PROVING GREEDY ALGORITHM GIVES THE OPTIMAL SOLUTION

1 Introduction

In order to show correctness of a greedy algorithm, we need to prove that the following two properties hold:

- **Greedy choice property**: We show greedy choice property holds to show that the greedy choice we make in our algorithm makes sense. We prove this property by showing that there is an optimal solution such that it contains the best item according to our greedy criterion.
- **Optimal substructure:** This means that the optimal solution to our problem S contains an optimal to subproblems of S.

2 Fractional Knapsack

In this problem, we have a set of items with values $v_1, v_2, ..., v_n$ and weights $w_1, w_2, ..., w_n$. We also have a knapsack weight capacity W. We want to fit items having maximum value in our knapsack without exceeding the weight capacity. Note that, we are allowed to split the items into smaller sizes take a fraction of an item.

2.1 Notation

- S: our problem (set of n items with weight w_i and value v_i)
- W: knapsack weight capacity
- X: our solution (set of items we pick to put into knapsack)
- V: the value of our solution (sum of values of all items in our knapsack)

2.2 Greedy solution

At each step, we pick the item with maximum value to weight ratio (v_i/w_i) and we put as much of that item as possible in our knapsack. We repeat this process until either the knapsack is full or we run out of items to pick.

2.3 Example from slides-revisited

Name of item	Total weight of item (lbs)	Price per lb	Total value
Vibranium	2	$50K	$100K
Gold	5	$10K	$50K
Silver	8	$8K	$64K

Let us use our notation for this example. For this example, S=(2,$100K),(5,$50K),(8,$64K). The knapsack capacity W is given as 10 lbs. Using the greedy strategy we have, we keep picking the items with maximum value to weight ratio, namely price per lb. Let us execute our greedy strategy on this example:

1. Since vibranium has the best price per lb value, we first put as much as vibranium as we can in our knapsack. Now, the knapsack capacity has capacity of 8 lbs left and the total value of the items in our knapsack is $100K.

2. We pick the second best item in terms of price per lb and we put as much as gold in our knapsack. Now, the capacity of 3 lbs is left in the knapsack and the total value of items in the knapsack is $150K.

3. Finally, we put 3 lbs of silver in the knapsack. Now, there is no space left in the knapsack and the total value is $174K.

4. The algorithm terminates as there is no more space left in the knapsack.

So, the V=$174K and X=(2,$100K),(5,$50K),(3,$24K). We cannot do better than this and it seems like our greedy strategy works for this problem. In fact, it does! However, we need to prove the two properties given in Section 1.

2.4 Prove Greedy Choice Property

Let us first work on proving greedy choice property for our example. We need to show that the optimal solution should contain as much of the item with maximum price per lb value as possible. In our case, it means that the optimal solution should contain as much of vibranium as possible.
We want to prove this statement by contradiction. We start with the assumption that there exists an optimal solution that contains less then 2 lbs of vibranium. For example, let us assume that the solution we claim to be optimal contains 1 lb of vibranium, 5 lbs of gold and 4 lbs of silver. The resulting solution has value $132K. But we know that we can do better if we replace 1 lb of silver with 1 lb of vibranium. This result contradicts with our starting assumption which states that the solution with 1 lb of vibranium is optimal.
Now that we have the intuition, let us try to generalize this intuition to come up with a formal proof. Let item i be the item with the maximum value to weight ratio (v/w). We want to show that the optimal solution contains as much of item i as possible.
We prove that this statement is true by contradiction. We start by assuming

that there is an optimal solution where we did not take as much of item i as possible and we also assume that our knapsack is full (If it is not full, just add more of item i!). Since item i has the highest value to weight ratio, there must exist an item j in our knapsack such that $\frac{v_j}{w_j} < \frac{v_i}{w_i}$. We can take item j of weight x from our knapsack and we can add item i of weight x to our knapsack (Since we take out x weight and put in x weight, we are stilll within capacity.). The change in value of our knapsack is $x \cdot \frac{v_i}{w_i} - x \cdot \frac{v_j}{w_j} = x \cdot \left(\frac{v_i}{w_i} - \cdot \frac{v_j}{w_j} \right) > 0$ because $\frac{v_j}{w_j} < \frac{v_i}{w_i}$. Therefore, we arrive at a contradiction because the "so-called" optimal solution in our starting assumption, can in fact be improved by taking out some of item j and adding more of item i. Hence, it is not optimal. □

2.5 Showing optimal substructure

Let us first show optimal substructure on our example. Recall that our problem S is {(2,\$100K),(5,\$50K),(8,\$64K)}, knapsack capacity W is 10, the value of our greedy solution is V=\$174K and the greedy solution X is {(2,\$100K),(5,\$50K),(3,\$24K)}. We want to show that this optimal solution X of problem S also gives us the optimal solution to subproblem S'. Let us define a specific subproblem. Assume that X is the optimal solution to problem S and the subproblem S' is (2,\$100K),(8,\$64K) (We discard golds), W'= 5 and X'=(2,\$100K),(3,\$24K) and V'=\$124K. Since S' is a subset of S, S' is a subproblem of S and since X' is a subset of X, the solution to subproblem S' is contained in the optimal solution to the original problem S. We now need to show that X' is the optimal solution to S'. We prove this by contradiction. Assume that X' is not the optimal solution to S' and we have another solution X" to S' that has a higher total value $V'' > V'$=\$124K. Then, we get a better solution to problem S by adding the discarded golds to X" and this solution has value $V'' + \$50K > +\$50K + \$124K = \$174K$, which is the value of our optimal solution to problem S. This is a contradiction because we assumed that X is the optimal solution to S with value V=\$174K.
Now, let us generalize this idea to give the formal proof. Assume that X is the optimal solution with value V to problem S with knapsack capacity W. Then, we want to prove that $X' = X - x_j$ is an optimal solution to the subproblem $S' = S - \{j\}$ and the knapsack capacity $W' = W - w_j$. We prove this by contradiction. We assume that X' is not optimal to S' and that we have another solution X" to S' that has a higher total value $V'' > V$. Then, $X'' \cup \{x_j\}$ is a solution to S with value $V'' + v_j > V' + v_j = V$. This is a contradiction because V is assumed to be optimal in the beginning. □