## Randomized Rounding for Online Set Cover

Instructor: Thomas Kesselheim
In this lecture, we will see how the fractional solutions that we got for the Set Cover LP can be turned into integral ones. We will do this in a randomized way by interpreting fractional values as probabilities.

## 1 Rounding Fractional Solutions for Set Cover

Our algorithm is as follows.
A priori, we choose for each $S \in \mathcal{S}$ a threshold $\theta_{S}$ uniformly from $[0,1]$.
Upon arrival of an element $e$, update $\left(x_{S}^{(t)}\right)_{S \in \mathcal{S}}$.
(a) Pick all sets $S$, for which $x_{S}^{(t)} \geq \frac{1}{2 \ln t} \theta_{S}$.
(b) If $e$ is still uncovered, choose one set $S$ from probability distribution defined by $\left(x_{S}^{(t)}\right)_{S: e \in S}$ and pick it. Note that this is possible since $\sum_{S: e \in S} x_{S}^{(t)} \geq 1$.
Theorem 4.1. Using a c-competitive algorithm for the fractional problem, the algorithm for the integral problem is $O(c \cdot \log m)$-competitive.
Lemma 4.2. For every element e, the probability that part (b) is executed is at most $\frac{1}{t^{2}}$.
Proof. Let $X_{S}^{(t)}=1$ if set $S$ has been picked until round $t$ because $x_{S}^{(t)} \geq \frac{1}{2 \ln t} \theta_{S}$, otherwise set $X_{S}^{(t)}=1$.

By design

$$
\operatorname{Pr}\left[X_{S}^{(t)}=1\right]=\operatorname{Pr}\left[x_{S}^{(t)} \geq \frac{1}{2 \ln t} \theta_{S}\right]=\operatorname{Pr}\left[\theta_{S} \geq(2 \ln t) x_{S}^{(t)}\right]=(2 \ln t) x_{S}^{(t)} .
$$

As $1-q \leq \exp (-q)$ for $q \in[0,1]$, this implies

$$
\operatorname{Pr}\left[X_{S}^{(t)}=0\right]=1-(2 \ln t) x_{S}^{(t)} \leq \exp \left(-(2 \ln t) x_{S}^{(t)}\right)
$$

Note that these choices are independent, so
$\operatorname{Pr}\left[\bigwedge_{S: e \in S} X_{S}^{(t)}=0\right]=\prod_{S: e \in S} \operatorname{Pr}\left[X_{S}^{(t)}=0\right] \leq \prod_{S: e \in S} \exp \left(-(\ln m) x_{S}^{(t)}\right)=\exp \left(-(2 \ln t) \prod_{S: e \in S} x_{S}^{(t)}\right) \leq \frac{1}{t^{2}}$.

Lemma 4.3. The expected cost due to set $S$ within the first $m$ rounds is at most $\left(\sum_{t=1}^{m} \frac{1}{t^{2}}+\right.$ $2 \ln m) c_{S} x_{S}^{(t)}$.
Proof. Let $X_{t, S}=1$ if set $S$ is chosen in part (b) statement of step $t$. Note that we have $\mathbf{E}\left[X_{t, S}\right] \leq \frac{1}{t^{2}} \cdot x_{S}^{(t)} \leq \frac{1}{t^{2}} \cdot x_{S}^{(m)}$.

The expected cost due to set $S$ within the first $m$ rounds is
$\mathbf{E}\left[\sum_{t=1}^{m} c_{S} X_{t, S}+c_{S} X_{S}^{(m)}\right]=\sum_{t=1}^{m} c_{S} \mathbf{E}\left[X_{t, S}\right]+c_{S} \mathbf{E}\left[X_{S}^{(m)}\right] \leq t \cdot c_{S} \sum_{t=1}^{m} \frac{1}{t^{2}} x_{S}^{(m)}+c_{S}(\ln m) x_{S}^{(m)}$.

Proof of Theorem 4.1. We consider the outcome after $m$ rounds. Let $x^{*}$ be an optimal offline fractional solution to the LP relaxation. As the fractional algorithm is c-competitive, we have $\sum_{S \in \mathcal{S}} c_{S} x_{S}^{(m)} \leq c \cdot \sum_{S \in \mathcal{S}} c_{S} x_{S}^{*}$. Furthermore, by Lemma 4.3, $\mathbf{E}[\operatorname{cost}(\operatorname{ALG}(\sigma))] \leq$ $\sum_{S \in \mathcal{S}}\left(\sum_{t=1}^{m} \frac{1}{t^{2}}+\ln m\right) c_{S} x_{S}^{(m)}$.

Note that $\sum_{t=1}^{\infty} \frac{1}{t^{2}}=\frac{\pi^{2}}{6}$. So overall, $\mathbf{E}[\operatorname{cost}(\operatorname{ALG}(\sigma))] \leq c \cdot\left(\frac{\pi^{2}}{6}+\ln m\right) \sum_{S \in \mathcal{S}} c_{S} x_{S}^{*} \leq$ $c \cdot\left(\frac{\pi^{2}}{6}+\ln m\right) \operatorname{cost}(\mathrm{OPT}(\sigma))$.

## 2 Ski Rental

Let us come back to the ski-rental problem. Suppose you want to go skiing. You can either rent skis (cost 1 per day) or buy them (cost $B$ once). The problem is that you do not what which option is cheaper because you do not know the number of days you will go skiing in advance.

This problem is a very simple special case of online set cover. There are $m$ elements ( $m$ is unknown beforehand). The set $\mathcal{S}$ consists of the singleton sets, with a cost of 1 each, and the set that covers all $m$ elements for a cost of $B$. In each round, we observe an element and have to make sure it is covered.

We can use the same algorithmic ideas as for set cover to design a randomized $\frac{e}{e-1}$-competitive algorithm. First, again we solve the fractional problem with a deterministic algorithm. It follows the multiplicative-weights ideas.

Theorem 4.4. There is a $\frac{e}{e-1}$-competitive deterministic algorithm for fractional Set Cover instances as derived from Ski Rental.

We will skip the proof here.
The randomized rounding follows the same idea we had before and is very simple here. We have a primal variable $x_{\text {buy }}$ that corresponds to buying the skis that is increased over time by the fractional algorithm. We draw $\theta$ uniformly from $[0,1]$ beforehand and buy the skis as soon as $x_{\text {buy }}^{(t)} \geq \theta$.

Theorem 4.5. Given a c-competitive algorithm for fractional Set Cover instances as derived from Ski Rental that never increases variables by more than is necessary. Then the above algorithm is a randomized c-competitive algorithm for Ski Rental.

Proof. Let $Z_{\text {buy }}$ be the random variable indicating the cost of buy skis within the first $m$ steps. We do so with probability $x_{\text {buy }}^{(m)}$. The expected cost from buying in the first $m$ steps is $\mathbf{E}\left[Z_{\text {buy }}\right]=B \cdot x_{\text {buy }}^{(m)}$.

Let $Z_{\text {rent }, t}$ be the random variable indicating the cost of renting skis in the $t$-th step. We rent skis with probability $1-x_{\text {buy }}^{(t)}$. Note that $x_{\text {buy }}^{(t)}+x_{\text {rent }, t}^{(t)}=1$ because this is the new constraint that has to be fulfilled and the algorithm does not increase the variables by more than is necessary. In other words, we rent skis with probability at most $x_{\mathrm{rent}, t}^{(t)}$. So, $\mathbf{E}\left[Z_{\mathrm{rent}, t}\right] \leq x_{\mathrm{rent}, t}^{(t)} \leq x_{\mathrm{rent}, t}^{(m)}$.

Overall, our expected cost is

$$
\mathbf{E}\left[Z_{\mathrm{buy}}+\sum_{t=1}^{m} Z_{\mathrm{rent}, t}\right] \leq B \cdot x_{\mathrm{buy}}^{(m)}+\sum_{t=1}^{m} x_{\mathrm{rent}, t}^{(m)} \leq c\left(B \cdot x_{\mathrm{buy}}^{*}+\sum_{t=1}^{m} x_{\mathrm{rent}, t}^{*}\right) \leq \operatorname{cost}(\mathrm{OPT}(\sigma)),
$$

where $x^{*}$ is an optimal offline solution.

