

COMP361

Assignment 3

Knapsack Problem

William Kilty

300473541

2:

```
calculate(Item[] items, int limit) {
    values[items.length+1][limit+1] <- 0 //initialise everything to zero
    for (item = 1; < values.length; ++i) {
        for (weight = 0; < values[item].length; ++w) {
            <- Barometer here (each time values[x][y] is assigned)
            values[item][weight] = weight >= items[item-1].weight ?
            Math.max(values[item-1][weight-items[item-1].weight] +
            items[item-1].value, values[item-1][weight]) :
            values[item-1][weight];
        }
    }
    return values[items.length][limit]
}
```

The meat of the algorithm is the part within both for-loops. This works on the following logic:

if the available weight is enough to hold the new item we take the largest of either:

- the best case without this item,
- the best case without the weight of this item with this item's value added

otherwise, we take the best case without this item

In other words, we either take the item, or we don't. We choose the option that gives us the biggest result. And since those two options are the only options, the largest of the two is the optimal answer.

So if we're taking the optimal answer at each step, we can be sure that each step beforehand that we're basing this step off of will be optimal too.

If we have a bag filled this way, the final value at the end of the loop will be a bag with the highest value possible - the optimal solution to the problem.

My test case generator works by creating a list of random length containing items with random weights and values, and then picking a random weight limit. This is then put through the algorithm and the results are analysed to check correctness.

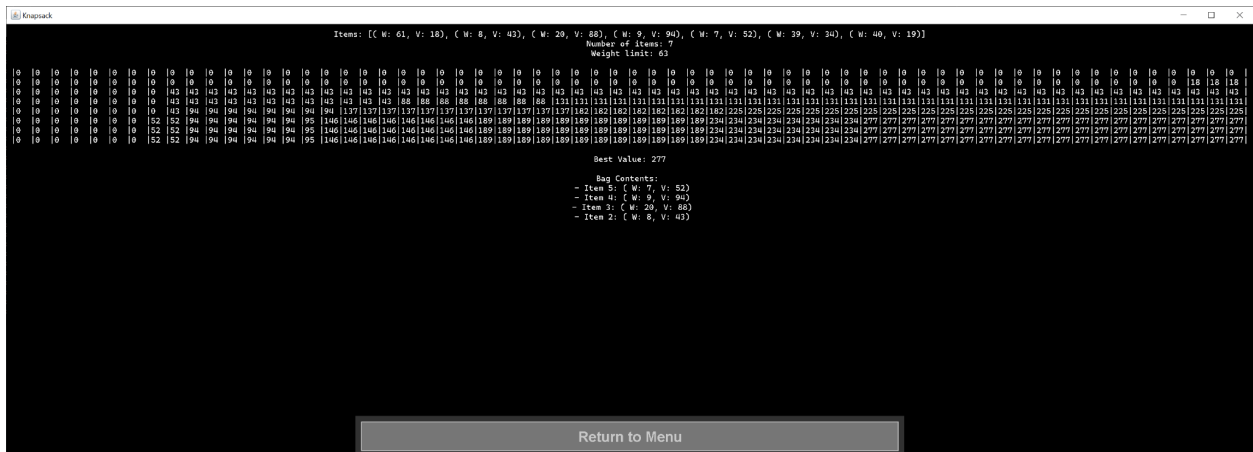
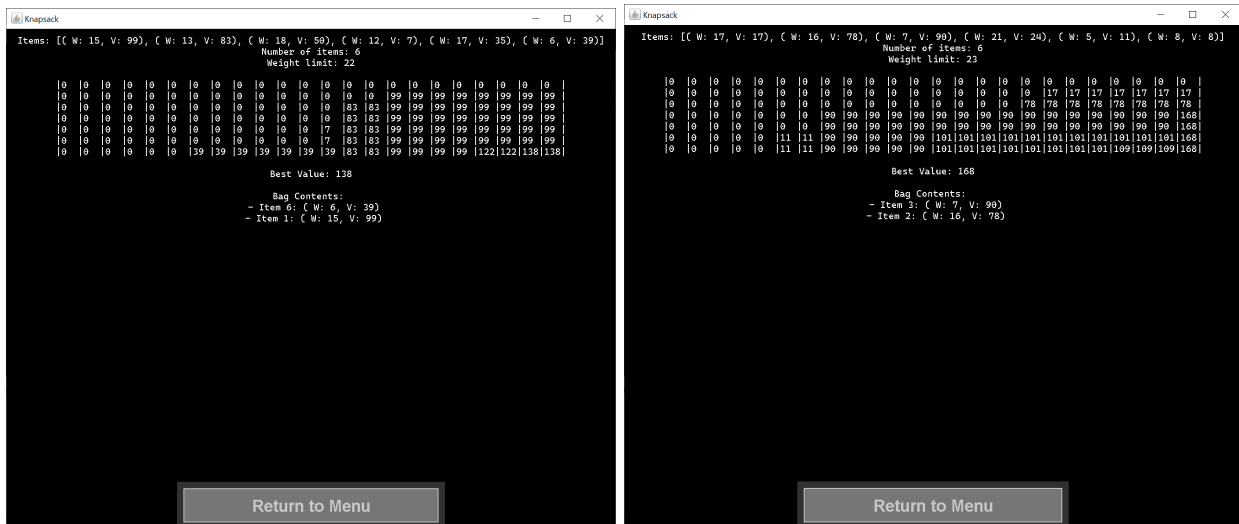
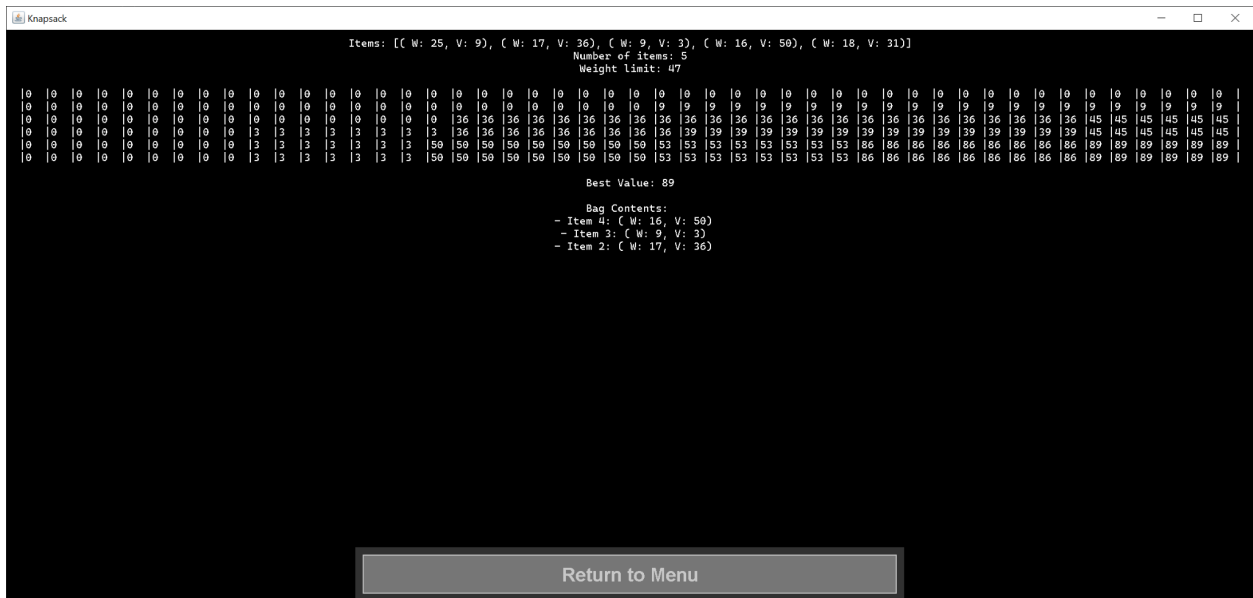
We can then check the barometer results against the inputs to determine complexity.

Our expected complexity should be $N*(M+1)$ where N is the size of the list of items, and M is the weight limit. This translates to Big O complexity of $O(m)$ if we assume that $\text{limit} \gg \text{number of items}$. (for an interesting problem we want most items to have a weight > 1 and for the bag to be able to hold more than 1 item)

Based on testing;

- Barometer: 371, Num items: 7, Weight Limit: 52
 - $7*(52+1) = 371$
- Barometer: 64, Num items: 4, Weight Limit: 15
 - $4*(15+1) = 64$
- Barometer: 553, Num items: 7, Weight Limit: 78
 - $7*(78+1) = 553$
- Barometer: 305, Num items: 5, Weight Limit: 60
 - $5*(60+1) = 305$

My estimation is spot on. And in each case, $\text{limit} \gg \text{number of items}$.
Complexity $O(m)$



3:

```
calculate(Item[] items, int limit) {
    sorted <- items.sort
    weight <- limit
    for (item = 0; < sorted.length; ++i) {
        <- Barometer here (each time we try to take an item)
        Item item <- sorted.get(i);
        if (item.weight <= weight) {
            Item take <- new Item(item.weight, item.value,
                Math.min(item.number, weight/item.weight));
            taking.add(take);
            weight-=take.weight*take.number;
        }
    }
}
```

The meat of the algorithm is the part within the for-loop. This works on the following logic:

if the available weight is enough to hold the new item type we take the smallest of either:

- the number of this item available
- the number of times this item can be taken regardless of how many there actually are

otherwise, we move on

This ensures we always take the most we can of each item, either by taking all there is to take, or by taking as many as we can hold.

The key strategy in this case is utilising a sorted list. This list is sorted by the ratio of values to weights, with the higher the ratio, the better the item is regarded. We assume that taking as many of the best item-type as we can will be better than any other option. We do not check any other options, and so we cannot be sure this is the case.

The greedy algorithm is not optimal. But it will provide the best guess possible given the assumption talked about above.

Once the algorithm has run, we end up with a list of all the items we took, and how many times we took each item. The sum of all the values in this list is the final answer for the problem.

My test case generator works by creating a list of random length containing items with random weights and values and number of instances, and then picking a random weight limit. This is then put through the algorithm and the results are analysed to check correctness.

We can then check the barometer results against the inputs to determine complexity.

Our expected complexity should be N where N is the size of the list of items. This translates to Big O complexity of $O(n)$. However, this is not the full picture... We will go under the assumption of $O(n)$ for now.

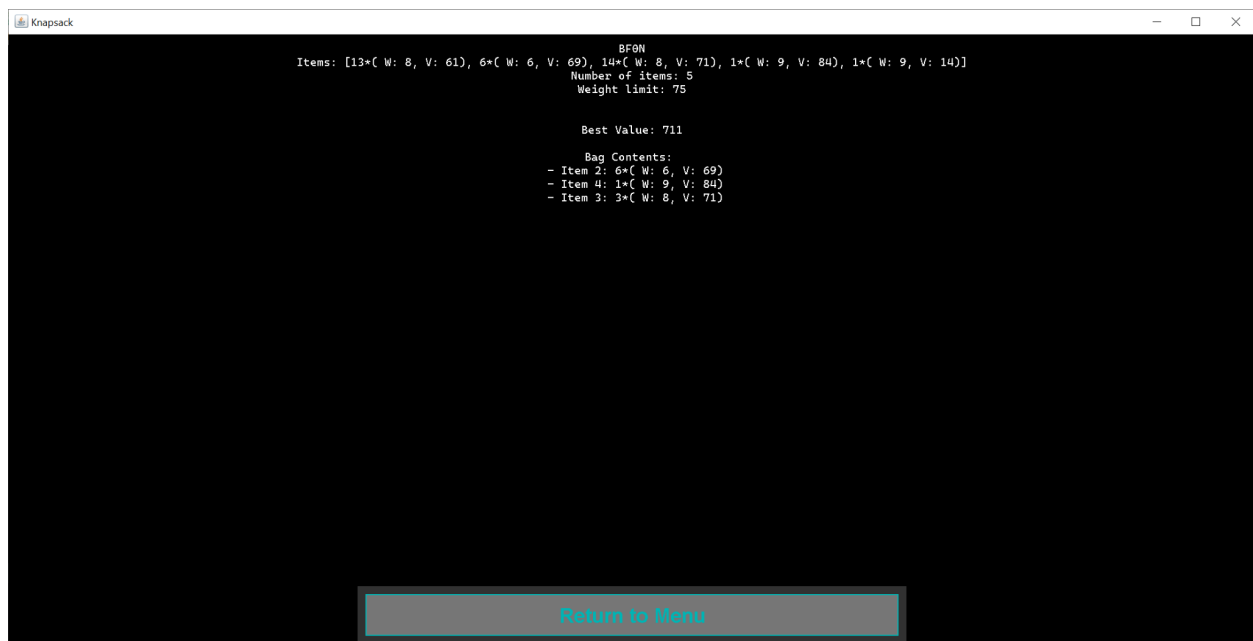
Based on testing;

- Barometer: 7, Num items: 7
- Barometer: 5, Num items: 5
- Barometer: 2, Num items: 2
- Barometer: 4, Num items: 4
- Barometer: 5, Num items: 5

My estimation is spot on. Complexity $O(n)$...

But this is not true, because there's another feature to the algorithm that has been glossed over. We have to first sort the list of items. The sorting method I used to do this is merge-sort. The complexity of merge sort has been discussed in lectures and previous assignments, so I know it to be $O(n \log n)$.

This complexity is greater than $O(n)$, so the complexity of the whole algorithm ends up being $O(n \log n)$.



```
Knapsack
BF0N
Items: [13*( W: 8, V: 61), 6*( W: 6, V: 69), 14*( W: 8, V: 71), 1*( W: 9, V: 84), 1*( W: 9, V: 14)]
Number of items: 5
Weight limit: 75

Best Value: 711

Bag Contents:
- Item 2: 6*( W: 6, V: 69)
- Item 4: 1*( W: 9, V: 84)
- Item 3: 3*( W: 8, V: 71)

Return to Menu
```



4:

```
calculate(Item[] items, int limit) {
    instances <- (each item : items) item.number // the total collection of items
    values[instances+1][limit+1] <- 0 //initialise everything to zero
    for (instance = 1; < values.length; ++ ) {
        for (weight = diag; < values[item].length; ++ ) {
            <- Barometer here (each time values[x][y] is assigned)
            values[instance][weight] = weight >= items[item-1].weight ?
            Math.max(values[instance-1][weight-items[item-1].weight] +
            items[item-1].value, values[instance-1][weight]) :
            values[instance-1][weight];
        }
    }
    return values[instances-1][limit]
}
```

The meat of the algorithm is the part within both for-loops. This works on the following logic:

if the available weight is enough to hold the new item we take the largest of either:

- the best case without this item,
- the best case without the weight of this item with this item's value added

otherwise, we take the best case without this item

In other words, we either take the item, or we don't. We choose the option that gives us the biggest result. And since those two options are the only options, the largest of the two is the optimal answer.

So if we're taking the optimal answer at each step, we can be sure that each step beforehand that we're basing this step off of will be optimal too.

If we have a bag filled this way, the final value at the end of the loop will be a bag with the highest value possible - the optimal solution to the problem.

This is essentially identical to part 2. We just include as many copies of each item as we have specified as item.number

My test case generator is the same as in part 3. It works by creating a list of random length, containing items with random weights and values and number of instances, and then picking a random weight limit. This is then put through the algorithm and the results are analysed to check correctness.

We can then check the barometer results against the inputs to determine complexity.

Our expected complexity should be $N \cdot (M+1)$ where N is the size of the list of instances, and M is the weight limit. This translates to Big O complexity of $O(nm)$

Based on testing;

- Barometer: 3700, Num instances: 50, Weight Limit: 73
 - $50 \cdot (73+1) = 3700$
- Barometer: 2320, Num instances: 29, Weight Limit: 79
 - $29 \cdot (79+1) = 2320$
- Barometer: 4324, Num instances: 46, Weight Limit: 93
 - $46 \cdot (93+1) = 4324$
- Barometer: 1836, Num instances: 34, Weight Limit: 53
 - $34 \cdot (53+1) = 1836$

My estimation is spot on.

Complexity $O(nm)$

The screenshot shows a Knack application window titled "Knack". The main content area displays a large grid of numbers, likely representing a complex calculation or data set. The grid is organized into rows and columns, with numbers ranging from 0 to 100. The numbers are arranged in a pattern that suggests a complex calculation or data set. At the bottom of the window, there is a "Return to Menu" button.

```

Number of items: 4
Number of instances: 34
Weight limit: 53

```

Best Value: 490
Return to Menu

```

Number of items: 3
Number of instances: 29
Weight limit: 74

```

Best Value: 589

Bag Contents:

- Item 4: 1x(C W: 9, V: 7
- Item 4: 1x(C W: 9, V: 7
- Item 4: 1x(C W: 9, V: 7
- **Return to Menu**
- Item 4: 1x(C W: 9, V: 7

5:

```
calculate(Item[] items, int limit) {
    items.sort //sort the items so we can do a best-first-search
    node[items.length+2] <- 0 //initialise everything to zero
    //node stores how many of each item we take, followed by the running total of
    //weight and value
    node = traverse(node)
    taking[i] <- node[i]*items[i] //take the number of each item as given in node
    return node[node.length-1] //return the best value from the traversal
}

traverse(int[] node) {
    <- Barometer here (each time we traverse the graph)
    If (visited(node)) return node
    bestNode = node;
    for (i = 0; < node.length-2; ++i) {
        newNode = node (with node[i]+=1)
        newV = newNode[length-2]+items[i].value
        newW = newNode[length-1]+items[i].weight

        If (pruning) continue

        bestNode = best of (traverse(newNode)) or (bestNode)
    }
    return bestNode
}
```

We sort the items, so we can do some simple heuristic work by default. Then we recursively traverse through the graph until we've found the node with the highest value in the graph, with a weight below the limit.

We prune based on whether or not the next node will break the weight limit or if we've run out of items to add in that dimension. This ensures we never break the rules.

We also prune based on if the path started by this traversal can possibly return a better result than what we already have. If not there's no point looking

Then we simply look at every node left, and take the one with the best value. The only way this would give us a sub-optimal answer is if we accidentally prune the best answer. However, we only prune branches we know do not contain the best answer. So our algorithm will always return an optimal answer.

My test case generator is the same as in part 3. It works by creating a list of random length, containing items with random weights and values and number of instances, and then picking a random weight limit. This is then put through the algorithm and the results are analysed to check correctness.

We can then check the barometer results against the inputs to determine complexity.

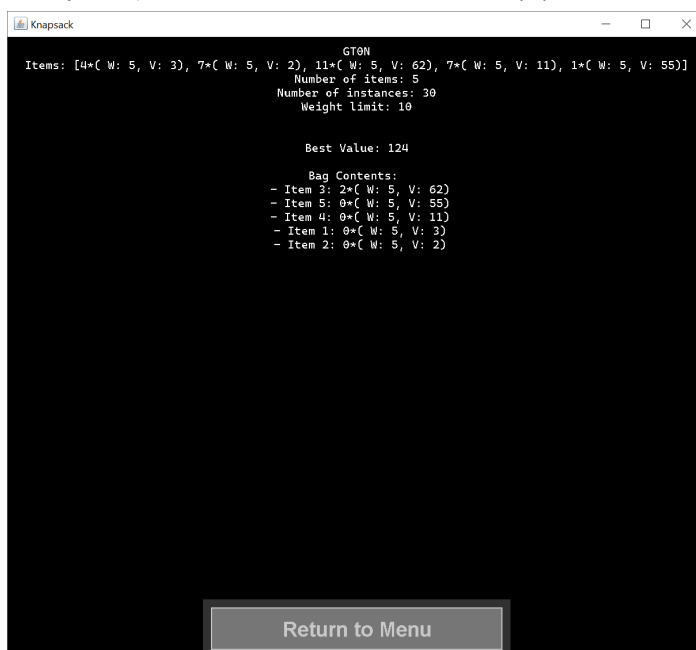
Our expected worst-case complexity should be every combination of instances of each item I^N where I is the number of instances of each item, and N is the number of items. This translates to Big O complexity of $O(i^n)$ In practice, it should be a lot lower than this as we do a lot of pruning.

Based on testing;

- Barometer: 9656, Num instances: 43, Num Items: 6
 - $(43/6 \approx 7)^6 \approx 117,649 > 9656$
- Barometer: 1037, Num instances: 28, Num Items: 4
 - $(28/4 = 7)^4 = 2401 > 1037$
- Barometer: 10625, Num instances: 41, Num Items: 5
 - $(41/5 \approx 8)^5 \approx 37074 > 10625$
- Barometer: 30, Num instances: 30, Num Items: 5
 - $(30/5 = 6)^5 = 7776 > 30$

My estimation is correct.

Complexity is somewhere lower than $O(i^n)$



```

Knapsack
-----
GT0N
Items: [4*( W: 5, V: 3), 7*( W: 5, V: 2), 11*( W: 5, V: 62), 7*( W: 5, V: 11), 1*( W: 5, V: 55)]
Number of items: 5
Number of instances: 30
Weight limit: 10

Best Value: 124

Bag Contents:
- Item 2: 2*( W: 5, V: 62)
- Item 5: 0*( W: 5, V: 55)
- Item 4: 0*( W: 5, V: 11)
- Item 1: 0*( W: 5, V: 3)
- Item 2: 0*( W: 5, V: 2)

Return to Menu
  
```

```
Knapsack

GT0N
Items: [8*( W: 7, V: 17), 3*( W: 5, V: 39), 9*( W: 6, V: 84), 1*( W: 6, V: 1), 7*( W: 7, V: 61), 10*( W: 5, V: 50)]
Number of items: 6
Weight limit: 37

Best Value: 594

Bag Contents:
- Item 3: 6*( W: 6, V: 84)
- Item 6: 0*( W: 5, V: 50)
- Item 5: 0*( W: 7, V: 61)
- Item 2: 0*( W: 5, V: 39)
- Item 1: 0*( W: 7, V: 17)
- Item 4: 0*( W: 6, V: 1)

Return to Menu
```

```
Knapsack

GT0N
Items: [14*( W: 7, V: 80), 12*( W: 10, V: 69), 2*( W: 8, V: 9), 1*( W: 9, V: 78)]
Number of items: 4
Weight limit: 89

Best Value: 960

Bag Contents:
- Item 1: 12*( W: 10, V: 69)
- Item 4: 0*( W: 9, V: 78)
- Item 2: 0*( W: 10, V: 69)
- Item 3: 0*( W: 8, V: 9)

Return to Menu
```

```
Knapsack

GT0N
Items: [6*( W: 11, V: 73), 6*( W: 7, V: 66), 14*( W: 9, V: 31), 2*( W: 5, V: 79), 13*( W: 10, V: 5), 3*( W: 8, V: 89)]
Number of items: 6
Weight limit: 80

Best Value: 828

Bag Contents:
- Item 4: 2*( W: 5, V: 79)
- Item 6: 3*( W: 8, V: 89)
- Item 2: 5*( W: 7, V: 66)
- Item 1: 1*( W: 11, V: 73)
- Item 3: 0*( W: 9, V: 31)
- Item 5: 0*( W: 10, V: 5)

Return to Menu
```

6:

1.

```
a. explore(Node) {  
    If (Node.isExplored) return List.of(Node);  
    Node.explore;  
    For (Edge : Node.neighbours) {  
        If (Edge.isExplored) continue;  
        Edge.explore;  
        List l = explore(Edge.node);  
        If (l != null) {  
            If (l.contains(Node)) print(l);//cycle  
            l.add(Node);  
            return l;  
        }  
    }  
    return null;  
}
```

b. This algorithm has a time complexity of $O(m+n)$.

This is because in the case of no cycles, the worst case scenario, we will explore every node connected in the graph. This is a complexity of n .

But for each node, we explore each of its neighbouring edges too - exactly once, for a complexity of $O(m)$

This all leads to a complexity of $O(n+m)$

2. Base case: one leaf, no nodes.

Each step, take all the leaf nodes, and turn them into nodes with two new leaf nodes for children.

Assuming step k follows the rules, with n nodes, and e leaves, step $k+1$ now has:

- Nodes, $N: n+e$
- Leaves, $E: 2*e$

We have, by assumption, $e=n+1$.

$$N = 2n+1$$

$$E = 2n+2$$

$$E=N+1$$

3. Assume G does contain an edge not present in T . By theorem in pg 19 of graphs-algorithms lecture slides, this edge, (x, y) for some $x, y \in V$, must hold that either x or y is an ancestor of the other in depth first search.

If we now look at BFS, we have two cases; x is an ancestor of y ; or y is an ancestor of x in DFS.

Case 1 (x is ancestor of y in DFS):

When we start calculating x in BFS, we look at each of x 's children, which includes y 's ancestor (from DFS), and y itself (from edge (x, y)). These are both added to the queue.

y will now be expanded at the same time as y 's ancestor. In other words, y 's ancestor will no longer be y 's ancestor in BFS.

DFS does not equal BFS. Contradiction.

Case 2 (y is ancestor of x in DFS):

This case is identical by symmetry.

DFS does not equal BFS. Contradiction

G cannot contain any edges not found in T .

$G=T$

4. To show NPC, we need to translate from one problem to the other. And show both are NP.

Start with 0-1 Knapsack:

Showing an answer is correct means adding all the weights of selected items. If the sum is $\leq W$, then the answer is correct. Summing is in P, so 0-1 knapsack is NP.

Now, translate from Subset Problem to 0-1 knapsack.

Let each value in the input set for Subset Problem be an Item, where weight = value = the value of the number in the input set.

Let the weight limit = the target number for the subset problem.

Now, running the problem as a 0-1 knapsack problem will return the correct answer if one exists - it will return the largest value. It will never overestimate, as it cannot exceed the weight limit, so the largest available answer is the correct one. If there is an answer, knapsack can find it.

0-1 knapsack is NPC.

Now, 0-N Knapsack:

Showing an answer is correct means adding all the weights of selected items. If the sum is $\leq W$, then the answer is correct. Summing is in P, so 0-1 knapsack is NP.

Now, translate 0-1 to 0-N. If we select $N=1$ for every item, 0-N = 0-1.

0-N knapsack is NPC.

5. To show NP-Hard, we just need to translate from one to the other.

Translating 0-N knapsack to integer linear programming:

Integer linear programming problem is to maximise linear function subject to constraints.

0-N knapsack is to maximise linear function subject to one constraint.

$\sum \text{Item profits}(A_p, B_p, C_p, \dots)$ subject to $\sum \text{Item weights}(A_w, B_w, C_w, \dots) \leq W$.

This is a linear programming problem.

Integer linear programming is NP-Hard.