

Analysis of task knapsack

Solutions: Dimitar Shapatov; Statement, analysis: Kiril Zashev

If during the contest you were wondering whether the formula is some kind of troll, for example that you are not required to find the most optimal solution for 100 points – it is actually not a troll, and that is perhaps the troll part in this task 😊).

Here are three possible approaches for solving this task:



Hill Climb

This is the most natural approach for an output-only task, and the following is only one possible idea. Initially we take some random set of items. On every single iteration of the hill climb we can take one random item from the taken items (the ones which currently participate in the answer) and remove it. After that we can put random items from the free elements (the ones which do not participate in the answer) into the set, while the weight is $\leq W$.

We revert to the previous state of the items if it is more optimal than the new one. There is one particularity in the implementation – how are we going to maintain dynamically the sets of the taken and the free items, so that we can also access random elements in them quickly? Most likely there is no structure which supports these operations in constant time, but we can do this in $O(\log N)$.

In the author solutions an `ordered_set` is used, taking advantage of the fact that it can find the k -th largest element in $O(\log N)$ (that is, the idea is to choose a random number for k). It turns out that such a solution does not work well with respect to the formula (it gets only about 10 points, even if it is run for a very long time). It is also possible to optimize the way in which we choose which item to remove, or to remove more than one item, but such solutions most likely still will not manage to get good points.



Fractional Knapsack

For those who have not heard of this knapsack – the idea is to sort the items by the ratio value/weight and to take them one after another until the weight becomes greater than W . This strategy breaks if we have to find the exact best answer, but it works comparatively well when we have to get close to the best answer. Unfortunately, again because of the nasty formula, this solution gives only about 10 points.



Fractional Knapsack + Hill Climb

The idea is that instead of starting the hill climb from a random set, we start from the set found by the fractional knapsack. This solution already has a greater chance of success, but the nasty thing in this task (and perhaps that is why it is good that it is in the easier day) is that the hill climb has to be run for a very long time in order to make a lot of points - if it is written well, it can reach about 80 points with about 2 hours of execution time.

The solution `knapsack_combined.zip` gets 100 points on the task, and it is a combination of slight variations of the solution above. A clean archive of the solution above is in `knapsack_best.zip` for 76.86 pts.