On the approximation ratio of Greedy Knapsack

Pekka Kilpeläinen
University of Eastern Finland
School of Computing (Kuopio)

October 1, 2010

Abstract

We justify the claim that the greedy algorithm for knapsack optimization provides 1/2-approximations, that is, solutions whose value is at least half of the optimum. Further, we see that this bound is tight. (**Source**: Section 13.2.2 in G. Brassard and P. Bratley, *Fundamentals of Algorithmics*, Prentice Hall, 1996.)

Let T be the volume of the knapsack, and $L = (a_1, \ldots, a_n)$ be a list of items sorted in decreasing unit-value order, that is,

$$value(a_i)/volume(a_i) \geq value(a_{i+1})/volume(a_{i+1})$$

for each $i \in \{1, 2, \ldots, n-1\}$. We can assume that each of the items fits by itself within volume T, since too large items are excluded by the `greedyKnapsack` procedure anyhow.

For any list of items A, let $value(A)$ and $volume(A)$ denote, respectively, the total value and the total volume of the items in A.

If $volume(L) \leq T$, the `greedyKnapsack` procedure gets an optimal load by taking all of L. Otherwise, let $a_j \in L$ be the first excluded item, that is, $volume((a_1, \ldots, a_{j-1})) \leq T$ but $volume((a_1, \ldots, a_j)) > T$. Let $K_{j-1} := (a_1, \ldots, a_{j-1})$ and $K_j := K_{j-1} \cup (a_j)$. Now K_j is an optimal load for $volume(K_j)$, since each unit of its space is used by a maximally valuable item that is available. Consider then an optimal load K^* for volume T. Since $T < volume(K_j)$, we must have

$$value(K^*) \leq value(K_j) \tag{1}$$

Let $a_{\max}$ be a maximally valuable item of L. The `greedyKnapsack` procedure computes a load K that includes K_{j-1} (and possibly some additional items). Then selecting K_{greedy} as the more valuable of the loads K and $(a_{\max})$, its value can be estimated as follows:

$$\begin{aligned}
\mathit{value}(K_{\text{greedy}}) &= \max\{\mathit{value}(K), \mathit{value}(a_{\max})\} \\
&\geq \max\{\mathit{value}(K_{j-1}), \mathit{value}(a_j)\} \\
&\geq (\mathit{value}(K_{j-1}) + \mathit{value}(a_j))/2 \\
&= \mathit{value}(K_j)/2 \\
&\geq \mathit{value}(K^*)/2 && \text{by inequality (1)} .
\end{aligned}$$

The bound is tight, that is, the error can get arbitrarily close to 50% of the optimum: Consider $L = (a_1, a_2, a_3)$ where $\mathit{value}(a_1) = 1 + \epsilon/2$ and $\mathit{volume}(a_1) = 1 + \epsilon/3$ for some $\epsilon > 0$. Let $\mathit{value}(a_j) = \mathit{volume}(a_j) = 1$ for $j \in \{2, 3\}$, and let the knapsack volume bound T be 2. The greedy strategy will choose $K_{\text{greedy}} := (a_1)$ with $\mathit{value}(K_{\text{greedy}}) = 1 + \epsilon/2$, while the value of the optimal load $K^* = (a_2, a_3)$ is 2, which equals $2 * \mathit{value}(K_{\text{greedy}}) - \epsilon$. So, the value of the optimal load can get arbitrarily close to twice the value of the greedily computed load, by choosing a sufficiently small value for ϵ.