CSC 373 Lecture 26

Announcements:

   So far four requests for TA office hour. Will announce TA office hour (starting this week) on web page.

   Test graded out 45 with 50 being maximum obtainable (and obtained).

Today

Answer to question about one constraint IP

• Continue IP/LP rounding.
  – f-frequency set cover
  – Start makespan on unrelated machines.
NP hardness of IP with one constraint

• Lets consider say a minimization problem in the form: 
  \[ \text{min} \ \sum c_i x_i \ \text{subject to a single constraint:} \ \sum a_i x_i \ \text{R} \ \ b_i \]  
  where R could be = or >=. We also have \( x \geq 0 \). Lets just consider the case that \( b \) and all \( a_i \) are positive integers.

• If R is =, then just to determine if there is any feasible solution is NP hard since we then have an integer (rather than 0-1) version of the subset sum problem. But the proof of the transformation of 3SAT to Subset-Sum also shows that the integer version is also NP-hard.

• If R is >=, then determining feasibility is easy. But if we want to minimize the objective \( \sum a_i x_i \) then we are again solving the integer Subset-Sum problem.
**Figure 34.19 of CLRS**

The reduction of 3-CNF-SAT to SUBSET-SUM. The formula in 3-CNF is $\phi = C_1 \land C_2 \land C_3 \land C_4$, where $C_1 = (x_1 \lor \neg x_2 \lor \neg x_3)$, $C_2 = (\neg x_1 \lor \neg x_2 \lor \neg x_3)$, $C_3 = (\neg x_1 \lor \neg x_2 \lor x_3)$, and $C_4 = (x_1 \lor x_2 \lor x_3)$. A satisfying assignment of $\phi$ is $(x_1 = 0, x_2 = 0, x_3 = 1)$. The set $S$ produced by the reduction consists of the base-10 numbers shown; reading from top to bottom, $S = \{1001001, 1000110, 100001, 101110, 10011, 11100, 1000, 2000, 100, 200, 10, 5, 0\}$. 

<table>
<thead>
<tr>
<th>$v'_1$</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v'_2$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$v_3$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$s_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$s'_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$s'_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$s'_4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$t$</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$S_

Set cover and $f$-frequency set cover

• We are given a collection of (possibly weighted) sets $C = \{S_1, \ldots, S_n\}$ over a universe $U$. The set cover problem is to find a minimal size (weight) subcollection $C'$ that covers all the elements in $U$.

• Set cover generalizes vertex cover and turns out to be hard to approximate (given well believed assumptions about NP) better than $H_m$ where $m = |U|$ is the size of the Universe. There is a natural greedy algorithm that will achieve an approximation of $H_d$ where $d = \max_i |S_i|$.

• $f$-frequency set cover and vertex cover as 2-frequency set cover problem with $U = E$ and sets $S_i = \{e \mid e$ adjacent to vertex $v_i\}$.
The IP/LP for f-frequency set cover

We have essentially the same IP/LP rounding algorithm for the f-frequency set cover problem. Minimize sum $w_i * x_i$ subj to

$$\text{Sum}_{i: \ u_j \ in \ S_i} x_i \geq 1 \ for \ each \ u_j \ in \ U; \ x_i \ in \ \{0,1\}.$$ 

The meaning is that $x_i = 1$ iff set $S_i$ is in the cover.

The LP relaxation is to relax the integrality condition to $x_i \geq 0$. Again, it follows that an optimal LP solution also satisfies $x_i \leq 1$.

Suppose $x^*$ is an LP optimum. We apply the naive rounding $x'_i = 1$ iff $x^*_i \geq 1/f$. 

•
IP/LP with a non naive rounding.

• The makespan problem for the unrelated machines model. The input consists of a given $m$ (the number of machines) and $n$ jobs $J_1,...,J_n$ where each job $J_j$ is represented by a vector $<p_{1j},p_{2j},...,p_{mj}>$ where $p_{ij}$ represents the processing time of job $J_j$ on machine $i$. WLG $m <= n$.

• We will sketch a 2-approximation IP/LP rounding algorithm. This is the best known poly time approximation and it is known that it is NP hard to achieve better than $3/2$ approximation even for the special case of the restrictive machines model for which every $p_{ij}$ is either some $p_j$ or infinity.

• Note: Unlike identical machines case, I do not know of any greedy or local search or DP O(1) approx alg.
In the IP formulation, the problem is:

\[
\begin{align*}
\text{minimize} \quad & t \\
\text{subject to} \quad & \sum_{1 \leq l \leq m} x_{i,j} = 1 \quad \text{for each job } J_j. \\
& \sum_{1 \leq j \leq n} p_{ij} x_{i,j} \leq t \quad \text{for each machine.} \\
& x_{i,j} \in \{0,1\} \quad \text{The intended meaning is that} \\
& x_{i,j} = 1 \quad \text{iff job } J_j \text{ is scheduled on machine } i.
\end{align*}
\]

The LP relaxation is that $0 \leq x_{ij} ; (\leq 1 \text{ implied})$

The integrality gap is unbounded! Consider one job with processing time $m$, which has $OPT = m$ and $OPT_{LP} = 1$. 
Getting around the integrality gap

• The IP must set $x_{i,j} = 0$ if $p_{i,j} > t$ whereas the fractional OPT does not have this constraint. We want to say for all $(i,j)$: “if $p_{i,j} > t$ then $x_{i,j} = 0$”

But this isn’t a linear constraint!

Since we are only hoping for a good approx, we can assume all $p_{ij}$ are integral. We can then use binary search to find the best LP bound $T$ by solving the search problem $LP(T)$ for fixed $T$ eliminating the objective function and then removing any $x_{i,j}$ having $p_{i,j} > T$. We clearly have that $IP-OPT >= T$. 