Algorithms: Design and Analysis, Part II

Approximation Algorithms for NP-Complete Problems

Analysis of a Dynamic Programming Heuristic for Knapsack

The Dynamic Programming Heuristic

Step 1: set $\hat{v}_i = \left\lfloor \frac{v_i}{m} \right\rfloor$ for every item i.

Step 2: compute optimal solution with respect to the $\hat{v}_i$'s using dynamic programming.

Plan for analysis:

(1) figure out how big we can take m, subject to achieving a $(1-\epsilon)$-approximation

(2) plug in this value of m to determine running time

Quiz

Question: Suppose we round v_i to the value $\tilde{v}_i$. Which of the following is true?

(A) $\hat{v}_i$ is between $v_i - m$ and v_i

(B) $\hat{v}_i$ is between v_i and $v_i + m$

(C) $m \cdot \hat{v}_i$ is between $v_i - m$ and v_i

(D) $m \cdot \hat{v}_i$ is between $v_i - 1$ and v_i

Accuracy Analysis I

From quiz: since we rounded down to the nearest multiple of m, $m \cdot \hat{v}_i \in [v_i - m, v_i]$ for each item i.

Thus: (1) $v_i \geq m \cdot \hat{v}_i$ (2) $m \cdot \hat{v}_i \geq v_i - m$

Also: if S^* = optimal solution to the original problem (with the original v_i's), and S = our heuristic's solution, then

(3) $$\sum_{i \in S} \hat{v}_i \geq \sum_{i \in S^*} \hat{v}_i$$

[since S is optimal for the $\hat{v}_i$'s]

(recall Step 2)

Accuracy Analysis II

S = our soln

S^* = optimal soln

$$\sum_{i\in S} v_i \overset{①}{\geq} m\sum_{i\in S}\hat{v}_i \overset{③}{\geq} m\sum_{i\in S^*}\hat{v}_i \overset{②}{\geq} \sum_{i\in S^*}(v_i - m)$$

contains at most n items

Thus: $$\sum_{i\in S} v_i \geq \left(\sum_{i\in S^*} v_i\right) - mn$$

Constraint: $$\sum_{i\in S} v_i \geq (1-\varepsilon)\cdot\sum_{i\in S^*} v_i$$

To achieve this: Choose m small enough that $mn \leq \varepsilon \sum_{i\in S^*} v_i$

unknown to algorithm, but definitely $\geq v_{max}$

Sufficient: set m so that $mn = \varepsilon v_{max}$

i.e., heuristic uses the value $m = \frac{\varepsilon v_{max}}{n}$

Running Time Analysis

Point: setting $m = \frac{\varepsilon v_{max}}{n}$ guarantees that value of our solution is $\geq (1-\varepsilon)\cdot$ value of optimal solution.

Recall: running time is $O(n^2 \hat{v}_{max})$.

Note: for every item i, $\hat{v}_i \leq \frac{v_i}{m} \leq \frac{v_{max}}{m} = \cancel{v_{max}} \cdot \frac{n}{\varepsilon \cancel{v_{max}}} = \frac{n}{\varepsilon}$

$\Rightarrow$ running time $= O(n^3/\varepsilon)$