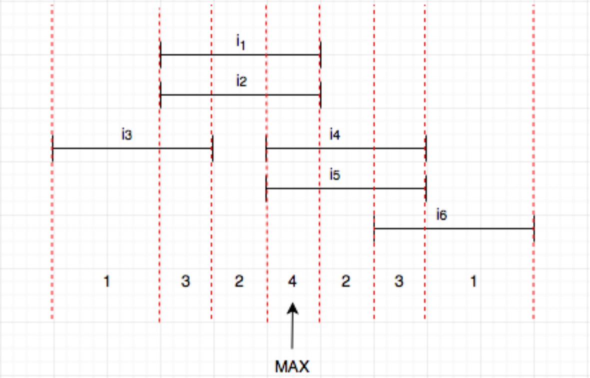
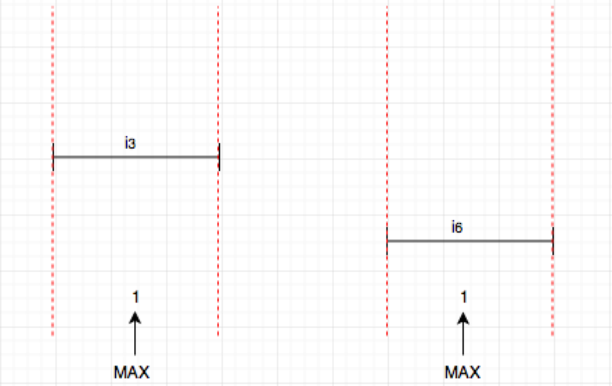
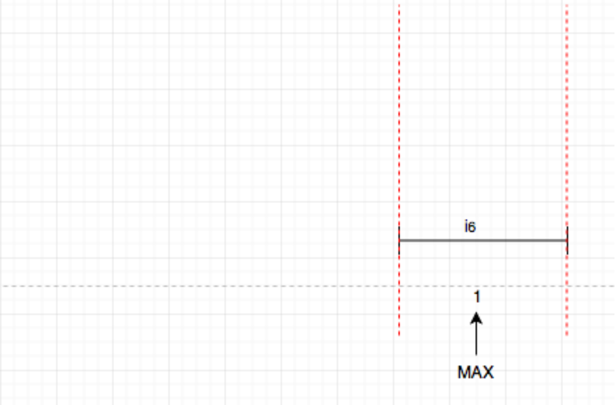
Q1-a
Considering the given intervals {i1, i2, i3, i4, i5, i6} are distributed as follows:
(numbers below indicate the number of intervals in the timeline)



Since the max number of intervals is 4, therefore, intervals i1, i2, i4 and i5 have been covered.

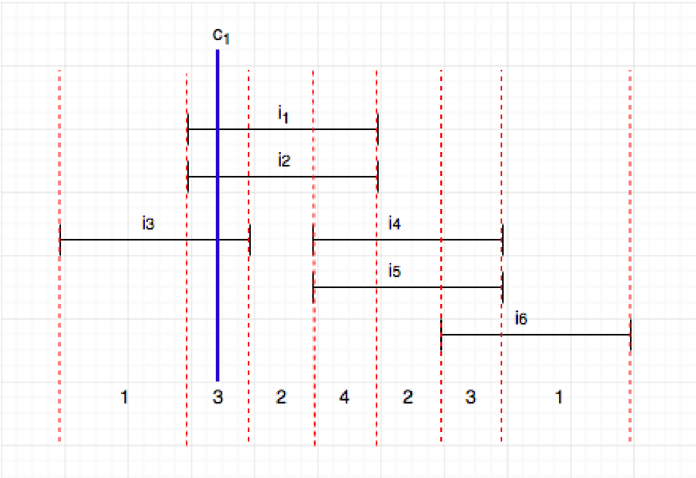


Time i3 and i6 both contain one interval, so by the algorithm, we would pick any one of a number from i3 or i6. Let's choose a number from i3, so interval i3 has been covered.

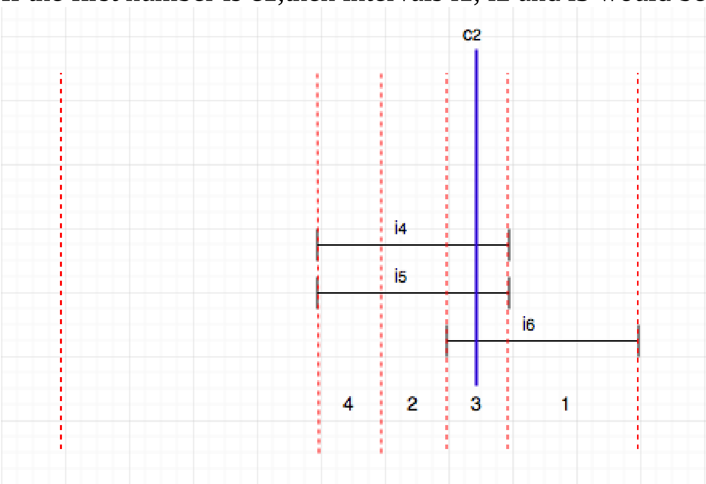


Then the largest number of intervals is 1, it needs a number from i6 to cover i6.
The minimum cover C would contain three numbers by the algorithm.
However, there exists a better choice of C which only contains two numbers.

Better Choice:



If the first number is c1,then intervals i1, i2 and i3 would be covered first.



Then the number c2 would cover all the rest intervals i4, i5 and i6.
The minimal cardinality cover would only contain two numbers c1 and c2, which is less than the size of the minimal cover C with the given algorithm in Q1-a.
Therefore, the algorithm in Q1-a does not always find a minimum cover.

Q1-b

Algorithm:

High Level Pseudocode:

Sort intervals in increasing finish time
A := NULL;
$F := -\infty$;
C := NULL;
for each interval i in sorted order, do
    If i overlaps no interval in A,
    then $A := A \cup \{i\}$;
    $C := C \cup \{f(i)\}$;
    F:= f(i);
return C

Proof for Correctness:

Claim 1: A is the max condinary feasible subset of *I* (proved in lecture)

A is get from the “earliest-finish-time-first” algorithm and we proved this claim in lecture.

Let m be the number of intervals in A
Claim 2: The size of C is at least m.

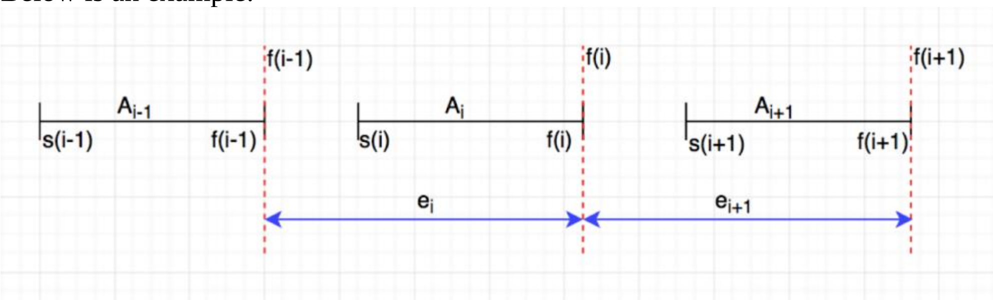
Basic Idea: since the m intervals in A are not overlapped with each other, there is at least one number c picked from each interval.

Proof by contradiction.
Suppose the size of C is k, where k < m.
Let the m intervals in A be the pigeons and k numbers in C be the pigeonholes. If the interval A_i in A contains the number c_x in C, then place the pigeon A_i into the pigeonhole c_x. There are m pigeons and k pigeonholes. By the assumption k < m, we can get there is at least one pigeonhole containing two pigeons, meaning that there is one number in C cover two intervals in A. This leads to a contradiction, since by Claim 1, A is feasible, so there is no two intervals overlapped in A.
Therefore k >= m, so the size of C is at least m.

Let s(i) and f(i) be the start-time and finish time of the ith interval in A(sorted by finished time)

Claim 3: For all intervals starting $\begin{cases} [0, f(i)], & i = 1 \\ (f(i-1), f(i)], & 1 < i \leq m \end{cases}$ contains number f(i).

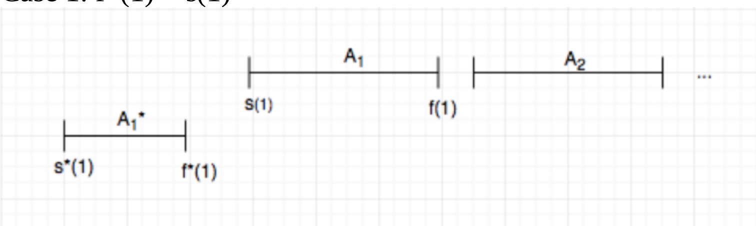
Below is an example.



Suppose A_{i-1}, A_i and A_{i+1} are intervals in the max condinary feasible subset A. Intervals starting with numbers in e_i (i.e in (f(i-1), f(i)]) would contain the number f(i).

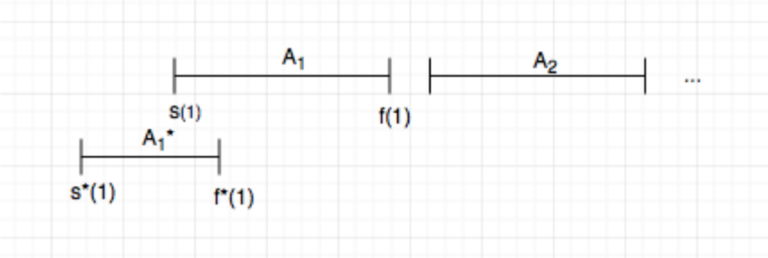
Prove by Induction

Basis: when i = 1,
Suppose there exists an interval A_1* in the given intervals but not in A (i.e in *I* - A) with start time s*(1) and finish time f*(1), where s*(1) <= f(1).
Dividing the proof of the claim into three parts by the different finish time.
Case 1: f*(1) < s(1)



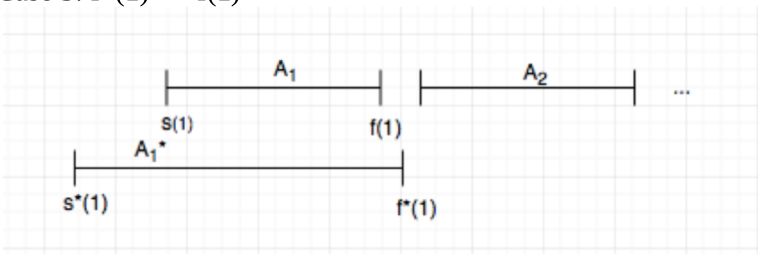
When f*(1) < s(1), A_1* and A_1 are not overlapped and A_1 is the first interval in A. Therefore, A_1* would not overlap with any intervals in A.
By Claim 1, A_1* is supposed to be added to A to make a max condinary feasible subset.
However, A_1* is not contained in A, which is a contradiction.
Therefore, Case 1 would not happen.

Case 2: $s(1) <= f^*(1) < f(1)$



By the algorithm, all the intervals would be sorted by the increasing finish time, A_1^* would be priory to A_1. In the first iteration, A_1^* would be selected first and added to the max condinary feasible subset A by the algorithm, since A_1^* has smaller finish time. However, A_1 is added to the max condinary feasible subset A instead of A_1^*, which is a contradiction.
Therefore, Case 2 would not happen.

Case 3: $f^*(1) >= f(1)$

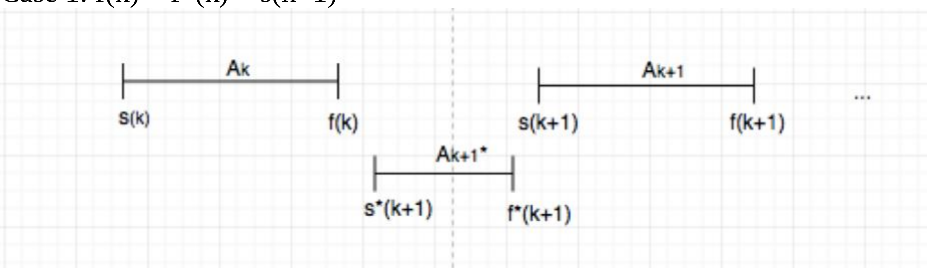


This is a possible situation.
Since A_1^* is an interval that starts before f(1) and ends after f(1), it would contain the number f(1), which follows the Claim 3.

Induction Hypothesis:
For all intervals starting $\begin{cases} [0, f(k)], & k = 1 \\ (f(k-1), f(k)], & 1 < k < m \end{cases}$ contains number f(k).
Want to show for all intervals starting in (f(k), f(k+1)] contains f(k+1).

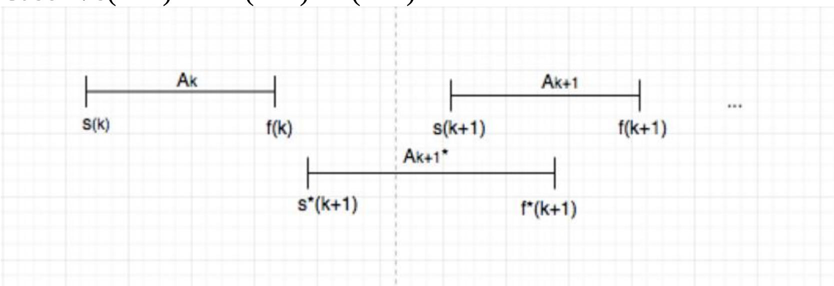
Induction Steps:
Suppose in the (k + 1)th iteration, interval A_{k+1} is added to the set A.
Considering all the intervals start with the number from the interval (f(k), f(k + 1)].
Let's choose an arbitrary interval $A^*(k + 1)$ with start-time $s^*(k + 1)$ and finish-time $f^*(k + 1)$, where $s^*(k + 1) \in (f(k), f(k + 1)]$

Case 1: $f(k) < f^*(k) < s(k+1)$



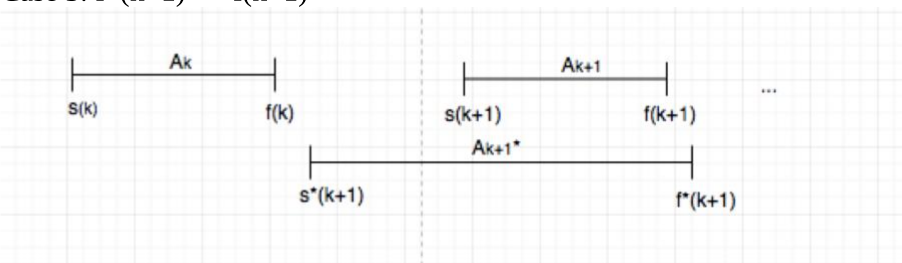
$s^*(k + 1) > f(k)$ => A_{k+1}^* starts after interval A_k
$f^*(k) < s(k+1)$ => A_{k+1}^* finishes before interval A_{k+1}
Since A_{k+1} is added right after A_k, there is no interval in A between A_k and A_{k+1}. Therefore, A_{k+1}^* would not overlap with any intervals in A.
By claim1 A_{k+1}^* should be added to A to make a max condinary feasible subset. However, A_k^* is not contained in the max condinary feasible subset A, which is a contradiction.
Therefore, Case 1 would not happen.

Case 2: s(k+1) <= f*(k+1) < f(k+1)



By the algorithm, all the intervals would be sorted by the increasing finish time, A_{k+1}^* would be priory to A_{k+1}.
By Claim 1, A is feasible, then none of the first k intervals overlapped with each other in A. Since s*(k + 1) > f(k), A_{k+1}^* starts after the kth interval in A. Therefore, A_{k+1}^* does not overlap with the first k intervals in A. In the (k+1)th iteration, A_{k+1}^* would be selected to the max condinary feasible subset A due to a smaller finish time in our assumption. However, $A_{k+1}$is added to the max condinary feasible subset A instead of A_{k+1}^*, which is a contradiction.
Therefore, Case 2 would not happen.

Case 3: f*(k+1) >= f(k+1)

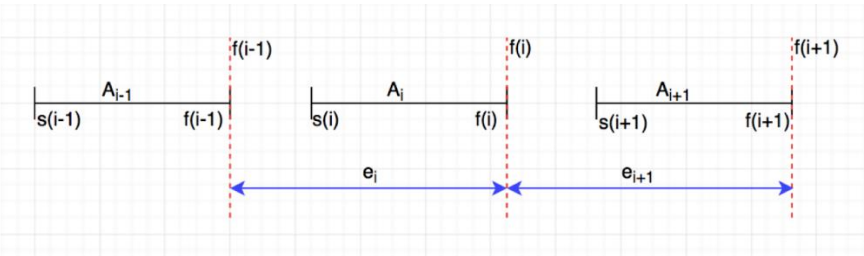


This is a possible situation.
Since A_{k+1}^* is an interval that starts before f(k+1) and ends after f(k+1), it would contain the number f(k+1), which follows the Claim 3.
Claim 3 holds for all 1<= k <= m.
Therefore, Claim 3 is true.

Claim 4: There is no interval that starts after the finish time of the last interval in A.

Proof by contradiction.
If the start time of an interval is larger than the finish time of the last interval in A, the interval should be added to A to make the max condinary feasible set A. However, the interval with start time larger than the finish time of the last interval in A does not appear in A, which is a contradiction. Therefore, there is no interval that starts after the finish time of the last interval in A.

Claim 5: C is the minimal cover.



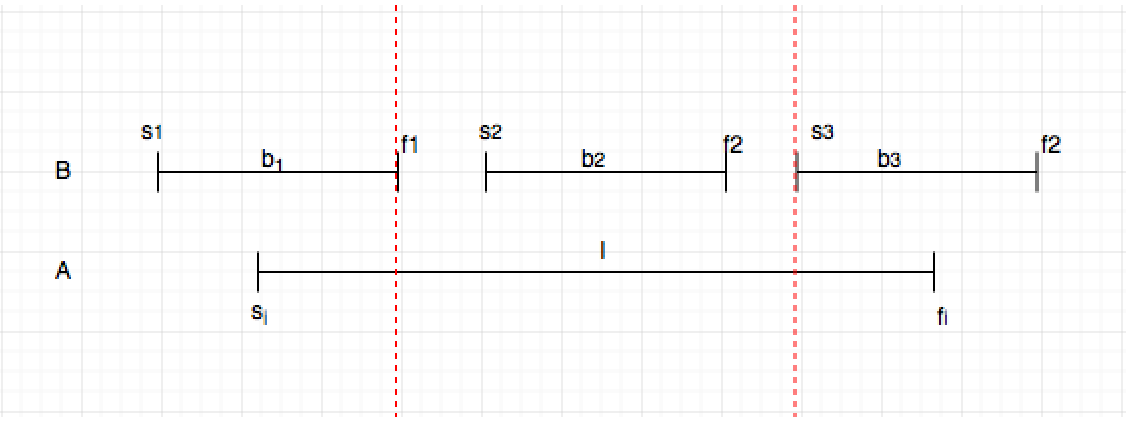
Let $\mathbf{e_i}$ = $\begin{cases} (f(i-1), f(i)] \text{ when } i = 2,3\ldots, m \\ [0, f(i)] \text{ . when } i = 1 \end{cases}$

By Claim 4, all the intervals would start before f(m). Therefore, all intervals in given set *I* should start within one of **e_i** where i = 1,2,...,m
By Claim 3, all the intervals start with numbers from **e_i** would be covered by f(i), where i = 1,2,...,m.
By Claim 2, there are at least m numbers in C. By algorithm, m finish time are added to the set C. Therefore, C is the minimal cover which has the fewest numbers of c and covers all the intervals. Q.E.D.

Q2

Let B be any optimal set

Claim 1: A is feasible
Claim 2: The maximum number of intervals in *B* with which *I* overlaps is 2.



Proof by contradiction:

Suppose I is an arbitrary interval in A and overlaps with 3 intervals in B (b1, b2 and b3).

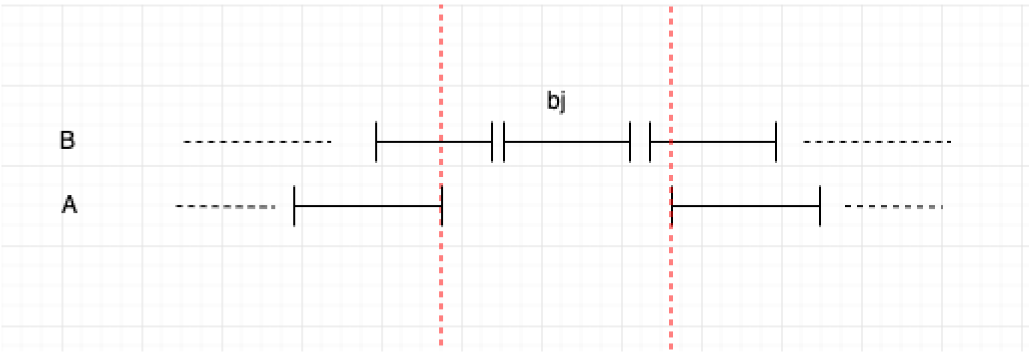
Let n(x) be the number of intervals in the given sequence *I* which x overlaps.
And the start time and finish time are labeled in the graph.

=> s2 >= f1 (B is feasible) >= si (I overlapped with b1) [1]
=> f2 <= s3 (B is feasible) <= fi (I overlapped with b3) [2]
=> n(b2) <= n(I) - 2 (by [1] and [2] and (b2 not overlapped with b1 and b3 but I did))
=> n(b2) < n(I)
=> then we should consider put b2 into A rather than I

Similar for I overlapped more than 3 intervals, there will be more b be considered before I

By contradiction, Claim 2 is True

Claim 3: ∀ interval b in B, b is overlapped by some intervals in A



Proof by contradiction

Suppose after algorithm terminated, there is an interval bj in B such that bj not overlaps any interval in A

=> bj is not removed from *I* since bj is not overlapped with any interval in A
=> *I* is not empty at least bj $\in$ *I*
=> algorithm will not stop

Similar for more than one interval in B that not overlapped with any interval in A

By contradiction, Claim 3 is True

Let |S| be the total number of intervals in set S

if |A| < |B|/2:
Max # of overlapped intervals in B by A = 2|A| < |B| By claim2
Which contradicts with claim 3

=> |A| >= |B|/2
=> |A| >= m/2 since B is an optimal set