Part I

Problems and Terminology
Problem Types

- **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 $G$?

- **Search Problem**: Find a solution if input is a YES input.
  Example: Given graph $G$, nodes $s$, $t$, find an $s$-$t$ path.

- **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.
A problem $\Pi$ consists of an infinite collection of inputs $\{I_1, I_2, \ldots, \}$. Each input is referred to as an instance.

The size of an instance $I$ is the number of bits in its representation.

For an instance $I$, $sol(I)$ is a set of feasible solutions to $I$. **Typical implicit assumption:** given instance $I$ and $y \in \Sigma^*$, there is an way to check if $y \in sol(I)$. In other words, problem is in NP.

For optimization problems each solution $s \in sol(I)$ has an associated value. **Typical implicit assumption:** given $s$, can compute value efficiently.
Problem Types

- **Decision Problem**: Given $I$ output whether $sol(I) = \emptyset$ or not.
- **Search Problem**: Given $I$, find a solution $s \in sol(I)$ if $sol(I) \neq \emptyset$.
- **Optimization Problem**: Given $I$,
  - Minimization problem. Find a solution $s \in sol(I)$ of minimum value
  - Maximization problem. Find a solution $s \in sol(I)$ of maximum value
- **Notation**: $opt(I)$: interchangeably (when there is no confusion) used to denote the value of an optimum solution or some fixed optimum solution.
Part II

Greedy Algorithms: Tools and Techniques
What is a Greedy Algorithm?

No real consensus on a universal definition.

Greedy algorithms:
- make decision incrementally in small steps
- without backtracking
- decision at each step is based on improving local or current state in a myopic fashion without paying attention to the global situation
- decisions often based on some fixed and simple priority rules
What is a Greedy Algorithm?

No real consensus on a universal definition.
What is a Greedy Algorithm?

No real consensus on a universal definition.

Greedy algorithms:

- make decision incrementally in small steps *without* backtracking
- decision at each step is based on improving *local or current* state in a myopic fashion *without* paying attention to the *global* situation
- decisions often based on some fixed and simple *priority* rules
Pros and Cons of Greedy Algorithms

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

Cons:
- Very often greedy algorithms don’t work. Easy to lull oneself into believing they work
- 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
Pros and Cons of Greedy Algorithms

Pros:

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

Cons:

- Very often greedy algorithms don’t work. Easy to lull oneself into believing they work
- Many greedy algorithms possible for a problem and no structured way to find effective ones
Pros and Cons of Greedy Algorithms

Pros:

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

Cons:

- Very often greedy algorithms don’t work. Easy to lull oneself into believing they work
- 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
Greedy Algorithm Types

Crude classification:

- **Non-adaptive**: fix some ordering of decisions apriori and stick with the order
- **Adaptive**: make decisions adaptively but greedily/locally at each step
Greedy Algorithm Types

Crude classification:

- **Non-adaptive:** fix some ordering of decisions apriori and stick with the order
- **Adaptive:** make decisions adaptively but greedily/locally at each step

Plan:

- See several examples
- Pick up some proof techniques
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
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

- Two jobs with overlapping intervals cannot both be scheduled!
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

- Two jobs with overlapping intervals cannot both be scheduled!

---

Diagram showing overlapping intervals and a sorted list of intervals to be scheduled.
Greedy Template

Initially $R$ is the set of all requests
$A$ is empty (* $A$ will store all the jobs that will be scheduled *)
while $R$ is not empty
  choose $i \in R$
  add $i$ to $A$
  remove from $R$ all requests that overlap with $i$
return the set $A$
Initially R is the set of all requests
A is empty (* A will store all the jobs that will be scheduled *)
while R is not empty
    choose i ∈ R
    add i to A
    remove from R all requests that overlap with i
return the set A

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

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

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

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

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

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

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

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

Figure: Counter example for earliest start time
Earliest Start Time

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

Figure: Counter example for earliest start time
Earliest Start Time

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

Figure: Counter example for earliest start time
Smallest Processing Time

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

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

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

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

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

---

Back Counter
Smallest Processing Time

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

Figure: Counter example for smallest processing time
Smallest Processing Time

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

Figure: Counter example for smallest processing time
Smallest Processing Time

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

Figure: Counter example for smallest processing time
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.
Fewest Conflicts

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

Process jobs in that have the fewest “conflicts” first.
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.

---

Back  Counter
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.

Figure: Counter example for fewest conflicts
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.

Figure: Counter example for fewest conflicts
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.

Figure: Counter example for fewest conflicts
Fewest Conflicts

Process jobs in that have the fewest “conflicts” first.

Figure: Counter example for fewest conflicts
Earliest Finish Time

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

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

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

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

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

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

---

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

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

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

Theorem

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

- **Correctness**: Clearly the algorithm returns a set of jobs that does not have any conflicts.
Proving Optimality

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

- For a set of requests $R$, let $O$ be the optimal set and let $A$ be the set returned by the greedy algorithm. Then $O = A$
Proving Optimality

- **Correctness**: Clearly the algorithm returns a set of jobs that does not have any conflicts.
- For a set of requests $R$, let $O$ be the optimal set and let $A$ be the set returned by the greedy algorithm. Then $O = A$? Not likely!

Diagram: [Illustration of intervals]
Proving Optimality

- **Correctness**: Clearly the algorithm returns a set of jobs that does not have any conflicts.
- For a set of requests $R$, let $O$ be the optimal set and let $A$ be the set returned by the greedy algorithm. Then $O = A$? Not likely!
Proving Optimality

- **Correctness:** Clearly the algorithm returns a set of jobs that does not have any conflicts.
- For a set of requests $R$, let $O$ be the optimal set and let $A$ be the set returned by the greedy algorithm. Then $O = A$? Not likely!
Proving Optimality

- **Correctness:** Clearly the algorithm returns a set of jobs that does not have any conflicts.

- For a set of requests $R$, let $O$ be the optimal set and let $A$ be the set returned by the greedy algorithm. Then $O = A$? Not likely!

Instead we will show that $|O| = |A|$.
Proof of Optimality: Key Lemma

**Lemma**

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 and $j_1$ be the interval in $O$ with the smallest finish time.
Proof of Optimality: Key Lemma

**Lemma**

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 and $j_1$ be the interval in $O$ with the smallest finish time. **Claim:** $i_1$ does not conflict with any interval in $O - \{j_1\}$. (proof later)
Proof of Optimality: Key Lemma

Lemma

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 and \( j_1 \) be the interval in \( O \) with the smallest finish time.

Claim: \( i_1 \) does not conflict with any interval in \( O - \{ j_1 \} \). (proof later)

- Form a new set \( O' \) by and adding \( i_1 \) to \( O \) and removing \( j_1 \).
- From claim, \( O' \) is a feasible solution (no conflicts).
- Since \( |O'| = |O| \), \( O' \) is also an optimum solution and it contains \( i_1 \).
Proof of Claim

Claim

\( i_1 \) does not conflict with any interval in \( O \) − \{\( j_1 \)\}.

Proof.

- Suppose \( j_h \in O \) − \{\( j_1 \)\} overlaps with \( i_1 \)
- Since \( i_1 \) has earliest finish time, \( j_h \) and \( i_1 \) overlap at \( f(i_1) \).
  Implies \( j_h \) starts no later than \( f(i_1) \)
- However \( f(j_h) \geq f(j_1) \geq f(i_1) \)
- Implies \( j_h \) and \( j_1 \) overlap at \( f(j_1) \) but \( j_1, j_h \) belong to solution \( O \) and should not overlap. A contradiction!

See figure in next slide.
Figure for proof of Claim

Figure: If $j_h$ conflicts with $i_1$ then they conflict at $f(i_1)$. But $f(i_1) \leq f(j_1) \leq f(j_h)$ implies $j_1$ and $j_h$ conflict.
Proof of Optimality of Earliest Finish Time First

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 $I'$.

\[
|G(I)| = 1 + |G(I')| \quad \text{(from Greedy description)}
\]
\[
\geq 1 + |O'| \quad \text{(By induction, $G(I')$ is optimum for $I'$)}
\]
\[
\geq |O|
\]
Implementation and Running Time

Initially R is the set of all requests
A is empty (* A 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 A
    remove from R all requests that overlap with i
return the set A
Implementation and Running Time

Initially R is the set of all requests
A is empty (* A 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
    if i does not overlap with requests in A
        add i to A
return the set A
Implementation and Running Time

Initially $R$ is the set of all requests
A is empty (* A will store all the jobs that will be scheduled *)
while $R$ is not empty
    choose $i \in R$ such that finishing time of $i$ is least
    if $i$ does not overlap with requests in $A$
        add $i$ to $A$
return the set $A$

- Pre-sort all requests based on finishing time. $O(n \log n)$ time
- Now choosing least finishing time is $O(1)$
Implementation and Running Time

Initially R is the set of all requests
A is empty (* A 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
    if i does not overlap with requests in A
        add i to A
return the set A

- Pre-sort all requests based on finishing time. \( O(n \log n) \) time
- Now choosing least finishing time is \( O(1) \)
- Keep track of the finishing time of the last request added to A. Then check if starting time of i later than that
- Thus, checking non-overlapping is \( O(1) \)
Implementation and Running Time

Initially R is the set of all requests
A is empty (* A 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
    if i does not overlap with requests in A
        add i to A
return the set A

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

- Interesting Exercise: smallest interval first picks at least half the optimum number of intervals.

- Instead of maximizing the total number of requests, associate value/weight with each job that is scheduled. Try to schedule jobs to maximize total value/weight. No greedy algorithm. Will be seen later in this course to illustrate dynamic programming.

- All requests need not be known at the beginning. Such online algorithms are a subject of research.
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.

![Diagram of intervals]

*Figure: A schedule requiring 4 classrooms*
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: A schedule requiring 4 classrooms](image1)

![Figure: A schedule requiring 3 classrooms](image2)
Greedy Algorithm

Initially R is the set of all requests
d = 0 (* number of classrooms *)
while R is not empty
    choose \( i \in R \)
    if \( i \) can be scheduled in some class-room \( k \leq d \)
        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?
Greedy Algorithm

Initially R is the set of all requests
d = 0 (* number of classrooms *)
while R is not empty
    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)
Depth of Lectures

Definition

- 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 $t$. 
Depth of Lectures

**Definition**
- 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 $t$.
- The depth of a set of lectures $R$ is the maximum number of lectures in conflict at any time.
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 $t$.

The depth of a set of lectures $R$ is the maximum number of lectures in conflict at any time.
Lemma

For any set $R$ of lectures, the number of class-rooms required is at least the depth of $R$. 
Depth and Number of Class-rooms

Lemma

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.
Number of Class-rooms used by Greedy Algorithm

Lemma

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

Proof.

Suppose the greedy algorithm uses more that $d$ rooms. Let $j$ be the first lecture that is scheduled in room $d + 1$. 
Number of Class-rooms used by Greedy Algorithm

**Lemma**

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

**Proof.**

- Suppose the greedy algorithm uses more than $d$ rooms. Let $j$ be the first lecture that is scheduled in room $d + 1$.
- Since we process lectures according to start times, there are $d$ lectures that start (at or) before $j$ and which are in conflict with $j$. 

Chekuri CS473ug
Lemma

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

Proof.

- Suppose the greedy algorithm uses more than $d$ rooms. Let $j$ be the first lecture that is scheduled in room $d + 1$.
- Since we process lectures according to start times, there are $d$ lectures that start (at or) before $j$ and which are in conflict with $j$.
- Thus, at the start time of $j$, there are at least $d + 1$ lectures in conflict, which contradicts the fact that the depth is $d$. □
Figure

The problem of interval scheduling involves scheduling a set of intervals on a single processor to minimize the maximum lateness. The goal is to find an optimal assignment of the intervals to the processor so that the lateness of the most delayed interval is minimized.

The algorithm for interval scheduling involves sorting the intervals by their start times, and then scheduling them in increasing order of their start times. If an interval overlaps with an already scheduled interval, it is not scheduled. Otherwise, it is scheduled immediately after the end of the previously scheduled interval.

Correctness

The correctness of the algorithm can be proven by induction. If the algorithm is correct for the first n-1 intervals, and the n-th interval does not overlap with any of the previously scheduled intervals, then the n-th interval can be scheduled immediately after the last scheduled interval. If the n-th interval overlaps with an already scheduled interval, it is not scheduled, and the algorithm is still correct.

Running Time

The running time of the algorithm is O(n log n), where n is the number of intervals. The sorting step takes O(n log n) time, and the scheduling step takes O(n) time.

Figure

The figure illustrates the process of scheduling intervals. The x-axis represents the time, and the y-axis represents the intervals. The solid lines represent the scheduled intervals, and the dashed line represents the interval of the n-th job that is not scheduled because it overlaps with an already scheduled interval.

- j: the interval of the n-th job
- s(j): the start time of the n-th job

The figure shows that there is no such job scheduled before j, and therefore, the n-th job is not scheduled. This is because the n-th job overlaps with the interval of the previously scheduled job.
Correctness

Observation

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

Theorem

The greedy algorithm is correct and uses the optimal number of class-rooms.
Initially R is the set of all requests

\[ d = 0 \text{ (* number of classrooms *)} \]

while R is not empty

\[ \text{choose } i \in R \text{ such that start time of } i \text{ is earliest} \]

if i can be scheduled in some class-room \( k \leq d \)

\[ \text{schedule lecture } i \text{ in class-room } k \]

else

\[ \text{allocate a new class-room } d+1 \text{ and schedule lecture } i \text{ in } d+1 \]

\[ d = d+1 \]
Implementation and Running Time

Initially R is the set of all requests
d = 0 (* number of classrooms *)
while R is not empty
    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
    Pre-sort according to start times. Picking lecture with earliest
    start time can be done in O(1) time.
Implementation and Running Time

Initially $R$ is the set of all requests

d = 0 (* number of classrooms *)

while $R$ is not empty

    choose $i \in R$ such that start time of $i$ is earliest
    if $i$ can be scheduled in some class-room $k \leq d$
        schedule lecture $i$ in class-room $k$
    else
        allocate a new class-room $d+1$ and schedule lecture $i$ in $d+1$
        $d = d+1$

• Pre-sort according to start times. Picking lecture with earliest start time can be done in $O(1)$ time.
Implementation and Running Time

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

- Pre-sort according to start times. Picking lecture with earliest start time can be done in $O(1)$ time.
- Keep track of the finish time of last lecture in each room.
- Checking conflict takes $O(d)$ time.
- Total time $= O(n \log n + nd)$
Implementation and Running Time

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

- Pre-sort according to start times. Picking lecture with earliest start time can be done in \(O(1)\) time.
- Keep track of the finish time of last lecture in each room.
- With priority queues, checking conflict takes \(O(\log d)\) time.
- Total time = \(O(n \log n + n \log d) = O(n \log n)\)
Scheduling to Minimize Lateness

- Given jobs with deadlines and processing times to be scheduled on a single resource.
- 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.
- The lateness of a job is $l_i = \max(0, f_i - d_i)$.
- Schedule all jobs such that $L = \max l_i$ is minimized.
Scheduling to Minimize Lateness

- Given jobs with deadlines and processing times to be scheduled on a single resource.
- 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.
- The lateness of a job is \( l_i = \max(0, f_i - d_i) \).
- Schedule all jobs such that \( L = \max l_i \) is minimized.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_i )</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>( d_i )</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

\[ l_1 = 2 \quad l_5 = 0 \quad l_4 = 6 \]

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A Simpler Feasibility Problem

- Given jobs with deadlines and processing times to be scheduled on a single resource.
- 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.
- Schedule all jobs such that each of them completes before its deadline (in other words $L = \max_i l_i = 0$).

**Definition**

A schedule is **feasible** if all jobs finish before their deadline.
Initially R is the set of all requests
curr-time = 0
while R is not empty
    choose i ∈ R
    curr-time = curr-time + t_i
    if (curr-time > d_i) then
        return ‘‘no feasible schedule’’
    end if
end while
return ‘‘found feasible schedule’’
Greedy Template

Initially $R$ is the set of all requests
curr-time = 0
while $R$ is not empty
    choose $i \in R$
    curr-time = curr-time + $t_i$
    if (curr-time > $d_i$) then
        return ‘‘no feasible schedule’’
    end if
end while
return ‘‘found feasible schedule’’

**Main task:** Decide the order in which to process jobs in $R$
Three Algorithms

- Shortest job first — sort according to $t_i$.
- Shortest slack first — sort according to $d_i - t_i$.
- Earliest deadline first — sort according to $d_i$. 
Three Algorithms

- Shortest job first — sort according to $t_i$.
- Shortest slack first — sort according to $d_i - t_i$.
- Earliest deadline first — sort according to $d_i$.

Counter examples for first two: exercise
Earliest Deadline First

**Theorem**

*Greedy with EDF rule for picking requests correctly decides if there is a feasible schedule.*
Earliest Deadline First

**Theorem**

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

Proof via an exchange argument.
Theorem

*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.
Earliest Deadline First

**Theorem**

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

Proof via an exchange argument.

**Lemma**

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

Definition

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

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

Proof: exercise.
Inversions

**Definition**

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**

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

Proof: exercise.
Main Lemma

Lemma

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

Proof Sketch.

Let $S$ be a schedule with minimum number of inversions.

- If $S$ has 0 inversions, done.
- Suppose $S$ has one or more inversions. By claim there are two adjacent jobs $i$ and $j$ that define an inversion.
- Swap positions of $i$ and $j$.
- New schedule is still feasible. (Why?)
- New schedule has one fewer inversion — contradiction!
Goal: schedule to minimize \( L = \max_i l_i \).
Goal: schedule to minimize $L = \max_i l_i$.

How can we use algorithm for simpler feasibility problem?
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 $\max_i l_i \leq L$?
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 $\max_i l_i \leq L$?

Yes! Set $d'_i = d_i + L$ for each job $i$. Use feasibility algorithm with new deadlines.
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 $\max_i l_i \leq L$?

Yes! Set $d'_i = d_i + L$ for each job $i$. Use feasibility algorithm with new deadlines.

How can we find minimum $L$?
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 \( \max_i l_i \leq L \)?

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!
Binary search for finding minimum lateness

$$L = L_{\text{min}} = 0$$

$$L_{\text{max}} = \sum_i t_i \quad \text{// why is this sufficient?}$$

While $$L_{\text{min}} < L_{\text{max}}$$ do

$$L = \left\lfloor \frac{(L_{\text{max}} + L_{\text{min}})}{2} \right\rfloor$$

check if there is a feasible schedule with lateness $$L$$

if ‘‘yes’’ then $$L_{\text{max}} = L$$

else $$L_{\text{min}} = L + 1$$

endwhile

return $$L$$
Binary search for finding minimum lateness

\[ L = L_{\min} = 0 \]
\[ L_{\max} = \sum_i t_i \quad // \text{why is this sufficient?} \]

While \( L_{\min} < L_{\max} \) do
  \[ L = \lfloor (L_{\max} + L_{\min})/2 \rfloor \]
  check if there is a feasible schedule with lateness \( L \)
  if ‘‘yes’’ then \( L_{\max} = L \)
  else \( L_{\min} = L + 1 \)
endwhile
return \( L \)

**Running time:** \( O(n \log n \cdot \log T) \) where \( T = \sum_i t_i \)

- \( O(n \log n) \) for feasibility test (sort by deadlines)
- \( O(\log T) \) calls to feasibility test in binary search
Do we need binary search?

What happens in each call?
EDF algorithm with deadlines $d'_i = d_i + L$. 
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 $L$!!!
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?
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
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 (see book).
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 + ti
    late = curr-time - di
    curr-late = max (late, curr-late)
return curr-late

Exercise: argue directly that above algorithm is correct (see book).

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

- **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.

- **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.

- **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.

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