



Corso Luigi Einaudi, 55 - Torino

Appunti universitari

Tesi di laurea

Cartoleria e cancelleria

Stampa file e fotocopie

Print on demand

Rilegature

NUMERO: 1428A -

ANNO: 2015

A P P U N T I

STUDENTE: Ruggieri

MATERIA: Programmazione e Gestione della Produzione.
Prof. Chiabert

Il presente lavoro nasce dall'impegno dell'autore ed è distribuito in accordo con il Centro Appunti.

Tutti i diritti sono riservati. È vietata qualsiasi riproduzione, copia totale o parziale, dei contenuti inseriti nel presente volume, ivi inclusa la memorizzazione, rielaborazione, diffusione o distribuzione dei contenuti stessi mediante qualunque supporto magnetico o cartaceo, piattaforma tecnologica o rete telematica, senza previa autorizzazione scritta dell'autore.

ATTENZIONE: QUESTI APPUNTI SONO FATTI DA STUDENTIE NON SONO STATI VISIONATI DAL DOCENTE.
IL NOME DEL PROFESSORE, SERVE SOLO PER IDENTIFICARE IL CORSO.

A01 - L'azienda

Un'**azienda** è un sistema che opera attivamente secondo strategie; quindi ci sono attività che vengono eseguite secondo una strategia (raggiungere determinati obiettivi). Le **strategie** sono linee di azione coordinate per il conseguimento di **obiettivi di gruppo**, di **business unit** e **funzionali**. Gli obiettivi sono collegati all'orizzonte temporale che può essere di **lungo, medio e breve periodo**. Gli obiettivi di gruppo sono quelli di lungo periodo, come ad esempio l'espansione o la diversificazione; gli obiettivi di business unit sono quelli che riguardano i prodotti e i mercati ed infine quelli funzionali sono quelli di breve periodo, che devono essere fatti giorno per giorno per mandare avanti un'azienda e riguardano il marketing o l'innovazione ecc.

L'azienda è un sistema complesso e quindi può essere suddiviso in sottosistemi come il **mercato** che cura la relazione esterna di domanda, immagine e competitività dell'azienda; il **settore tecnologie/risorse** che prevede acquisizione e mantenimento delle risorse aziendali, delle attività produttive e logistiche ed in particolare, di nostro interesse, quello della **produzione**, il quale è un sottosistema importantissimo di tecnologie/risorse; il sottosistema **economico finanziario** che riguarda la rappresentazione dei fenomeni misurabili in termini monetari ed infine il sottosistema **informativo** ovvero la comunicazione e distribuzione dell'informazione.

Un'azienda può essere analizzata in vari modi: in particolare può essere fatta un'analisi funzionale che ci permette di capire come i processi decisionali influiscono sull'azienda stessa oppure operativa per capire esattamente quali sono le risorse oppure gli stati in cui l'azienda deve muoversi per poter arrivare ad una certa soluzione o prospettiva (analizza la struttura fisica). L'analisi dell'azienda è basata sul fatto che i tempi, o meglio, gli orizzonti temporali sono differenti. Quindi gli obiettivi possono essere eseguiti nel breve termine (schedulazione di dettaglio, percorsi alternativi ecc.), medio termine (forza lavoro, piani straordinari, livelli di inventario ecc.) e lungo termine (aumenti di capacità su siti, inserire o rimuovere prodotti, modificare processi ecc. - management di alto livello). Il tutto può essere riassunto nella seguente tabella:

	Orizzonte temporale	Periodo	Frequenza rigenerazione	Livello dettaglio	Esempi
Strategico	Lungo termine	Anni - Decenni	Quadrimestrale - Annuale	Poco accurato	Decisioni finanziarie, Strategie marketing, Progettazione prodotto, Scelta processi tecnologici, Definizione capacità, Localizzazione impianti, Contratti fornitori, Sviluppo personale, Qualità.
Tattico	Medio termine	Settimane - Anni	Settimanale - Mensile	Accurato	Schedulazione lavoro, Assegnazione personale, Manutenzione preventiva, Promozione vendite, Decisioni acquisti
Operativo	Breve termine	Ore - Settimane	Tempo reale	Molto accurato	Controllo flusso materiali, Assegnazione operai, Attrezzaggio macchine, Controllo processo, Controllo qualità, Riparazioni macchinari

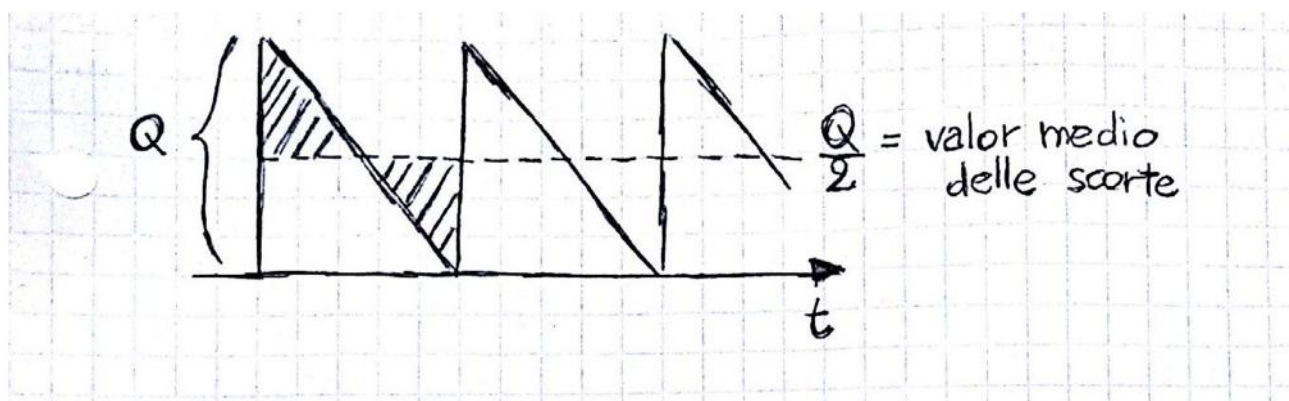
PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE

Prof. Chiabert Paolo

B01 - Gestione delle scorte

La gestione delle scorte riguarda i materiali accumulati in magazzino al fine di garantire una corretta alimentazione dei flussi produttivi. Essa è importante perché ha un'enorme rilevanza economica: essa ha un ruolo nel ciclo economico, ha peso rispetto al PIL nazionale e implica costi logistici. La rilevanza economica non è l'unica, vi è anche una rilevanza aziendale e quindi aumento dell'attivo e degli indici di indebitamento (dovute al pagamento in anticipo delle scorte) e diminuzione dei profitti e degli indici di redditività. Le scorte hanno quindi un impatto diretto nell'azienda. L'indice con cui misuro le scorte e l'efficienza della gestione delle stesse è il **turnover** dato dal rapporto tra fatturato e valore medio delle scorte ($turnover = \frac{fatturato}{valore\ medio\ delle\ scorte}$). Il valore standard che esso deve avere va da 8 a massimo 10.

Per calcolare il valore medio delle scorte bisogna capire come funziona il magazzino. Analizzando graficamente cosa accade, possiamo notare che se l'ordine effettuato è pari a Q , le scorte presenti, che verranno utilizzate per soddisfare la domanda, ogni volta scenderanno fino a toccare lo zero sull'asse delle ascisse. Il valore medio è quindi pari a $Q/2$.



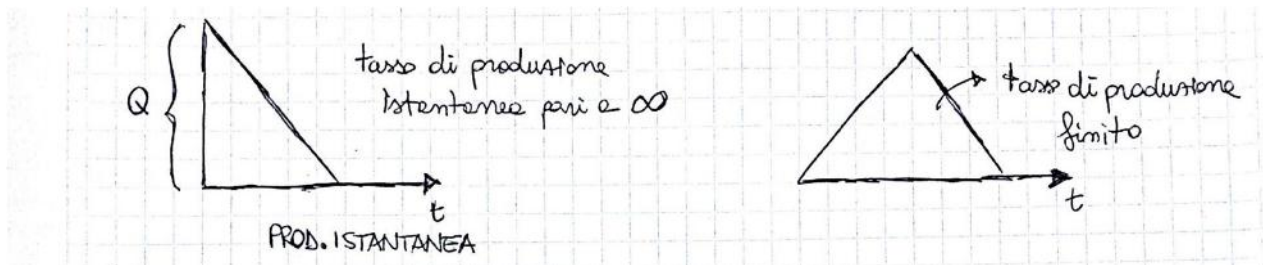
L'**indice di giacenza media** è pari invece a $\frac{1}{turnover}$. Quando vengono gestite le scorte bisogna essenzialmente decidere come osservare le scorte, scegliere quando ordinare, capire la dimensione dell'ordine per ogni prodotto e per ogni componente. A che cosa servono però le scorte in azienda: esiste una **risposta economica** ovvero **costi di transazione** in cui l'azienda cerca di distribuire i costi su un numero maggiore di prodotti; per **prudenza/scorte di sicurezza** dove l'azienda si cautela dal punto di vista delle quantità di prodotti; oppure **speculazione** in cui l'azienda si cautela dal punto di vista del prezzo dei prodotti. Dal punto di vista della **risposta aziendale** abbiamo **scorte di ciclo**, **scorte di congestione** (nel caso in cui qualche macchina dovesse bloccarsi), **scorte di sicurezza**, **scorte di anticipazione** (materiale disponibile in magazzino ben prima di avviare il ciclo), **scorte di flusso**, **scorte di disaccoppiamento** (magazzino comune tra due macchinari); in sostanza esistono tantissimi motivi che mi indicano che devo avere scorte in magazzino. Per quanto riguarda la tipologia delle scorte, abbiamo **raw materials (materie prime)**, **work in process (semilavorati)**, **finished goods (prodotti finiti)**, **spare parts (parti di ricambio)**, per ogni tipologia dobbiamo pensare come osservarli. Abbiamo due metodi, il **continuous review system (Q system)** e il **periodic review system (P system)**: nel primo il livello di un prodotto viene

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE

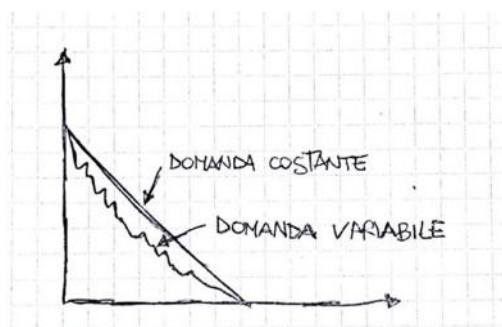
Prof. Chiabert Paolo

Le grandezze da considerare sono il **tempo T**, la **domanda del prodotto D**, la **quantità Q** che vogliamo ordinare del prodotto (se facciamo ordini piccoli avremo meno materiale in magazzino e andare spesso dal fornitore, viceversa avremo magazzini più grandi), il **costo unitario del prodotto** indicato con **c**, il **costo di ordinazione o di setup** indicato con **A**, il **costo unitario h di stoccaggio**; esso è diviso in due parti: la parte di holding cost (spese per poter conservare e contenere il prodotto in azienda) e il **costo finanziario unitario di stoccaggio r (interesse finanziario)**. In genere $h = h + r \cdot c$. Tramite le grandezze così definite possiamo determinare la dimensione ottimale dell'ordine. Per poter calcolare la dimensione ottimale, bisogna però fare delle ipotesi semplificative ovvero:

1. **la produzione è istantanea**: nel momento in cui mi arriva il primo prodotto dell'ordine arriva anche l'ultimo; questo non è sempre vero poiché non ho un tasso di produzione infinito (nella realtà è finito). Quindi non vi sono vincoli di capacità produttiva e il lotto viene prodotto istantaneamente;



2. **la consegna è immediata**: ovvero nel momento in cui ordino mi arriva la merce, noi immaginiamo che non vi siano ritardi tra produzione e disponibilità per soddisfare la domanda;
3. **la domanda è deterministica**: non vi è incertezza sulla quantità o sulla tempistica della domanda, ma la domanda è per definizione stocastica con una sua variabilità;
4. **la domanda è costante nel tempo**: quindi rappresentabile tramite linea retta;



5. **il costo di ordinazione è costante**: non dipende dalla dimensione del lotto o dallo stato del sistema produttivo;
6. **analisi di un singolo prodotto**: consideriamo che non ci sia interferenza tra prodotti differenti;
7. **il prezzo di acquisto è costante**: non dipende dalla quantità di prodotto né dall'istante di tempo in cui ordino;

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE

Prof. Chiabert Paolo

Se $Q_1 = 2Q^*$ allora: $TC_1 = cD + \sqrt{2ADh} + \frac{1}{4}\sqrt{2ADh}$

Se $Q_2 = \frac{1}{2}Q^*$ allora: $TC_2 = cD + \sqrt{2ADh} + \frac{1}{4}\sqrt{2ADh}$

$$\frac{Y(Q)}{Y(Q^*)} = \frac{A\frac{D}{Q} + h\frac{Q}{2}}{A\frac{D}{Q^*} + h\frac{Q^*}{2}} = \frac{1}{2}\left(\frac{Q^*}{Q} + \frac{Q}{Q^*}\right)$$

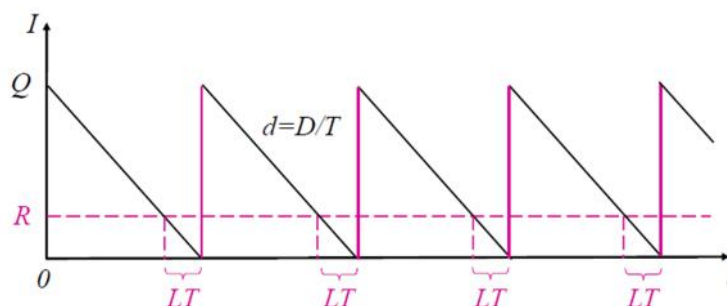
Se $Q_1 = 2Q^*$ allora: $\frac{Y(Q_1)}{Y(Q^*)} = \frac{1}{2}\left(\frac{Q^*}{2Q^*} + \frac{2Q^*}{Q^*}\right) = \frac{5}{4}$

Se $Q_2 = \frac{1}{2}Q^*$ allora: $\frac{Y(Q_2)}{Y(Q^*)} = \frac{1}{2}\left(\frac{Q^*}{Q^*/2} + \frac{Q^*/2}{Q^*}\right) = \frac{5}{4}$

Se noi quindi raddoppiamo o dimezziamo la quantità rispetto a Q^* il nostro costo varierà soltanto del 25%. Quindi una variazione del 100% sulla quantità comporta aumento del 25% dei costi di gestione ottimali.

Le ipotesi che ci hanno portato a definire il modello dell'EOQ possono anche cadere. Anche se ciò avvenisse, il modello continua a funzionare.

Una delle ipotesi fatte ad esempio è che la consegna è immediata. Ciò non è vero perché vi è un certo tempo prima che la merce arrivi, che tecnicamente viene chiamato **lead time**. La soluzione per far funzionare l'EOQ in presenza di lead time è l'emissione dell'ordine quando il magazzino raggiunge un certo livello di riordino. Si tratta di stabilire questo **livello di riordino R** dato da $R = \frac{D}{T}LT$ dove $\frac{D}{T} = d$ e rappresenta il tasso di domanda.



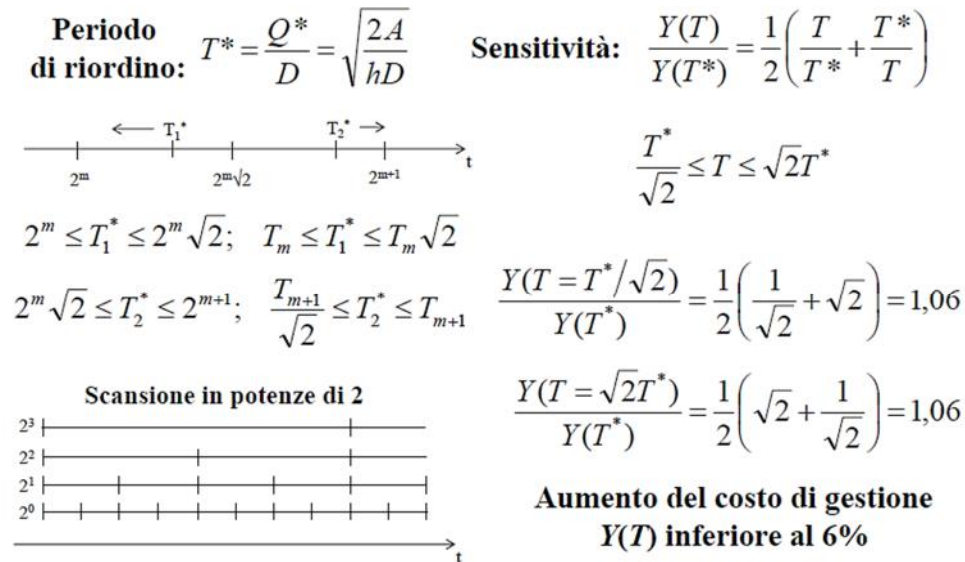
Il funzionamento prevede che appena raggiunto il livello R, l'azienda emette l'ordine e dopo un periodo definito dal lead time avviene la consegna del materiale. Le condizioni da rispettare sono però la presenza di domanda costante e LT costante, altrimenti si genera una rottura di stock. La giacenza è il livello di inventario del materiale fisicamente presente a magazzino; la disponibilità invece è utilizzata per identificare l'attraversamento del livello di riordino ed è dato da giacenza + ordine fornitore - ordine cliente. Definito il periodo di riordino ottimale come $T^* = \frac{Q^*}{D}$, allora se il lead time è maggiore di T^* ($T^* < LT$) allora il

calcolo del livello di riordino viene fatto tramite $R = \frac{D}{T} \cdot LT - \left[\frac{\frac{D}{T} \cdot LT}{EOQ} \right] \cdot EOQ$.

Un'altra delle ipotesi che potrebbe crollare è quella di produzione istantanea. Siamo nel modello dell'**EMQ (Economic Manufacturing Quantity)**. Il costo totale TC sarà dato da $TC = cD + A\frac{D}{Q} + h\frac{I_m}{2}$ con $I_m/2$ il livello medio di magazzino. I_m possiamo ottenerlo

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE
 Prof. Chiabert Paolo

In ultima osservazione, un'altra delle ipotesi che potrebbe crollare è che gli ordini tra differenti prodotti non si influenzano tra loro. Si ha il **multi-product settings**. L'azienda in questo caso adotta una politica di emissione dell'ordine basata sulla scansione in potenze di 2. Il vantaggio è che se si sa che ogni periodo, scandito dalla potenza di 2, il settore acquisti deve emettere l'ordine. Se il periodo di riordino ottimale non è una potenza di 2 (si trova nell'intervallo tra due potenze) approssimo alla potenza più vicina.



L'errore che viene commesso, analizzando la sensitività della funzione $Y(T)$, è pari al 6%, e di conseguenza vi è un aumento del costo di gestione che può essere inferiore al 6%.

▪ **GESTIONE DELLE SCORTE - DOMANDA VARIABILE**

Si è prima supposto che la domanda fosse costante, ovvero la stessa in ogni istante di tempo, ma non esiste una domanda costante (solo se avesse bassa variabilità, allora potrebbe essere vantaggioso considerarla in questo modo). La domanda reale è variabile e ad esempio con un buon sistema di previsione possiamo effettivamente capire come varia nel tempo.

Nel caso di domanda variabile è utile ed interessante capire come l'intervallo di tempo viene suddiviso. L'orizzonte temporale viene suddiviso in **time bucket**, che possono essere mesi, settimane, giorni, in modo da discretizzare il problema per poi essere risolto mediante PL, programmazione dinamica oppure tramite euristiche. Di conseguenza la relazione sotto espressa riguardante il magazzino:

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE
Prof. Chiabert Paolo

soddisfare la domanda non cambia per cui il costo di produzione può essere trascurato (essendo costante).

- **LOT FOR LOT**

L'euristica Lot-For-Lot prevede che in ogni time bucket si ordina una quantità pari alla domanda nello stesso time bucket. Quindi ordino esattamente ciò che mi serve. Questa è quella più semplice ed è utilizzata nel caso di domanda intermittente.

- **FIXED ORDER QUANTITY**

Con questa regola, viene definito un lotto costante denominato proprio FOQ da ordinare in ogni time bucket in cui non è possibile soddisfare la domanda. Devo ordinare quindi sempre la stessa quantità. Come calcolare questa quantità fissa da ordinare? Se non abbiamo ordini dall'esterno, la quantità fissa da ordinare si calcola utilizzando una stima dell'EOQ, mediante la formula:

$$FOQ = \sqrt{\frac{2A\bar{D}}{h}} \quad \text{dove} \quad \bar{D} = \frac{1}{T} \sum_{t=1}^T D_t$$

Usiamo la formula dell'EOQ mettendo in essa la domanda media. In linea di principio la FOQ così calcolata dovrebbe minimizzarci il costo di inventario. Si utilizza quando è vantaggioso avere ordini della stessa dimensione.

- **PERIODIC ORDER QUANTITY**

Si definisce, in questa euristica, un intervallo di riordino (POQ) e si ordina poi la quantità necessaria a soddisfare la domanda in quell'intervallo di tempo. Il POQ è anche detto periodo di copertura. Per stimare il periodo di copertura possiamo utilizzare la seguente relazione:

$$POQ = \frac{EOQ}{\bar{D}} \quad \text{dove} \quad EOQ = \sqrt{\frac{2A\bar{D}}{h}} \quad \text{e} \quad \bar{D} = \frac{1}{T} \sum_{t=1}^T D_t$$

Essa rappresenta il viceversa della regola FOQ. Viene utilizzata quando è vantaggioso avere ordini periodici.

- **SILVER MEAL**

Calcolo il costo necessario a soddisfare la domanda di un numero crescente di time bucket fino a quando i costi non divengono crescenti. L'idea che sta sotto il silver meal è la seguente: noi abbiamo costi di ordinazione, che vengono pagati ogni volta che emetto l'ordine, e poi costi di stoccaggio che crescono quanti più time bucket voglio soddisfare con un solo ordine; si tratta quindi di cercare il minimo dei costi medi di stoccaggio al variare

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE
 Prof. Chiabert Paolo

Orizzonte di pianificazione →

	1	2	3	4	5	...	
Ultimo tb t di produzione ↓	1	A_1	$A_1+h_1D_2$	$A_1+h_1D_2+(h_1+h_2)D_3$	$A_1+h_1D_2+(h_1+h_2)D_3+(h_1+h_2+h_3)D_4$		
	2	$T_{C1}^*+A_2$	$T_{C1}^*+A_2+h_2D_3$	$T_{C1}^*+A_2+h_2D_3+(h_2+h_3)D_4$			
	3		$T_{C2}^*+A_3$	$T_{C2}^*+A_3+h_3D_4$			
	4			$T_{C3}^*+A_4$			
	5						
	...						
	T_C^*	T_{C1}^*	T_{C2}^*	T_{C3}^*	T_{C4}^*		
	t^*	1	t_2^*	t_3^*	t_4^*		

Time bucket di produzione ottimale Costo ottimale

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

bisogni mantenere stabili questi due valori oppure modificarli. Possiamo utilizzare le relazioni $I_t = I_{t-1} + Q_t - D_t$ per l'inventario e $W_t = W_{t-1} + H_t - L_t$ per la forza lavoro (H sono assunzioni e L licenziamenti). Abbiamo poi l'identificazione dei vincoli fisici e politici, l'identificazione dei costi (ordinari, straordinari, assunzioni, licenziamenti, ritardi consegne) e infine la preparazione di un piano ammissibile, implementarlo e poi aggiornarlo in base al periodo di pianificazione. Questo vuol dire risolvere un problema MILP di programmazione lineare oppure usare il metodo del tableau.

Vediamo un esempio:

L'azienda ColoriFelici produce e vende all'ingrosso e al minuto prodotti per verniciatura. La domanda ha un andamento fortemente stagionale e presenta un picco nel terzo trimestre. Il direttore di produzione richiede il miglior piano di produzione per l'anno in corso e fornisce le seguenti informazioni:

Domanda		Magazzino		Costi (unitari in Eur)	
I trim	300000	Iniziale	250000	Produzione ordinaria	1
II trim	850000	Finale	300000	Produzione straordinaria	1,50
III trim	1500000			Produzione esterna	1,90
IV trim	350000			Stoccaggio (per trim)	0,30

Vincoli		Capacità produttiva (migliaia di latte)				
In ogni trimestre		Trimestre				
			I trim	II trim	III trim	IV trim
Straor. $\leq 0,2$ Ordin.		Orario ordinario	450	450	750	450
No backorders		Orario straordinario	90	90	150	90
No stockout.		Produzione esterna	200	200	200	200

Si tratta di un problema MILP che verrà strutturato così:

Parametri

h costo unitario di inventario (€/periodo); I_0, I_T livelli iniziale e finale magazzino
r costo unitario produzione ordinaria (€); R_t capacità produzione ordinaria in t
c costo unitario produzione straordinaria (€); O_t capacità produzione straordinaria in t
s costo unitario produzione esterna (€); S_t capacità produzione esterna in t
 D_t domanda prevista in t

Variabili

Q_{Rt} Produzione ordinaria in t
 Q_{Ot} Produzione straordinaria in t
 Q_{St} Produzione esterna in t

$$\min \sum_{t=1}^T h I_t + r Q_{Rt} + c Q_{Ot} + s Q_{St}$$

$$s.t. \quad I_t = I_{t-1} + Q_{Rt} + Q_{Ot} + Q_{St} - D_t \quad t = 1, \dots, T-1$$

$$I_0 = I(0), \quad I_T = I(T), \quad I_t \geq 0 \quad t = 1, \dots, T$$

$$0 \leq Q_{Rt} \leq R_t \quad t = 1, \dots, T$$

$$0 \leq Q_{Ot} \leq O_t \quad t = 1, \dots, T$$

$$0 \leq Q_{St} \leq S_t \quad t = 1, \dots, T$$

Modello

Quello che si vuole fare è la minimizzazione della somma dei costi totali di inventario, costi totali di produzione ordinaria (costo unitario di prod. ordinaria moltiplicato per la quantità di prod. ordinaria), costi totali di produzione straordinaria (costo unitario di prod. straordinaria moltiplicato per la quantità di prod. straordinaria) e costi totali di produzione esterna (costo unitario di prod. esterna moltiplicato per la quantità di prod. esterna). Nei vincoli del modello abbiamo che la dinamica del magazzino è sempre la stessa ovvero livello precedente più le varie quantità che riesco a mettere dentro durante l'orizzonte

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Il problema sarà di minimizzazione dei soli costi totali di stoccaggio. Non ho nessun costo di produzione perché se il costo di produzione non cambia nel tempo noi sappiamo esattamente sempre quanto ci costa produrre. Anche se dovessi distribuire la produzione nel tempo, il costo di produzione dato da $\sum_{t=1}^T cD_t$ risulta essere sempre costante e anche aggiungerlo nel problema di ottimizzazione non lo influenza. Questo costo qui si dice **costo affondato**. Sarebbe diverso nel caso in cui ci fosse un c_t quindi variabile ad esempio durante l'anno: è ovvio che in quel caso cercherei di produrre il più possibile dove il costo è minore.

$$\begin{aligned} \min & \sum_{t=1}^T \sum_{i=1}^m h_i I_{i,t} \\ \text{s.t.} & I_{i,t} = I_{i,t-1} + Q_{i,t} - D_{i,t}, \quad I_{i,t} \geq 0 \quad \forall i, \forall t \\ & I_{i,0} = \tilde{I}_{i,0}, \quad I_{i,T} = \tilde{I}_{i,T} \\ & \sum_{i=1}^m Q_{i,t} t_{i,j} \leq t_{Maxj} \quad \forall j, \forall t \\ & Q_{i,t} \geq 0 \quad \forall i, \forall t \end{aligned}$$

Ciò che vado a stoccare è sempre soggetta alla legge dinamica del magazzino. Quando noi produciamo nel time bucket t una quantità Q_i dobbiamo utilizzare una certa risorsa $t_{i,j}$ e di conseguenza questa variabile moltiplicata per la quantità mi dice l'impiego della risorsa per produrre nel time bucket t Q unità del prodotto i e questa sommatoria (perché consideriamo tutti i prodotti) deve essere minore di t_{Maxj} ovvero la disponibilità nel time bucket. Gli altri vincoli sono nello standard che le varie variabili devono essere ≥ 0 .

MODELLO DI PIANIFICAZIONE AGGREGATA (APP) 2

Un'azienda deve predisporre un piano di produzione utilizzando n risorse per produrre m prodotti su un orizzonte temporale T . Non sono ammessi *lost sales* e *backorders*, ma si può utilizzare lavoro *straordinario*. Che cosa vuol dire lavoro straordinario? Vuol dire che la disponibilità totale nel time bucket può crescere perché si tiene conto anche dello straordinario che però ha un costo da tenere in considerazione.

Insiemi: $i=1,..,m$ prodotti; $j=1,..,n$ risorse; $t=1,..,T$ time bucket

Parametri:

- $D_{i,t}$ domanda per i in t ;
- h_i costo di stoccaggio di i nel time bucket;
- $t_{i,j}$ utilizzo di j per unità di i ;
- t_{Maxj} disponibilità di j nel time bucket;
- o_j costo di straordinario su j nel time bucket;
- o_{Maxj} disponibilità straordinario su j nel time bucket

Variabili:

- $I_{i,t}$ magazzino di i in t ;
- $Q_{i,t}$ produzione di i in t ;
- $O_{j,t}$ straordinario su j in t ;

Il modello che può essere scritto è il seguente:

$$\begin{aligned} \min & \sum_{t=1}^T \sum_{i=1}^m h_i I_{i,t} + \sum_{t=1}^T \sum_{j=1}^n o_j O_{j,t} \\ \text{s.t.} & I_{i,t} = I_{i,t-1} + Q_{i,t} - D_{i,t}, \quad I_{i,t} \geq 0 \quad \forall i, \forall t; \quad \tilde{I}_{i,0}, \tilde{I}_{i,T} \end{aligned}$$

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Bisogna quindi valutare il problema della perdita del cliente che potrebbe non voler aspettare la produzione.

MODELLO DI PIANIFICAZIONE AGGREGATA (APP) 4

Un'azienda deve predisporre un piano di produzione utilizzando n risorse per produrre m prodotti su un orizzonte temporale T . Le risorse sono *fisse* e sono ammesse *lost sales*.

Per la prima volta ammettiamo di non poter completamente soddisfare la domanda e quindi avere *lost sales*. Attenzione utilizzare solo il prezzo di vendita per massimizzare i profitti genera una visione del problema un po' miope. Infatti questo potrebbe essere un vantaggio per un eventuale concorrente poiché, anche se questo consente di risparmiare, ogni vendita persa rappresenta un danno per l'azienda. Per cui è importante inserire anche un costo di mancata vendita. Non è vero infatti dire che se non si vende non costa nulla ma almeno in termini di immagine danneggia l'azienda. Un modello attento deve tener conto di tutto ciò.

Insiemi: $i=1,..,m$ prodotti; $j=1,..,n$ risorse; $t=1,..,T$ time bucket

Parametri:

- $D_{i,t}$ domanda per i in t ;
- h_i costo di stoccaggio di i nel tb;
- $t_{i,j}$ utilizzo di j per unità di i ;
- t_{Maxj} disponibilità di j nel tb;
- P_i prezzo di vendita di i ;
- p_i costo di mancata vendita di i ;

Variabili:

- $Q_{i,t}$ produzione di i in t ;
- $I_{i,t}$ magazzino di i in t ;
- $V_{i,t}$ vendita di i in t ;

La nostra funzione obiettivo non può essere una minimizzazione dei costi perché in questo caso la minimizzazione potrebbe essere interpretata come il non produrre nulla; quindi non vendo nulla e di conseguenza non ho costi e perdite dovuti alle vendite mancate. Quindi questo è semplicemente un problema di massimizzazione dei profitti.

$$\max \sum_{t=1}^T \sum_{i=1}^m P_i V_{i,t} - \sum_{t=1}^T \sum_{i=1}^m h_i I_{i,t} - \sum_{t=1}^T \sum_{i=1}^m p_i (D_{i,t} - V_{i,t})$$

$$s.t. \quad I_{i,t} = I_{i,t-1} + Q_{i,t} - V_{i,t}, \quad I_{i,t} \geq 0 \quad \forall i, \forall t; \quad \tilde{I}_{i,0}, \tilde{I}_{i,T}$$

$$\sum_{i=1}^m Q_{i,t} t_{i,j} \leq t_{Maxj} \quad \forall j, \forall t$$

$$Q_{i,t} \geq 0 \quad \forall i, \forall t$$

$$0 \leq V_{i,t} \leq D_{i,t} \quad \forall i, \forall t$$

La nostra funzione obiettivo comprende la differenza tra il profitto (prezzo di vendita moltiplicato per le quantità vendute), i costi di stoccaggio e i costi di mancata vendita dati dal costo di mancata vendita moltiplicato la differenza tra la domanda e le vendite. Anche nella dinamica del magazzino variano le cose poiché non bisogna tener conto della domanda ma di ciò che abbiamo tolto per vendere e quindi bisogna considerare le vendite. Poi vi sono i soliti vincoli di produzione massima e di non negatività. Abbiamo in ultimo un altro vincolo ovvero che le vendite devono essere minori della domanda (non si può vendere di più di quanto la domanda chiede nel time bucket).

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Variabili

$Q_{i,t}$ produzione di i in t ;
 $I_{i,t}$ magazzino di i in t ;
 $S_{i,t} = \{1 \text{ lancio } i \text{ in } t, 0 \text{ altrimenti}\}$;

Bisogna decidere quanto produrre, quanto stoccare e se produrre il prodotto nel time bucket. Si suppone di soddisfare completamente la domanda.

$$\min \sum_{t=1}^T \sum_{i=1}^m h_i I_{i,t} + \sum_{t=1}^T \sum_{i=1}^m \sum_{j=1}^n c s_{i,j} S_{i,t}$$

$$s.t. \quad I_{i,t} = I_{i,t-1} + Q_{i,t} - D_{i,t}, \quad I_{i,t} \geq 0 \quad \forall i, \forall t \quad \tilde{I}_{i,0}, \tilde{I}_{i,T}$$

$$\sum_{i=1}^m Q_{i,t} t_{i,j} + \sum_{i=1}^m t s_{i,j} S_{i,t} \leq t_{Maxj} \quad \forall j, \forall t$$

$$0 \leq Q_{i,t} \leq M S_{i,t} \quad \forall j, \forall t$$

$$S_{i,t} \in \{0,1\} \quad \forall i, \forall t$$

La gestione del magazzino è sempre la stessa. Ciò che cambia è la gestione della produzione perché deve comprendere anche il setup. L'unica cosa a cui bisogna fare attenzione è che se non facciamo setup, non si può produrre il prodotto. Nel caso in cui si faccia setup verrà aggiunto il vincolo sulla quantità da produrre dato da $M \cdot S_{i,t}$.

MODELLO DI LOT-SIZING 2

In ogni time bucket è possibile produrre sulla risorsa un solo prodotto i e si vuole minimizzare il numero di setup. La quantità prodotta nel time bucket è la massima possibile e la sequenza di produzione dipende dai costi di set-up necessari a predisporre la risorsa per la lavorazione di quel determinato prodotto.

Qui si produce il massimo possibile della capacità dopo il setup.

Insiemi: $i=1,..,m$ prodotti ; $j=1,..,n$ risorse ; $t=1,..,T$ time bucket;

Parametri:

$D_{i,t}$ domanda per i in t ; h_i costo di stoccaggio di i nel tb ;
 Q_{Maxi} capacità produzione di i nel tb ; cs_j costo di set-up di i ;

Variabili

$I_{i,t}$ magazzino di i in t ; $S_{i,t} = \{1 \text{ setup } i \text{ in } t, 0 \text{ altrimenti}\}$;
 $X_{i,t} = \{1 \text{ produco } i \text{ in } t, 0 \text{ altrimenti}\}$;

$$\min \sum_{t=1}^T \sum_{i=1}^m h_i I_{i,t} + \sum_{t=1}^T \sum_{i=1}^m c s_i S_{i,t}$$

$$s.t. \quad I_{i,t} = I_{i,t-1} + Q_{Maxi} X_{i,t} - D_{i,t}, \quad I_{i,t} \geq 0 \quad \forall i, \forall t$$

$$\sum_{i=1}^m X_{i,t} \leq 1 \quad \forall t$$

$$X_{i,t} - X_{i,t-1} \leq S_{i,t}, \quad S_{i,t} \geq 0 \forall i, \forall t$$

$$X_{i,t} \in \{0,1\} \quad \forall i, \forall t$$

Il vincolo ≤ 1 prevede che venga prodotto un solo prodotto. Inoltre se faccio setup mi devo assicurare che lo faccio per produrre solo un prodotto diverso dal precedente.

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

$$\min \sum_{t=1}^T \sum_{i=1}^m (h_i I_{i,t} + c s_i S_{i,t}) + \sum_{t=1}^T \sum_{j=1}^n (h_j I_{j,t} + c r_j R_{j,t})$$

$$s.t. \quad I_{i,t} = I_{i,t-1} + Q_{i,t} - D_{i,t}, \quad I_{i,t} \geq 0 \quad \forall i, \forall t$$

$$\sum_{i=1}^m Q_{i,t} t_i \leq t_{Max}, \quad 0 \leq Q_{i,t} \leq M S_{i,t} \quad \forall i, \forall t$$

$$I_{j,t} = I_{j,t-1} + A_{j,t} - \sum_{i=1}^m a_{i,j} Q_{i,t}, \quad I_{j,t} \geq 0 \quad \forall j, \forall t$$

$$0 \leq A_{j,t} \leq M R_{j,t} \quad \forall j, \forall t$$

$$S_{i,t} \in \{0,1\} \quad \forall i, \forall t; \quad R_{j,t} \in \{0,1\} \quad \forall j, \forall t$$

MODELLO DI QUALITA'

Un'azienda vuole produrre m prodotti di l qualità utilizzando n risorse in un tempo T . E' ammesso il downgrade dei prodotti.

Insiemi: $i=1,..,m$ prodotti, $j=1,..,n$ risorse, $k=1,..,l$ classi di qualità,
 $t=1,..,T$ time bucket

Dati

$D_{i,k,t}$ domanda di i classe k in t ;
 $t_{i,j}$ utilizzo di j per unità di i ; t_{Maxj} disponibilità di j nel tb
 h_i costo di stoccaggio di i nel tb; c_i costo unitario di produzione di i
 b_{ik} frazione di i appartenente a k

Variabili

$Q_{i,t}$ produzione di i in t
 $I_{i,k,t}$ magazzino di i classe k in t
 $V_{i,k,k',t}$ vendita di i classe k come i classe k' in t

$$\min \sum_{t=1}^T \sum_{i=1}^m \sum_{k=1}^l h_i I_{i,k,t} + \sum_{t=1}^T \sum_{i=1}^m c_i Q_{i,t}$$

$$s.t. \quad I_{i,k,t} = I_{i,k,t-1} + b_{i,k} Q_{i,t} - \sum_{k'=k}^l V_{i,k,k',t} \quad \forall i, \forall k, \forall t$$

$$I_{i,k,t} \geq 0 \quad \forall i, \forall k, \forall t$$

$$\sum_{i=1}^m Q_{i,t} t_{i,j} \leq t_{Maxj}, \quad Q_{i,t} \geq 0 \quad \forall i, \forall j, \forall t$$

$$\sum_{k'=1}^k V_{i,k,k',t} = D_{i,k,t}, \quad V_{i,k,k',t} \geq 0 \quad \forall i, \forall k, \forall k', \forall t$$

Il costo di produzione dato da $\sum_{t=1}^T \sum_{i=1}^m c_i Q_{i,t}$ è da considerare solo se il costo unitario di produzione c_i è variabile nel tempo. Se esso fosse costante, allora sarebbe un costo affondato e la funzione obiettivo si riduce a:

$$\min \sum_{t=1}^T \sum_{i=1}^m \sum_{k=1}^l h_i I_{i,k,t}$$

➤ **VEDERE ESEMPI DI MILP sulle slides delle lezioni**

Programmazione e gestione della Produzione
Prof. Paolo Chiabert

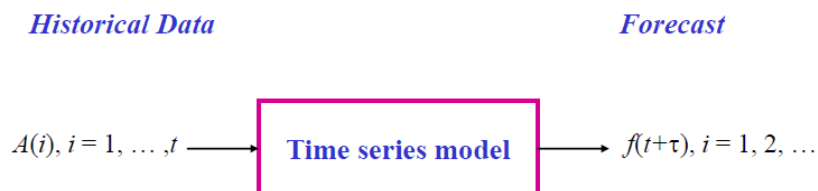
1. **Indagine sulle intenzioni degli acquirenti:** il venditore chiede agli acquirenti il numero di prodotti che essi desiderano acquistare durante il prossimo periodo;
2. **Test di marketing:** metodo preferibile quando l'azienda mette in commercio un nuovo tipo di prodotto sul mercato (fa innovazione);
3. **Sales Force Composite:** gli acquirenti ritengono che i direttori delle aree di vendita conoscano meglio di chiunque altro la domanda dei prodotti;
4. **Metodo esecutivo:** la società costituisce un comitato per le previsioni e questo dovrà riunirsi e discutere le previsioni;
5. **Metodo Delphi:** i membri che fanno parte del comitato sono sia esterni che interni all'azienda. Un moderatore fa sì che avvengano i contatti tra loro. Esso prepara i dati e le sue previsioni gli invia ai vari membri. I membri a loro volta inviano le loro stime che vengono analizzate. Questo accade finché non si è d'accordo sulle stesse previsioni.

➤ **METODI QUANTITATIVI**

Utilizzano i dati che abbiamo dal passato per predire il futuro, cercano di ignorare il rumore ovvero la variabilità dei dati e si basano su delle procedure standard di previsione. Esse si basano su approcci casuali come analisi di regressione oppure sulle serie temporali che saranno quelle di nostro interesse.

SERIE DI TEMPO

Esse sono modelli che si basano su informazioni che riguardano le osservazioni in ingresso in un determinato periodo t denotate con $A(t)$ e ci danno in uscita delle previsioni $f(t + \tau)$ con $t + \tau$ il periodo su cui si vogliono proiettare le previsioni. Quelle che noi andremo a calcolare a partire dai dati $A(t)$ in ingresso è una stima della domanda nel periodo t e questa verrà proiettata nel periodo $t + \tau$; se noi abbiamo un trend la nostra proiezione dovrà tener conto anche del trend $T(t)$ che pensiamo si manifesti in quel periodo.



$A(i)$ = observation in period $i, \quad i = 1, \dots, t$
 t = current period
 $f(t + \tau)$ = forecast for period $t + \tau$
 $F(t)$ = smoothed estimate as of period t
 $T(t)$ = smoothed trend as of period t

Il primo modello prende il nome di **Moving Average**, in cui supponiamo che non ci sia trend perché se il trend è in crescita la stima risulterà sempre in ritardo, se in decrescita allora risulterà in anticipo e poi uguale peso per tutte le m osservazioni fatte nel passato.

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Anche qui vi è il bisogno di condizioni iniziali come $F(1) = A(1)$ e $T(2) = A(2) - A(1)$.

Come ultimo metodo abbiamo il **metodo di Winter per le serie stagionali**. Dapprima abbiamo stimato la nostra domanda $F(t)$, poi il trend $T(t)$ e adesso introduciamo un fattore stagionale. Moltissimi prodotti possono avere una componente di stagionalità di cui bisogna tener conto. La componente di stagionalità può essere in genere moltiplicativa o additiva. L'insieme dei dati in ingresso presenta un andamento ciclico e quindi un qualcosa che si ripete nel tempo; in questo caso il metodo di Winter dà dei buoni risultati. La nostra serie stagionale deve ripetersi per N periodi ma almeno 3. Il fattore di stagionalità rappresenta una media tra le quantità al di sopra o al di sotto della domanda.

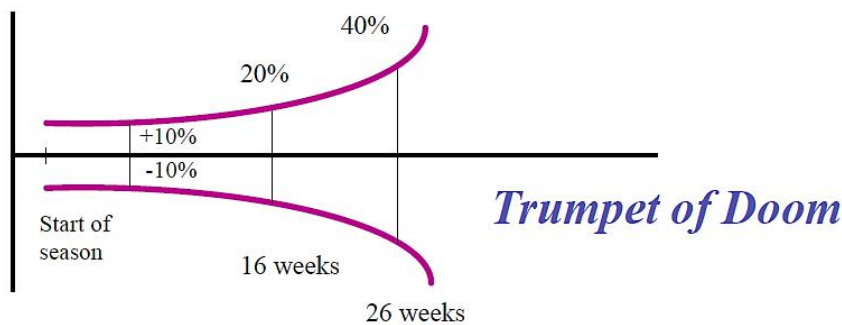
$$\text{The series: } F(t) = \alpha(A(t)/c(t-N)) + (1-\alpha)(F(t-1) + T(t-1))$$

$$\text{The trend: } T(t) = \beta(F(t) - F(t-1)) + (1-\beta)T(t-1)$$

$$\text{The seasonal factors: } c(t) = \gamma(A(t)/F(t)) + (1-\gamma)c(t-N)$$

$$\text{The forecast: } f(t+\tau) = (F(t) + \tau T(t))c(t+\tau-N), \quad t+\tau = N+1, \dots, 2N$$

Bisogna individuare un metodo per sapere se la nostra stima segue bene la domanda oppure no e questo metodo non può essere altro che calcolare l'errore tra il dato rilevato e la stima fatta e che ne tenga conto man mano che il nostro metodo di previsione sta lavorando. In genere quando più il periodo di stima aumenta più le previsioni potrebbero essere sbagliate, come indica la "tromba degli errori".



Possiamo fare delle stime assolute degli errori con alcuni metodi. In genere nelle stime degli errori avremo valori positivi che si compensano con valori negativi. Si dividono in **errori assoluti** e **errori relativi**: I primi sono il **BIAS** che indica se le previsioni non superano ($BIAS < 0$) o superano ($BIAS > 0$) la domanda:

$$BIAS = \sum_{t=1}^n (f(t) - A(t)) / n$$

il **MAD o Mean Absolute Deviation** che indica l'accuratezza delle previsioni in base all'evoluzione della domanda:

$$MAD = \sum_{t=1}^n |f(t) - A(t)| / n$$

Programmazione e gestione della Produzione
Prof. Paolo Chiabert

ADV03 - DOMANDA STOCASTICA

Finora abbiamo supposto che la domanda sia fissa e nota. Che cosa succede quando la domanda è variabile? Come spesso succede, non si ha una conoscenza deterministica della domanda e quindi ho a che fare con la **variabilità** della stessa, la quale avrà un suo valore medio e una sua deviazione standard. Queste due variabili che influenzano la domanda potrebbero portare alla fine ad avere una domanda più alta rispetto agli ordini e quindi ho **stockout** e di conseguenza un mancato guadagno. Può anche verificarsi il caso contrario ovvero domanda più bassa rispetto a quella prevista e quindi ho **overstock** ovvero ho più materiale in magazzino. Questo può essere molto pericoloso quando ho prodotti deperibili e quindi anche con l'overstock bisogna tener conto di eventuali costi. Per tener conto di questi costi di stockout e overstock è possibile fare dei modelli matematici.

NEWSBOY VENDOR MODEL

Il **problema del giornalaio** è un modello matematico usato in caso di domanda stocastica. Il problema prevede che il giornalaio debba vendere tutti i giornali nella giornata perché se vende meno ha un overstock e ovviamente deve restituirli; ogni copia non venduta quindi genera un costo di overstock. Se invece avesse avuto la possibilità di vendere di più, ma non aveva più copie, ha un costo legato a non aver sfruttato a pieno il mercato (stockout). Il problema è determinare la quantità Q che mi dà la massima probabilità di avere meno costi legati a stockout e overstock.

Le variabili del modello sono le seguenti:

- X domanda (in unità); è una variabile casuale;
- $G(X) = P(X \leq x)$ funzione di distribuzione cumulativa della domanda;
- $g(x) = \frac{d}{dx} G(X)$ funzione di densità della domanda;
- μ media della domanda;
- σ deviazione standard della domanda
- c_o costo di overstock (costo di acquisto - reso);
- c_s costo di stockout (prezzo di vendita - costo di acquisto)
- Q quantità da ordinare ovvero la variabile decisionale.

Assumiamo le seguenti ipotesi per sviluppare il modello: i prodotti sono separabili, la pianificazione è fatta per il singolo periodo (l'effetto della decisione attuale sui periodi futuri è trascurabile), la domanda è casuale (caratterizzata da una distribuzione di probabilità nota), le consegne vengono effettuate in anticipo e i costi in eccesso o in difetto sono lineari.

Se si produce una quantità Q e la domanda che abbiamo è pari a X , il numero di unità in eccesso sarà dato da:

$$\text{Units over} = \max\{Q - X, 0\}$$

cioè se $Q \geq X$ allora le unità in eccesso sono $Q-X$ altrimenti 0 per via della carenza di unità. Possiamo calcolare il valore atteso come:

$$E[\text{units over}] = \int_0^{\infty} \max\{Q - x, 0\} g(x) dx = \int_0^Q (Q - x) g(x) dx$$

Analogamente per le unità in difetto:

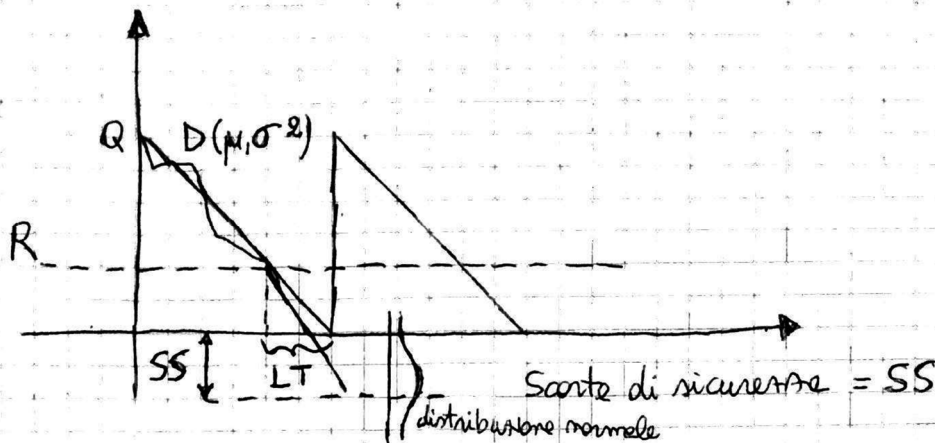
Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

MODELLO (Q,R) (dal libro Factory Physics)

Nel nostro magazzino, quando l'inventario raggiunge un determinato livello di scorte indicato dal punto di riordino R, bisogna effettuare un ordine di Q pezzi per riempire di nuovo il magazzino. Il problema è determinare infatti i valori di Q ed R. Valori alti di Q faranno in modo di avere pochi ordini l'anno e livelli medi di inventario alti; valori bassi di Q porteranno, invece, ad avere un basso livello di inventario medio ma molti ordini all'anno da effettuare. Per quanto riguarda il punto di riordino R, si ha che un punto di riordino alto si tradurrà in un inventario alto e bassa probabilità di stockout, al contrario di un punto di riordino basso, che ridurrà le scorte in magazzino a scapito anche di una maggior probabilità di stockout. Questo perché il punto di riordino R influisce sulla probabilità di avere rotture di stock.

Quindi Q ed R possono interagire in termini di effetti sulle scorte, sulla produzione o sulla frequenza degli ordini e sui servizi ai clienti. Bisogna ricordare che la quantità di riordino Q (definita dall'EOQ) colpisce il ciclo delle scorte in magazzino, mentre il punto di riordino R colpisce le scorte di sicurezza. In sostanza il **modello (Q,R)** fa interagire questi due tipi di parametri.

La domanda è stocastica, quindi caratterizzata da una distribuzione con media e varianza.



1. La scorta di sicurezza non serve;
2. La scorta di sicurezza costa ma copre un rischio;
3. Scorta di sicurezza è dimensionata sulla variabilità della domanda, non sul valore medio

Domande nell'unità di tempo = $d(\mu, \sigma^2)$

$$D_{LT} = \sum_{i=1}^{LT} d(t) \Rightarrow \begin{cases} \mu_{D_{LT}} = d \cdot LT \\ \sigma_{D_{LT}}^2 = \sigma_D^2 \cdot LT \end{cases}$$

NEL CASO IN CUI \rightarrow lead time
 LT È VARIABILE

proprietà
 della media e
 varianza

$$D_{LT} = \sum_{i=1}^{LT} d(t)$$

$$LT \sim N(\mu_{LT}, \sigma_{LT}^2)$$

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

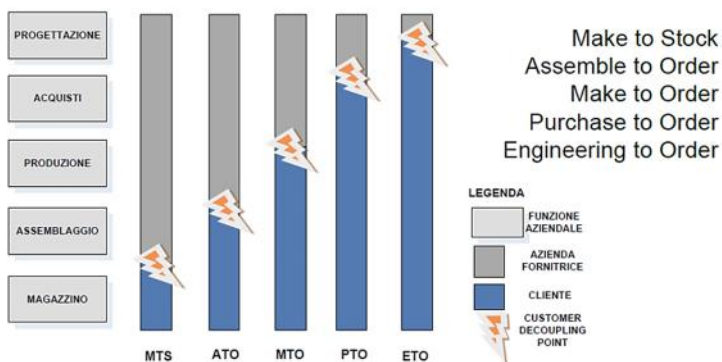
C01/D01 - DOMANDA INDIPENDENTE (MPS) E SISTEMA MRP

DOMANDA INDIPENDENTE MPS

Il **Master Production Schedule** ovvero l'**MPS** rappresenta il piano di produzione di uno stabilimento. Tramite questo meccanismo si sceglie cosa produrre di prodotto finito in ogni time bucket. Una volta fatto questo piano bisogna verificare però che esso sia fattibile. La pianificazione viene verificata tramite **RCCP** ovvero il **Rough-Cut Capacity Planning** che verifica la disponibilità di risorse rispetto al piano di produzione. Vediamo nel dettaglio tutti questi meccanismi.

Bisogna realizzare l'MPS a seconda dei sistemi produttivi che abbiamo.

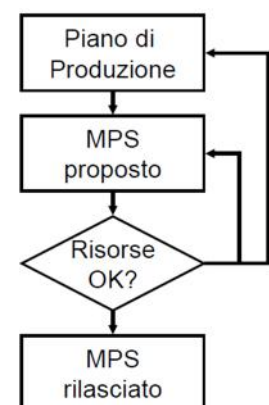
Customer Decoupling Point (Wortmann)



Nello schema possiamo vedere le varie strategie e vari reparti produttivi analizzando i punti di disaccoppiamento del cliente. Questo punto di disaccoppiamento rappresenta la parte del sistema produttivo in cui si manifesta la domanda del cliente. A seconda del punto in cui si manifesta la domanda del cliente abbiamo diversi sistemi produttivi. Parliamo di **Make to Stock** quando la domanda si manifesta a livello magazzino, ovvero produciamo per

immagazzinare (ad esempio in supermercato) e in questo caso vi sono numerosi prodotti finiti. Nel sistema **Assemble to Order** vengono prodotti tutti i sotto-assemblati e poi appena si manifesta la domanda del cliente vengono messi insieme (caso della produzione di automobili). Si scivola nel sistema **Make to Order** quando si vuole una personalizzazione del prodotto finito ovvero la classica produzione per commessa. Nel **Purchase to Order** gli acquisti vengono invece stabiliti dalla domanda del cliente e infine l'**Engineering to Order** dove la progettazione del prodotto è basata sul cliente. A seconda del tipo di strategia la pianificazione viene fatta a livelli diversi.

Il **master production schedule** definisce, a partire da una valutazione dal piano di produzione aggregata, le quantità di prodotti finiti che devono essere prodotti nei vari time bucket. Perché non possiamo fare una copia di ciò che è contenuto nell'APP (Aggregate Production Planning)? Primo perché cambia l'orizzonte temporale; infatti nell'APP l'orizzonte temporale è annuale mentre qui dobbiamo vedere mese per mese o addirittura settimana per settimana la quantità da produrre. Secondo perché bisogna notare che le capacità dello stabilimento sono limitate e quindi non è possibile ad esempio produrre nel singolo mese ciò che è la domanda per l'intero anno dato dall'APP; quindi devo suddividere la produzione nel tempo. L'MPS rappresenta un vero e proprio nuovo piano di produzione. Dopodiché si controlla con l'RCCP se questo piano è fattibile; se fattibile si rilascia l'MPS viceversa si va ad agire a livello delle quantità e se non si riesce anche in questo caso bisogna riprogrammare l'APP che è non fattibile.



Se lavoro con un sistema Make to Stock abbiamo un'azienda focalizzata sul prodotto; abbiamo quindi una bassa varietà di prodotto standardizzati a partire da numerose materie prime. Di conseguenza il nostro MPS

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

ordino la prima quantità MPS fino alla successiva quantità MPS per soddisfare eventuali nuovi ordini da parte dei clienti. Il calcolo è dato da $ATP_1 = I_0 + MPS_t - \sum_{t'=1}^{next\ MPS_t} O_{t'}$ nel primo periodo perché devo tener conto del materiale in magazzino mentre negli altri periodi è pari a $ATP_t = MPS_t - \sum_{t'=t}^{next\ MPS_t} O_{t'}$. Posso accettare ordini se e solo se questi non rendono negativo l'Available to Promise. Si può avere Available to Promise poiché noi pianifichiamo anche in base al massimo tra le previsioni e la domanda e quindi se pianifico in base alle previsioni e se esse sono corrette si può essere sicuri che prima o poi gli ordini arriveranno. Infatti quando gli ordini sono inferiori alle previsioni, vuol dire che ho del materiale in più pronto per soddisfare eventuali ordini aggiuntivi. Questo consente all'azienda di essere reattivi nei confronti della domanda del cliente. Il vantaggio sta nel fatto di poter gestire gli ordini senza intaccare il piano di produzione dell'azienda. Vediamo un esempio:

Calcolo ATP

t	1	2	3	4	5	6	7	8
F(t)	20	20	20	20	40	40	40	40
O(t)	23	15	8	4	0	0	0	0
I(t)	22	2	62	42	2	42	2	42
MPS			80			80		80
ATP	7		68			80		80

$ATP(1) = 45 - (23+15) = 7$ posso ancora accettare ordini per 7 pezzi

$ATP(3) = 80 - (8+4) = 68$ posso ancora accettare ordini per 68 pezzi

$ATP(6) = 80 - 0 = 80$ $ATP(8) = 80 - 0 = 80$

Supponiamo di avere questi ordini:

1) 5 pezzi per la settimana 2 2) 38 pezzi per la settimana 5

3) 24 pezzi per la settimana 3 4) 15 pezzi per la settimana 4

Gli accettiamo oppure no?

Ordine 1: 5 pezzi per la settimana 2 $\Rightarrow O_t(2) = 15+5=20$
 Posso accettare poiché $ATP(1)=7$ $ATP=7-5=2$

Ordine 2: 38 pezzi per la settimana 5 $\Rightarrow O_t(5) = 0+38=38$
 Posso accettare poiché $ATP=68$ $ATP=68-38=30$

Ordine 3: 24 pezzi, settimana 3 $\Rightarrow O_t(3) = 24+8=32$
 Posso accettare poiché $ATP=30$ $ATP=30-24=6$

Ordine 4: 15 pezzi, settimana 4
 Rifiuto perché $ATP=6$ oppure 8 aggiungendo anche il periodo precedente

In linea di principio è possibile posticipare per non rischiare di essere invenduto ma nella realtà bisogna usarlo nelle settimane in cui è stato approvato.

Available to Promise

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

WS	Time	1	2	3	4	5
	End Item					
10	1	400	400	400	800	800
	2	1600	1600	1600	800	800
	3	400	800	1600	800	400
	Totale	2400	2800	3600	2400	2000
20	1	400	400	400	800	800
	2	0	0	0	0	0
	3	800	1600	3200	1600	800
	Totale	1200	2000	3600	2400	1600
30	1	200	200	200	400	400
	2	800	800	800	400	400
	3	600	1200	2400	1200	600
	Totale	1600	2200	3400	2000	1400

I Master production schedule per i prodotti finiti 1, 2 e 3 non sono realizzabili. I problemi sono nella 3^a settimana in cui si supera la disponibilità. Occorrerà anticipare o posticipare la produzione prevista nel periodo 3.

ESERCIZIO COMPLETO:



Esempio MPS

Il prospetto descrive la situazione di un prodotto il cui lotto di produzione è di 50 unità.

Quantity on hand	Week									
	1	2	3	4	5	6	7	8	9	10
5										
Ft (Forecast)	20	10	40	10	0	0	40	20	30	10
Ot (Booked Orders)	30	20	5	8	0	0	0	0	0	0
It (Projected on hand inventory)	25	5	-35	-45	-45	-45	-85	-105	-135	-145
MPSt (MPS Quantity)	50									
APTt (Available to promise inventory)	5									

- Calcolare il Master production schedule
- Determinare le quantità di Available to promise inventory
- Decidere se accettare o rifiutare i seguenti ordini:

Order Quantity Week Due

1 15 6
 2 4 2
 3 32 3

- Valutare la fattibilità del piano. La disponibilità delle risorse Lathe, Drill e Mill è di 40 h/week e il loro impiego per la produzione di un'unità di prodotto (in ore) è riportato in tabella

	Lathe	Drill	Mill
Product	1	0,5	2

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Nuovo MPS ottenuto:

Quantity on hand 5	Week									
	1	2	3	4	5	6	7	8	9	10
Ft (Forecast)	20	10	40	10	0	0	40	20	30	10
Ot (Booked Orders)	30	24	5	8	0	15	0	0	0	0
It (Projected on hand inventory)	25	1	11	1	1	-14	-4	26	-4	-14
MPS _t (MPS Quantity)	50		50				50	50		
APT _t (Available to promise inventory)	1		22				50	50		

• Valutazione dell' RCCP

Product	Lathe	Drill	Mill
	1	0,5	2

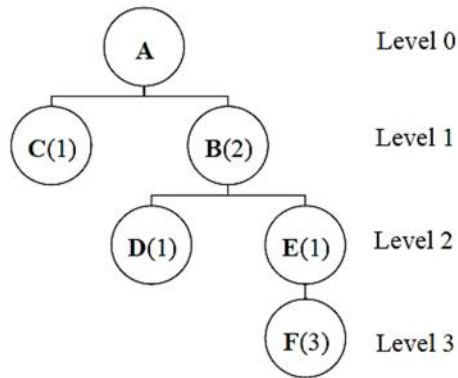
Disponibilità = 40 h/settimana

t	1	2	3	4	5	6	7	8	9	10
MPS	50	0	50	0	0	0	50	50	0	0
Lathe	50*1=50	0	50	0	0	0	50	50	0	0
Drill	50*0,5=25	0	25	0	0	0	25	25	0	0
Mill	50*2=100	0	100	0	0	0	100	100	0	0
Disp.	40	40	40	40	40	40	40	40	40	40

Mill supera la disponibilità \Rightarrow occorre anticipare o posticipare la produzione

MPS non ammissibile

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert



0	Prodotto A	1
Level	Description	Quantity
.1	B	2
...2	D	1
...2	E	1
....3	F	3
.1	C	1

Nella situazione seguente, il prodotto finito, che è indicato con A, è composto da 1 unità di C e 2 unità di B. B è composto a sua volta da 1 unità di D e 1 di E. Infine E è composto da 3 unità di F. Partiamo quindi dal basso per assemblare il prodotto finito. La stessa informazione è contenuta nella tabella. Il **low-level code** indicato è l'indice del livello più basso della BOM in cui compare il componente; sarà quello il livello in cui verrà pianificato solo una volta il materiale corrispondente.

L'altro elemento fondamentale è l'**inventory record**. Per ciascun materiale si ha un inventory record che si prolunga nel tempo secondo l'orizzonte di pianificazione. Esso è formato dai seguenti componenti:

- **Gross requirements GR(t)** che rappresentano i fabbisogni lordi ovvero esattamente la quantità che ci viene richiesta del materiale; essa è derivata dall'MPS nel caso del prodotto finito mentre da tutti i piani di produzione dei livelli superiori (inventory record superiori) nel caso dei componenti;
- **Scheduled receipts** che sono delle consegne attese, già rilasciate ma non ancora completati nel momento in cui si pianifica l'attuale inventory record;
- **Adjusted Scheduled receipts SR(t)** sono gli scheduled receipts cosiddetti "aggiustati" ovvero sono gli SR(t) che vengono anticipati o posticipati al posto di ordinare un nuovo materiale per poter venire incontro alle esigenze di produzione; utilizzo ciò che è già presente nella nostra officina di produzione per soddisfare la domanda;
- **Projected on-hand inventory POH(t)** che rappresenta la giacenza, ovvero la quantità disponibile, in ogni time bucket. Esso è dato da $I(t) = I(t - 1) + SR(t) + PR(t) - GR(t)$.
- **Net requirements N(t)** che sono i fabbisogni netti, ovvero ciò che serve effettivamente per soddisfare la nostra domanda, tenendo conto che arriva il materiale che serve. Esso è pari a

$$\begin{cases} 0 & \text{se } t < t^* \\ -I(t) & \text{se } t = t^*; \\ GR(t) & \text{se } t > t^* \end{cases}$$
- **Planned orders receipts PR(t)** ovvero gli ordini del materiale pianificati per mantenere il livello del POH al di sopra del livello di sicurezza. Si esegue quella che è l'operazione di lot-sizing;
- **Planned orders releases PO(t)** che sono gli ordini rilasciati e traslati temporalmente affinché questi vengano consegnati nella data prevista, a seconda del Lead Time.

Item X	1	2	3	4	5	6	7	8
Gross requirements								
Scheduled receipts								
Projected on-hand inventory								
Net requirements								
Planned receipts								
Planned order releases								



ESEMPIO – MRP (step 2)

Pianificazione livello 2

Item	Gamba	Gamba	t=0	1	2	3	4	5	6
I(0)	0	Gross requirement - $D(t)$		0	0	400	0	0	0
SS	0	Scheduled receipt							
		Adjusted Scheduled receipt - $SR(t)$							
Lot type	POQ	Projected on-hand inventory - $I(t)$	0	0	0	0	0	0	0
Lot value	2	Net requirement - $N(t)$		0	0	400	0	0	0
		Planned order receipt - $PR(t)$				400			
Lead time	2	Planned order release - $PO(t)$	0	400	0	0	0	0	0

Item	Bullone	Bullone	t=0	1	2	3	4	5	6
I(0)	0	Gross requirement - $D(t)$		0	0	400	200	0	0
SS	0	Scheduled receipt							
		Adjusted Scheduled receipt - $SR(t)$							
Lot type	POQ	Projected on-hand inventory - $I(t)$	0	0	0	200	0	0	0
Lot value	3	Net requirement - $N(t)$		0	0	400	0	0	0
		Planned order receipt - $PR(t)$				600			
Lead time	2	Planned order release - $PO(t)$	0	600	0	0	0	0	0

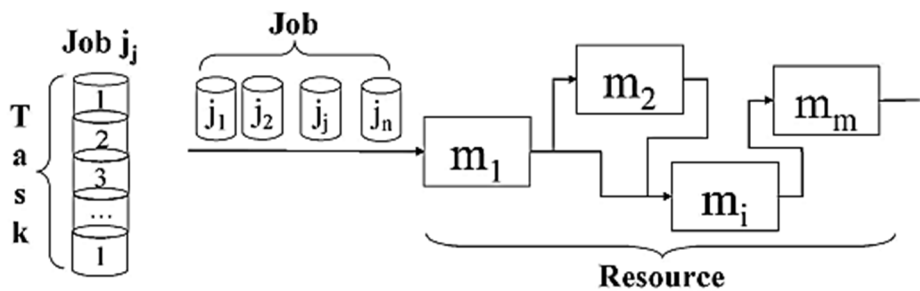
L'MRP è un sistema decisamente complicato da tenere sotto controllo anche se algoritmicamente è banale. Esso deve rispondere a numerosi cambiamenti. Il problema è come e quando aggiornarlo. Bisogna ripianificare perché ci potrebbero essere nuovi ordini nell'MPS, ritardi di completamento degli ordini, problemi di qualità dei prodotti ad esempio un numero di pezzi rotti oppure modifiche sulla distinta base (BOM). Il problema è come tenere conto di queste modifiche. Una soluzione potrebbe essere il **Regenerative MRP** che prevede di ricalcolare ogni volta l'intero MRP a partire dall'MPS ed esplodendo tutti i livelli della BOM; questa soluzione richiede molte risorse di calcolo ma elimina errori precedenti. Una seconda soluzione potrebbe essere il **Net Change MRP** che conserva la pianificazione iniziale e ricalcola l'MRP solo dei prodotti modificati; è ovviamente più veloce ma deve essere rigenerato interamente in modo periodico. Una terza soluzione è il **Top Down Planning** ovvero applica l'MRP tenendo conto dei dati modificati, MPS e scheduled receipts, e ricalcola il piano; può generare piani inammissibili, deve minimizzare i lead time e non riconosce i piani MPS non ammissibili. In ultima soluzione abbiamo un **Bottom Up Replanning** che utilizza un algoritmo di pegging, ovvero esplorare al contrario l'albero della distinta base, e i firm planned order (producono materiali indipendentemente da ordini cliente e lead time) per guidare la ripianificazione.

La complessità dell'MRP non sta quindi nell'algoritmo ma sono i dati che possono dar luogo a delle non linearità che danno luogo al **nervosismo** del sistema; ad esempio avevamo un piano che funzionava perfettamente e a causa di cambiamenti nell'MPS il piano diventa infattibile. In sostanza piccoli cambiamenti nell'MPS possono portare a grossi cambiamenti nella fattibilità del piano.

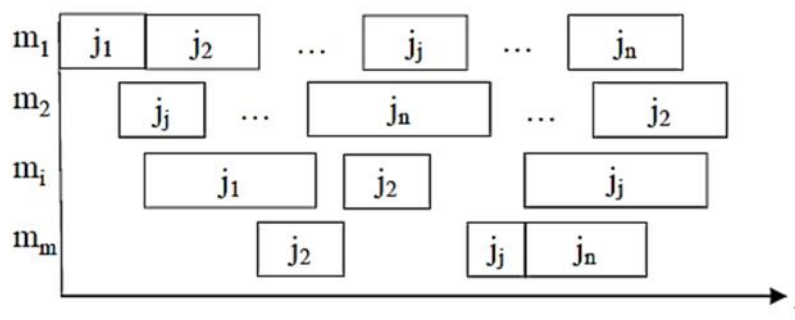
Il **pegging** associa i fabbisogni lordi agli ordini rilasciati al livello superiore della BOM (single pegging) o all'ordine cliente (full pegging). In pratica ci dice quale prodotto o quale ordine non riusciamo a soddisfare. Per far ciò abbiamo bisogno di questo algoritmo in grado di risalire a contrario la distinta base. Per poter risalire, ad ogni ordine dobbiamo capire da dove proviene ai livelli superiori fino ad arrivare alla domanda indipendente che abbiamo pianificato. E' uno strumento per gestire le non conformità o la non fattibilità del nostro piano di produzione.

E01 - SCHEDULAZIONE

Un problema di **schedulazione** è un problema che riguarda l'allocazione di risorse, normalmente scarse, e che devo usare nel tempo per l'esecuzione di compiti. Sono problemi decisionali di allocazione ottimale di risorse scarse, ovvero il tempo, alle attività. Essi inoltre sono problemi di breve e brevissimo periodo. In genere sono sviluppati a livello di singola risorsa o singola macchina, in cui sappiamo disponibilità della macchina, quantità, tempo impiegato ecc. Nel caso della schedulazione si lavora a capacità finita; finora infatti non si era tenuto conto di questo fattore. Per risolvere problemi di schedulazione noi lavoriamo su **job** che sono l'insieme di operazioni da svolgere su una sola risorsa oppure sui **task** che rappresentano invece le singole operazioni di un job fatte sulla risorsa. Ovviamente poi abbiamo a che fare con le **risorse** e quindi le macchine che effettuano le operazioni. La soluzione di questi problemi può avere diversi gradi di dettaglio: la **sequenza** è l'ordine con cui i job devono essere eseguiti su ciascuna macchina; dalla sequenza passiamo poi allo **schedule** ovvero al piano effettivo che comprende anche gli istanti di inizio e fine lavorazione del job; infine possiamo anche avere una **dispatching list** ossia uno schedule in cui è indicata anche la risorsa da utilizzare. Man mano che si scende nel dettaglio aumenta la quantità di informazione. Queste soluzioni possono essere ottenute in modo **statico** nel senso che tutti i job sono disponibili all'inizio del problema e sappiamo esattamente quali sono (schedulazione classica) oppure ottenute in modo **dinamico** i cui i nuovi job possono arrivare nel tempo; qui non sono noti a priori in cui istante per istante dobbiamo vedere quali sono da allocare.



Il risultato della pianificazione è un **diagramma di Gantt**. E' una rappresentazione grafica dell'assegnazione dei job alle risorse nel tempo.



Le ipotesi fatte per risolvere un problema di schedulazione sono le seguenti: la stessa risorsa non può effettuare più di un'operazione per volta, questo naturalmente non è vero perché basta pensare all'operazione di stampaggio che producono più pezzi per volta; ogni job deve essere lavorato da una risorsa per volta, anche questo non è vero perché su un singolo job possiamo avere più macchine che ci lavorano su; dopodiché i vincoli di precedenza dei job devono essere sempre rispettati. Poiché nella realtà i problemi sono complessi, si cerca di renderli semplici per poterli svolgere con questi strumenti. Vi sono infatti ulteriori ipotesi semplificative ovvero tutti i job sono disponibili all'inizio del problema, i tempi di processo sono deterministici e non dipendono dalla sequenza (nessun problema di setup), le risorse sono

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

avremo tutti i vantaggi a posticipare o anticipare la produzione, sono casi particolari però nella realtà. Esempi di misure equivalenti possono essere i seguenti:

$$I_i \Leftrightarrow C_{\max} \quad \text{Perché differiscono di una costante: } I_i = C_{\max} - \sum_j p_{ij}$$

$$CL(C_j) \Leftrightarrow CL(F_j) \Leftrightarrow CL(L_j) \quad \text{Perché differiscono di una costante: } C_j = F_j + r_j = L_j + d_j$$

$$\min(L_{\max}) \Leftrightarrow \min(T_{\max}) \quad \text{Perché con job in ritardo } L_j = T_j, \text{ senza job in ritardo } L_j < T_j = 0$$

Il risultato che effettivamente mi interessa della soluzione al problema di schedulazione riguarda il raggiungimento di determinate funzioni obiettivo ad esempio la **minimizzazione del massimo tempo di completamento o makespan** C_{\max} ($C_{\max} = \max_j C_{\max} \rightarrow \min C_{\max}$); altra misura è la **massima lateness** cioè $L_{\max} = \max_j L_j$ ed è una misura che indica il ritardo del job che è maggiormente in ritardo rispetto alla sua data di consegna; abbiamo ancora la **massima tardiness** data da $T_{\max} = \max(0, L_{\max})$; infine abbiamo il **tempo di completamento totale** che può essere anche pesato dato da $\sum_j C_j$ oppure $\sum_j w_j C_j$ se pesato e se diviso per il numero di job ottengo il tempo medio.

Un problema di schedulazione può essere classificato mediante la notazione di Graham che si basa su tre campi ovvero α , β e γ

$$\alpha / \beta / \gamma$$

α sistema e numero di macchine
 β eventuali caratteristiche particolari
 γ funzione obiettivo

Esempi:

1 / / $\sum_j C_j$ Problema su macchina singola che richiede la minimizzazione del tempo di completamento totale (Risolto con la regola SPT)

F2 / / C_{\max} Problema di flow shop su due macchine che richiede la minimizzazione del makespan (Risolto con l'algoritmo di Johnson)

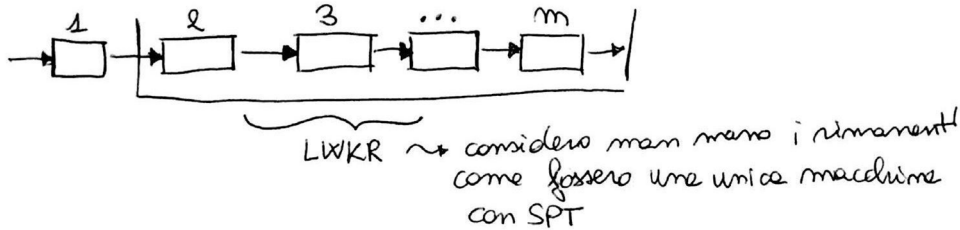
1/prec/ L_{\max} Problema su macchina singola con vincoli di precedenza tra i job che richiede la minimizzazione della massima lateness (Risolto con l'algoritmo di Lawler)

Nel campo β , dove sono indicate le eventuali caratteristiche particolari, possiamo trovare **tempi di setup** s_{jk} che indicano il tempo necessario a riattrezzare la macchina per eseguire il job j dopo aver lavorato il job k , **preemption** ovvero la possibilità di interrompere l'esecuzione del job j per eseguire un job k avente maggiore priorità, **vincoli di precedenza** nell'esecuzione dei job ecc.

La soluzione di un problema di schedulazione può essere determinata mediante metodi **esatti** o **euristici**: sono **esatti** quei metodi **costruttivi**, come le regole di priorità o algoritmi per problemi specifici dove la soluzione viene costruita, o **enumerativi** come il branch & bound; sono **euristici** quei metodi che rappresentano delle regole di buon senso che ci permettono di avere una soluzione ottimale o non necessariamente ottimale, tra i quali troviamo i metodi **iterativi** come le metaeuristiche oppure quelli **one-shot**, che risolvono il problema man mano che si presenta, come le regole di priorità generali.

Quando costruiamo un algoritmo, per fare un gran numero di operazioni, esso impiegherà un certo tempo. A seconda della tipologia di problema che abbiamo, possiamo costruire i diversi algoritmi, alcuni più rapidi altri meno rapidi ma nella realtà esistono due grosse categorie di problemi: quelli che ammettono algoritmi

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert



La **MWRK** o **Most Work Remaining** è la generalizzazione della LPT quando ho più macchine sulla linea che voglio utilizzare con una regola statica. In pratica definisce la priorità di un job come direttamente proporzionale alla somma dei tempi di processo rimanenti. Ordina secondo valori non decrescenti di $\sum_{k=i}^m p_{kj}$. Fornisce una soluzione non ottima al problema $m // C_{max}$.

ALGORITMO DI JOHNSON

Fornisce la soluzione ottima al problema $F_2 // C_{max}$, sotto l'ipotesi che il buffer tra la prima e la seconda macchina abbia dimensione infinita. Il procedimento per eseguire l'algoritmo è il seguente:

Passo 1. Suddividere i job negli insiemi J_1 e J_2

$$J_1 = \{\text{job } j : p_{1j} < p_{2j}\}, \quad J_2 = \{\text{job } j : p_{2j} < p_{1j}\}$$

Passo 2. Ordinare i job negli insiemi J_1 e J_2

Ordinare J_1 con SPT; Ordinare J_2 con LPT

La sequenza finale sarà data dall'unione delle due sotto-sequenze J_1 e J_2 . Questa sequenza finale è quella che minimizza il makespan.

Il procedimento algoritmico invece può essere descritto nel seguente modo:

Passo 1. Porre $k = 1$ e $l = n$ numero di job

Passo 2. Cercare $\min_{i,j} \{p_{ij}\}$

Se $\min_{i,j} \{p_{ij}\} = \min_j \{p_{1j}\}$ mettere j in posizione k e $k = k + 1$

Se $\min_{i,j} \{p_{ij}\} = \min_j \{p_{2j}\}$ mettere j in posizione l e $l = l - 1$

Passo 3. Se la sequenza non è completa, tornare al Passo 2.

ALGORITMO DI CAMPBELL

Per risolvere un problema di tipo $F_m // C_{max}$ l'algoritmo di Campbell fornisce una soluzione non ottima a patto che i buffer tra le macchine abbiano dimensione infinita. La sequenza migliore è data risolvendo $m-1$ problemi di Johnson. Se non ho buffer tra le macchine il job rischia di rimanere nella macchina in cui è stato lavorato; bisogna avere quindi la possibilità di mettere i job nel buffer e quando la macchina successiva è libera, fargli iniziare la lavorazione. Come funziona l'algoritmo?

Passo 1. Per $k = 1, \dots, m-1$ creare le macchine:

$$M1^k = \left\{ \bigcup_{l=1}^k M_l \right\}, \quad P1_j^k = \sum_{l=1}^k p_{jl} \quad M2^k = \left\{ \bigcup_{l=k+1}^m M_l \right\}, \quad P2_j^k = \sum_{l=k+1}^m p_{jl}$$

Ad esempio nel caso di 3 macchine dette M_1 , M_2 ed M_3 le macchine che potremo creare le macchine M_1 e M_2+M_3 oppure M_1+M_2 e M_3 (non è possibile invertire l'ordine della lavorazione).

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Passo 3. Porre il job j^* in posizione N nella sequenza finale e aggiornare :

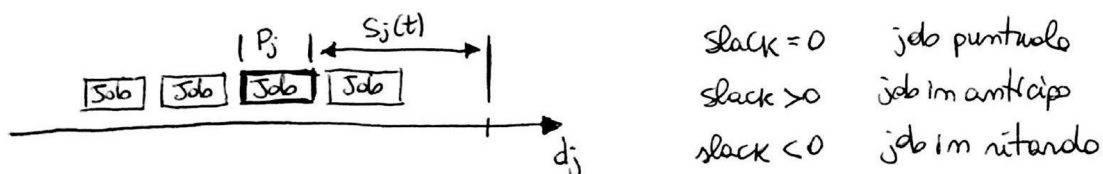
$$N = N - \{j^*\}$$

$$P(N) = P(N) - p_{j^*}$$

Passo 4. Se $N = \emptyset$ stop, altrimenti torna al Passo 2.

REGOLE DI PRIORITA' DINAMICHE

Finora ci siamo occupati di regole statiche, ma potrebbe capitare di non sapere i job in arrivo e quando questi arrivano. Inoltre non sempre è possibile trovare un algoritmo che ci fornisce una soluzione mediamente buona. Per poter trovare la priorità dei job nel caso di regole dinamiche ed avere la giusta sequenza, bisogna in qualche modo misurare questa urgenza del job e lo facciamo attraverso lo slack. La decisione di anticipare di Δ unità di tempo la lavorazione di un job determina dei vantaggi e comporta dei costi. Lo **slack** non è altro che il tempo che ci rimane per poter consegnare il job ed è dato da $S_j(t) = d_j - p_j - t$. Sulla base dello slack possiamo stabilire un **vantaggio** oppure un **costo**. Data la posizione di un job, lo spostamento del job comporta un vantaggio e un costo di spostamento. Il vantaggio è dato da $Vantaggio = w_j \cdot u(S_j(t)) \cdot \Delta$ dove u è una funzione di urgenza e Δ l'anticipo di tempo, mentre il costo è dato da $Costo = (R \cdot p_j) \cdot I \cdot \Delta$ dove R è il costo della risorsa per unità di tempo; abbiamo il costo perché il fatto di aver anticipato il job vuol dire che stiamo utilizzando una risorsa che non era stata schedulata e quindi bisogna pagarla. La **priorità** del job è data dalla differenza tra il vantaggio e il costo diviso il tempo di processo del job ed è pari a $Priorità \Pi_j = \frac{w_j}{p_j} u(S_j(t))$.



A seconda della funzione di urgenza scelta si ottengono diverse regole di priorità:

$$WSPT : u(S_j(t)) = 1$$

$$TARDY JOB : u(S_j(t)) = U(S_j(t))$$

$$LEAST SLACK TIME : u(S_j(t)) = S_j(t)$$

$$COVERT : u(S_j(t)) = \delta \left(1 - \frac{\delta(S_j(t))}{\bar{p}} \right)$$

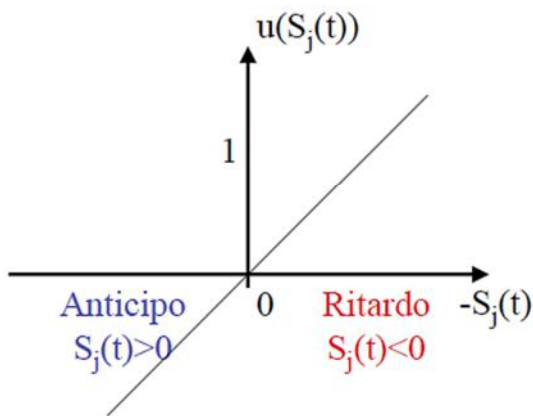
$$CRITICAL RATIO : u(S_j(t)) = \frac{1}{1 + S_j(t)/p_j}$$

REGOLE DI PRIORITA' DINAMICHE: SLACK

Abbiamo definito lo slack come da $S_j(t) = d_j - p_j - t$; esso è positivo quando siamo in anticipo e negativo in caso contrario; andando a rappresentarlo graficamente, si decide di rappresentarlo sull'asse delle ascisse l'opposto affinché risulta esserci un anticipo a sinistra e un ritardo a destra. Quando abbiamo schedulazioni che utilizzano regole di priorità, non siamo normalmente su problemi a macchina singola ma su delle linee che hanno più macchine. In questo caso lo slack deve essere calcolato come $S_j(t) = d_j - p_{ij} -$

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

$$\Pi_j = \frac{w_j}{p_j} u(S_j(t)) = \frac{w_j}{p_j} (-S_j(t))$$



$$Slack = S_j(t) = d_j - p_j - t$$

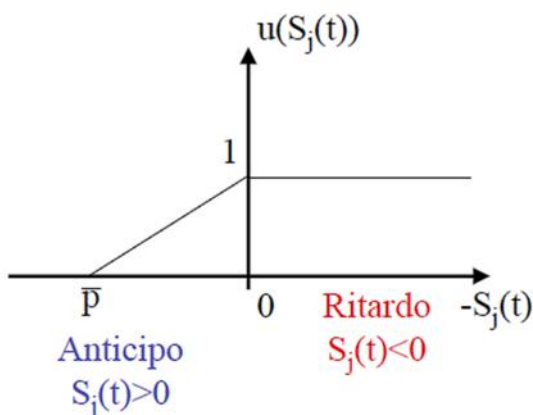
REGOLA COVERT

Questa regola ha una funzione lineare a tratti (come la tardyjob) ma consente di gestire meglio il passaggio del job nella fase di ritardo. La priorità è data da:

$$\Pi_j = \frac{w_j}{p_j} \delta \left(1 - \frac{\delta(S_j(t))}{\bar{p}} \right)$$

con $\bar{p} = \frac{1}{N} \sum_j p_j$ che rappresenta il tempo di processo medio dei job che concorrono ad accedere alla risorsa, cioè i job da schedulare secondo la regola e $\delta(x) = \begin{cases} x & \text{se } x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$. Questa può anche essere scritta nella forma

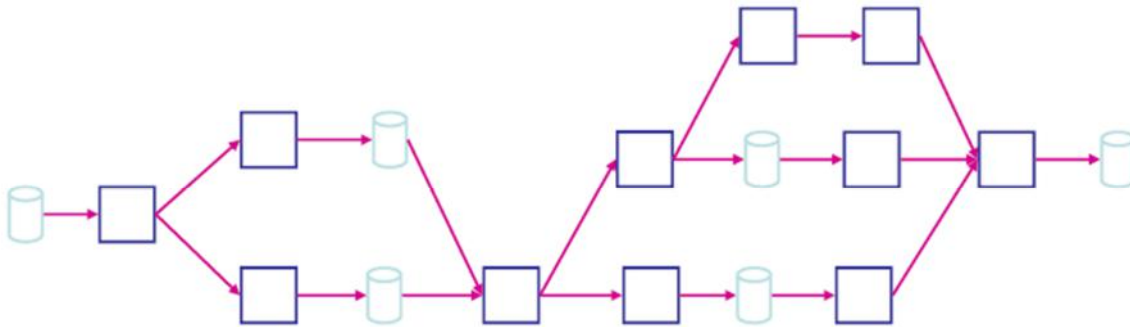
$$\Pi_j = \begin{cases} 0 & \text{se } S_j(t) \geq \bar{p} \\ \frac{w_j}{p_j} \left(1 - \frac{S_j(t)}{\bar{p}} \right) & \text{se } 0 \leq S_j(t) \leq \bar{p} \\ \frac{w_j}{p_j} & \text{se } S_j(t) \leq 0 \end{cases}$$



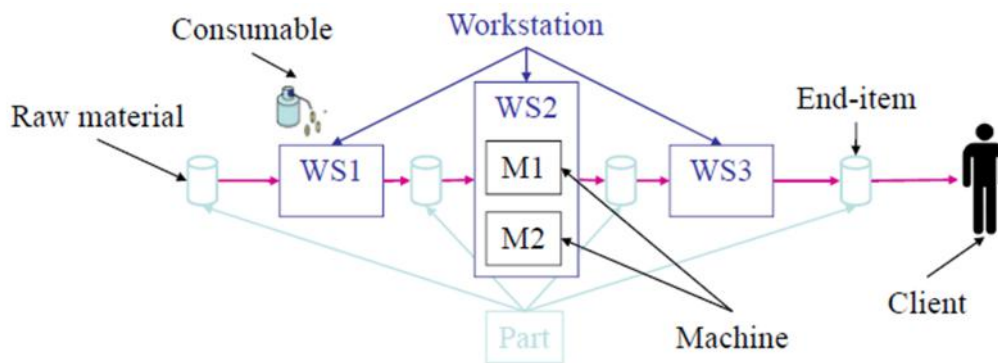
$$Slack = S_j(t) = d_j - p_j - t$$

FF1 - FACTORY PHYSICS APPROACH - LINE

Questo approccio adottato riguarda l'analisi delle leggi fisiche di un sistema di produzione. Un **sistema di produzione** è una rete di processi e punti di stoccaggio attraverso cui le parti fluiscono secondo uno scopo definito. La rete è costituita da **routing** o **linee** che, a loro volta, sono costituite da processi; una stessa macchina può appartenere a routing diversi, così come potrebbe accadere anche per i punti di stoccaggio.

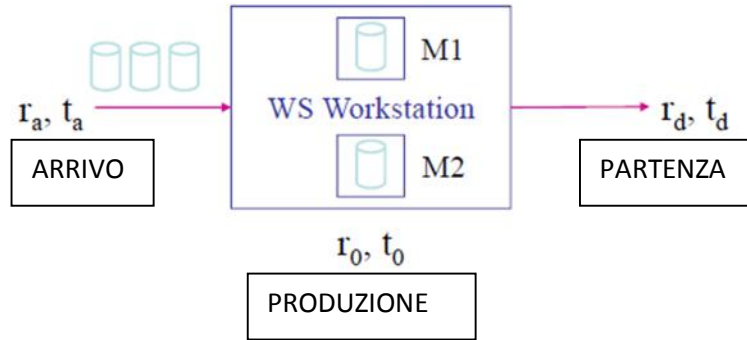


Nello schema i singoli quadratini rappresentano le workstation mentre invece i cilindri sono i punti di stoccaggio. L'obiettivo è capire come funziona il sistema di produzione, ossia come la rete lavora e come i prodotti fluiscono lungo questa rete. Lo studio della fisica della fabbrica, infatti, si concentra sull'analisi della rete e dei flussi lungo le linee. Per l'analisi le condizioni sono semplificate rispetto alla realtà; si suppone che il sistema sia in regime stazionario. Analizziamo in dettaglio il sistema di produzione.

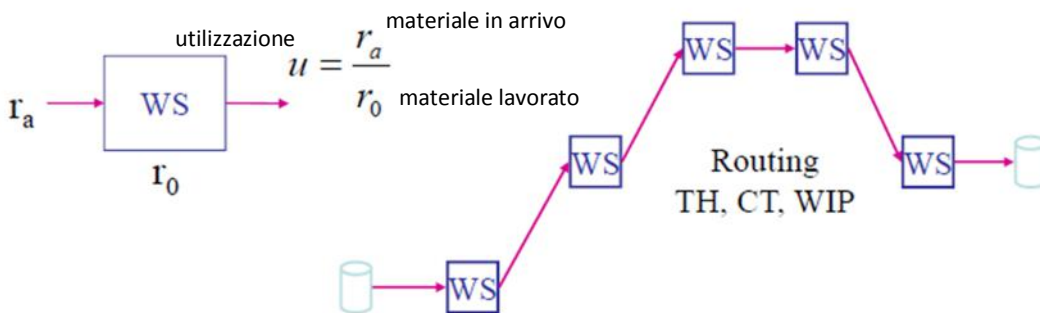


Una **workstation** è l'insieme di una o più macchine che svolgono la stessa funzione; in caso di più macchine, quindi, queste sono parallele, con le stesse caratteristiche, altrimenti appartengono a stazioni diverse. Possiamo poi notare i **punti di stoccaggio** o **buffer** dei pezzi che possono essere **materie prime** (materiale grezzo acquistato all'esterno), **componenti**, **sotto-assemblati** e **prodotti finiti** (parte venduta al cliente); tutti questi componenti sono **parti**, ovvero oggetti che possiamo identificare nella distinta base del nostro prodotto. Esistono poi altri elementi necessari per la lavorazione, ma che non fanno parte della distinta base perché non entrano all'interno del prodotto, e sono i **consumable** come ad esempio il liquido lubro-refrigerante per i macchinari, gli utensili; la loro mancanza ci blocca la produzione e bisogna gestirli nella **bill of manufacturing**.

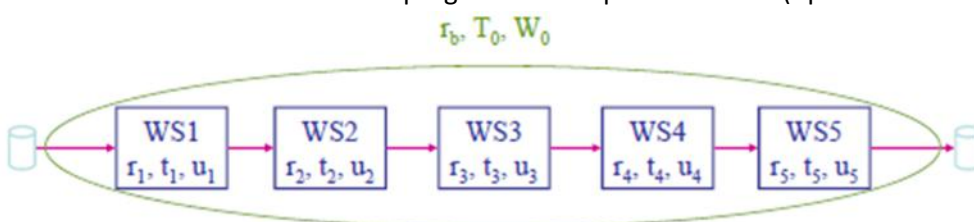
Programmazione e gestione della Produzione
 Prof. Paolo Chiabert



I parametri che conosciamo della nostra stazione di lavoro sono quelli che ci descrivono i processi che agiscono sulla stazione di lavoro. Individuiamo tre differenti processi. Il primo è quello di **arrivo** definito da un **tasso di arrivo medio** r_a e un **tempo medio di interarrivo** t_a (il reciproco di t_a è pari a r_a ; in formule $r_a = 1/t_a$) e quindi quante parti mi arrivano nell'unità di tempo dalla stazione e il tempo tra due arrivi delle parti dalla stazione; il secondo processo è quello di produzione in cui abbiamo un **tasso di produzione** r_0 e un **tempo medio di processo** t_0 ; questi due processi interagiscono tra di loro per generare il **tasso medio di partenza** r_d e il **tempo medio di interpartenza** t_d ; quindi quante parti mi escono nell'unità di tempo dalla stazione e il tempo tra due partenze delle parti dalla stazione.



A seconda della tipologia di processo che stiamo utilizzando, e quindi make to order o make to stock, il soddisfacimento dei nostri clienti può essere definito da due indicatori: nel primo caso abbiamo il **service level** o **livello di servizio** che rappresenta la probabilità di soddisfare il cliente ovvero la $P[CT \leq LT]$; nel secondo caso abbiamo il **fill rate** che è la frazione di ordini soddisfatta dallo stoccaggio di materiale, ad esempio quante volte il cliente trova il materiale cercato. Altri concetti importanti sono la **capacità produttiva** che è il limite superiore del TH di un processo produttivo e quindi quanto la macchina può produrre e l'**utilizzo** ovvero la frazione di tempo in cui la workstation non risulta inattiva a causa della mancanza di parti da lavorare, dato dal rapporto tra il tasso di arrivo e il tasso di produzione. Non è possibile avere un'utilizzazione del 100% nel mondo reale e tanto più facciamo aumentare l'utilizzazione tanto più difficile diventa gestire il sistema. Tutti i sistemi di produzione chiedono un'utilizzazione che sia in progettazione superiore al 95% (tipicamente 98%-98.5%).



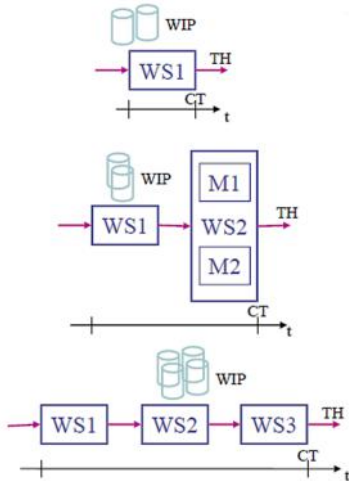
Come misuriamo le prestazioni di una linea? Nei nostri sistemi tutto ciò che entra nella linea esce, ammesso che non vi siano flussi in ingresso o in uscita; quindi il tasso di arrivo sarà identico, sempre come valor medio, su

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

SE CAMBIA IL MIX DI PRODUZIONE CAMBIA IL COLLO DI BOTTIGLIA

Bisogna andare a vedere dove c'è l'utilizzazione maggiore ed individuare il collo di bottiglia. => sarà la macchina che richiede attenzione nel calcolo delle prestazioni:

Matematicamente quindi non è la macchina più lenta che costituisce il collo di bottiglia (nel 99% dei casi è così) ma dal punto di vista ingegneristico è la macchina con la più alta utilizzazione.



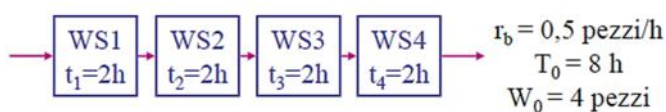
La **LEGGE DI LITTLE** ci dice che la quantità di pezzi che abbiamo su un qualsiasi sistema di produzione ben definito risulta essere uguale al prodotto tra il throughput, ovvero quanto il sistema può produrre, e il tempo ciclo, ovvero il tempo necessario affinché il prodotto attraversa tutto il sistema di produzione; in termini matematici abbiamo $WIP = TH \cdot CT$. Questa è la legge fondante di qualsiasi sistema di produzione.

BENCHMARKING

Il **benchmarking** consiste nella valutazione della prestazione della linea, ovvero come sta funzionando il nostro sistema in termini di TH, WIP e CT, con dati reperiti all'esterno o all'interno della linea, e confrontare queste valutazioni con gli obiettivi che ci vogliamo prefiggere. Esistono due modi per fare benchmarking: quello **esterno** dove si confrontano i dati di due linee simili, dove i problemi sono il reperimento dei dati e della disomogeneità dei dati delle due linee e quindi la non confrontabilità dei dati; quello **interno**, dove si definiscono le prestazioni della linea, quindi quali sono le migliori e le peggiori prestazioni che la linea può avere, e si confrontano con i dati misurati sulla linea stessa; eventuali miglioramenti della linea potrebbero incrementare le prestazioni calcolate. Le nostre misurazioni si andranno a piazzare in qualche punto tra il **best case**, ovvero le migliori prestazioni della linea, e il **worst case**, ovvero le peggiori. Questo piazzamento indica il funzionamento, in linea di massima, della linea di produzione e vedere se bisogna effettuare miglioramenti.

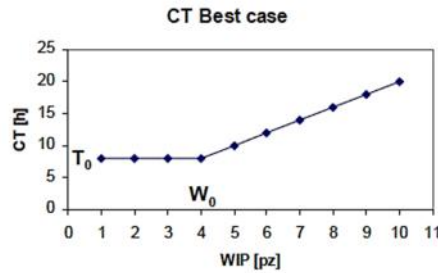
Analisi di Best Case

Il best case indica l'assenza di variabilità. Vediamo un esempio:



Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

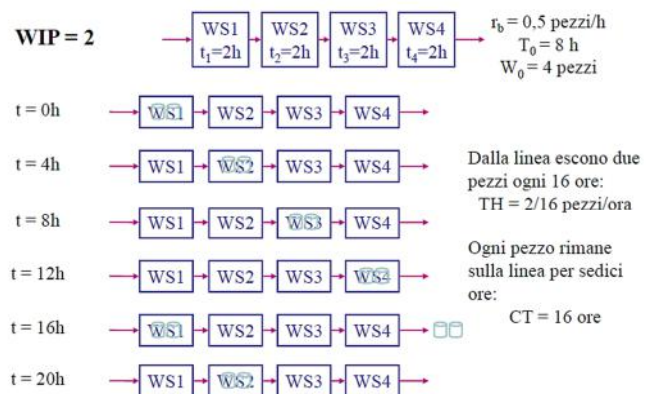
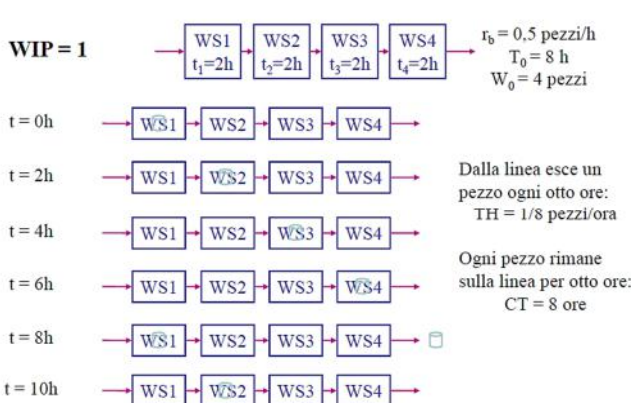
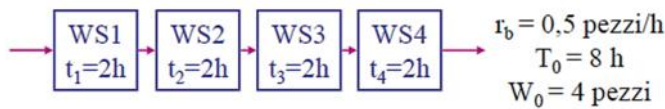
$$CT_{BEST}(w) = \begin{cases} T_0 & \text{se } w \leq W_0 \\ \frac{w}{r_b} & \text{se } w > W_0 \end{cases}$$



Siamo sicuri in questo caso che la nostra linea non potrà mai comportarsi meglio di quanto si è visto. Il punto di lavoro della linea sarà nell'area sottesa dal grafico (nel TH sotto il grafico, nel CT sopra il grafico).

Analisi di Worst Case

La condizione peggiore si ha quando i tempi di processo sono variabili in modo deterministico; si ha quindi massima variabilità deterministica. Se i tempi sono molto variabili, stiamo lavorando male ossia che il processo non è ripetibile, ci sono molti guasti, la manodopera non è addestrata correttamente a svolgere il loro lavoro ecc. La situazione peggiore potrebbe essere quella in cui i tempi di processo aumentano man mano che metto dei pezzi nella mia linea. Non possiamo modificare il tempo di processo del singolo pezzo perché altrimenti verrebbero cambiati i parametri della linea. Bisogna fare in modo che il tempo cresca senza cambiare i parametri della linea, e questo lo si fa inserendo i batch ovvero i lotti di produzione. In pratica, nel worst case, si ha la situazione in cui ciascun pezzo viene messo in un lotto di produzione che non si può spostare nella fase di lavorazione successiva fin quando non si è completata la lavorazione di tutti i pezzi che sono nel lotto. Vediamo un esempio:



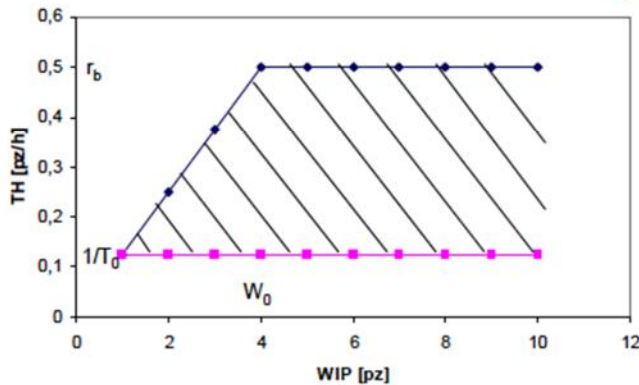
Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Best case performance

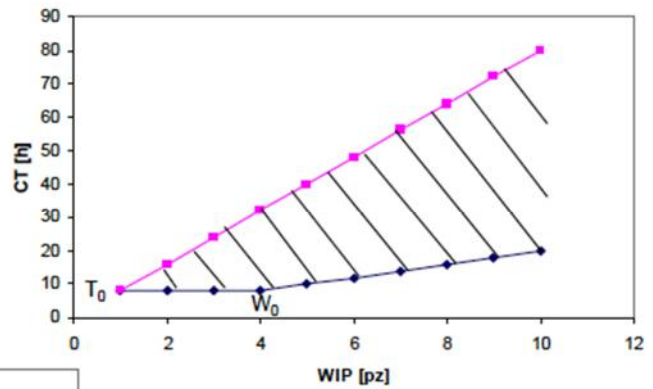
$$TH_{BEST}(w) = \begin{cases} \frac{w}{T_0} & \text{se } w \leq W_0 \\ r_b & \text{se } w > W_0 \end{cases}$$

$$CT_{BEST}(w) = \begin{cases} T_0 & \text{se } w \leq W_0 \\ \frac{w}{r_b} & \text{se } w > W_0 \end{cases}$$

Throughput



Cycle time



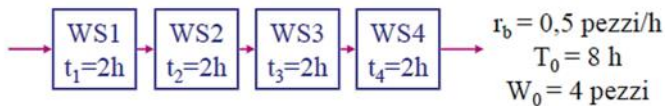
Worst case performance

$$TH_{WORST}(w) = \frac{1}{T_0}$$

$$CT_{WORST}(w) = w \cdot T_0$$

Practical Worst Case

Supponiamo di stabilire che nella nostra linea data da:



Abbiamo un WIP pari a 3 pezzi (WIP = 3). Vi sono 20 possibili stati del sistema per la disposizione dei pezzi e una occupazione media di ogni workstation di 15/20 pezzi; possiamo quindi creare un vettore di stato: esso ci dice esattamente il numero di pezzi che possiamo avere sulle macchine.

	Stato	Vettore	Stato	Vettore	Stato	Vettore	Stato	Vettore	
Worst case	1	(3,0,0,0)	6	(2,0,1,0)	11	(1,0,2,0)	16	(0,0,1,2)	
	2	(0,3,0,0)	7	(2,0,0,1)	12	(0,1,2,0)	Best case	17	(1,1,1,0)
	3	(0,0,3,0)	8	(1,2,0,0)	13	(0,0,2,1)		18	(1,1,0,1)
	4	(0,0,0,3)	9	(0,2,1,0)	14	(1,0,0,2)		19	(1,0,1,1)
5	(2,1,0,0)	10	(0,2,0,1)	15	(0,1,0,2)	20		(0,1,1,1)	

Come possiamo descrivere una situazione come questa per poter tirare fuori dei valori? Bisogna fare delle ipotesi, come:

1. la linea è bilanciata (stessa utilizzazione e tempi di processo analoghi);
2. workstation composte da una sola macchina (macchine singole);

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

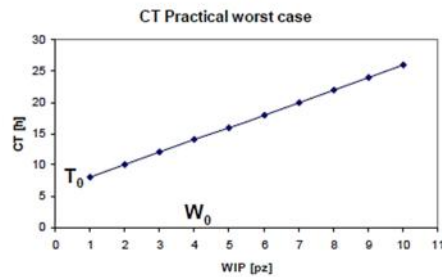
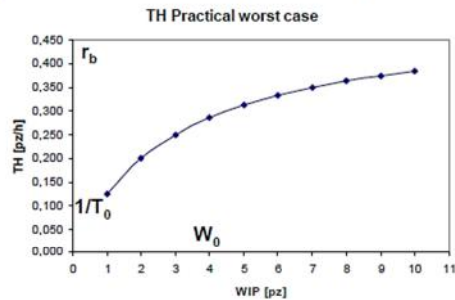
Il CT è inversamente proporzionale alla capacità di lavoro delle macchine ed è direttamente proporzionale al numero di pezzi che ci sono nel sistema. Riassumendo avremo che:



Benchmarking: Practical worst case

$$TH_{PWC}(w) = \frac{w}{W_0 + w - 1} r_b$$

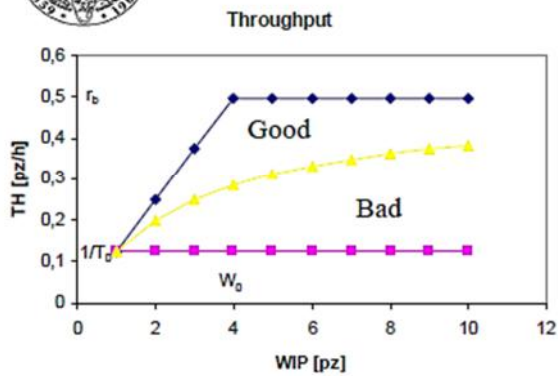
$$CT_{PWC}(w) = T_0 + \frac{w-1}{r_b}$$



Mettendo insieme tutti i casi, ci ritroveremo ad avere un area **good**, dove noi stiamo lavorando bene, e un area **bad**, dove non stiamo lavorando bene e bisogna investire per migliorare la situazione.



Benchmarking interno



Best case

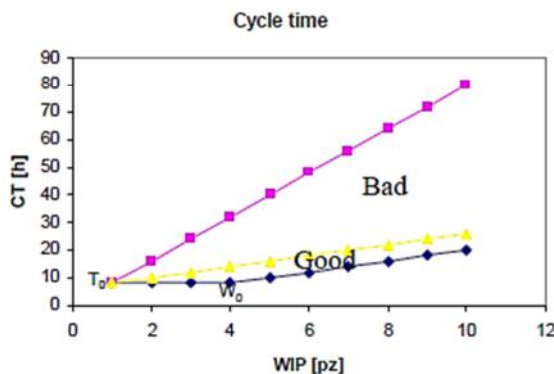
$$TH_{BEST}(w) = \begin{cases} \frac{w}{T_0} & se\ w \leq W_0 \\ r_b & se\ w > W_0 \end{cases}$$

$$CT_{BEST}(w) = \begin{cases} T_0 & se\ w \leq W_0 \\ \frac{w}{r_b} & se\ w > W_0 \end{cases}$$

Worst case

$$TH_{WORST}(w) = \frac{1}{T_0}$$

$$CT_{WORST}(w) = w \cdot T_0$$



Practical worst case

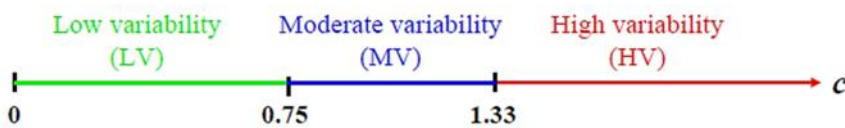
$$TH_{PWC}(w) = \frac{w}{W_0 + w - 1} r_b$$

$$CT_{PWC}(w) = T_0 + \frac{w-1}{r_b}$$

FF2 - FACTORY PHYSICS APPROACH - WORKSTATION

La **variabilità** rappresenta la non uniformità di una classe di entità; consiste essenzialmente nello scostamento dalla regolarità e dalla predicibilità del comportamento del sistema. Può essere **deterministica** o **stocastica**: è deterministica nel **worst case**, dove è completamente predicibile, conseguenza di un controllo inadeguato del sistema, mentre è stocastica nel **practical worst case** perché è dovuta alla casualità. Gli elementi che possono essere caratterizzati dalla variabilità sono dimensioni fisiche del pezzo, tempi di processo, tempi di riparazione dei guasti, tempi di setup ed altri elementi ancora. Questi elementi possono essere caratterizzati tramite una variabile casuale; in statistica sappiamo che ogni variabile casuale è caratterizzata da una distribuzione (esponenziale, normale, di Poisson) e da una media μ e una varianza σ^2 . In una variabile casuale, la **grandezza** è misurata statisticamente tramite la media, mentre invece, la **variabilità** è misurata statisticamente attraverso la varianza o la deviazione standard. Si possono definire dei coefficienti adimensionali per classificare la variabilità di una variabile casuale: il **coefficiente di variazione quadratico (SCV)** definito come $c^2 = \frac{\sigma^2}{\mu^2}$, che è quello più utilizzato, e il coefficiente di variazione (CV) dato da $c = \frac{\sigma}{\mu}$. Per confrontare tra loro le variabili casuali si classificano i sistemi secondo la variabilità espressa attraverso il coefficiente di variazione:

- Bassa variabilità $c < 0,75$ LV
- Media variabilità $0,75 < c < 1,33$ MV
- Alta variabilità $c \geq 1,33$ HV



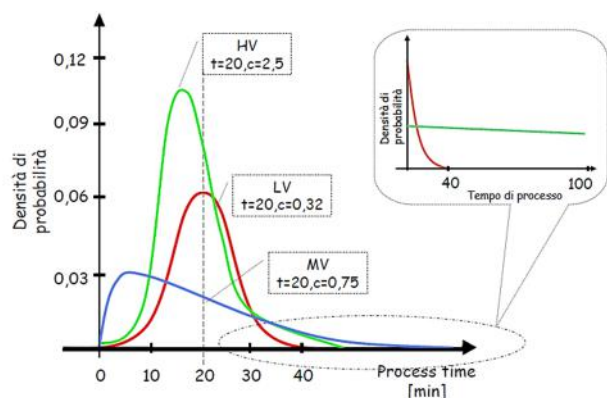
Le prestazioni del benchmarking interno sono invece Low Variability nel caso in cui ci troviamo tra Best Case e Practical Worst Case, Mean Variability nel Practical Worst Case ed in infine High Variability tra Practical Worst Case e Worst Case. Di solito assumiamo un coefficiente di variabilità pari a 1, per comodità, nel caso di variabilità media. Vediamo un esempio:

Si consideri una macchina M1 caratterizzata da un tempo di processo con distribuzione normale con media $t=20$ minuti.

• E' a alta, media o bassa variabilità?

• Sperimentalmente si osserva che se:

- $\sigma = 6,3$ minuti $\rightarrow c = 6,3 / 20 \approx 0,32 < 0,75$ \rightarrow LV
- $\sigma = 15$ minuti $\rightarrow c = 15 / 20 = 0,75$ \rightarrow MV
- $\sigma = 64$ minuti $\rightarrow c = 64 / 20 = 3,2 > 1,33$ \rightarrow HV



Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

• **NON-PREEMPTIVE OUTAGES**

Sono essenzialmente dovuti a setups, ossia gli attrezzaggi, o sostituzione di utensili e consentono la conclusione del job. Si assume che la probabilità di eseguire un setup dopo ogni pezzo sia la stessa.

N_s = numero medio di pezzi processati dalla macchina tra un setup e un altro

t_s = tempo medio di setup

c_s = coefficiente di variazione del tempo di setup

Il setup incrementa il tempo effettivo di processo di ciascun job della quantità $\frac{t_s}{N_s}$. Si ha quindi:

$$t_e = t_0 + \frac{t_s}{N_s}$$

La varianza del tempo effettivo di processo è pari a:

$$\sigma_e^2 = \underbrace{\sigma_0^2}_{\text{Variabilità naturale}} + \underbrace{\frac{\sigma_s^2}{N_s}}_{\text{Variabilità tempo di setup}} + \underbrace{\frac{(N_s - 1)}{N_s^2} t_s^2}_{\text{Interazione fra i vari setup}}$$

Mentre invece il coefficiente di variazione quadratico del tempo effettivo di processo è dato da:

$$c_e^2 = \frac{\sigma_e^2}{t_e^2}$$

• **REWORKING FOR QUALITY**



Siamo nella situazione in cui a valle del nostro sistema di produzione c'è un sistema in grado di rilevare il controllo della qualità del prodotto. Caratterizziamo il processo individuando la probabilità di produrre una parte difettosa ed indicandola con p . Il controllo di qualità, qualora dovesse rilevare il difetto, richiede la lavorazione del pezzo finché questa difettosità non viene eliminata. Le rilavorazioni incrementano il tempo effettivo di processo, riducono l'utilizzazione della stazione, aumentano la variabilità e quindi congestionano il sistema. Dimostriamo questa variazione del tempo di processo: supponiamo che il processo di rilavorazione continui all'infinito (dal punto di vista analitico) fino a quando non si elimina la difettosità.

Programmazione e gestione della Produzione
 Prof. Paolo Chiabert

Su una singola workstation possiamo avere tre differenti processi: **un processo di arrivo** (parametri indicati con il pedice a ovvero arrive), **un processo di produzione** (parametri indicati con il pedice e ovvero effective) e un **processo di partenza** (parametri indicati con il pedice d ovvero departure). Il nostro problema è capire come questi processi si mischiano tra loro. Il processo di uscita è creato dall'interazione del processo di arrivo e quello di produzione.

Low variability arrivals



Un basso coefficiente di variazione dei tempi tra gli arrivi indica arrivi regolari e uniformemente spazati.

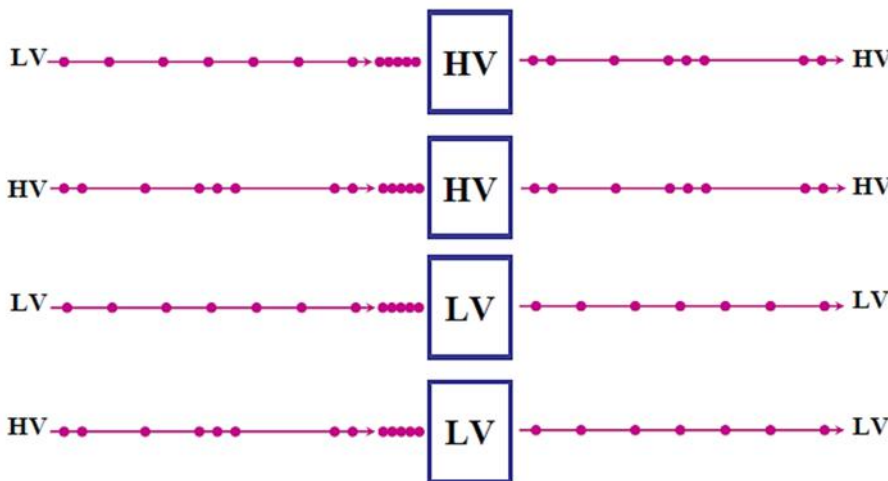
High variability arrivals



Viceversa, un alto coefficiente di variazione dei tempi tra gli arrivi indica arrivi irregolari con una distribuzione irregolare.

Sappiamo quindi che il processo di arrivo può avere alta o bassa variabilità, così come anche il processo di produzione. Siamo interessati a capire come le variabilità di questi due processi si combinano e si propagano per dar origine al processo di partenza.

Stazione ad alta utilizzazione

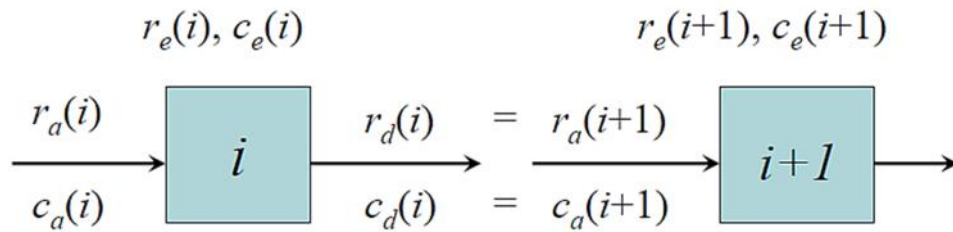


- $r_a = 1/t_a$ tasso di arrivo
- t_a = tempo tra gli arrivi
- σ_a = deviazione standard del tempo tra gli arrivi
- c_a = coeff. di variazione del tempo tra gli arrivi

Nel primo caso abbiamo un alta utilizzazione, ovvero quasi pari ad 1 ovvero 0,99. Nella prima linea abbiamo in arrivo bassa variabilità mentre il processo di produzione ha alta variabilità, in uscita avremo quindi un processo con alta variabilità. Nella seconda linea abbiamo in arrivo alta variabilità mentre il processo di produzione ha anch'esso alta variabilità, in uscita avremo quindi un processo con alta variabilità. Nella terza linea abbiamo in arrivo bassa variabilità

mentre il processo di produzione ha anch'esso bassa variabilità, in uscita quindi avremo un processo con bassa variabilità. Infine, nella quarta linea abbiamo in arrivo alta variabilità mentre il processo di produzione ha bassa variabilità, in uscita avremo quindi un processo con bassa variabilità. **Il processo di uscita è quindi una copia del processo di produzione.**

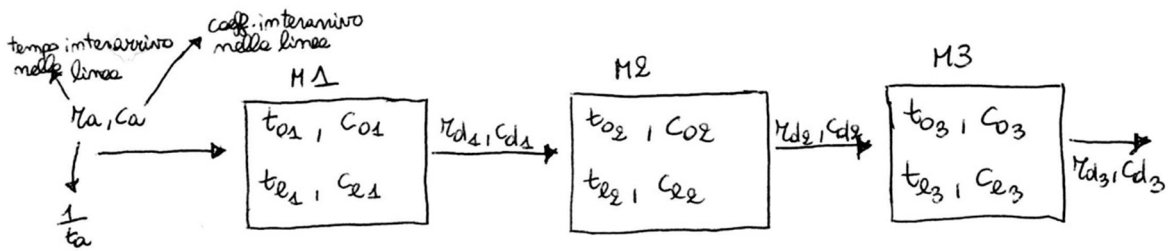
Programmazione e gestione della Produzione
 Prof. Paolo Chiabert



Se tra le due stazioni non è presente un magazzino si ha:

$$\begin{aligned} t_a(i+1) &= t_d(i) \\ c_a(i+1) &= c_d(i) \\ TH &= r_a \end{aligned}$$

Se $u < 1$ allora il sistema è stabile, perché la capacità di produzione è superiore alla capacità dei pezzi da lavorare; se $u > 1$ allora il sistema è instabile, ovvero arrivano più pezzi rispetto a quelli che posso lavorare e si sta creando un accumulo di pezzi. Se il sistema è stabile, posso tranquillamente calcolare i processi di partenza; viceversa non posso andare avanti senza buffer, che consentono alla linea di produrre abbassando il throughput.



t_{oi} = tempo naturale di processo nella stazione i
 c_{oi} = coeff. variazione naturale del tempo di processo della WS i
 A causa di guasti, setup e rilavorazioni abbiamo per ogni WS i
 t_{ei} = tempo effettivo di processo
 c_{ei} = coeff. variazione del tempo effettivo di processo

$$u_i = \frac{\mu_a \cdot t_{ei}}{m_i} \Rightarrow \begin{cases} \text{Se } u < 1 \\ \text{Sistema} \\ \text{stabile} \end{cases} \Rightarrow \begin{aligned} \mu_d &= \mu_a \\ c_d^2 &= u^2 c_e^2 + (1-u^2) c_a^2 \quad \text{se } m=1 \\ c_d^2 &= 1 + (1-u^2) \frac{(c_a^2 - 1) + u^2 (c_e^2 - 1)}{\sqrt{m}} \quad \text{se } m \geq 1 \end{aligned}$$

Programmazione e gestione della Produzione
Prof. Paolo Chiabert

Le performance, invece, possono essere misurate da:

- p_n che rappresenta la probabilità che vi siano n job alla stazione;
- CT_q ovvero il tempo di attesa che si presume si spenda in coda;
- CT ossia il tempo di attesa che si presume si spenda nella stazione;
- WIP cioè il work in process medio alla stazione;
- WIP_q ovvero il work in process medio in coda;
- TH che è il throughput del sistema;
- Stato del sistema che rappresenta il numero di pezzi contenuti nel sistema, con la probabilità degli stati si possono calcolare le performance della macchina.

Essendo sistemi molto complessi, per un sistema a coda è stata scelta la seguente notazione, nota come **notazione di Kendall**.

Notazione di Kendall: **A / B / m / b**

- **A** distribuzione dei tempi di arrivo:
 - D = distribuzione deterministica (costante)
 - M = distribuzione esponenziale (Markoviana)
 - G = distribuzione generale (normale, uniforme,..)
- **B** distribuzione dei tempi di processo
 - D = distribuzione deterministica (costante)
 - M = distribuzione esponenziale (Markoviana)
 - G = distribuzione generale (normale, uniforme,..)
- **m** numero di macchine nella stazione
- **b** massimo numero i job ammissibili nel sistema

MODELLO **M/M/1/∞**

Questo modello indica tempi di arrivo distribuiti esponenzialmente, tempi di processo anch'essi distribuiti esponenzialmente, una sola macchina nella stazione e quindi macchina singola ed infine coda di dimensione illimitata. I dati che ci interessano sono i seguenti: il tempo medio tra gli arrivi è dato da t_a e un tasso degli arrivi pari a $r_a = \frac{1}{t_a}$, con una variabilità c_a media pari a 1 perché markoviana (esponenziale); il tempo medio di processo effettivo è t_e e un tasso degli arrivi pari a $r_e = \frac{1}{t_e}$, poiché siamo su macchina singola, altrimenti sarebbe $r_e = m \frac{1}{t_e}$, con una variabilità c_e media pari a 1; infine abbiamo il numero n di job presenti nel sistema (essendo un sistema senza memoria è interessante il dato in un certo istante). Le ipotesi che possiamo fare sono p_n la probabilità che il sistema a regime si trovi nello stato n -esimo, ossia vi sono n job nelle WS e che i job arrivano uno per volta, prodotti uno per volta e che determinano il passaggio di stato da $n-1$ a n oppure da n a $n-1$; questo vuol dire che se sul sistema ho n pezzi e ne arriva uno nuovo, lo stato sale a $n+1$,

Programmazione e gestione della Produzione
Prof. Paolo Chiabert

Questo deve essere utilizzato per determinare una delle performance del sistema, per poi calcolare le altre tramite la legge di Little. Quello che ci è più comodo da calcolare è il WIP ossia il numero medio di pezzi che ho sulla macchina. Esso è quindi il prodotto tra il numero di pezzi n in ciascuno degli stati possibili per la probabilità di trovarsi esattamente in quello stato.

$$\begin{aligned} WIP &= \sum_{m=0}^{\infty} m \cdot p_m = \sum_{m=0}^{\infty} m \cdot u^m \cdot p_0 = \sum_{m=0}^{\infty} m \cdot u^m (1-u) = (1-u) \sum_{m=0}^{\infty} m \cdot u^m = \\ &= u(1-u) \sum_{m=0}^{\infty} m u^{m-1} = u \cdot \frac{1}{(1-u)^2} = \frac{u}{1-u} \end{aligned}$$

⇓

$$\frac{d}{du} \sum_{m=0}^{\infty} u^m = \frac{d}{du} \left(\frac{1}{1-u} \right) = \frac{1}{(1-u)^2}$$

(teorema di derivazione sotto il segno di serie)

Da cui:
$$WIP = \frac{u}{1-u}$$

Le misure di prestazione peggiorano al crescere dell'utilizzazione e del tempo di processo effettivo ed esplodono quando u si avvicina ad 1. Infatti se abbiamo un sistema di produzione nuovo, con un' elevata utilizzazione c'è questa esplosione del WIP perché c'è una forte variabilità.

Naturalmente tutto il ragionamento fatto è vero solo se il sistema è stabile ovvero $u < 1$.

Possiamo applicare la legge di Little per ricavare le performance del sistema.

- $WIP = \frac{u}{1-u}$
- $TH = r_a$
- $CT = \frac{WIP}{TH} = \frac{t_e}{(1-u)}$
- $CT_q = CT - t_e = \frac{u}{(1-u)} t_e$
- $WIP_q = \frac{u^2}{1-u}$
- $u = \frac{r_a}{r_e} < 1$

MODELLO G/G/1/∞

Avere una distribuzione esponenziale nella realtà dei tempi vuol dire avere una somma di fenomeni completamente differenti; ad esempio ricevere ordini da clienti sparsi in tutto il mondo, tra loro non correlati per avere una distribuzione esponenziale dei tempi di arrivo. Le altre distribuzioni non rientrano nel caso precedente, perché non danno luogo a processi markoviani. Possiamo creare un modello generale per tutte le distribuzioni. Tramite qualche semplificazione possiamo avere l'**equazione di Kingsman** o **Relazione VUT**

Programmazione e gestione della Produzione
Prof. Paolo Chiabert

Se $u \neq 1$ la somma di ordine b della serie geometrica è:

$$\sum_{n=0}^b u^n = \frac{1-u^{b+1}}{1-u}$$

$$P_m = u^m P_0$$

$$b \text{ stati: } P_0 + P_1 + P_2 + P_3 \dots + P_b$$

$$P_0 (1 + u + u^2 + \dots + u^b) = 1$$

$$P_0 \sum_{m=0}^b u^m = 1 \Rightarrow P_0 \left(\sum_{m=0}^{\infty} u^m - \sum_{m=b+1}^{\infty} u^m \right) = 1$$

$$\Rightarrow P_0 \left(\frac{1}{1-u} - \sum_{m=0}^{\infty} u^{m+b+1} \right) =$$

$$= P_0 \left(\frac{1}{1-u} - u^{b+1} \cdot \frac{1}{1-u} \right) =$$

$$\Rightarrow P_0 \left(\frac{1-u^{b+1}}{1-u} \right) = 1 \Rightarrow P_0 \left(\frac{1-u^{b+1}}{1-u} \right) = 1$$

$$P_0 = \frac{1-u}{1-u^{b+1}} \quad \rightarrow \quad P_m = \frac{u^m(1-u)}{1-u^{b+1}}$$

Se $u = 1$ gli stati sono equiprobabili:

$$\underbrace{P_0 + P_0 + P_0 + \dots + P_0}_{(1+b) \text{ volte}} = 1$$

per cui:

$$P_0 = \frac{1}{1+b}$$

Se l'utilizzazione è maggiore di 1, butto via del lavoro, perché potrei avere un tempo ciclo così lungo da superare il lead time che mi è concesso; per cui si preferisce servire meno clienti, ovvero fare in modo che il TH sia inferiore a quello che è il tasso di arrivo (posso farlo con il buffer) pur di mantenere breve il CT.

Vediamo come calcolare le prestazioni del sistema; qui si considera invece del WIP il WIP parziale indicato con WIPP.

Ricordiamo che anche qui c_a e c_e sono pari a 1

Programmazione e gestione della Produzione
Prof. Paolo Chiabert

Throughput: dalle formule precedenti:

$$TH \approx \frac{1-u\rho^{b-1}}{1-u^2\rho^{b-1}} r_a$$

Work in process: per buffer piccoli il WIP tenderà ad avvicinarsi alla dimensione del buffer $b-1$; per buffer grandi si avvicinerà a WIP_{nb} :

$$WIP < \min\{WIP_{nb}, b-1\}$$

Tempo ciclo: utilizzando la legge di Little si può ottenere una stima del limite inferiore di CT :

$$CT > \frac{\min\{WIP_{nb}, b-1\}}{TH}$$

MODELLO $G/G/1/b$: Caso $u > 1$

Nel caso in cui l'utilizzazione è maggiore di 1, vuol dire essenzialmente che $r_a > r_e$. Il WIP è bloccato e quindi abbiamo infiniti pezzi in coda. Di conseguenza, il WIP in assenza di blocco è calcolabile se pensiamo di invertire il flusso del materiale e quindi trasformare il processo di produzione in processo di arrivo e viceversa. In tal caso r_a sarà il nostro tasso di partenza e r_e il nostro tasso di arrivo. L'utilizzazione della stazione nella linea rovesciata diviene $1/u$ che è minore di 1.

$$WIP_{Rnb} = \left(\frac{c_a^2 + c_e^2}{2} \right) \left(\frac{1/u^2}{1-1/u} \right) + \frac{1}{u}$$

Utilizzando WIP_{Rnb} si può calcolare un'utilizzazione corretta ρ per la linea rovesciata:

$$\rho_R = \frac{WIP_{Rnb} - 1/u}{WIP_{Rnb}}$$

Ed infine utilizzare $\rho = 1/\rho_R$ per calcolare il TH a partire dal modello $M/M/1/b$:

$$TH \approx \frac{1-u\rho^{b-1}}{1-u^2\rho^{b-1}} r_a$$

Throughput: dalle formule precedenti:

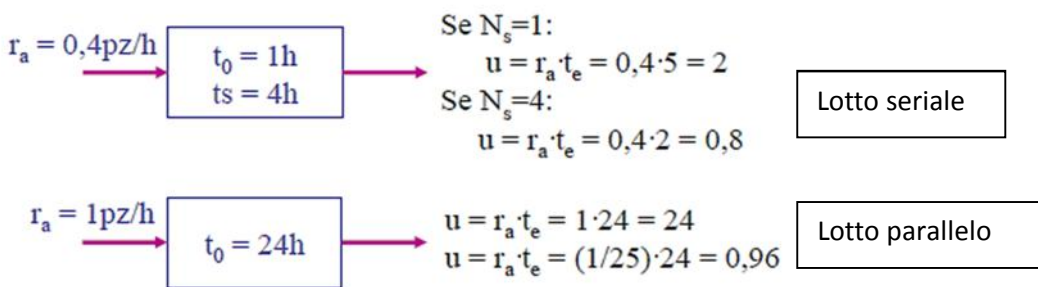
$$TH \approx \frac{1-u\rho^{b-1}}{1-u^2\rho^{b-1}} r_a$$

Work in process: per buffer piccoli il WIP tenderà ad avvicinarsi alla dimensione del buffer $b-1$; per buffer grandi si avvicinerà a WIP_{nb} :

$$WIP < \min\{WIP_{nb}, b-1\}$$

FF3 - FACTORY PHYSICS APPROACH - BATCHING

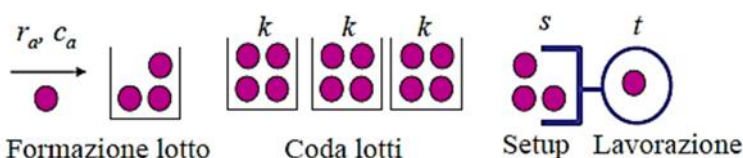
Si è visto già in precedenza come l'inserimento di lotti porta dalla situazione di Best Case, dove non c'è variabilità, alla situazione di Worst Case, dove vi è la massima variabilità deterministica. Quindi l'impiego di lotti è un'importante sorgente di variabilità. Introducendo i lotti cambia enormemente la variabilità del processo di arrivo, perché il primo job impiega del tempo per arrivare ma poi gli altri sono già lì con tempo nullo; pur avendo il tempo medio di arrivo lo stesso valore del singolo pezzo, la variabilità risulta essere molto elevata e di conseguenza le performance diminuiscono. Perché fare il batching? Semplicemente perché quando abbiamo a che fare con tempi di setup elevati dobbiamo riuscire a spalmarli su un numero elevato di job e far sì che l'utilizzazione sia minore di 1, con la conseguente stabilità del sistema. Questo è il caso di lotto seriale. Esiste anche il caso di lotto parallelo o di trasferimento.



Nell'ottica occidentale, si tende ad impiegare i lotti per ottimizzare i costi, mentre nell'ottica orientale si è utilizzato un criterio opposto diminuendo i tempi di setup, come ad esempio la **SMED (Single Minute Exchange Die)**. Queste soluzioni però a volte possono essere costose, ecco perché si utilizzano i lotti.

Nel dettaglio si hanno **lotti di processo**, che riguardano la stazione, e **lotti di trasferimento** dove l'insieme delle parti viene accumulato prima di essere trasferito alla stazione successiva. I lotti di processo si dividono in **lotto seriale** e **lotto parallelo**: il lotto seriale è l'insieme di parti di una famiglia di prodotti lavorate singolarmente prima che la stazione sia attrezzata per lavorare altri prodotti, cioè sono tutti i pezzi che vengono realizzati tra un setup e quello successivo; viceversa il lotto parallelo è l'insieme di parti che vengono lavorate simultaneamente da una stazione quale un forno o una fornace. **Attenzione a non confondere il lotto parallelo con le macchine parallele che lavorano pezzi.** Normalmente si tende ad avere un lotto di processo e un lotto di trasferimento uguali per non dover separare le parti; tuttavia quando sono diversi si parla di **lot splitting**.

LOTTO DI PROCESSO SERIALE



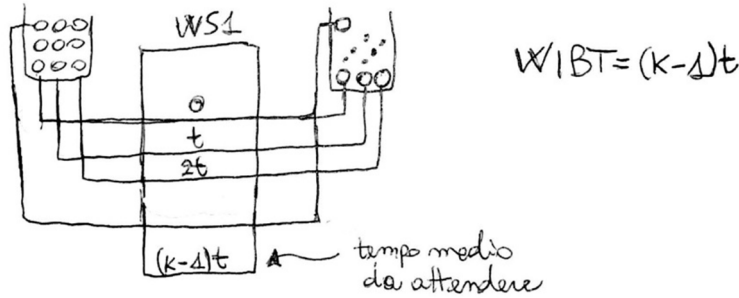
Esiste un processo in cui arrivano le parti che entrano per formare dei lotti di k parti. Una volta che il lotto è pronto, viene messo in coda ed inizia l'effettivo tempo ciclo delle parti; quindi un lotto entra in coda, dovrà aspettare gli altri davanti a lui fin quando non riuscirà ad arrivare di fronte alla macchina. Successivamente la macchina riconoscerà il tipo di lotto, viene fatto il setup e poi si cominciano a lavorare le parti una per

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE
 Prof. Paolo Chiabert

Tempo ciclo del lotto: $CT_{nosplit} = CT_q + s + WIBT_{nosplit} + t$

Tempo ciclo in coda del lotto: $CT_q = \left(\frac{1+c_e^2}{2}\right)\left(\frac{u}{1-u}\right)t_e$

Wait In Batch Time (Nosplit): $(k-1)t$

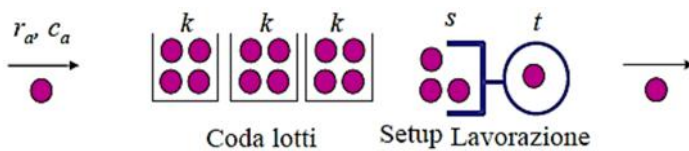


Ciascuna parte aspetta $(k-1)t$ istanti di tempo perché deve uscire e rientrare nel lotto.

$$CT_{nosplit} = CT_q + s + (k-1)t + t = CT_q + s + kt$$

Mettendo insieme tutti i tempi otteniamo:

Nel caso **split**, invece, le parti proseguono una per una e non devono riformare il lotto. E' una cosa che accade molto spesso in una linea di produzione. All'uscita non si riforma il lotto.



Tempo ciclo del lotto: $CT_{split} = CT_q + s + WIBT_{split} + t$

Tempo ciclo in coda del lotto: $CT_q = \left(\frac{1+c_e^2}{2}\right)\left(\frac{u}{1-u}\right)t_e$

Wait In Batch Time (Split): $\left(\frac{k-1}{2}\right)t$

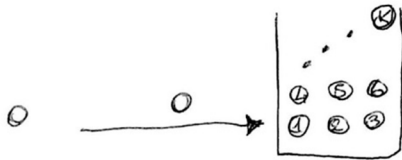
Quello che cambia è il WIBT; risulta essere più piccolo perché non devono aspettare altro dopo la lavorazione, come la formazione del lotto (si ottiene applicando la progressione aritmetica). Di conseguenza si ha:

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE
 Prof. Paolo Chiabert

Possiamo calcolare i parametri per capire se siamo in grado di soddisfare il cliente:

Il tempo ciclo è composto da **WTBT (Wait to Batch Time)** ossia il tempo per comporre il lotto, il tempo ciclo in coda del lotto e infine il tempo di processo della singola parte che è identico a quello del lotto.

Tempo ciclo del lotto: $CT = WTBT + CT_q + t$



tutte devono aspettare $(k-1)t$

$$\frac{1}{k} \sum_{m=0}^{k-1} m t_a = \left(\frac{k-1}{2}\right) t_a \quad \text{per progressione aritmetica}$$

$$\Rightarrow \frac{1}{k} \cdot \frac{k}{2} \frac{(k-1)}{t_a} = \frac{1}{t_a} = \kappa_a$$

Wait To Batch Time: $\left(\frac{k-1}{2}\right) \frac{1}{\kappa_a}$

Tempo ciclo in coda del lotto: $CT_q = \left(\frac{c_a^2/k + c_e^2}{2}\right) \left(\frac{u}{1-u}\right) t$

Come otteniamo c_a^2 dei lotti?

$$c_a^2 = \frac{\sigma_a^2}{t_a^2}$$

κ_a = tasso arrivo parti

$\frac{\kappa_a}{k}$ = tasso di arrivo olei lotti

t_a = tempo intervallo olei lotti

$$\frac{1}{\kappa_a} = \frac{k}{\kappa_a} = t_a \cdot k \quad \Rightarrow \quad \text{essendo } \frac{1}{\kappa_a} = t_a$$

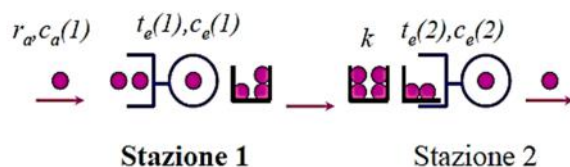
σ_a deviazione standard di $t_a \Rightarrow \sigma_a^2$ = varianza di t_a delle parti

Poi che abbiamo k parti che danno luogo al tempo di arrivo dei lotti si ha che

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE

Prof. Paolo Chiabert

STAZIONE 1:



$$\text{Tempo ciclo: } CT = \left(\frac{c_a^2(1) + c_e^2(1)}{2} \right) \left(\frac{u(1)}{1-u(1)} \right) t_e(1) + t_e(1)$$

$$\text{Wait To Batch Time: } \frac{k-1}{2} \frac{1}{r_a(1)} = \frac{k-1}{2 u(1)} t_e(1)$$

$$CT(1) = \left(\frac{c_a^2(1) + c_e^2(1)}{2} \right) \frac{u(1)}{1-u(1)} t_e(1) + t_e(1) + \frac{k-1}{2 u(1)} t_e(1)$$

Intervallo di tempo tra l'uscita delle singole parti: t_a

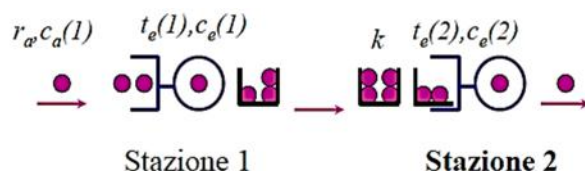
$$\text{SCV tempo di uscita delle parti: } c_d^2(1) = (1-u^2(1)) c_a^2(1) + u^2(1) c_e^2(1)$$

Intervallo di tempo tra l'uscita dei lotti: $k t_a$

$$\text{SCV tempo di uscita dei lotti: } c_a^2(2) = \frac{k c_d^2(1) t_a^2}{k^2 t_a^2} = \frac{c_d^2(1)}{k}$$

La varianza del tempo di uscita del lotto è la somma delle varianze del tempo di uscita delle singole parti.

STAZIONE 2:



Tempo di processo del lotto: $k t_e(2)$

$$\text{SCV del tempo di processo del lotto: } \frac{k c_e^2(2) t_e^2(2)}{k^2 t_e^2(2)} = \frac{c_e^2(2)}{k}$$

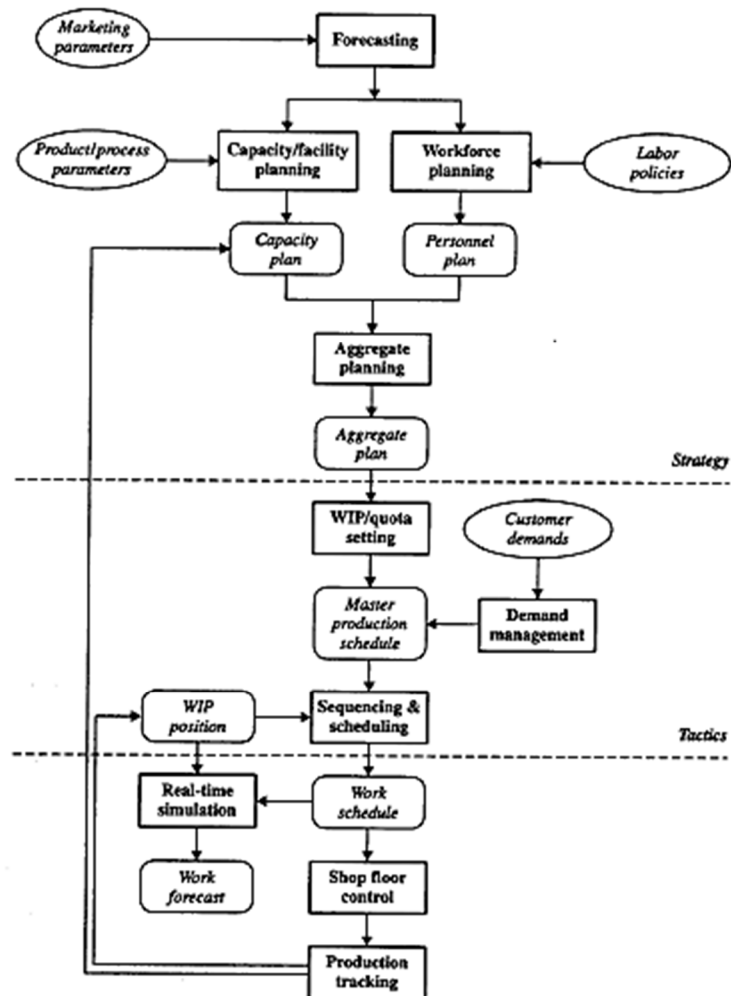
La varianza del tempo di processo del lotto è la somma delle varianze dei tempi di processo delle singole parti.

$$\text{Wait In Batch Time: } \frac{k-1}{2} t_e(2)$$

$$CT(2) = \frac{c_d^2(1)/k + c_e^2(2)/k}{2} \frac{u(2)}{1-u(2)} k t_e(2) + \frac{k-1}{2} t_e(2) + t_e(2)$$

G01 - SISTEMI PULL E JUST IN TIME

I **sistemi pull** sono stati introdotti per la prima volta nel mondo orientale, per la precisione in Giappone. Nel sistema push io ordino la produzione; faccio uscire il materiale dal magazzino, lo faccio lavorare e poi lo riporto in magazzino o lo consegno al cliente in base al tipo di sistema utilizzato (make to stock o make to order). Il sistema pull invece si basa sull'autorizzazione al lavoro, ovvero si vede qual è la quantità di merce che vi è in magazzino e se questa è inferiore al valore stabilito per soddisfare la domanda, allora autorizzo la stazione ad avviare la produzione. Questa stazione poi dovrà a sua volta richiedere il materiale per lavorare, svuotando i magazzini a monte, i quali raggiungeranno anch'essi un determinato livello inferiore a quanto stabilito (WIP che voglio avere) e quindi bisogna produrre o riempire questi magazzini e così via fino a risalire alle materie prime. **Questo è un sistema pull perché la produzione è tirata dalla stazione a**



valle verso la stazione a monte, mentre nel sistema push è la stazione a monte che spinge la produzione verso la stazione a valle. Di conseguenza un sistema push si studia da monte a valle; un sistema pull si studia da valle a monte. Questo piccolo cambiamento di paradigma fa sì che i sistemi siano completamente differenti; infatti questo ha generato un vero e proprio metodo diverso per il controllo e la pianificazione della produzione. In particolare, ciò che è differente non è la parte di input della produzione (APP) ma la parte sottostante. In un sistema push si andava direttamente al MPS e al MRP; qui invece si stabiliscono dapprima due elementi, che sono il **WIP** voluto all'interno della linea e la **quota setting**, ovvero la quota di produzione per periodo o per giornata. Solo dopo aver stabilito queste quantità si calcola il Master Production Schedule e poi da questo si estrapola direttamente la schedulazione delle operazioni lungo la linea di produzione; questo perché l'MPS ci dice quanti pezzi verranno sottratti al magazzino prodotti finiti e questo basta ad innescare la produzione di tutti i semilavorati lungo la linea. Quindi, in sostanza, non bisogna più pianificare in anticipo, ma semplicemente si stabilisce la quantità che si vuole avere di WIP e una volta che questo scende, autorizzo le stazioni a produrre la quantità per raggiungere di nuovo il livello. Inoltre, se controllo il WIP nella linea, non può più accadere un'esplosione di work in process; quello che posso avere è che il WIP non sia più sufficiente a soddisfare il TH e in quel caso posso vedere come sta funzionando la linea, misuro il CT e vedo se aumentare o diminuire il WIP. In pratica, ciò che si ottiene è un sistema più smooth, molto più osservabile e controllabile perché il WIP è più facile da controllare rispetto al TH.

PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE
 Prof. Paolo Chiabert

4500 auto a settimana : $\frac{4500}{5} = 900$ auto al giorno
 $\frac{900}{2} = 450$ auto ogni turno di 8 ore
 $\frac{450}{8} = 56,25$ auto all'ora
 $\frac{56,25}{60} = 0,9375 = 0,94$ auto al minuto

$\frac{4500}{4800} = 0,94$ auto al minuto \Rightarrow per produrre 1 auto sono necessari
 $\hookrightarrow 5 \cdot 2 \cdot 8 \cdot 60 = 4800$ $\frac{1}{0,94} = 1,067$ minuti

SOLUZIONE 1: Produco 200 berline, 150 coupé, 100 station wagon
 UNA VOLTA AL GIORNO (8h)

$200 \cdot 1,067 + 150 \cdot 1,067 + 100 \cdot 1,067 = 480,15 \text{ min} = 8 \text{ ore} \Rightarrow$ ciclo 8h
 Magazzini elevati, domande dipendente fortemente variabile

SOLUZIONE 2: Produco 40 berline, 30 coupé, 20 station wagon
 CINQUE VOLTE PER TURNO (5 · 1,6)h

$40 \cdot 1,067 + 30 \cdot 1,067 + 20 \cdot 1,067 = 96,03 \text{ min} = 1,6 \text{ ore} \Rightarrow$ ciclo 1,6h
 Magazzini grandi, domande dipendente variabile

SOLUZIONE 3: Produco 4 berline, 3 coupé e 2 station wagon
 CINQUANTA VOLTE PER TURNO (50 · 9,6) min

$4 \cdot 1,067 + 3 \cdot 1,067 + 2 \cdot 1,067 = 9,6 \text{ min} \Rightarrow$ ciclo 9,6 min
 Magazzini bassi, domande dipendente poco variabile

SOLUZIONE 4: Produco nell'ordine B, W, B, C, B, C, B, W, C
 CINQUANTA VOLTE PER TURNO (50 · 9,6) min

One Piece Flow \Rightarrow lotto di un pezzo \Rightarrow ciclo 9,6 min

Magazzini minimi, domande dipendente costante

MPS smooth

Nell'ultima soluzione abbiamo raggiunto quello che prende il nome di **One Piece Flow**, ovvero il lotto contenente un solo pezzo.

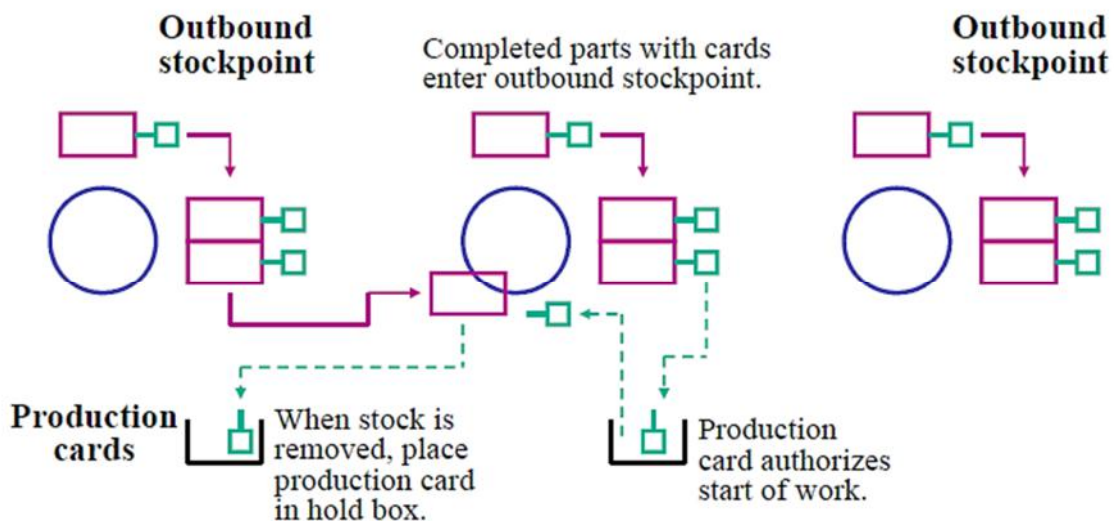
PROGRAMMAZIONE E GESTIONE DELLA PRODUZIONE Prof. Paolo Chiabert

È il lavoratore che svolge alcuni compiti di manutenzione ordinaria sulla macchina per mantenerla in piena efficienza. Ultimo fondamento del JIT è il **miglioramento continuo**. Quest'ultimo è il motivo di forza della filosofia Just In Time. Il miglioramento continuo persegue l'eccellenza e la massima competitività. Una delle metodologie consiste nel ridurre periodicamente il livello delle scorte: ciò stressa il sistema produttivo e fa emergere i problemi che vengono segnalati, analizzati e risolti in modo da migliorare il processo di produzione.

KANBAN

Operativamente, il JIT viene implementato tramite i **Kanban**, ovvero i cartellini. Sviluppato da Toyota, indica che avendo un cartellino nella bacheca di produzione, si è autorizzati a produrre quel determinato materiale. Il materiale verrà prodotto nella quantità indicata sul cartellino, con il codice previsto dal cartellino e lo metterò nel contenitore che deve contenere quel materiale ed infine gli applico il cartellino. Quel cartellino è adesso associato ad un contenitore pieno e quindi non si è più autorizzati a produrre, ammesso che non ci siano altri cartellini. Questa è quella che si chiama autorizzazione alla produzione. Quando andiamo ad utilizzare il materiale, il cartellino viene staccato quindi dal contenitore e rimandato alla stazione che deve produrre quel determinato materiale; quest'ultima lo vede, riempie il contenitore e lo rimanda dove necessario, andando avanti di questo passo. Non è possibile lavorare o movimentare i materiali se non è presente un cartellino di produzione o trasporto. I cartellini, inoltre, consentono un controllo completo e dettagliato dei flussi di materiale e determinano il livello di **WIP** all'interno del sistema. I sistemi possono essere a **singolo cartellino**, a **due cartellini**, a **contenitore** (senza cartellini) o anche **senza contenitore**.

KANBAN - 1 CARTELLINO



Nella figura i cerchi indicano le stazioni di lavoro mentre i rettangoli i magazzini. Quando una macchina realizza un prodotto, questo ha un suo contenitore e un suo cartellino, e una volta finita la produzione, viene messo nel magazzino prodotti finiti. Si parte dalla stazione più a valle.

Il cliente arriva e richiede il prodotto finito, estraendo il materiale dal contenitore. Se il contenitore si svuota bisogna staccare il cartellino e metterlo nella bacheca dei cartellini di produzione. Sul cartellino ci sono le informazioni sulla produzione. L'operatore che vede il cartellino e nota che deve produrre quel