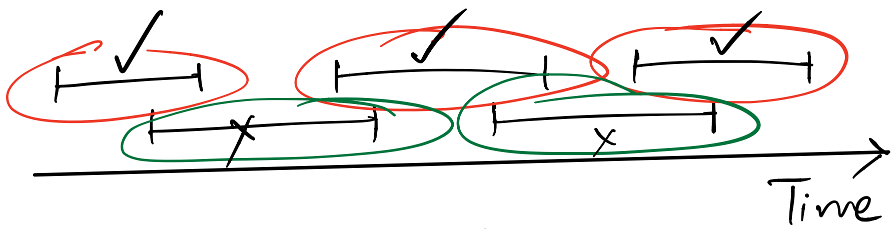
# Greedy Algorithms:-

- No fixed recipe - choice is very much problem dependent - unlike D&C / DP.
- Runtime analysis is typically easy.
- Proof of Correctness - Highly Non-trivial!

**Example I:** Maximum Non-overlapping Set of Intervals

I/P:

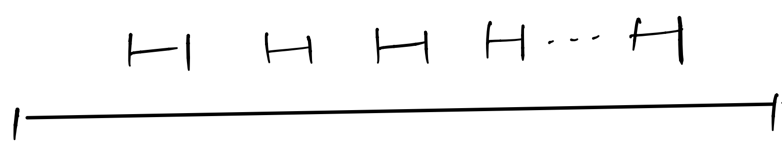



Given $n$ intervals $(s_1, f_1), (s_2, f_2), \ldots (s_n, f_n)$.

O/P: Max cardinality subset of non-overlapping intervals.

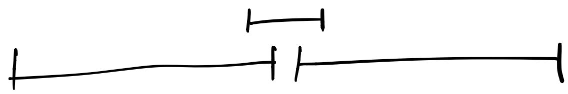
Attemp I:
- Pick the interval with min. start time
- Remove incompatible ones - continue till possible.

Bad Example:



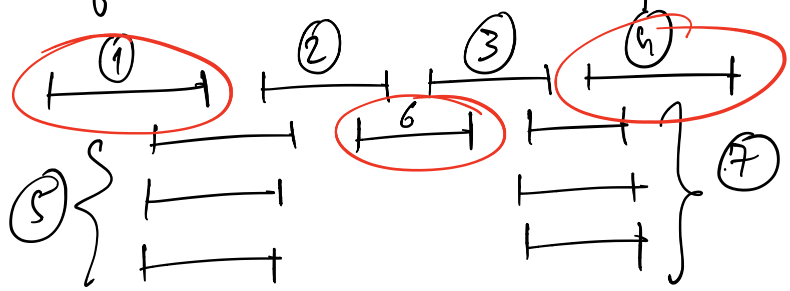
Attempt 2: Pick greedily by non-decreasing order of "length" of intervals.

Bad Example:



Attempt 3: Pick the interval which has minimum number of conflicts in the remaining set.

Bad Example:




Finally, the correct one:

- Sort intervals in increasing order of finish times. - let the set be R.
- Till $R = \emptyset$,
  Pick the next interval in R.
  Remove all conflicting intervals.
- Return R.

Thm: Earliest finish time first is Optimal.

Pf: let $s(j), f(j)$ be the start and end times for interval $j$. Let ALG be the set returned by Greedy and OPT be the optimal set.

We prove $|OPT| = |ALG|$.

Claim (Greedy stays ahead):-

Assume that both sets are sorted by start or end time (They are the same $\because$ the sets are compatible)

Suppose: $OPT = \{ j_1, j_2, j_3, \dots, j_m \}$ ; $m \geq k$

$ALG = \{ j'_1, j'_2, j'_3, \dots, j'_k \}$

Consider the first $r$ jobs in both sets.

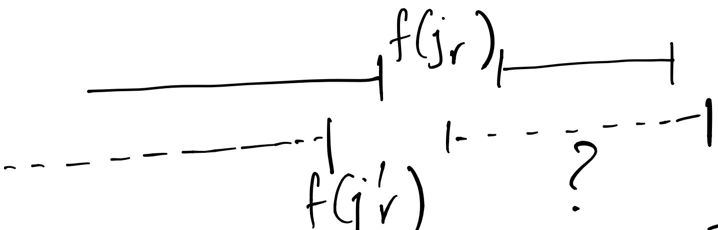
<u>Claim</u>: $f(j'_r) \leq f(j_r)$.

Pf: by induction on $r$.

<u>Base case</u>: <u>$r=1$</u>; $f(j'_1) \leq f(j_1)$ by greedy choice.

Suppose hypotheses true for $r$.

$f(j'_{r+1})$, $f(j_{r+1})$




Suppose $f(j'_{r+1}) > f(j_{r+1})$. But $j_{r+1}$ is compatible with $j'_r$, since $f(j'_r) < f(j_r) \leq s(j_{r+1})$, by induction hypothesis.

$\Rightarrow$ Greedy should have picked $j_{r+1}$ instead. ☒

<u>Proof of Theorem</u>: By contradiction. Suppose $m > k$.

$\Rightarrow$ ~~Our~~ Greedy stopped at $j'_k$

Consider $j_{k+1}$ – the $(k+1)^{th}$ job in OPT.

We claim that $j_{k+1}$ is compatible with ALG.

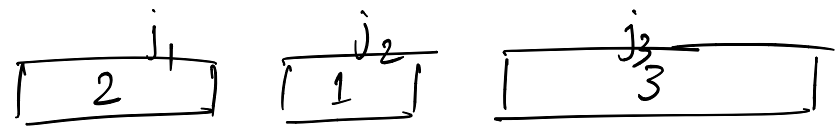
This is because $S(j_{k+1}) > f(j_k) \geq f(j'_k)$

(by the above claim)

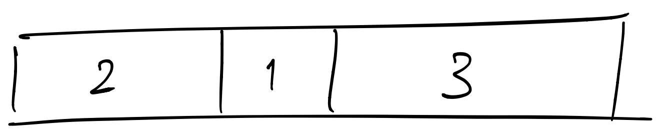
$\Rightarrow$ ALG should have picked $j_{k+1}$ as well.

## Example II: Scheduling to Minimize Completion Time

Jobs:-




Suppose:-




Completion time of job :- $C(j_1) = 2$

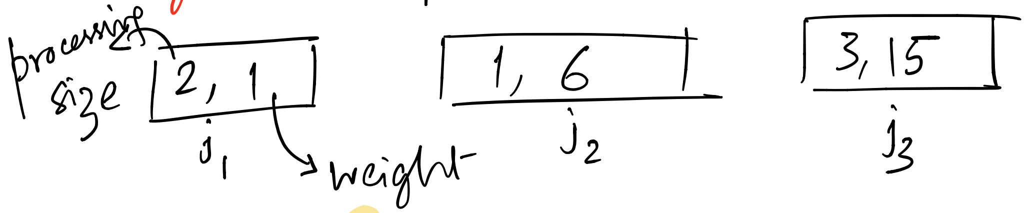
$C(j_2) = 3$

$C(j_3) = 6$

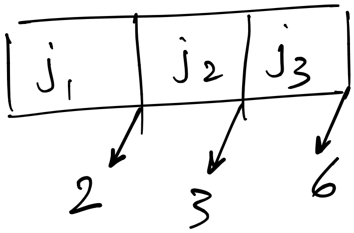
$C(j_1) + C(j_2) + C(j_3) = 11$ for this schedule.

I/P: Set of $n$ jobs $\{j_1, j_2, \ldots, j_n\}$

O/P: Schedule to Minimize Sum of Weighted Completion Times.

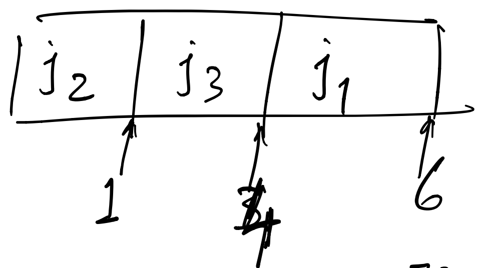


min. $\sum_j W_j \cdot C_j$ [$W_j$ : weight of j] : Weighted completion Time



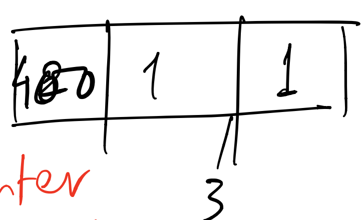
$= 2\times 1 + 3\times 6 + 6\times 15 = 110$

Better:



$= 1\times 6 + 4\times 15 + 1\times 6 = 72$

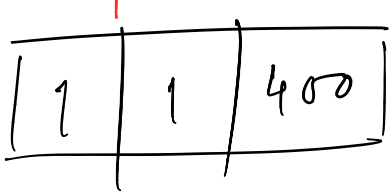
Highest Weight First: (1, 1), (1, 1), (400, 2)



: ~~400×2 + 3×1 + 4×1~~

$400\times 2 + 401\times 1 + 402\times 1$
$= 800 + 401 + 402 = 1603$

Counter Example



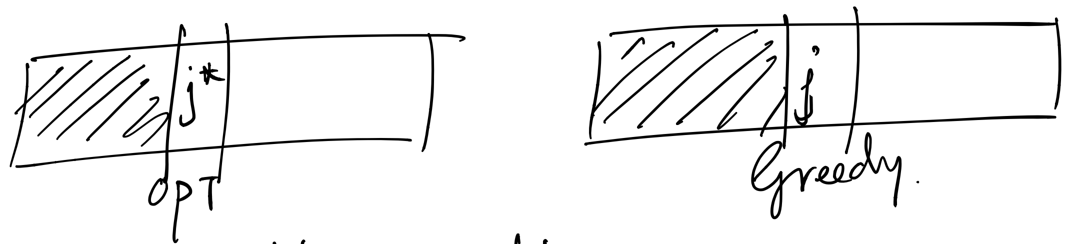
$1\times 1 + 2\times 1 + 402\times 2 = 804 + 3$
$= 807!$

- Correct Algorithm: Sort in non-increasing order of $w_j/p_j$. Schedule according to this.

- Thm: Greedy according to Highest Density First is Optimal.

Pf: (Exchange Argument).

Suppose $\exists$ an optimal schedule.




$$\frac{w_j}{p_j} \geq \frac{w_{j^*}}{p_{j^*}}$$

Claim. $\exists$ a pair of 2 adjacent jobs $j'$ and $j''$ in OPT such that $j'$ is scheduled before $j''$, but

$$\frac{w_{j'}}{p_{j'}} < \frac{w_{j''}}{p_{j''}}$$