Chapter 12

Greedy Algorithms

OLD CS 473: Fundamental Algorithms, Spring 2015
March 3, 2015

12.1 Problems and Terminology

12.2 Problem Types

12.2.0.1 Problem Types

(A) **Decision Problem**: Is the input a YES or NO input?
Example: Given graph G, nodes s, t, is there a path from s to t in G?
(B) **Search Problem:** Find a *solution* if input is a YES input.
Example: Given graph G, nodes s, t, find an s-t path.
(C) **Optimization Problem:** Find a *best* solution among all solutions for the input.
Example: Given graph G, nodes s, t, find a shortest s-t path.

12.2.0.2 Terminology

(A) A **problem** Π consists of an *infinite* collection of inputs $\{I_1, I_2, \ldots, \}$. Each input is referred to as an **instance**.
(B) The **size** of an instance I is the number of bits in its representation.
(C) I: instance. $sol(I)$: set of **feasible solutions** to I.
(D) *Implicit assumption:* given I, $y \in \Sigma^*$, one can check (efficiently!) if $y \in sol(I)$.
(E) $\implies$ Problem is in **NP**. (More on this later in the course.)
(F) Optimization problems:
$\forall$ solution $s \in sol(I)$ has associated **value**.
(G) *Implicit assumption:* given s, can compute value efficiently.

12.2.0.3 Problem Types

Given instance I...
(A) **Decision Problem**: Output whether $sol(I) = \emptyset$ or not.

(B) **Search Problem:** Compute solution $s \in sol(I)$ if $sol(I) \neq \emptyset$.
(C) **Optimization Problem:** Given I,
    (A) Minimization problem. Find solution $s \in sol(I)$ of min value.
    (B) Maximization problem. Find solution $s \in sol(I)$ of max value.
    (C) Notation:
    $opt(I)$: denote the value of an optimum solution or some fixed optimum solution.

12.3 Greedy Algorithms: Tools and Techniques

12.3.0.4 What is a Greedy Algorithm?

No real consensus on a universal definition.

Greedy algorithms:

(A) Do the right thing. Locally.
(B) Make decisions incrementally in small steps *no backtracking*.
(C) Decision at each step based on improving *local or current* state in myopic fashion. ... without considering the *global* situation.
(D) ***myopia***: lack of understanding or foresight.
(E) Decisions often based on some fixed and simple *priority* rules.

12.3.0.5 Pros and Cons of Greedy Algorithms

Pros:

(A) Usually (too) easy to design greedy algorithms
(B) Easy to implement and often run fast since they are simple
(C) Several important cases where they are effective/optimal
(D) Lead to first-cut heuristic when problem not well understood

Cons:

(A) **Very often** greedy algorithms don't work.
(B) Easy to lull oneself into believing they work
(C) Many greedy algorithms possible for a problem and no structured way to find effective ones.

CS 473: Every greedy algorithm needs a proof of correctness

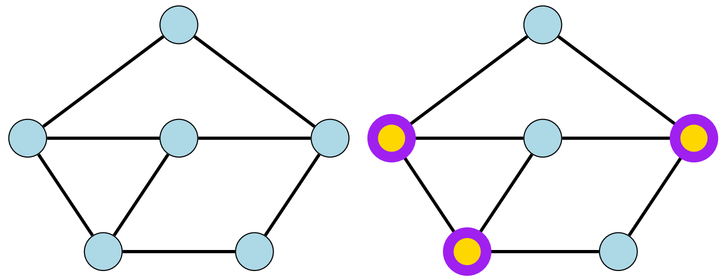
12.3.0.6 Greedy Algorithm Types

(A) Crude classification:
    (A) **Non-adaptive:** fix ordering of decisions a priori and stick with it.
    (B) **Adaptive:** make decisions adaptively but greedily/locally at each step.
(B) Plan:
    (A) See several examples
    (B) Pick up some proof techniques

12.3.0.7 Vertex Cover

Given a graph $\mathsf{G} = (\mathsf{V}, \mathsf{E})$, a set of vertices S is:

(A) A ***vertex cover*** if every $e \in E$ has at least one endpoint in S.



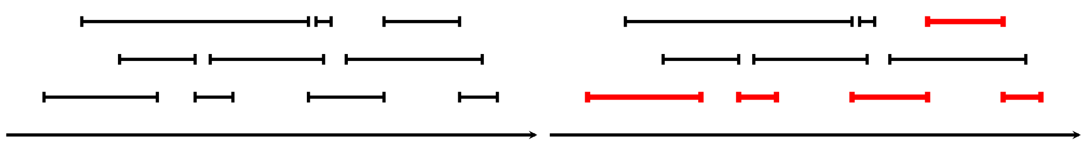
Natural algorithms for computing vertex cover?

12.4 Interval Scheduling

12.4.0.8 Interval Scheduling

Problem 12.4.1 (Interval Scheduling).
Input: *A set of jobs with start and finish times to be scheduled on a resource (example: classes and class rooms).*
Goal: *Schedule as many jobs as possible*

(A) Two jobs with overlapping intervals cannot both be scheduled!



12.4.1 The Algorithm

12.4.1.1 Greedy Template

R is the set of all requests
X is empty (* X will store all the jobs that will be scheduled *)
while R is not empty **do**
    <2->choose $i \in R$
    add i to X
    remove from R all requests that overlap with i
return the set X

Main task: Decide the order in which to process requests in R

12.4.1.2 Earliest Start Time

Process jobs in the order of their starting times, beginning with those that start earliest.

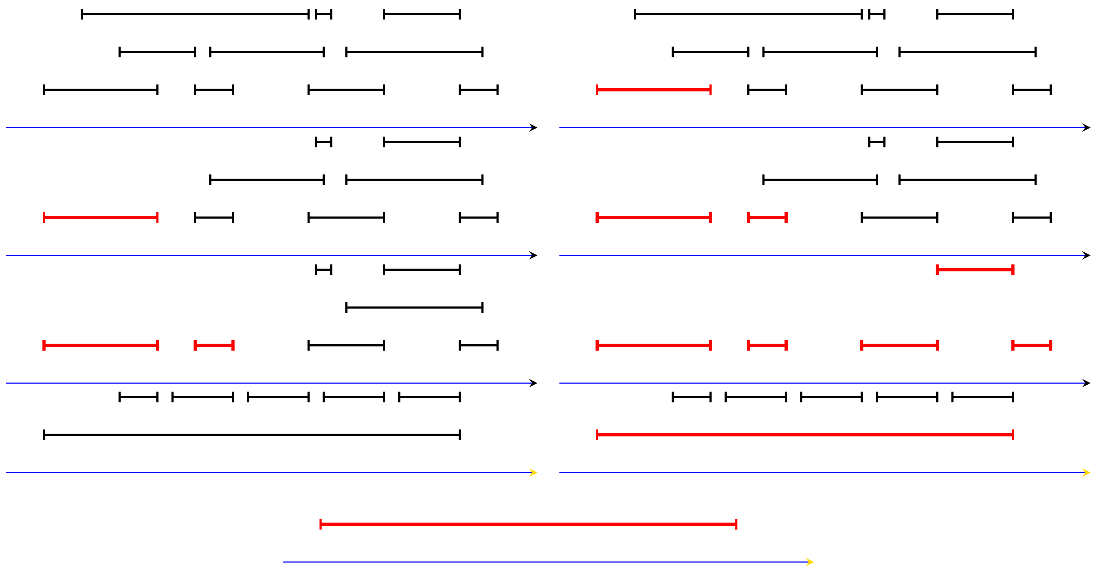
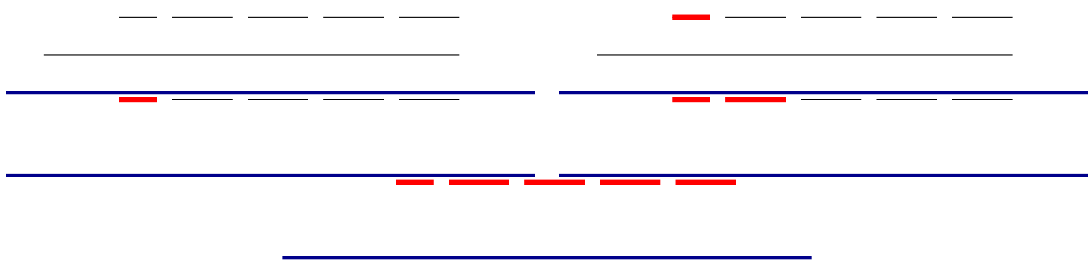


Figure 12.1: Counter example for earliest start time



12.4.1.3 Smallest Processing Time

Process jobs in the order of processing time, starting with jobs that require the shortest processing.

12.4.1.4 Fewest Conflicts

Process jobs in that have the fewest "conflicts" first.

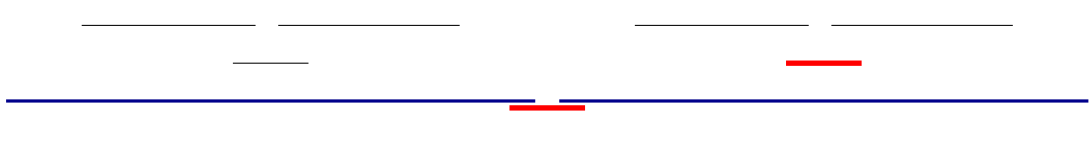


Figure 12.2: Counter example for smallest processing time

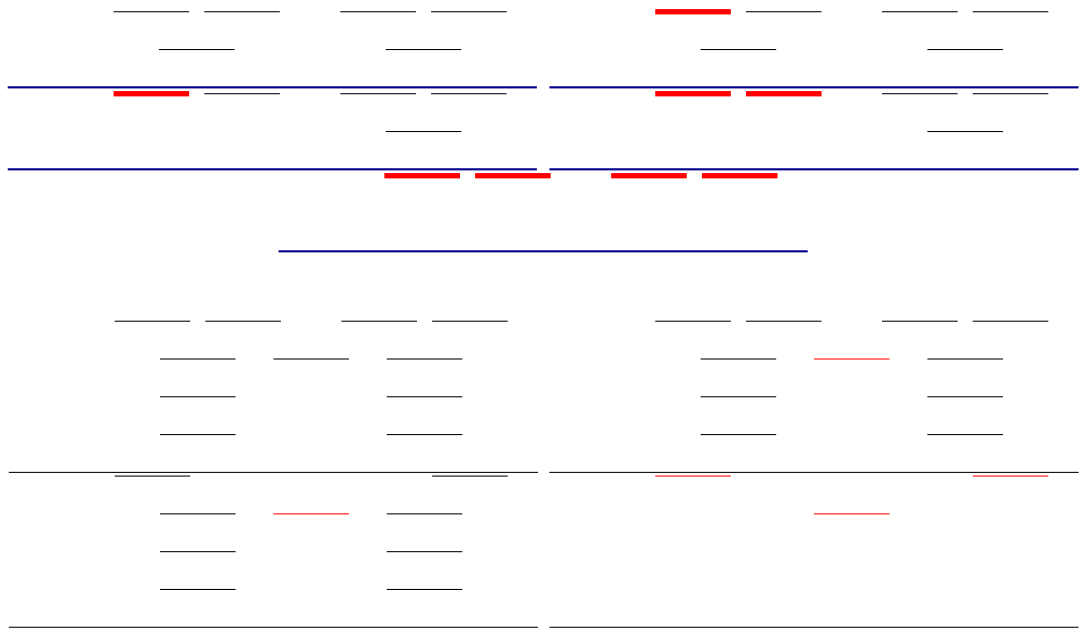


Figure 12.3: Counter example for fewest conflicts

12.4.1.5 Earliest Finish Time

Process jobs in the order of their finishing times, beginning with those that finish earliest.

12.4.2 Correctness

12.4.2.1 Optimal Greedy Algorithm

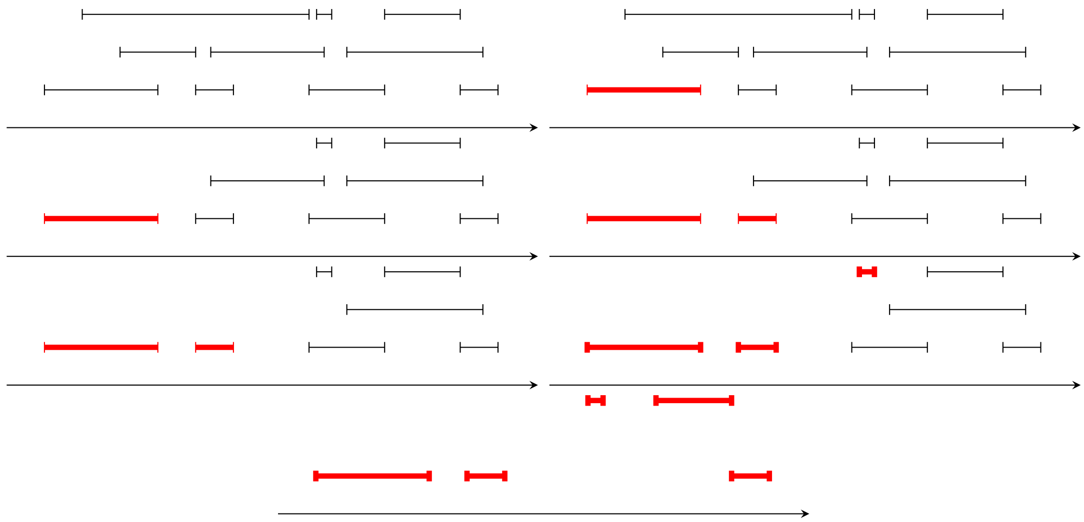
```
R is the set of all requests
X is empty (* X will store all the jobs that will be scheduled *)
while R is not empty
    choose i ∈ R such that finishing time of i is least
    add i to X
    remove from R all requests that overlap with i
return X
```

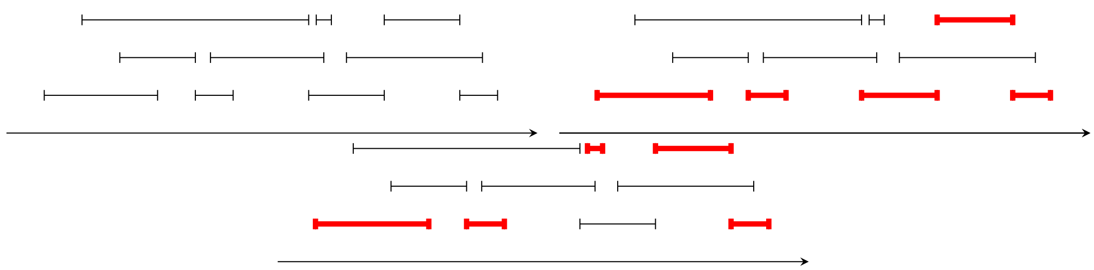
Theorem 12.4.2. *The greedy algorithm that picks jobs in the order of their finishing times is optimal.*

12.4.2.2 Proving Optimality

(A) **Correctness:** Clearly the algorithm returns a set of jobs that does not have any conflicts

(B) For a set of requests R, let O be an optimal set and let X be the set returned by the greedy algorithm. Then $O = X$? Not likely! Instead we will show that $|O| = |X|$





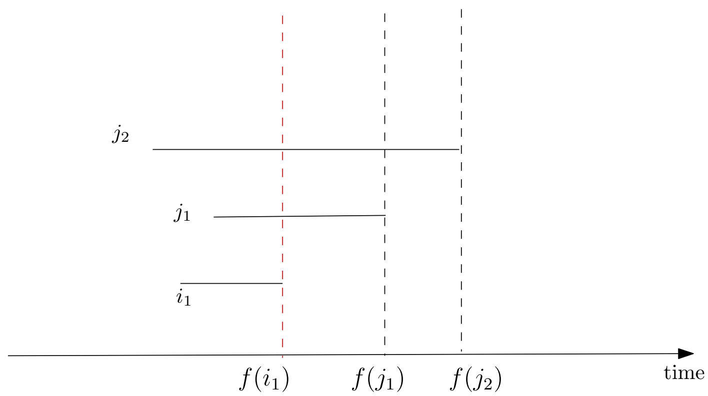


Figure 12.4: Since i_1 has the earliest finish time, any interval that conflicts with it does so at $f(i_1)$. This implies j_1 and j_2 conflict.

12.4.2.3 Proof of Optimality: Key Lemma

Lemma 12.4.3. *Let i_1 be first interval picked by Greedy. There exists an optimum solution that contains i_1.*

Proof: Let O be an *arbitrary* optimum solution. If $i_1 \in O$ we are done.

Claim: If $i_1 \notin O$ then there is exactly one interval $j_1 \in O$ that conflicts with i_1. (proof later)

(A) Form a new set O' by removing j_1 from O and adding i_1, that is $O' = (O - \{j_1\}) \cup \{i_1\}$.
(B) From claim, O' is a *feasible* solution (no conflicts).
(C) Since $|O'| = |O|$, O' is also an optimum solution and it contains i_1.

■

12.4.2.4 Proof of Claim

Claim 12.4.4. *If $i_1 \notin O$ then there is exactly one interval $j_1 \in O$ that conflicts with i_1.*

Proof:
(A) Suppose $j_1, j_2 \in O$ such that $j_1 \neq j_2$ and both j_1 and j_2 conflict with i_1.
(B) Since i_1 has earliest finish time, j_1 and i_1 overlap at $f(i_1)$.
(C) For same reason j_2 also overlaps with i_1 at $f(i_1)$.
(D) Implies that j_1, j_2 overlap at $f(i_1)$ contradicting the feasibility of O.
See figure in next slide. ■

12.4.2.5 Figure for proof of Claim

12.4.2.6 Proof of Optimality of Earliest Finish Time First

Proof:[Proof by Induction on number of intervals] **Base Case:** $n = 1$. Trivial since Greedy picks one interval.
Induction Step: Assume theorem holds for $i < n$.

Let I be an instance with n intervals
I': I with i_1 and all intervals that overlap with i_1 removed
$G(I), G(I')$: Solution produced by Greedy on I and I'
From Lemma, there is an optimum solution O to I and $i_1 \in O$.
Let $O' = O - \{i_1\}$. O' is a solution to I'.

$$\begin{aligned} |G(I)| &= 1 + |G(I')| \quad \text{(from Greedy description)} \\ &\leq 1 + |O'| \quad \text{(By induction, } G(I') \text{ is optimum for } I'\text{)} \\ &= |O| \end{aligned}$$

■

12.4.3 Running Time

12.4.3.1 Implementation and Running Time

```
Initially R is the set of all requests
X is empty (* X will store all the jobs that will be scheduled *)
while R is not empty
    <3>choose i ∈ R such that finishing time of i is least
    <4>if i does not overlap with requests in X
        add i to X
    <5>remove i from R
return the set X
```

(A) Presort all requests based on finishing time. $O(n \log n)$ time
(B) Now choosing least finishing time is $O(1)$
(C) Keep track of the finishing time of the last request added to A. Then check if starting time of i later than that
(D) Thus, checking non-overlapping is $O(1)$
(E) Total time $O(n \log n + n) = O(n \log n)$

12.4.4 Extensions and Comments

12.4.4.1 Comments

(A) Interesting Exercise: smallest interval first picks at least half the optimum number of intervals.
(B) All requests need not be known at the beginning. Such *online* algorithms are a subject of research.

12.4.5 Interval Partitioning

12.4.6 The Problem

12.4.6.1 Scheduling all Requests

Input A set of lectures, with start and end times

Goal Find the minimum number of classrooms, needed to schedule all the lectures such two lectures do not occur at the same time in the same room.

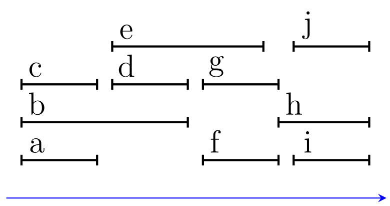


Figure 12.5: A schedule requiring 4 classrooms

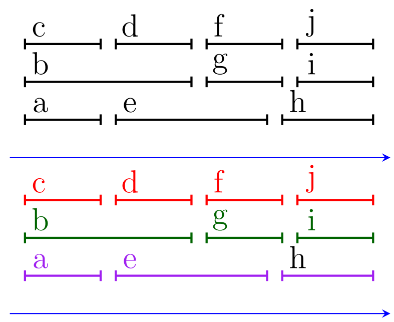


Figure 12.6: A schedule requiring 3 classrooms

12.4.7 The Algorithm

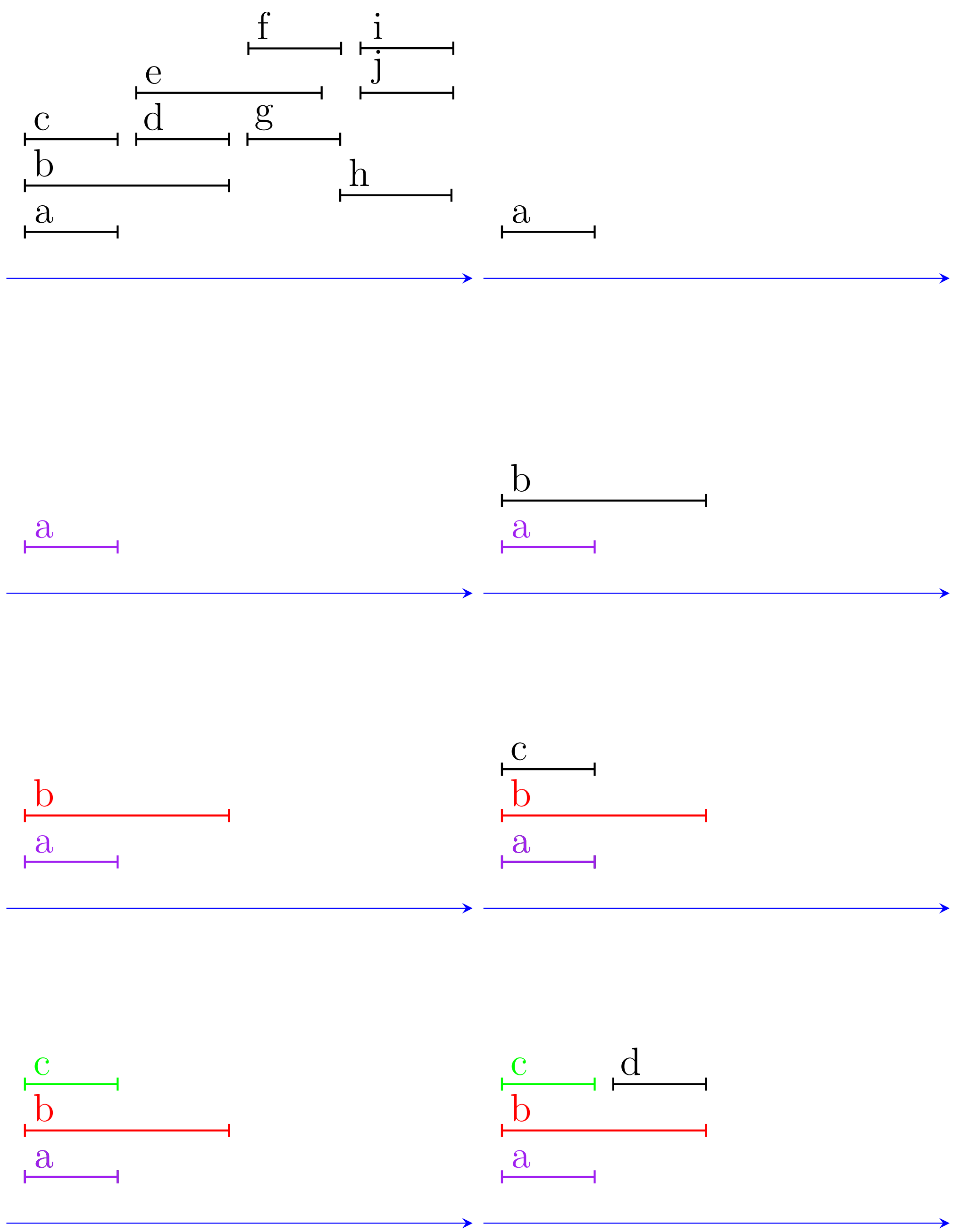
12.4.7.1 Greedy Algorithm

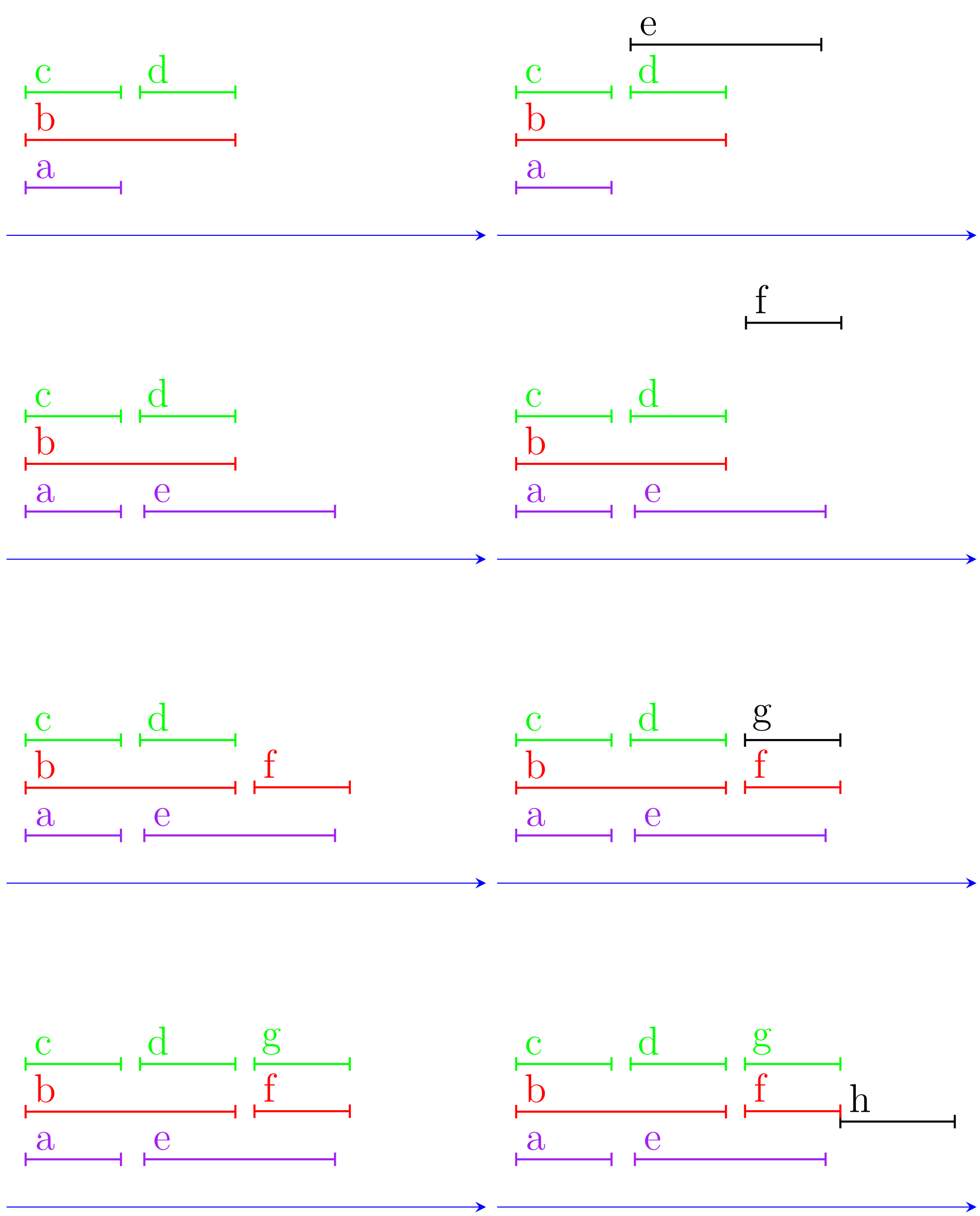
```
Initially R is the set of all requests
d = 0 (* number of classrooms *)
while R is not empty do
    choose i ∈ R such that start time of i is earliest
    if i can be scheduled in some class-room k ≤ d
        schedule lecture i in class-room k
    else
        allocate a new class-room d + 1
                and schedule lecture i in d + 1
        d = d + 1
```

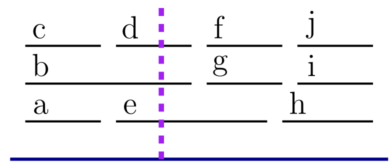
What order should we process requests in? According to start times (breaking ties arbitrarily)

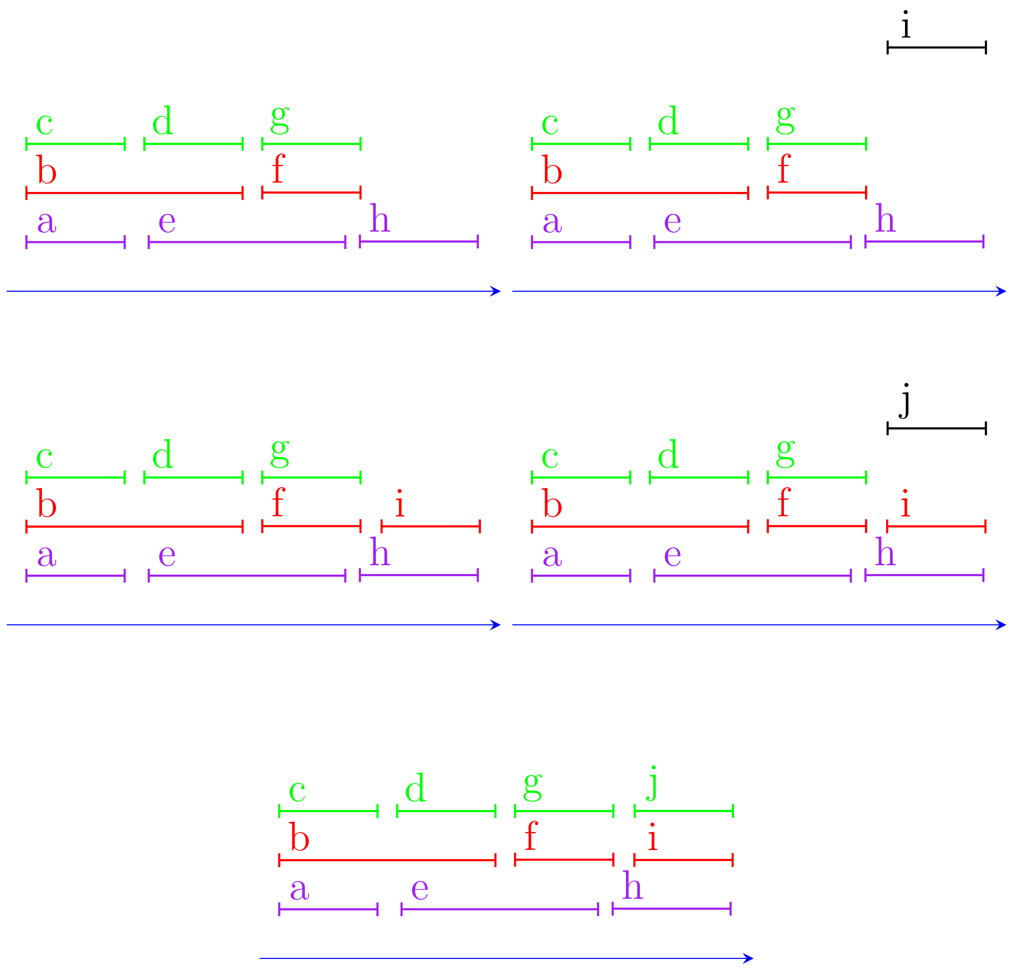
12.4.8 Example of algorithm execution

12.4.8.1 “Few things are harder to put up with than a good example.” – Mark Twain




c d b a
e c d b a
c d b a e
f c d b a e
c d b f a e
g c d f b a e
c d g b f a e
c d g b f h a e



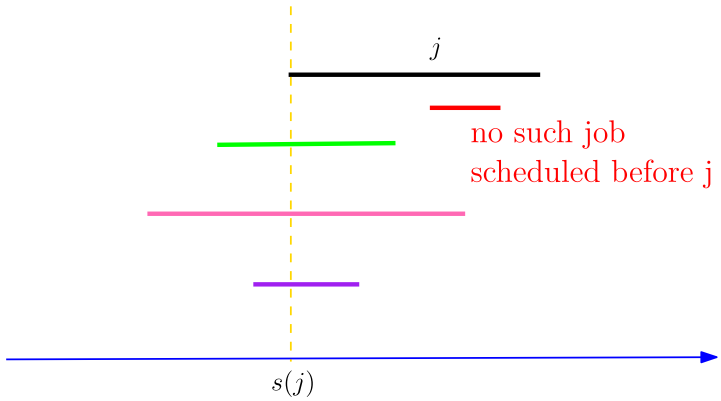


12.4.9 Correctness

12.4.9.1 Depth of Lectures

Definition 12.4.5. *(A) For a set of lectures R, k are said to be **in conflict** if there is some time t such that there are k lectures going on at time t.*

*(B) The **depth** of a set of lectures R is the maximum number of lectures in conflict at any time.*



12.4.9.2 Depth and Number of Class-rooms

Lemma 12.4.6. *For any set R of lectures, the number of class-rooms required is at least the depth of R.*

Proof: All lectures that are in conflict must be scheduled in different rooms. ■

12.4.9.3 Number of Class-rooms used by Greedy Algorithm

Lemma 12.4.7. *Let d be the depth of the set of lectures R. The number of class-rooms used by the greedy algorithm is d.*

Proof:

(A) Suppose the greedy algorithm uses more that d rooms. Let j be the first lecture that is scheduled in room $d+1$.

(B) Since we process lectures according to start times, there are d lectures that start (at or) before j and which conflict with j.

(C) Thus, *at the start time of* j, there are at least $d+1$ lectures in conflict, which contradicts the fact that the depth is d.

■

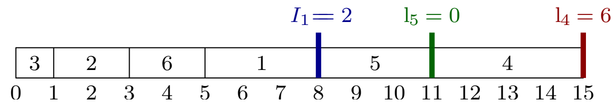
12.4.9.4 Figure

12.4.9.5 Correctness

Observation 12.4.8. *The greedy algorithm does not schedule two overlapping lectures in the same room.*

Theorem 12.4.9. *The greedy algorithm is correct and uses the optimal number of class-rooms.*

	1	2	3	4	5	6
t_i	3	2	1	4	3	2
d_i	6	8	9	9	14	15



12.4.10 Running Time

12.4.10.1 Implementation and Running Time

```
Initially R is the set of all requests
d = 0 (* number of classrooms *)
while R is not empty
    <1-2>choose i ∈ R such that start time of i is earliest
    <3->if i can be scheduled in some class-room k ≤ d
        schedule lecture i in class-room k
    else
        allocate a new class-room d + 1 and schedule lecture i in d + 1
        d = d + 1
```

(A) Presort according to start times. Picking lecture with earliest start time can be done in $O(1)$ time.
(B) Keep track of the finish time of last lecture in each room.
(C) ¡4¿Checking conflict takes $O(d)$ time.¡5¿With priority queues, checking conflict takes $O(\log d)$ time.
(D) Total time ¡4¿$= O(n \log n + nd)$¡5¿$= O(n \log n + n \log d) = O(n \log n)$

12.5 Scheduling to Minimize Lateness

12.5.1 The Problem

12.5.1.1 Scheduling to Minimize Lateness

(A) Given jobs with deadlines and processing times to be scheduled on a single resource.
(B) If a job i starts at time s_i then it will finish at time $f_i = s_i + t_i$, where t_i is its processing time. d_i: deadline.
(C) The lateness of a job is $l_i = \max(0, f_i - d_i)$.
(D) Schedule all jobs such that $L = \max l_i$ is **minimized**.

12.5.1.2 A Simpler Feasibility Problem

(A) Given jobs with deadlines and processing times to be scheduled on a single resource.
(B) If a job i starts at time s_i then it will finish at time $f_i = s_i + t_i$, where t_i is its processing time.

(C) Schedule all jobs such that each of them completes before its deadline (in other words $L = \max_i l_i = 0$).

Definition 12.5.1. *A schedule is* ***feasible*** *if all jobs finish before their deadline.*

12.5.2 The Algorithm

12.5.2.1 Greedy Template

```
Initially R is the set of all requests
curr_time = 0
while R is not empty do
    &lt;2->choose i ∈ R
    curr_time = curr_time + t_i
    if (curr_time > d_i) then
        return ``no feasible schedule''

return ``found feasible schedule''
```

Main task: Decide the order in which to process jobs in R

12.5.2.2 Three Algorithms

(A) Shortest job first — sort according to t_i.
(B) Shortest slack first — sort according to $d_i - t_i$.
(C) EDF = Earliest deadline first — sort according to d_i.
Counter examples for first two: exercise

12.5.2.3 Earliest Deadline First

Theorem 12.5.2. *Greedy with EDF rule for picking requests correctly decides if there is a feasible schedule.*

Proof via an exchange argument.
Idle time: time during which machine is not working.

Lemma 12.5.3. *If there is a feasible schedule then there is one with no idle time before all jobs are finished.*

12.5.2.4 Inversions

Definition 12.5.4. *A schedule S is said to have an* ***inversion*** *if there are jobs i and j such that S schedules i before j, but $d_i > d_j$.*

Claim 12.5.5. *If a schedule S has an inversion then there is an inversion between two* adjacently *scheduled jobs.*

Proof: exercise.

12.5.2.5 Main Lemma

Lemma 12.5.6. *If there is a feasible schedule, then there is one with no inversions.*

Proof:[Proof Sketch] Let S be a schedule with minimum number of inversions.
(A) If S has 0 inversions, done.
(B) Suppose S has one or more inversions. By claim there are two adjacent jobs i and j that define an inversion.
(C) Swap positions of i and j.
(D) New schedule is still feasible. (Why?)
(E) New schedule has one fewer inversion — contradiction!

∎

12.5.2.6 Back to Minimizing Lateness

Goal: schedule to minimize $L = \max_i l_i$.

How can we use algorithm for simpler feasibility problem?
Given a lateness bound L, can we check if there is a schedule such that $\max_i l_i \le L$?
Yes! Set $d'_i = d_i + L$ for each job i. Use feasibility algorithm with new deadlines.
How can we find *minimum* L? Binary search!

12.5.2.7 Binary search for finding minimum lateness

```
L = L_min = 0
L_max = Σ_i t_i  // why is this sufficient?
While L_min < L_max do
    L = ⌊(L_max + L_min)/2⌋
    check if there is a feasible schedule with lateness L
    if ''yes'' then L_max = L
    else L_min = L + 1
end while
return L
```

Running time: $O(n \log n \cdot \log T)$ where $T = \sum_i t_i$
(A) $O(n \log n)$ for feasibility test (sort by deadlines)
(B) $O(\log T)$ calls to feasibility test in binary search

12.5.2.8 Do we need binary search?

What happens in each call?
EDF algorithm with deadlines $d'_i = d_i + L$.

Greedy with EDF schedules the jobs in the same order for all L!!!
Maybe there is a direct greedy algorithm for minimizing maximum lateness?

12.5.2.9 Greedy Algorithm for Minimizing Lateness

```
Initially R is the set of all requests
curr_time = 0
curr_late = 0
while R is not empty
    choose i ∈ R with earliest deadline
    curr_time = curr_time + t_i
    late = curr_time − d_i
    curr_late = max(late, curr_late)
return curr_late
```

Exercise: argue directly that above algorithm is correct

Can be easily implemented in $O(n \log n)$ time after sorting jobs.

12.5.2.10 Greedy Analysis: Overview

(A) **Greedy's first step leads to an optimum solution.** Show that there is an optimum solution leading from the first step of Greedy and then use induction. Example, Interval Scheduling.

(B) **Greedy algorithm stays ahead.** Show that after each step the solution of the greedy algorithm is at least as good as the solution of any other algorithm. Example, Interval scheduling.

(C) **Structural property of solution.** Observe some structural bound of every solution to the problem, and show that greedy algorithm achieves this bound. Example, Interval Partitioning.

(D) **Exchange argument.** Gradually transform any optimal solution to the one produced by the greedy algorithm, without hurting its optimality. Example, Minimizing lateness.

12.5.2.11 Takeaway Points

(A) Greedy algorithms come naturally but often are incorrect. A proof of correctness is an absolute necessity.

(B) *Exchange* arguments are often the key proof ingredient. Focus on why the first step of the algorithm is correct: need to show that there is an optimum/correct solution with the first step of the algorithm.

(C) Thinking about correctness is also a good way to figure out which of the many greedy strategies is likely to work.

Bibliography