# CSE 221: Algorithms Greedy algorithms

## Mumit Khan Fatema Tuz Zohora

Computer Science and Engineering BRAC University

#### References

- Jon Kleinberg and Éva Tardos, Algorithm Design. Pearson Education, 2006.
- Michael T. Goodrich and Roberto Tamassia, Data Structures and Algorithms in Java, Fourth Edition. John Wiley & Sons, 2006.
- T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, Second Edition. The MIT Press, September 2001.

Last modified: November 13, 2012



# Contents

- Greedy algorithms
  - Introduction
  - Interval scheduling problem
  - Scheduling all Intervals problem
  - Fractional knapsack problem
  - Coin changing problem
  - What problems can be solved by greedy approach?
  - Conclusion

# Greedy algorithms

- Introduction
- Interval scheduling problem
- Scheduling all Intervals problem
- Fractional knapsack problem
- Coin changing problem
- What problems can be solved by greedy approach?
- Conclusion

# Greedy design strategy

Greed . . . is good. Greed is right. Greed works.

Greed . . . is good. Greed is right. Greed works.



Gordon Gekko (played by Michael Douglas), in the 1987 movie Wall Street.

Greed . . . is good. Greed is right. Greed works.



Gordon Gekko (played by Michael Douglas), in the 1987 movie Wall Street.

#### Basic idea:

Solves the problem step by step.

Greedy algorithms

# Greed . . . is good. Greed is right. Greed works.



Gordon Gekko (played by Michael Douglas), in the 1987 movie Wall Street.

- Solves the problem step by step.
- At each step, pick the choice which looks best based on **some criteria**(i.e. *greedily*) at that moment given the information currently available.

Greed ... is good. Greed is right. Greed works.



Gordon Gekko (played by Michael Douglas), in the 1987 movie Wall Street.

- Solves the problem step by step.
- At each step, pick the choice which looks best based on some criteria(i.e. greedily) at that moment given the information currently available.
- Main challenge is to determine the criteria on which the next item will be selected

# Greedy design strategy

Greed . . . is good. Greed is right. Greed works.



Gordon Gekko (played by Michael Douglas), in the 1987 movie Wall Street.

- Solves the problem step by step.
- At each step, pick the choice which looks best based on **some criteria**(i.e. *greedily*) at that moment given the information currently available.
- Main challenge is to determine the criteria on which the next item will be selected
- Often leads to very efficient solutions to optimization problems.

# Greedy design strategy

Greed . . . is good. Greed is right. Greed works.



Gordon Gekko (played by Michael Douglas), in the 1987 movie Wall Street.

- Solves the problem step by step.
- At each step, pick the choice which looks best based on **some criteria**(i.e. *greedily*) at that moment given the information currently available.
- Main challenge is to determine the criteria on which the next item will be selected
- Often leads to very efficient solutions to optimization problems.
- However, not all problems have greedy solutions.

# **Optimal Solution**

#### What is an Optimal Solution?

- Given a problem, more than one solution exist
- One of the solution is the best based on some given constraints, that solution is called the optimal solution

#### What is Global Optimal Solution?

Optimal Solution to the main problem

#### What is local Optimal Solution?

Optimal Solution to the subproblems

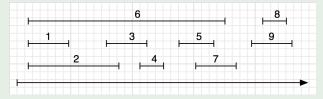
# Greedy algorithms

- Introduction
- Interval scheduling problem
- Scheduling all Intervals problem
- Fractional knapsack problem
- Coin changing problem
- What problems can be solved by greedy approach?
- Conclusion

## Definition (Interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , find the largest set  $A \subseteq I$  such that the members of A are non-conflicting.

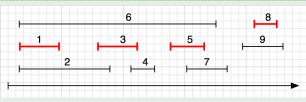
### Example



## Definition (Interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , find the largest set  $A \subseteq I$  such that the members of A are non-conflicting.

## Example

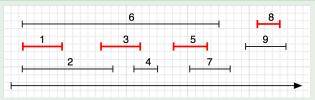


$$A = \{1, 3, 5, 8\}, \quad |A| = 4.$$

### Definition (Interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , find the largest set  $A \subseteq I$  such that the members of A are non-conflicting.

### Example



$$A = \{1, 3, 5, 8\}, |A| = 4.$$

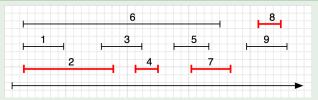
#### Question

Is this the only "correct" answer?

### Definition (Interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , find the largest set  $A \subseteq I$  such that the members of A are non-conflicting.

### Example



$$A = \{1, 3, 5, 8\}, \quad |A| = 4.$$

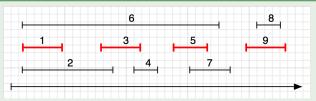
#### Question

How about  $\{2, 4, 7, 8\}$ ?

### Definition (Interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , find the largest set  $A \subseteq I$  such that the members of A are non-conflicting.

### Example



$$A = \{1, 3, 5, 8\}, |A| = 4.$$

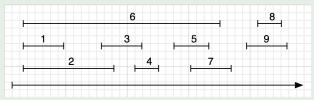
#### Question

How about  $\{2,4,7,8\}$ ?  $\{1,3,5,9\}$ ?

## Definition (Interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , find  $A \subseteq I$  such that the members of A are non-conflicting and |A| is maximized.

### Example



$$A = \{1, 3, 5, 8\}, |A| = 4.$$

#### Question

 $\{1,3,5,8\}$ ?  $\{2,4,7,8\}$ ?  $\{1,3,5,9\}$ ? ... How many?

A solution must be one of the subsets of the set of intervals.

A solution must be one of the subsets of the set of intervals.

• Enumerate all possible configurations (i.e., all possible subsets of the intervals).

#### A solution must be one of the subsets of the set of intervals.

- Enumerate all possible *configurations* (i.e., all possible subsets of the intervals).
- ② Go through the set of subsets and remove the ones that have one or more conflicting schedules.

#### A solution must be one of the subsets of the set of intervals.

- Enumerate all possible *configurations* (i.e., all possible subsets of the intervals).
- ② Go through the set of subsets and remove the ones that have one or more conflicting schedules.
- Open Pick (any one of) the largest subset from the ones that survive.

#### A solution must be one of the subsets of the set of intervals.

- Enumerate all possible *configurations* (i.e., all possible subsets of the intervals).
- ② Go through the set of subsets and remove the ones that have one or more conflicting schedules.
- Open Pick (any one of) the largest subset from the ones that survive.

### Complexity

• There are  $2^n - 1$  non-empty subsets, one or more of which may be a feasible solution.

#### A solution must be one of the subsets of the set of intervals.

- Enumerate all possible *configurations* (i.e., all possible subsets of the intervals).
- On through the set of subsets and remove the ones that have one or more conflicting schedules.
- Open Pick (any one of) the largest subset from the ones that survive.

#### Complexity

- There are  $2^n 1$  non-empty subsets, one or more of which may be a feasible solution.
- Each feasible solution must be scanned for conflict, which takes O(n) time.

# A solution must be one of the subsets of the set of intervals.

- Enumerate all possible *configurations* (i.e., all possible subsets of the intervals).
- ② Go through the set of subsets and remove the ones that have one or more conflicting schedules.
- Open Pick (any one of) the largest subset from the ones that survive.

### Complexity

- There are  $2^n 1$  non-empty subsets, one or more of which may be a feasible solution.
- Each feasible solution must be scanned for conflict, which takes O(n) time.
- The algorithm runs in  $\Theta(n2^n)$  time

#### A solution must be one of the subsets of the set of intervals.

- Enumerate all possible *configurations* (i.e., all possible subsets of the intervals).
- ② Go through the set of subsets and remove the ones that have one or more conflicting schedules.
- Open Pick (any one of) the largest subset from the ones that survive.

#### Complexity

- There are  $2^n 1$  non-empty subsets, one or more of which may be a feasible solution.
- Each feasible solution must be scanned for conflict, which takes O(n) time.
- The algorithm runs in  $\Theta(n2^n)$  time  $\Rightarrow$  an exponential time algorithm!

To compute the maximal set of intervals that can be scheduled, the basic idea is to:

• Use a "simple" rule (or strategy) to select the first interval  $i_1$ to be accepted.

- **①** Use a "simple" rule (or strategy) to select the first interval  $i_1$ to be accepted.
- ② Once  $i_1$  is accepted, remove from consideration all intervals the conflict with  $i_1$ .

#### Basic steps

- Use a "simple" rule (or strategy) to select the first interval  $i_1$ to be accepted.
- ② Once  $i_1$  is accepted, remove from consideration all intervals the conflict with  $i_1$ .
- **3** Select the second interval  $i_2$  to be accepted, and remove all the intervals that conflict with  $i_2$ .

#### Basic steps

- **①** Use a "simple" rule (or strategy) to select the first interval  $i_1$ to be accepted.
- ② Once  $i_1$  is accepted, remove from consideration all intervals the conflict with  $i_1$ .
- **3** Select the second interval  $i_2$  to be accepted, and remove all the intervals that conflict with  $i_2$ .
- And so on until there are no more requests remain.

To compute the maximal set of intervals that can be scheduled, the basic idea is to:

- Use a "simple" rule (or strategy) to select the first interval  $i_1$ to be accepted.
- ② Once  $i_1$  is accepted, remove from consideration all intervals the conflict with  $i_1$ .
- **3** Select the second interval  $i_2$  to be accepted, and remove all the intervals that conflict with  $i_2$ .
- And so on until there are no more requests remain.

# Key challenge

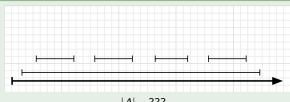
How to choose the "simple" rule to select the next interval that leads to an optimal solution?

#### Strategy 1. Earliest First

The idea is to start using the resource as early as possible.

- Sort the intervals by starting time, breaking ties arbitrarily.
- Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

#### Example



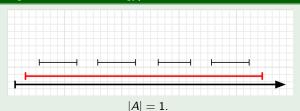
$$|A| = ???$$
.

#### Strategy 1. Earliest First

The idea is to start using the resource as early as possible.

- Sort the intervals by starting time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

#### Example (using *Earliest First* strategy)



#### Strategy 1. Earliest First

The idea is to start using the resource as early as possible.

- Sort the intervals by starting time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

#### Example (using an optimal strategy)



# Strategy 1. Earliest First

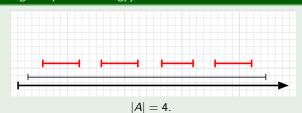
The idea is to start using the resource as early as possible.

Designing a greedy algorithm (continued)

- Sort the intervals by starting time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

This strategy does not lead to an optimal solution.

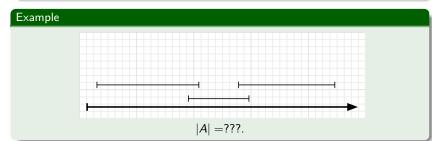
#### Example (using an optimal strategy)



## Strategy 2. Shortest First

The Earliest First strategy failed perhaps because it missed the shorter intervals, which would accommodate more intervals.

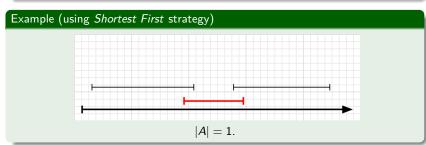
- Sort the intervals by length, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.



## Strategy 2. Shortest First

The Earliest First strategy failed perhaps because it missed the shorter intervals, which would accommodate more intervals.

- Sort the intervals by length, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.



## Strategy 2. Shortest First

The Earliest First strategy failed perhaps because it missed the shorter intervals, which would accommodate more intervals.

- Sort the intervals by length, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

# Example (using an optimal strategy) |A| = 2.

## Strategy 2. Shortest First

The Earliest First strategy failed perhaps because it missed the shorter intervals, which would accommodate more intervals.

- Sort the intervals by length, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

This strategy does not lead to an optimal solution.

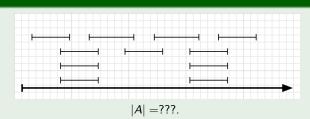
# Example (using an optimal strategy) |A| = 2.

## Strategy 3. Least-conflict First

The Shortest First strategy failed perhaps because the shorter ones had more conflicts, and ruled out too many intervals in the process.

- Sort the intervals by the number of other intervals which conflict with it.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

## Example



## Strategy 3. Least-conflict First

The Shortest First strategy failed perhaps because the shorter ones had more conflicts, and ruled out too many intervals in the process.

- Sort the intervals by the number of other intervals which conflict with it.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

# Example (using *Shortest First* strategy) |A| = 3.

## Strategy 3. Least-conflict First

The Shortest First strategy failed perhaps because the shorter ones had more conflicts, and ruled out too many intervals in the process.

- Sort the intervals by the number of other intervals which conflict with it.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

# Example (using an optimal strategy) |A| = 4.

## Strategy 3. Least-conflict First

The Shortest First strategy failed perhaps because the shorter ones had more conflicts, and ruled out too many intervals in the process.

- Sort the intervals by the number of other intervals which conflict with it.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

This strategy does not lead to an optimal solution.

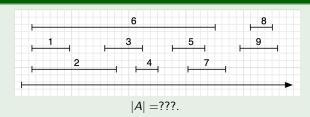
# Example (using an optimal strategy) |A| = 4.

#### Strategy 4. Finish First

The idea is to free up the resource as early as possible.

- Sort the intervals by the finishing time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

## Example

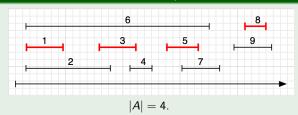


#### Strategy 4. Finish First

The idea is to free up the resource as early as possible.

- Sort the intervals by the finishing time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

## Example (using optimal Finish First strategy)



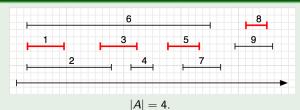
#### Strategy 4. Finish First

The idea is to free up the resource as early as possible.

- Sort the intervals by the finishing time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

This strategy is the one that works.

## Example (using optimal Finish First strategy)



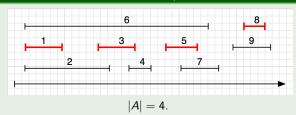
## Strategy 4. Finish First

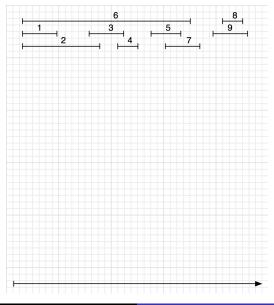
The idea is to free up the resource as early as possible.

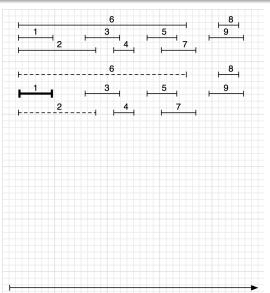
- Sort the intervals by the finishing time, breaking ties arbitrarily.
- 2 Pick the first one, removing it from the list along with all the intervals that conflict with it.
- Repeat Step 2, until the list is empty.

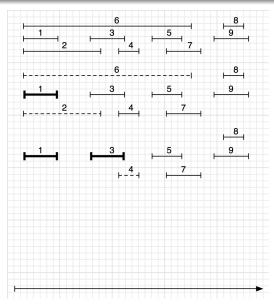
This strategy is the one that works. But can you prove that it works?

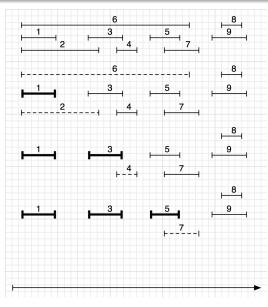
## Example (using optimal Finish First strategy)

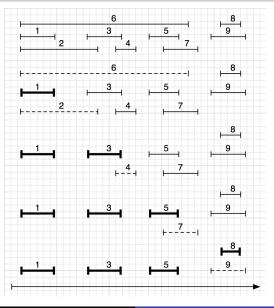


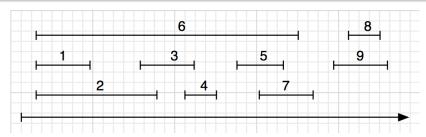




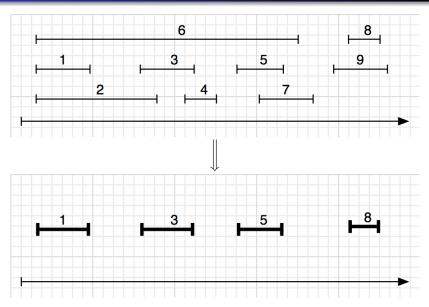








# Interval scheduling in action (continued)



# An $O(n \lg n)$ greedy algorithm for interval scheduling

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
    R = Sorted requests in order of finishing times such that f_i \leq f_j when i < j.
    Create an array S[1...n] with starting times such that S[i] contains s_i.
    A = \{R_1\}
                                               > select first interval from sorted list
    f = f_1
5
    while there are more intervals in S to look at
6
          do j = first interval for which s_i > f
              A \leftarrow A \cup \{i\}
              f \leftarrow f_i
8
9
    return A
```

# An $O(n \lg n)$ greedy algorithm for interval scheduling

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
    R = Sorted requests in order of finishing times such that f_i \leq f_j when i < j.
    Create an array S[1...n] with starting times such that S[i] contains s_i.
    A = \{R_1\}
                                               > select first interval from sorted list
   f = f_1
    while there are more intervals in S to look at
6
          do j = first interval for which s_i > f
              A \leftarrow A \cup \{i\}
              f \leftarrow f_i
8
9
    return A
```

#### **Analysis**

• The sorting step in takes  $O(n \lg n)$  time.

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
    R = Sorted requests in order of finishing times such that f_i \leq f_j when i < j.
    Create an array S[1...n] with starting times such that S[i] contains s_i.
    A = \{R_1\}
                                               > select first interval from sorted list
   f = f_1
    while there are more intervals in S to look at
6
          do j = first interval for which s_i > f
              A \leftarrow A \cup \{i\}
              f \leftarrow f_i
8
9
    return A
```

#### **Analysis**

- The sorting step in takes  $O(n \lg n)$  time.
- Creating the starting time array S[1...n] takes O(n) time.

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
    R = Sorted requests in order of finishing times such that f_i \leq f_j when i < j.
    Create an array S[1...n] with starting times such that S[i] contains s_i.
    A = \{R_1\}
                                               > select first interval from sorted list
   f = f_1
    while there are more intervals in S to look at
6
          do j = first interval for which s_i > f
              A \leftarrow A \cup \{i\}
              f \leftarrow f_i
8
9
    return A
```

## **Analysis**

- The sorting step in takes  $O(n \lg n)$  time.
- Creating the starting time array S[1..n] takes O(n) time.
- The single pass through the array S takes O(n) time

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
    R = Sorted requests in order of finishing times such that f_i \leq f_j when i < j.
    Create an array S[1...n] with starting times such that S[i] contains s_i.
    A = \{R_1\}
                                               > select first interval from sorted list
   f = f_1
    while there are more intervals in S to look at
6
          do j = first interval for which s_i > f
              A \leftarrow A \cup \{i\}
              f \leftarrow f_i
8
```

## **Analysis**

return A

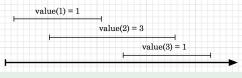
9

- The sorting step in takes O(n lg n) time.
- Creating the starting time array S[1..n] takes O(n) time.
- The single pass through the array S takes O(n) time
- An  $O(n \lg n)$  time algorithm for a problem with a natural search space of  $O(n2^n)$ .

## Definition (Weighted interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , with associated weights  $W = \{w_i\}$ , find  $A \subseteq I$  such that the members of A are non-conflicting and the total weight  $\sum_{i \in A} w_i$  is maximized.

## Example

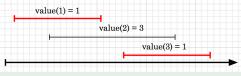


$$|A| = ???$$
,  $\sum_{i \in A} w_i = ???$ .

## Definition (Weighted interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , with associated weights  $W = \{w_i\}$ , find  $A \subseteq I$  such that the members of A are non-conflicting and the total weight  $\sum_{i \in A} w_i$  is maximized.

## Example (using our greedy strategy)

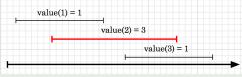


$$|A|=2$$
,  $\sum_{i\in A} w_i=2$ .

## Definition (Weighted interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , with associated weights  $W = \{w_i\}$ , find  $A \subseteq I$  such that the members of A are non-conflicting and the total weight  $\sum_{i \in A} w_i$  is maximized.

## Example (using an optimal strategy)

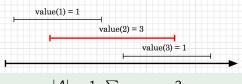


$$|A|=1$$
,  $\sum_{i\in A} w_i=3$ .

## Definition (Weighted interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , with associated weights  $W = \{w_i\}$ , find  $A \subseteq I$  such that the members of A are non-conflicting and the total weight  $\sum_{i \in A} w_i$  is maximized.

## Example (using an optimal strategy)



$$|A| = 1$$
,  $\sum_{i \in A} w_i = 3$ .

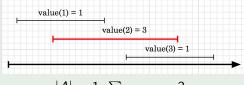
## Hmmm...

There is no greedy solution for the weighted interval scheduling problem!

## Definition (Weighted interval scheduling problem)

Given a set of schedules  $I = \{I_i\}$ , with associated weights  $W = \{w_i\}$ , find  $A \subseteq I$  such that the members of A are non-conflicting and the total weight  $\sum_{i \in A} w_i$  is maximized.

## Example (using an optimal strategy)



$$|A| = 1$$
,  $\sum_{i \in A} w_i = 3$ .

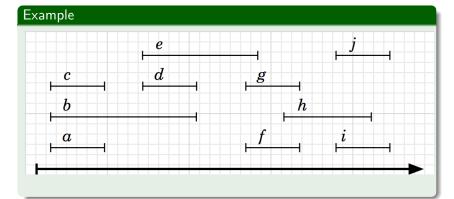
## Hmmm...

There is no greedy solution for the weighted interval scheduling problem! Why? (see Greedy Choice property later)

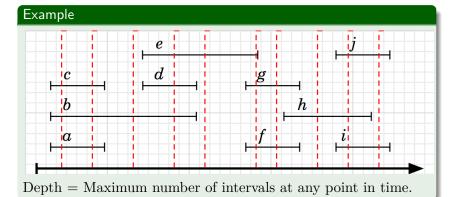
## Greedy algorithms

- Introduction
- Interval scheduling problem
- Scheduling all Intervals problem
- Fractional knapsack problem
- Coin changing problem
- What problems can be solved by greedy approach?
- Conclusion

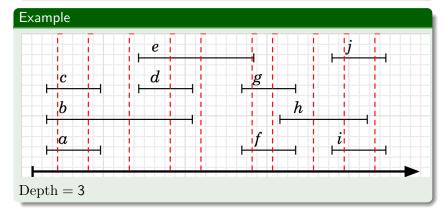
## Definition



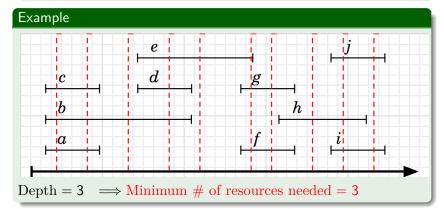
## Definition



## Definition



## Definition



# A greedy algorithm for scheduling all intervals

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
    R = Sorted requests in order of starting times, breaking ties
    arbitrarily, such that s_i \leq s_i when i < j.
    m \leftarrow 0 > the optimal number of resources needed to schedule R
3
    while R \neq \emptyset
4
          do reg = extract the next element in R
5
              if there is a resource j with no interval conflicting with req
6
                then schedule interval reg on resource j
                else
                       m \leftarrow m + 1 \triangleright allocate a new resource
9
                       schedule interval reg on resource m
```

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
   R = Sorted requests in order of starting times, breaking ties
    arbitrarily, such that s_i \leq s_i when i < j.
    m \leftarrow 0 > the optimal number of resources needed to schedule R
3
    while R \neq \emptyset
4
          do reg = extract the next element in R
5
              if there is a resource j with no interval conflicting with req
6
                then schedule interval reg on resource j
                else
                       m \leftarrow m + 1 \triangleright allocate a new resource
9
                       schedule interval reg on resource m
```

## Complexity

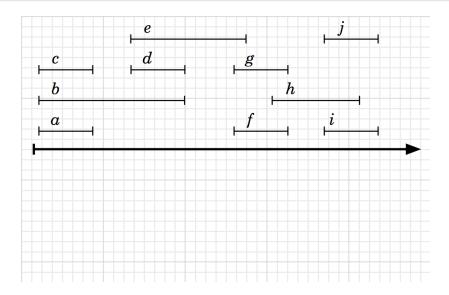
$$T(n) = O(n \lg n).$$

```
SCHEDULE-INTERVALS(I) \triangleright I = \{I_i\}, I_i = (s_i, f_i)
   R = Sorted requests in order of starting times, breaking ties
    arbitrarily, such that s_i \leq s_i when i < j.
    m \leftarrow 0 > the optimal number of resources needed to schedule R
3
    while R \neq \emptyset
4
          do reg = extract the next element in R
5
              if there is a resource j with no interval conflicting with req
6
                then schedule interval reg on resource j
                else
                       m \leftarrow m + 1 \triangleright allocate a new resource
9
                       schedule interval reg on resource m
```

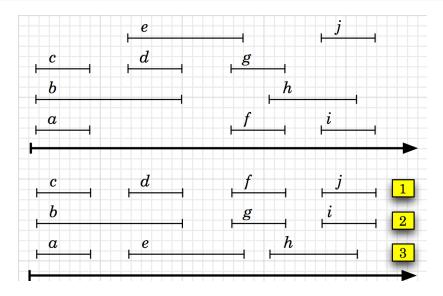
### Complexity

$$T(n) = O(n \lg n)$$
. Prove it.

# Scheduling all intervals in action



# Scheduling all intervals in action



#### Contents

### Greedy algorithms

- Introduction
- Interval scheduling problem
- Scheduling all Intervals problem
- Fractional knapsack problem
- Coin changing problem
- What problems can be solved by greedy approach?
- Conclusion

#### Definition (fractional knapsack problem)

Given a set S of n items, such that each item i has a positive benefit  $b_i$  and a positive weight  $w_i$ , the goal is to find the maximum-benefit subset that does not exceed a given weight W, allowing for fractional items.

### Definition (fractional knapsack problem)

Given a set S of n items, such that each item i has a positive benefit  $b_i$  and a positive weight  $w_i$ , the goal is to find the maximum-benefit subset that does not exceed a given weight W, allowing for fractional items.

• Taking fraction  $x_i$  of each item i, such that  $0 \le x_i \le w_i$  for each  $i \in S$ , and  $\sum_{i \in S} x_i \le W$ .

### Definition (fractional knapsack problem)

Given a set S of n items, such that each item i has a positive benefit  $b_i$  and a positive weight  $w_i$ , the goal is to find the maximum-benefit subset that does not exceed a given weight W, allowing for fractional items.

- Taking fraction  $x_i$  of each item i, such that  $0 \le x_i \le w_i$  for each  $i \in S$ , and  $\sum_{i \in S} x_i \le W$ .
- Benefit  $b_i$  is accumulated when whole object i is taken. So, for taking fraction  $x_i$  of item i is then  $b_i(x_i/w_i)$

### Definition (fractional knapsack problem)

Given a set S of n items, such that each item i has a positive benefit  $b_i$  and a positive weight  $w_i$ , the goal is to find the maximum-benefit subset that does not exceed a given weight W, allowing for fractional items.

- Taking fraction  $x_i$  of each item i, such that  $0 \le x_i \le w_i$  for each  $i \in S$ , and  $\sum_{i \in S} x_i \le W$ .
- Benefit  $b_i$  is accumulated when whole object i is taken. So, for taking fraction  $x_i$  of item i is then  $b_i(x_i/w_i)$
- Maximum-benefit subset is then maximizing  $\sum_{i \in S} b_i(x_i/w_i)$ .

#### Key question

• What strategy to use to select the next item (and the amount of it)?

- What strategy to use to select the next item (and the amount of it)?
- At least three different measures one can attempt to optimize when determining which object to select next

- What strategy to use to select the next item (and the amount of it)?
- At least three different measures one can attempt to optimize when determining which object to select next
- Total Profit

- What strategy to use to select the next item (and the amount of it)?
- At least three different measures one can attempt to optimize when determining which object to select next
- Total Profit
- Capacity Used

- What strategy to use to select the next item (and the amount of it)?
- At least three different measures one can attempt to optimize when determining which object to select next
- Total Profit
- Capacity Used
- Ratio of accumulated profit to capacity used

#### Key question

- What strategy to use to select the next item (and the amount of it)?
- At least three different measures one can attempt to optimize when determining which object to select next
- Total Profit
- Capacity Used
- Ratio of accumulated profit to capacity used

Let's see which measure to optimize...

Item	Benefit	Weight
Α	25	18 kg
В	15	10 kg
С	24	15 kg

Capacity W=20

Item	Benefit	Weight
Α	25	18 kg
В	15	10 kg
С	24	15 kg

Capacity W=20

First Strategy: Greedy method using Profit as it's measure. At each step it will choose an object that increases the profit the most.

• Capacity W = 20

- Capacity W = 20
- A has the maximum benefit 25. So select this. Remaining Capacity: 20 - 18 = 2

- Capacity W = 20
- A has the maximum benefit 25. So select this. Remaining Capacity: 20 - 18 = 2
- Then select object C as it has the second maximum benefit.

- Capacity W=20
- A has the maximum benefit 25. So select this. Remaining Capacity: 20 - 18 = 2
- Then select object C as it has the second maximum benefit.
- However, C's weight 15 > Current Capacity 2.

- Capacity W = 20
- A has the maximum benefit 25. So select this. Remaining Capacity: 20 - 18 = 2
- Then select object C as it has the second maximum benefit.
- However, C's weight 15 > Current Capacity 2.
- So select fraction of C:  $x_C = 2$ . Remaining Capacity: 2 - 2 = 0

- Capacity W = 20
- A has the maximum benefit 25. So select this. Remaining Capacity: 20 - 18 = 2
- Then select object C as it has the second maximum benefit.
- However, C's weight 15 > Current Capacity 2.
- So select fraction of C:  $x_C = 2$ . Remaining Capacity: 2 - 2 = 0
- Accumulated profit by adding  $x_C = 2$  to the knapsack is: 24(2/15)=3.2

- Capacity W = 20
- A has the maximum benefit 25. So select this. Remaining Capacity: 20 - 18 = 2
- Then select object C as it has the second maximum benefit.
- However, C's weight 15 > Current Capacity 2.
- So select fraction of C:  $x_C = 2$ . Remaining Capacity: 2 - 2 = 0
- Accumulated profit by adding  $x_C = 2$  to the knapsack is: 24(2/15)=3.2
- So, finally, total benefit = 25 + 3.2 = 28.2

 Note here. If we select C as first element and B as second element then,

- Note here. If we select C as first element and B as second element then.
- Total benefit= 24 + 15(5/10) = 31.5
- So, previous solution is not the optimal one.
- Thus, First Strategy fails.

**Second Strategy:** Greedy method using **Capacity** as it's measure will, at each step choose an object that increases the capacity the least.

• B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10

- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.

- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.
- However, C's weight 15 > Current Capacity 10.

- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.
- However, C's weight 15 > Current Capacity 10.
- So select fraction of C:  $x_C = 10$ . Remaining Capacity: 10 - 10 = 0

- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.
- However, C's weight 15 > Current Capacity 10.
- So select fraction of C:  $x_C = 10$ . Remaining Capacity: 10 - 10 = 0
- Accumulated profit by adding  $x_C = 10$  to the knapsack is: 24(10/15)=16

- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.
- However, C's weight 15 > Current Capacity 10.
- So select fraction of C:  $x_C = 10$ . Remaining Capacity: 10 - 10 = 0
- Accumulated profit by adding  $x_C = 10$  to the knapsack is: 24(10/15)=16
- So, total benefit = 15 + 16 = 31.

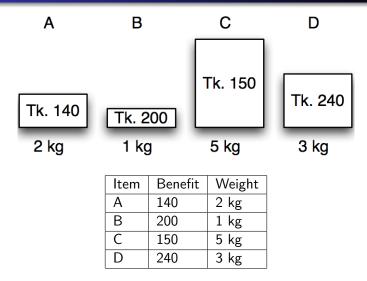
- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.
- However, C's weight 15 > Current Capacity 10.
- So select fraction of C:  $x_C = 10$ . Remaining Capacity: 10 - 10 = 0
- Accumulated profit by adding  $x_C = 10$  to the knapsack is: 24(10/15)=16
- So, total benefit = 15 + 16 = 31. Again, it is not the optimal one.

- B has the least weight 10. So select this. Remaining Capacity: 20 - 10 = 10
- Then select object C as it has the second least weight.
- However, C's weight 15 > Current Capacity 10.
- So select fraction of C:  $x_C = 10$ . Remaining Capacity: 10 - 10 = 0
- Accumulated profit by adding  $x_C = 10$  to the knapsack is: 24(10/15)=16
- So, total benefit = 15 + 16 = 31. Again, it is not the optimal one.
- Thus, second strategy fails.

**Third Strategy:** Strives to achieve a balance between the rate at which profit increases and the rate at which capacity is used.

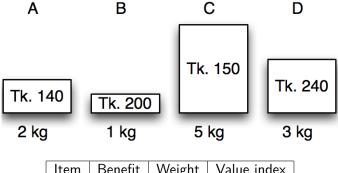
- At each step, include the object which has the maximum profit per unit of capacity used.
- That means, objects are considered in order of the ratio  $b_i/w_i$ .

# Fractional knapsack in action



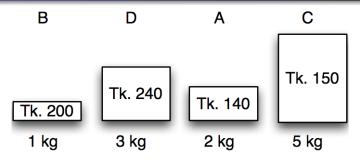
Calculate benefit/kg – the value index.

### Fractional knapsack in action



Item	Benefit	Weight	Value index
Α	140	2 kg	70
В	200	1 kg	200
С	150	5 kg	30
D	240	3 kg	80

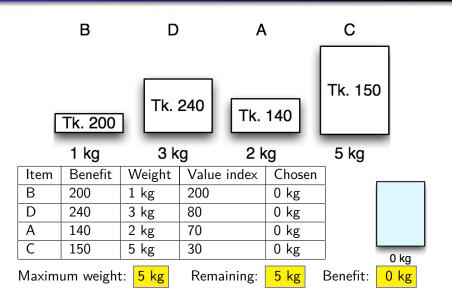
Sort by non-increasing value index.

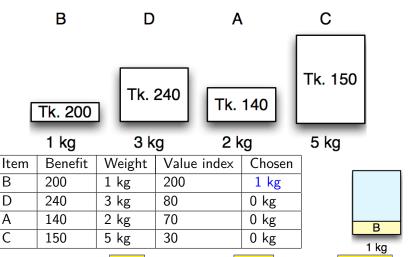


Item	Benefit	Weight	Value index
В	200	1 kg	200
D	240	3 kg	80
Α	140	2 kg	70
С	150	5 kg	30

Maximum weight:

5 kg



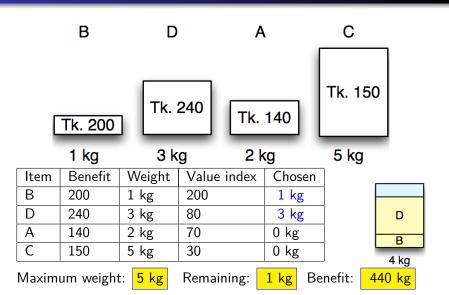


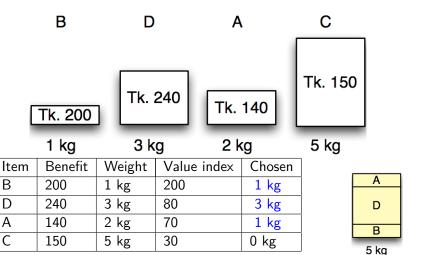
Maximum weight:

Remaining:

Benefit:

200 kg



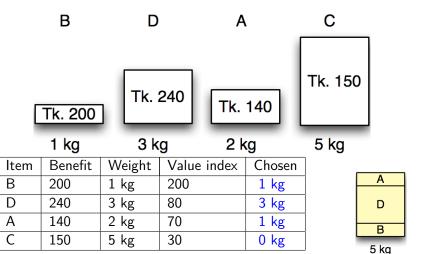


Maximum weight:

Remaining:

Benefit:

510 kg



Maximum weight:

Remaining:

Benefit:

510 kg

# Fractional knapsack greedy algorithm

```
FRACTIONAL-KNAPSACK(S, W) \triangleright S = \{(w_i, b_i)\}
     for each item i \in S
           do x_i \leftarrow 0 \Rightarrow amount of item i chosen (0 \le x \le w_i)
 3
               v_i \leftarrow b_i/w_i
                                                Sort the items in non-increasing order of the ratio b_i/w_i
 5
     w \leftarrow 0
     while w < W
           do i = \text{extract from } S the item with highest value index

    □ greedy choice

 8
               if w + w_i < W
 9
                 then x_i = w_i
10
                 else x_i = W - w > \text{fill up the remaining with } i
11
               w \leftarrow w + x_i
12
     return x \mapsto x_i contains amount of item i chosen
```

# Fractional knapsack greedy algorithm

```
FRACTIONAL-KNAPSACK(S, W) \triangleright S = \{(w_i, b_i)\}
     for each item i \in S
           do x_i \leftarrow 0 \triangleright amount of item i chosen (0 \le x \le w_i)
 3
               v_i \leftarrow b_i/w_i
                                                 Sort the items in non-increasing order of the ratio b_i/w_i
 5
     w \leftarrow 0
    while w < W
           do i = \text{extract from } S the item with highest value index

    □ greedy choice

 8
               if w + w_i < W
 9
                 then x_i = w_i
10
                 else x_i = W - w \triangleright fill up the remaining with i
11
               w \leftarrow w + x_i
12
     return x > x_i contains amount of item i chosen
```

### Complexity

$$T(n) = O(n \lg n).$$

CSE 221: Algorithms Licensed under

# Fractional knapsack greedy algorithm

```
FRACTIONAL-KNAPSACK(S, W) \triangleright S = \{(w_i, b_i)\}
     for each item i \in S
           do x_i \leftarrow 0 \triangleright amount of item i chosen (0 \le x \le w_i)
 3
               v_i \leftarrow b_i/w_i
                                                 Sort the items in non-increasing order of the ratio b_i/w_i
 5
     w \leftarrow 0
    while w < W
           do i = \text{extract from } S the item with highest value index

    □ greedy choice

 8
               if w + w_i < W
 9
                 then x_i = w_i
10
                 else x_i = W - w \triangleright fill up the remaining with i
11
               w \leftarrow w + x_i
12
     return x > x_i contains amount of item i chosen
```

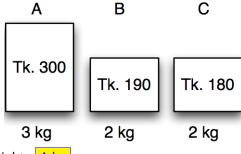
### Complexity

$$T(n) = O(n \lg n)$$
. Prove it.

Licensed under @

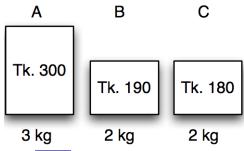
Exactly the same as the Fractional Knapsack Problem, except that fractional quantities are not allowed.

Exactly the same as the Fractional Knapsack Problem, except that fractional quantities are not allowed.



Maximum weight:

Exactly the same as the Fractional Knapsack Problem, except that fractional quantities are not allowed.

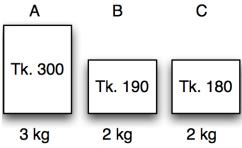


Maximum weight:

Greedy solution: item A

Benefit:

Exactly the same as the Fractional Knapsack Problem, except that fractional quantities are not allowed.



Maximum weight:

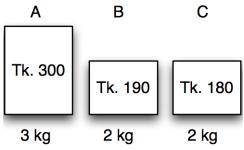
Greedy solution: item A

Optimal solution: items B and C

Benefit: 300

Benefit:

Exactly the same as the Fractional Knapsack Problem, except that fractional quantities are not allowed.



Maximum weight:

Greedy solution: item A

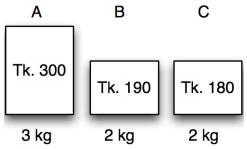
Optimal solution: items B and C

Benefit: 300

Benefit:

The 0/1 Knapsack Problem does not have a greedy solution!

Exactly the same as the Fractional Knapsack Problem, except that fractional quantities are not allowed.



Maximum weight:

Greedy solution: item A

Optimal solution: items B and C

Benefit: 300

Benefit:

The 0/1 Knapsack Problem does not have a greedy solution! Why?

### Greedy algorithms

- Introduction
- Interval scheduling problem
- Scheduling all Intervals problem
- Fractional knapsack problem
- Coin changing problem
- What problems can be solved by greedy approach?
- Conclusion

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

① Choose 2 25 coins, so remaining is 73 - 2 \* 25 = 23

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

① Choose 2 25 coins, so remaining is 73 - 2 \* 25 = 23

2 Choose 2 10 coins, so remaining is 23 - 2 \* 10 = 3

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

- ① Choose 2 25 coins, so remaining is 73 2 \* 25 = 23
- 2 Choose 2 10 coins, so remaining is 23 2 \* 10 = 3
- Choose 0 5 coins, so remaining is 3

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

- ① Choose 2 25 coins, so remaining is 73 2 \* 25 = 23
- 2 Choose 2 10 coins, so remaining is 23 2 \* 10 = 3
- Choose 0 5 coins, so remaining is 3
- Choose 3.1 coins, so remaining is 3 1 \* 3 = 0

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

- ① Choose 2 25 coins, so remaining is 73 2 \* 25 = 23
- 2 Choose 2 10 coins, so remaining is 23 2 \* 10 = 3
- Choose 0 5 coins, so remaining is 3
- Choose 3.1 coins, so remaining is 3 1 \* 3 = 0

Solution (and it's optimal):  $2 \times 25 + 2 \times 10 + 3 \times 1 = 7$  coins.

#### Definition

Given coin denominations in  $\{C\}$ , make change for a given amount A with the minimum number of coins.

#### Example

Coin denominations,  $C = \{25, 10, 5, 1\}$  Amount to change, A = 73

- ① Choose 2 25 coins, so remaining is 73 2 \* 25 = 23
- 2 Choose 2 10 coins, so remaining is 23 2 \* 10 = 3
- Choose 0 5 coins, so remaining is 3
- Choose 3.1 coins, so remaining is 3 1 \* 3 = 0

Solution (and it's optimal):  $2 \times 25 + 2 \times 10 + 3 \times 1 = 7$  coins.

#### Key question

Does a greedy approach always produce the optimal solution?

Coin denominations,  $C = \{12, 5, 1\}$  Amount to change, A = 15

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

Example (using greedy strategy)

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

① Choose 1 12 coins, so remaining is 15 - 1 \* 12 = 3

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

- Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

- ① Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Solution: 4 coins.

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

- ① Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Solution: 4 coins.

### Example (using optimal strategy)

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

- ① Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Solution: 4 coins.

#### Example (using optimal strategy)

• Choose 0 12 coins, so remaining is 15

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

- ① Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Solution: 4 coins.

#### Example (using optimal strategy)

- Choose 0 12 coins, so remaining is 15
- 2 Choose 3 5 coins, so remaining is 15 3 \* 5 = 0

Coin denominations,  $C = \{12, 5, 1\}$ Amount to change, A = 15

### Example (using greedy strategy)

- ① Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Solution: 4 coins.

#### Example (using optimal strategy)

- Choose 0 12 coins, so remaining is 15
- 2 Choose 3 5 coins, so remaining is 15 3 \* 5 = 0

Solution: 3 coins.

Coin denominations,  $C = \{12, 5, 1\}$  Amount to change, A = 15

### Example (using greedy strategy)

- Choose 1 12 coins, so remaining is 15 1 \* 12 = 3
- 2 Choose 3 1 coins, so remaining is 3 1 \* 3 = 0

Solution: 4 coins.

### Example (using optimal strategy)

- ① Choose 0 12 coins, so remaining is 15
- ② Choose 3 5 coins, so remaining is 15 3 \* 5 = 0

Solution: 3 coins.

#### Key observation

Correctness depends on the choice of coins, so greedy strategy does not provide a general solution to this problem!

### Contents

- Greedy algorithms
  - Introduction
  - Interval scheduling problem
  - Scheduling all Intervals problem
  - Fractional knapsack problem
  - Coin changing problem
  - What problems can be solved by greedy approach?
  - Conclusion

• There is no general of knowing whether a problem can be solved by a greedy algorithm.

- There is no general of knowing whether a problem can be solved by a greedy algorithm.
- If a problem has the following properties, then it's likely to have a greedy solution.

- There is no general of knowing whether a problem can be solved by a greedy algorithm.
- If a problem has the following properties, then it's likely to have a greedy solution.
  - Greedy choice property If the global optimal solution can be reached by making locally optimal choices, then it has the greedy choice property.

- There is no general of knowing whether a problem can be solved by a greedy algorithm.
- If a problem has the following properties, then it's likely to have a greedy solution.
  - Greedy choice property If the global optimal solution can be reached by making locally optimal choices, then it has the greedy choice property.
  - Subproblem optimality If the optimal solution to the entire problem contain optimal solution to the subproblems, then it has the subproblem optimality property.

• Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.
- The problem must have at least the following properties:
  - Greedy choice property

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.
- The problem must have at least the following properties:
  - Greedy choice property
  - Subproblem optimality

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.
- The problem must have at least the following properties:
  - Greedy choice property
  - Subproblem optimality
- The algorithm must be rigorously proven to be correct!

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.
- The problem must have at least the following properties:
  - Greedy choice property
  - Subproblem optimality
- The algorithm must be rigorously proven to be correct!
- Except for a few select problems, it is far better to use Dynamic Programming to solve such optimization problems.

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.
- The problem must have at least the following properties:
  - Greedy choice property
  - Subproblem optimality
- The algorithm must be rigorously proven to be correct!
- Except for a few select problems, it is far better to use Dynamic Programming to solve such optimization problems.
- So why study greedy algorithms?

- Greedy algorithms often lead to polynomial time-solution for an exponential-time problem.
- However, not all optimization problems can be solved by the greedy strategy.
- The problem must have at least the following properties:
  - Greedy choice property
  - Subproblem optimality
- The algorithm must be rigorously proven to be correct!
- Except for a few select problems, it is far better to use Dynamic Programming to solve such optimization problems.
- So why study greedy algorithms? Because there are very efficient provably correct greedy algorithms for many common problems (wait till we study graph algorithms).