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: *min sum c_i x_i subject to a single constraint: sum a_i x_i R b_i* where R could be = or >=. We also have ***x* >=** 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

[redacted]								
v_1'	=	1	0	0	0	1	1	0
[redacted]								
v_2'	=	0	1	0	1	1	1	0
v_3	=	0	0	1	0	0	1	1
[redacted]								
s_1	=	0	0	0	1	0	0	0
s_1'	=	0	0	0	2	0	0	0
[redacted]								
s_2'	=	0	0	0	0	2	0	0
s_3	=	0	0	0	0	0	1	0
[redacted]								
s_4	=	0	0	0	0	0	0	1
s_4'	=	0	0	0	0	0	0	2
t	=	1	1	1	4	4	4	4

Figure 34.19 The reduction of 3-CNF-SAT to SUBSET-SUM. The formula in 3-CNF is $\phi = C_1 \wedge C_2 \wedge C_3 \wedge C_4$, where $C_1 = (x_1 \vee \neg x_2 \vee \neg x_3)$, $C_2 = (\neg x_1 \vee \neg x_2 \vee \neg x_3)$, $C_3 = (\neg x_1 \vee \neg x_2 \vee x_3)$, and $C_4 = (x_1 \vee x_2 \vee x_3)$. A satisfying assignment of ϕ is $\langle x_1 = 0, x_2 = 0, x_3 = 1\rangle$. The set S produced by the reduction consists of the base-10 numbers shown; reading from top to bottom, $S = \{1001001, 1000110, 100001, 101110, 10011, 11100, 1000, 2000, 100, 200, 10, 20, 1, 2\}$. The

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

$Sum_{\{i: u_j \in S_i\}} x_i >= 1$ *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 $x_i >= 0$. Again, it follows that an optimal LP solution also satisfies $x_i <=1$.

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

•

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 $J_1,\ldots,J_n$ where each job J_j is represented by a vector $<p_{1j},p_{2j},\ldots,p_{mj}>$ 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:

minimize t subject to

$\sum_{1 \le I \le m} x_{i,j} = 1$ for each job J_j.

$\sum_{1 \le j \le n} p_{ij} x_{i,j} \le t$ for each machine.

$x_{i,j}$ in {0,1} The intended meaning is that $x_{i,j} = 1$ iff job J_j is scheduled on machine i.

The LP relaxation is that $0 \le x_{ij}$; (<=1 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 $IP\text{-}OPT >= T$.