that C_j is satisfied:

$$\geq 1 - 4^{-z_j^*}$$

$$\geq 0 + \left(\tfrac{3}{4} - 0\right) z_j^*, \quad \text{by Fact 5.9}$$

$$= \tfrac{3}{4} z_j^*$$

$\square$

Ex: $(x_1 \vee x_2) \wedge (x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee x_2) \wedge (\bar{x}_1 \vee \bar{x}_2)$

$w_1 = w_2 = w_3 = w_4 = 1$

OPT $= 3$

$y_1 = y_2 = y_3 = \frac{1}{2} \Rightarrow Z = 4$

Hence, the integrality gap for the IP problem for MAXSAT is at most

$$\frac{OPT}{Z} = \frac{3}{4}$$

On the other hand, the proof that $Rand_f$ is a $3/4$-approx. alg shows that for any instance of MAXSAT, there is a sol. to the IP version which is at least $3/4$ as good as the opt. sol. to the LP problem.

Hence, the integrality gap is at least $3/4$.

The upper bound of $3/4$ on the integrality gap shows that no alg. whose approx. guarantee is based on a comparison to the opt. LP sol. can have a better approx. guarantee than $3/4$.