

DM865 – Spring 2020  
Heuristics and Approximation Algorithms

## Single Machine Problems

Marco Chiarandini

Department of Mathematics & Computer Science  
University of Southern Denmark

# Outline

Dispatching Rules  
Single Machine Algorithms  
Local Search  
Parallel Machine Models

1. Dispatching Rules
2. Single Machine Algorithms
3. Local Search
4. Parallel Machine Models  
CPM/PERT

# Outline

**Dispatching Rules**  
Single Machine Algorithms  
Local Search  
Parallel Machine Models

1. Dispatching Rules
2. Single Machine Algorithms
3. Local Search
4. Parallel Machine Models  
CPM/PERT

# Dispatching rules

Distinguish **static** and **dynamic** rules.

- Service in random order (SIRO)
- Earliest release date first (ERD=FIFO)
  - tends to min variations in waiting time
- Earliest due date (EDD)
- Minimal slack first (MS)
  - $j^* = \arg \min_j \{\max(d_j - p_j - t, 0)\}$ .
  - tends to min due date objectives (T,L)

- (Weighted) shortest processing time first (WSPT)
  - $j^* = \arg \max_j \{w_j / p_j\}$ .
  - tends to min  $\sum w_j C_j$  and max work in progress
- Longest processing time first (LPT)
  - balance work load over parallel machines
- Shortest setup time first (SST)
  - tends to min  $C_{max}$  and max throughput
- Least flexible job first (LFJ)
  - eligibility constraints

- Critical path (CP)
  - first job in the CP
  - tends to min  $C_{max}$
- Largest number of successors (LNS)
- Shortest queue at the next operation (SQNO)
  - tends to min idleness of machines

# Dispatching Rules in Scheduling

|  | RULE | DATA        | OBJECTIVES                            |
|--|------|-------------|---------------------------------------|
| Rules Dependent<br>on Release Dates<br>and Due Dates | ERD  | $r_j$       | Variance in Throughput Times          |
|  | EDD  | $d_j$       | Maximum Lateness                      |
|  | MS   | $d_j$       | Maximum Lateness                      |
| Rules Dependent<br>on Processing<br>Times            | LPT  | $p_j$       | Load Balancing over Parallel Machines |
|  | SPT  | $p_j$       | Sum of Completion Times, WIP          |
|  | WSPT | $p_j, w_j$  | Weighted Sum of Completion Times, WIP |
|  | CP   | $p_j, prec$ | Makespan                              |
|  | LNS  | $p_j, prec$ | Makespan                              |
| Miscellaneous  | SIRO | -           | Ease of Implementation                |
|  | SST  | $s_{jk}$    | Makespan and Throughput               |
|  | LFJ  | $M_j$       | Makespan and Throughput               |
|  | SQNO | -           | Machine Idleness                      |

When dispatching rules are optimal?

|    | RULE    | DATA        | ENVIRONMENT   |
|----|---------|-------------|---|
| 1  | SIRO    | —           | —   |
| 2  | ERD     | $r_j$       | $1 \mid r_j \mid \text{Var}(\sum(C_j - r_j)/n)$             |
| 3  | EDD     | $d_j$       | $1 \parallel L_{\max}$                                      |
| 4  | MS      | $d_j$       | $1 \parallel L_{\max}$                                      |
| 5  | SPT     | $p_j$       | $Pm \parallel \sum C_j; Fm \mid p_{ij} = p_j \mid \sum C_j$ |
| 6  | WSPT    | $w_j, p_j$  | $Pm \parallel \sum w_j C_j$                                 |
| 7  | LPT     | $p_j$       | $Pm \parallel C_{\max}$                                     |
| 8  | SPT-LPT | $p_j$       | $Fm \mid \text{block}, p_{ij} = p_j \mid C_{\max}$          |
| 9  | CP      | $p_j, prec$ | $Pm \mid prec \mid C_{\max}$                                |
| 10 | LNS     | $p_j, prec$ | $Pm \mid prec \mid C_{\max}$                                |
| 11 | SST     | $s_{jk}$    | $1 \mid s_{jk} \mid C_{\max}$                               |
| 12 | LFJ     | $M_j$       | $Pm \mid M_j \mid C_{\max}$                                 |
| 13 | LAPT    | $p_{ij}$    | $O2 \parallel C_{\max}$                                     |
| 14 | SQ      | —           | $Pm \parallel \sum C_j$                                     |
| 15 | SQNO    | —           | $Jm \parallel \gamma$                                       |



# Composite dispatching rules

Why composite rules?

- Example:  $1 \parallel \sum w_j T_j$ :
  - WSPT, optimal if due dates are zero
  - EDD, optimal if due dates are loose
  - MS, tends to minimize  $T$

► The efficacy of the rules depends on instance factors

## Instance characterization

- Job attributes: {weight, processing time, due date, release date}
- Machine attributes: {speed, num. of jobs waiting, num. of jobs eligible}
- Possible instance factors:

- $1 \mid \mid \sum w_j T_j$

$$\theta_1 = 1 - \frac{\bar{d}}{C_{max}} \quad (\text{due date tightness})$$

$$\theta_2 = \frac{d_{max} - d_{min}}{C_{max}} \quad (\text{due date range})$$

- $1 \mid s_{jk} \mid \sum w_j T_j$

$$(\theta_1, \theta_2 \text{ with estimated } \hat{C}_{max} = \sum_{j=1}^n p_j + n\bar{s})$$

$$\theta_3 = \frac{\bar{s}}{\bar{p}} \quad (\text{set up time severity})$$

- $1 \mid \mid \sum w_j T_j$ , dynamic apparent tardiness cost (ATC)

$$l_j(t) = \frac{w_j}{p_j} \exp \left( -\frac{\max(d_j - p_j - t, 0)}{K \bar{p}} \right)$$

- $1 \mid s_{jk} \mid \sum w_j T_j$ , dynamic apparent tardiness cost with setups (ATCS)

$$l_j(t, l) = \frac{w_j}{p_j} \exp \left( -\frac{\max(d_j - p_j - t, 0)}{K_1 \bar{p}} \right) \exp \left( \frac{-s_{jk}}{K_2 \bar{s}} \right)$$

after job  $l$  has finished.

# Outline

Dispatching Rules  
**Single Machine Algorithms**  
Local Search  
Parallel Machine Models

1. Dispatching Rules
2. Single Machine Algorithms
3. Local Search
4. Parallel Machine Models  
CPM/PERT

# Outlook

$1 \mid \mid \sum w_j C_j$  : weighted shortest processing time first is optimal

$1 \mid \mid \sum_j U_j$  : Moore's algorithm

$1 \mid prec \mid L_{max}$  : Lawler's algorithm, backward dynamic programming in  $O(n^2)$  [Lawler, 1973]

$1 \mid \mid \sum h_j(C_j)$  : dynamic programming in  $O(2^n)$

$1 \mid \mid \sum w_j T_j$  : local search and dynasearch

$1 \mid r_j, (prec) \mid L_{max}$  : branch and bound

$1 \mid s_{jk} \mid C_{max}$  : in the special case, Gilmore and Gomory algorithm  
 optimal in  $O(n^2)$

$1 \mid \mid \sum w_j T_j$  : column generation approaches

# Summary

## Single Machine Models:

- $C_{max}$  is sequence independent
- if  $r_j = 0$  and  $h_j$  is monotone non decreasing in  $C_j$  then optimal schedule is nondelay and has no preemption.

$$1 \mid \mid \sum w_j C_j$$

[Total weighted completion time]

### Theorem

*The weighted shortest processing time first (WSPT) rule is optimal.*

Extensions to  $1 \mid prec \mid \sum w_j C_j$

- in the general case strongly NP-hard
- **chain** precedences:  
process first chain with highest  $\rho$ -factor up to, and included, job with highest  $\rho$ -factor.
- polytime algorithm also for tree and sp-graph precedences

Extensions to  $1 \mid r_j, prmp \mid \sum w_j C_j$

- in the general case strongly NP-hard
- preemptive version of the WSPT if equal weights
- however,  $1 \mid r_j \mid \sum w_j C_j$  is strongly NP-hard



$$1 \parallel \sum_j U_j$$

[Number of tardy jobs]

- [Moore, 1968] algorithm in  $O(n \log n)$ 
  - Add jobs in increasing order of due dates
  - If inclusion of job  $j^*$  results in this job being completed late discard the scheduled job  $k^*$  with the longest processing time
- $1 \parallel \sum_j w_j U_j$  is a knapsack problem hence NP-hard

# Dynamic programming

Procedure based on divide and conquer

**Principle of optimality** the completion of an optimal sequence of decisions must be optimal

- Break down the problem into stages at which the decisions take place
- Find a recurrence relation that takes us backward (forward) from one stage to the previous (next)
- Typical technique: labelling with dominance criteria

(In scheduling, backward procedure feasible only if the makespan is schedule independent, eg, single machine problems without setups, multiple machines problems with identical processing times.)

# 1 | $prec$ | $h_{max}$

- $h_{max} = \max\{h_1(C_1), h_2(C_2), \dots, h_n(C_n)\}$ ,  $h_j$  regular
- special case: 1 |  $prec$  |  $L_{max}$  [maximum lateness]
- solved by backward dynamic programming in  $O(n^2)$

[Lawler, 1978]

$J$  set of jobs already scheduled;

$J^c$  set of jobs still to schedule;

$J' \subseteq J^c$  set of schedulable jobs

Step 1: Set  $J = \emptyset$ ,  $J^c = \{1, \dots, n\}$  and  $J'$  the set of all jobs with no successor

Step 2: Select  $j^*$  such that  $j^* = \arg \min_{j \in J'} \{h_j(\sum_{k \in J^c} p_k)\}$ ;  
 add  $j^*$  to  $J$ ; remove  $j^*$  from  $J^c$ ; update  $J'$ .

Step 3: If  $J^c$  is empty then stop, otherwise go to Step 2.

- For 1 | |  $L_{max}$  Earliest Due Date first
- 1 |  $r_i$  | is instead strongly NP-hard

# Summary

$1 \mid \mid \sum w_j C_j$  : weighted shortest processing time first is optimal

$1 \mid \mid \sum_j U_j$  : Moore's algorithm

$1 \mid prec \mid L_{max}$  : Lawler's algorithm, backward dynamic programming in  $O(n^2)$  [Lawler, 1973]

$1 \mid \mid \sum h_j(C_j)$  : dynamic programming in  $O(2^n)$

$1 \mid r_j, (prec) \mid L_{max}$  : branch and bound

$1 \mid \mid \sum w_j T_j$  : local search and dynasearch

$1 \mid \mid \sum w_j T_j$  : IP formulations, column generation approaches

$1 \mid s_{jk} \mid C_{max}$  : in the special case, Gilmore and Gomory algorithm  
 optimal in  $O(n^2)$

Multicriteria

$$1 \parallel \sum h_j(C_j)$$

A lot of work done on  $1 \parallel \sum w_j T_j$   
 [single-machine total weighted tardiness]

- $1 \parallel \sum T_j$  is hard in ordinary sense, hence admits a pseudo polynomial algorithm (dynamic programming in  $O(n^4 \sum p_j)$ )
- $1 \parallel \sum w_j T_j$  strongly NP-hard (reduction from 3-partition)

# 1 || $\sum h_j(C_j)$

- generalization of  $\sum w_j T_j$  hence strongly NP-hard
- (forward) dynamic programming algorithm  $O(2^n)$

$J$  set of jobs already scheduled;

$$V(J) = \sum_{j \in J} h_j(C_j)$$

Step 1: Set  $J = \emptyset$ ,  $V(j) = h_j(p_j)$ ,  $j = 1, \dots, n$

Step 2:  $V(J) = \min_{j \in J} (V(J - \{j\}) + h_j(\sum_{k \in J} p_k))$

Step 3: If  $J = \{1, 2, \dots, n\}$  then  $V(\{1, 2, \dots, n\})$  is optimum,  
 otherwise go to Step 2.

# Outline

Dispatching Rules  
Single Machine Algorithms  
**Local Search**  
Parallel Machine Models

1. Dispatching Rules
2. Single Machine Algorithms
3. Local Search
4. Parallel Machine Models  
CPM/PERT

$$1 \mid \mid \sum h_j(C_j)$$

### Local search

- search space (solution representation)
- initial solution
- neighborhood function
- evaluation function
- step function
- termination predicate

### Efficient implementations

- A. Incremental updates
- B. Neighborhood pruning



$$1 \mid \mid \sum h_j(C_j)$$

## Neighborhood updates and pruning

- Interchange neigh.: size  $\binom{n}{2}$  and  $O(|i - j|)$  evaluation each
  - first-improvement:  $\pi_j, \pi_k$ 

$p_{\pi_j} \leq p_{\pi_k}$  for improvements,  $w_j T_j + w_k T_k$  must decrease because jobs in  $\pi_j, \dots, \pi_k$  can only increase their tardiness.

$p_{\pi_j} \geq p_{\pi_k}$  possible use of auxiliary data structure to speed up the computation
  - best-improvement:  $\pi_j, \pi_k$ 

$p_{\pi_j} \leq p_{\pi_k}$  for improvements,  $w_j T_j + w_k T_k$  must decrease at least as the best interchange found so far because jobs in  $\pi_j, \dots, \pi_k$  can only increase their tardiness.

$p_{\pi_j} \geq p_{\pi_k}$  possible use of auxiliary data structure to speed up the computation
- Swap: size  $n - 1$  and  $O(1)$  evaluation each
- Insert: size  $(n - 1)^2$  and  $O(|i - j|)$  evaluation each
 

But possible to speed up with systematic examination by means of swaps: an interchange is equivalent to  $|i - j|$  swaps hence overall examination takes  $O(n^2)$

## Dynasearch

- two interchanges  $\delta_{jk}$  and  $\delta_{lm}$  are independent  
if  $\max\{j, k\} < \min\{l, m\}$  or  $\min\{l, k\} > \max\{l, m\}$ ;
- the dynasearch neighborhood is obtained by a series of independent interchanges;
- it has size  $2^{n-1} - 1$  (the number of subsets of  $n - 1$  pairwise jobs);
- but a best move can be found in  $O(n^3)$  searched by dynamic programming;
- it yields in average better results than the interchange neighborhood alone.

**Table 1** Data for the Problem Instance

| Job $j$               | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------------------|---|---|---|---|---|---|
| Processing time $p_j$ | 3 | 1 | 1 | 5 | 1 | 5 |
| Weight $w_j$          | 3 | 5 | 1 | 1 | 4 | 4 |
| Due date $d_j$        | 1 | 5 | 3 | 1 | 3 | 1 |

**Table 2** Swaps Made by Best-Improve Descent

| Iteration | Current Sequence | Total Weighted Tardiness |
|-----------|------------------|--------------------------|
|           | 1 2 3 4 5 6      | 109                      |
| 1         | 1 2 3 5 4 6      | 90                       |
| 2         | 1 2 3 5 6 4      | 75                       |
| 3         | 5 2 3 1 6 4      | 70                       |

**Table 3** Dynasearch Swaps

| Iteration | Current Sequence | Total Weighted Tardiness |
|-----------|------------------|--------------------------|
|           | 1 2 3 4 5 6      | 109                      |
| 1         | 1 3 2 5 4 6      | 89                       |
| 2         | 1 5 2 3 6 4      | 68                       |
| 3         | 5 1 2 3 6 4      | 67                       |

- state  $(k, \pi)$
- $\pi_k$  is the partial sequence at state  $(k, \pi)$  that has  $\min \sum wT$
- $\pi_k$  is obtained from state  $(i, \pi)$

$$\begin{cases} \text{appending job } \pi(k) \text{ after } \pi(i) & i = k - 1 \\ \text{appending job } \pi(k) \text{ and interchanging } \pi(i+1) \text{ and } \pi(k) & 0 \leq i < k - 1 \end{cases}$$

- $F(\pi_0) = 0; \quad F(\pi_1) = w_{\pi(1)} (p_{\pi(1)} - d_{\pi(1)})^+;$

$$F(\pi_k) = \min \begin{cases} F(\pi_{k-1}) + w_{\pi(k)} (C_{\pi(k)} - d_{\pi(k)})^+, \\ \min_{1 \leq i < k-1} \{ F(\pi_i) + w_{\pi(k)} (C_{\pi(i)} + p_{\pi(k)} - d_{\pi(k)})^+ + \\ + \sum_{j=i+2}^{k-1} w_{\pi(j)} (C_{\pi(j)} + p_{\pi(k)} - p_{\pi(i+1)} - d_{\pi(j)})^+ + \\ + w_{\pi(i+1)} (C_{\pi(k)} - d_{\pi(i+1)})^+ \} \end{cases}$$

- The best choice is computed by recursion in  $O(n^3)$  and the optimal series of interchanges for  $F(\pi_n)$  is found by backtrack.
- Local search with dynasearch neighborhood starts from an initial sequence, generated by Apparent Tardiness Cost, and at each iteration applies the best dynasearch move, until no improvement is possible (that is,  $F(\pi_n^t) = F(\pi_n^{(t-1)})$ , for iteration  $t$ ).
- Speedups:
  - pruning with considerations on  $p_{\pi(k)}$  and  $p_{\pi(i+1)}$
  - maintainig a string of late, no late jobs
  - $h_t$  largest index s.t.  $\pi^{(t-1)}(k) = \pi^{(t-2)}(k)$  for  $k = 1, \dots, h_t$  then  $F(\pi_k^{(t-1)}) = F(\pi_k^{(t-2)})$  for  $k = 1, \dots, h_t$  and at iter  $t$  no need to consider  $i < h_t$ .

Dynasearch, refinements:

- [Grosso et al. 2004] add insertion moves to interchanges.
- [Ergun and Orlin 2006] show that dynasearch neighborhood can be searched in  $O(n^2)$ .

### Performance:

- exact solution via branch and bound feasible up to 40 jobs  
[Potts and Wassenhove, Oper. Res., 1985]
- exact solution via time-indexed integer programming formulation used to lower bound in  
branch and bound solves instances of 100 jobs in 4-9 hours [Pan and Shi, Math. Progm., 2007]
- dynasearch: results reported for 100 jobs within a 0.005% gap from optimum in less than 3  
seconds [Grosso et al., Oper. Res. Lett., 2004]

# Summary

$1 \mid \mid \sum w_j C_j$  : weighted shortest processing time first is optimal

$1 \mid \mid \sum_j U_j$  : Moore's algorithm

$1 \mid prec \mid L_{max}$  : Lawler's algorithm, backward dynamic programming in  $O(n^2)$  [Lawler, 1973]

$1 \mid \mid \sum h_j(C_j)$  : dynamic programming in  $O(2^n)$

$1 \mid \mid \sum w_j T_j$  : local search and dynasearch

$1 \mid r_j, (prec) \mid L_{max}$  : branch and bound

$1 \mid \mid \sum w_j T_j$  : column generation approaches



# Outline

Dispatching Rules  
Single Machine Algorithms  
Local Search  
**Parallel Machine Models**

1. Dispatching Rules
2. Single Machine Algorithms
3. Local Search
4. Parallel Machine Models  
CPM/PERT

$Pm \mid \mid C_{max}$   
 (without preemption)

$P_{\infty} \mid prec \mid C_{max}$  CPM

$Pm \mid \mid C_{max}$  List scheduling, approximation ratio:  $2 - \frac{1}{n}$

$Pm \mid \mid C_{max}$  LPT heuristic, approximation ratio:  $\frac{4}{3} - \frac{1}{3m}$

$Rm \mid \mid \sum_j w_j C_j$  unrelated machines, local search with indirect solution representation, SWPT is optimal on  $1 \mid \mid \sum_j w_j C_j$ .

# Project Planning – Critical Path Method

| <b>Milwaukee General Hospital Project</b> |                                  |                              |                        |
|---|----------------------------------|------------------------------|------------------------|
| <b>Activity</b>                           | <b>Description</b>               | <b>Immediate Predecessor</b> | <b><i>Duration</i></b> |
| A   | Build internal components        | –                            | 2                      |
| B   | Modify roof and floor            | –                            | 3                      |
| C   | Construct collection stack       | A                            | 2                      |
| D   | Pour concrete and install frame  | A,B                          | 4                      |
| E   | Build high-temperature burner    | C                            | 4                      |
| F   | Install pollution control system | C                            | 3                      |
| G   | Install air pollution device     | D,E                          | 5                      |
| H   | Inspect and test                 | F,G                          | 2                      |

# Project Planning – Critical Path Method

Whenever a job has been completed, start all jobs whose predecessors have been completed.

Forward procedure

- $EST_j$  earliest starting time
- $EFT_j$  earliest finishing time

Backward procedure

- $LST_j$  latest starting time
- $LFT_j$  latest finishing time

$$EST_j = \max_{k:k \rightarrow j} EFT_k$$

$$LCT_j = \min_{k:j \rightarrow k} LST_k$$

$EST_j < LST_j$  slack job

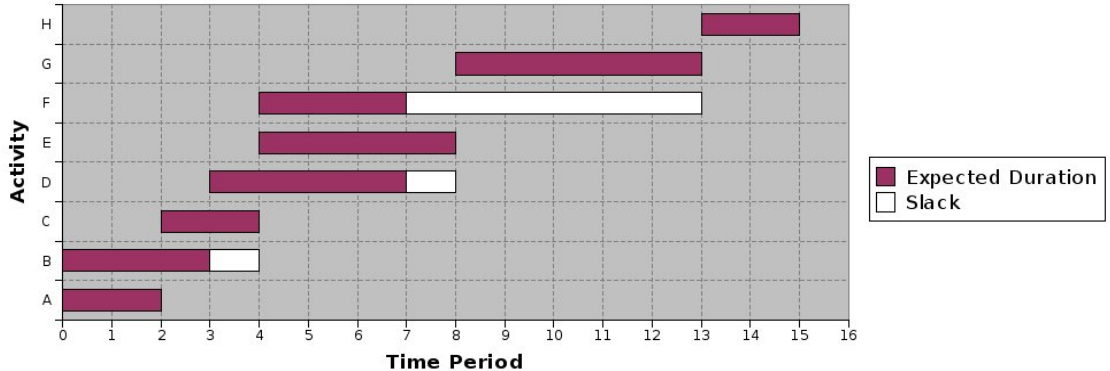
$EST_j = LST_j$  critical job

# Project Planning – Critical Path Method

| <b>Milwaukee General Hospital Project</b> |                                  |                       |          |     |           |     |     |       |
|---|----------------------------------|-----------------------|----------|-----|-----------|-----|-----|-------|
| Activity                                  | Description                      | Immediate Predecessor | Duration | EST | EFT       | LST | LFT | Slack |
| A   | Build internal components        | -                     | 2        | 0   | 2         | 0   | 2   | 0     |
| B   | Modify roof and floor            | -                     | 3        | 0   | 3         | 1   | 4   | 1     |
| C   | Construct collection stack       | A                     | 2        | 2   | 4         | 2   | 4   | 0     |
| D   | Pour concrete and install frame  | A,B                   | 4        | 3   | 7         | 6   | 10  | 3     |
| E   | Build high-temperature burner    | C                     | 4        | 4   | 8         | 6   | 10  | 2     |
| F   | Install pollution control system | C                     | 3        | 4   | 7         | 10  | 13  | 6     |
| G   | Install air pollution device     | D,E                   | 5        | 8   | 13        | 8   | 13  | 0     |
| H   | Inspect and test                 | F,G                   | 2        | 13  | 15        | 13  | 15  | 0     |
| <i>Expected project duration</i>          |                                  |                       |          |     | <b>15</b> |     |     |       |

# Project Planning – Critical Path Method

**Gantt Chart**



# Project Planning – Program Evaluation and Review

| Milwaukee General Hospital Project |                                  |                       | Expected                  |     |     |                              |     |       | Time Estimates |   |    | Activity Variance |
|------------------------------------|----------------------------------|-----------------------|---------------------------|-----|-----|------------------------------|-----|-------|----------------|---|----|-------------------|
| Activity                           | Description                      | Immediate Predecessor | $(a+4m+b)/6$              | EST | EFT | LST                          | LFT | Slack | a              | m | b  | $((b-a)/6)^2$     |
| A                                  | Build internal components        | -                     | 2                         | 0   | 2   | 0                            | 2   | 0     | 1              | 2 | 3  | 0.1111            |
| B                                  | Modify roof and floor            | -                     | 3                         | 0   | 3   | 1                            | 4   | 1     | 2              | 3 | 4  | 0.1111            |
| C                                  | Construct collection stack       | A                     | 2                         | 2   | 4   | 2                            | 4   | 0     | 1              | 2 | 3  | 0.1111            |
| D                                  | Pour concrete and install frame  | A,B                   | 4                         | 3   | 7   | 4                            | 8   | 1     | 2              | 4 | 6  | 0.4444            |
| E                                  | Build high-temperature burner    | C                     | 4                         | 4   | 8   | 4                            | 8   | 0     | 1              | 4 | 7  | 1.0000            |
| F                                  | Install pollution control system | C                     | 3                         | 4   | 7   | 10                           | 13  | 6     | 1              | 2 | 9  | 1.7778            |
| G                                  | Install air pollution device     | D,E                   | 5                         | 8   | 13  | 8                            | 13  | 0     | 3              | 4 | 11 | 1.7778            |
| H                                  | Inspect and test                 | F,G                   | 2                         | 13  | 15  | 13                           | 15  | 0     | 1              | 2 | 3  | 0.1111            |
|                                    |                                  |                       | Expected project duration |     | 15  | Variance of project duration |     |       | 3.1111         |   |    |                   |

# Project Planning – Program Evaluation and Review

- $a_l, a_m, a_u$  parameters for optimistic, most likely and pessimistic times.

$$\mu = \frac{a_l + 4a_m + a_u}{6}$$

$$\sigma = \frac{a_u - a_l}{6}$$

- independent events
- duration project = critical path duration

$$E[D_P] = \sum_i E[X_i]$$

$$\sigma^2[D_P] = \sum_i \sigma^2[X_i]$$

- $D_P$  is Gaussian