Introduzione alle reti di code nei sistemi manifatturieri

Alessandro Agnetis

1. Introduzione

Questi brevi appunti hanno lo scopo di fornire un'introduzione alle reti di code e al loro utilizzo nella modellistica dei sistemi manifatturieri. L'esposizione è in alcuni punti volutamente priva di sottigliezze formali, in quanto si è preferito concentrare l'attenzione sulla comprensione dei concetti e sul loro significato modellistico. Così, alcune dimostrazioni sono riportate solo per alcuni casi particolari, oppure viene presentata una giustificazione formale e concettuale.

Le reti di code sono state (e sono tutt'ora) oggetto di studi approfonditi, sia dal punto di vista modellistico che specificamente matematico; tuttavia i primi settori applicativi in cui si è fatto uso delle reti di code sono stati prevalentemente i sistemi di commutazione e quelli di elaborazione. Solo dalla fine degli anni '70 si è cominciata a studiare la possibilità di applicare queste metodologie anche ai sistemi produttivi: in questi sistemi le stazioni corrispondono a centri di lavorazione, o celle costituite da più centri in parallelo, e i clienti sono tipicamente dei job, cui spesso è possibile associare fisicamente dei pezzi in lavorazione.

In questa dispensa si suppongono noti i concetti di base relativi alle file d'attesa, ovvero: processi di nascita e morte, sistemi *M/M/1*, *M/M/s*. La trattazione seguirà per grandi linee quella di Askin e Standridge (1993); non mancheranno però alcuni approfondimenti rispetto a quel testo. Nel seguito indicheremo con *stazione* il singolo sistema di servizio, in generale costituito da *s serventi*. Per indicare tutte le grandezze già definite nell'ambito delle file d'attesa, useremo le stesse notazioni impiegate da De Julio e La Bella (1975).

Una *rete di code* è l'interconnessione di un insieme di *M stazioni* (indicate con $j=1,\dots,M$), ciascuna costituita da uno o più *serventi* e una *coda*. La stazione j consiste di s_j serventi identici in parallelo. Tutte le grandezze riferite alla stazione j saranno indicate con il pedice j. Ad esempio, N_j indicherà il valore atteso del numero di clienti nella stazione j.

In un sistema di servizio costituito da una sola fila d’attesa, dopo aver ricevuto il servizio ciascun cliente ritorna in una popolazione, spesso considerata infinita. In una rete di code, in generale, i clienti in uscita da una stazione vanno a disporsi in coda a un'altra, scelta eventualmente con criteri probabilistici. Se sono possibili ingressi nella rete dall'esterno e uscite verso l'esterno, la rete si dice *aperta*; se invece ciò non è possibile, e dunque il numero di clienti che fluiscono nella rete rimane costante, allora la rete si dice *chiusa*.

Prima di addentrarci nell’analisi delle reti di code vogliamo richiamare o stabilire alcuni risultati che saranno utili nel seguito. Nel seguito tratteremo sempre di stazioni il cui tempo di servizio è distribuito esponenzialmente (con parametro μ), e costituite da s serventi. Cominciamo con il ricordare un risultato che dovrebbe essere del tutto familiare:

Teorema 1. *Se il numero degli arrivi in un intervallo di tempo di lunghezza* t *ha distribuzione poissoniana, il tempo d'interarrivo alla stazione ha distribuzione esponenziale*.

DIM. — Si veda De Julio e La Bella (1975).

Teorema 2. *La somma di variabili aleatorie poissoniane di parametro* λ_1, λ_2, $\ldots\lambda_n$ *rispettivamente è una variabile aleatoria poissoniana di parametro* $\lambda = \lambda_1+\lambda_2+\ldots+\lambda_n$.

DIM. — Per semplicità, esponiamo la dimostrazione per il caso $n = 2$. L’estensione al caso generale è ovvia. In tal caso, si considerino due processi poissoniani di parametro λ_1 e λ_2, e siano $P'_{k-\delta}(t)$ e $P''_{\delta}(t)$ le corrispondenti probabilità di avere, all'istante t, rispettivamente $k-\delta$ e δ individui. La probabilità di avere complessivamente un numero di individui pari a k, è evidentemente pari alla somma:

$$\sum_{\delta=0}^{k} [P'_{k-\delta}(t)P''_{\delta}(t)]$$

ossia

$$\sum_{\delta=0}^{k}\left[\frac{(\lambda_1 t)^{k-\delta}}{(k-\delta)!}e^{-\lambda_1 t}\,\frac{(\lambda_2 t)^{\delta}}{\delta\,!}e^{-\lambda_2 t}\right]$$

da cui, ricordando la formula della potenza di un binomio, si ha la tesi:

$$P_k(t) = \frac{((\lambda_1+\lambda_2)t)^k}{k\,!}\,e^{-(\lambda_1+\lambda_2)t}$$

Una proprietà molto rilevante delle funzioni di densità di probabilità è che, se la variabile aleatoria y è pari alla *somma* delle due variabili aleatorie indipendenti x_1 e x_2, la sua densità di probabilità $p_Y(y)$ è data dalla *convoluzione* delle due densità di probabilità $p_{X_1}(x_1)$ e $p_{X_2}(x_2)$. Inoltre, ricordiamo che la trasformata di Laplace della convoluzione di due funzioni è uguale al prodotto delle trasformate di Laplace delle due funzioni.

Come noto, il tempo di interarrivo è una variabile aleatoria pari al tempo che intercorre tra due successivi arrivi ad un sistema di servizio. Analogamente, si definisce *tempo d'inter-uscita* la variabile aleatoria avente valore pari al tempo che intercorre tra uscite consecutive di clienti da un sistema di servizio. Un risultato molto importante per la caratterizzazione del processo delle uscite è dato dal seguente teorema.

Teorema 3. (Burke) *Il processo delle uscite da un sistema M/M/s con coda di capacità infinita è un processo poissoniano di parametro* λ *uguale a quello che caratterizza il processo degli arrivi a tale sistema*.

DIM. — Per semplicità, esporremo la dimostrazione per il solo caso $s = 1$. In base al Teorema 1, dimostreremo che il tempo di inter-uscita t_u da un sistema *M/M/*1 è distribuito esponenzialmente con parametro λ. Si consideri un cliente che esce dal sistema a un generico istante t. Possono presentarsi due casi: *i*) all'istante t nel sistema c'è *almeno* un altro cliente, oppure *ii*) all'istante t il sistema è vuoto. Nel caso *i*), la prossima uscita si avrà dopo un tempo pari al tempo di servizio di questo cliente, e dunque la densità di probabilità $p_u(t_u)$ del tempo di inter-uscita è pari a

$$p_u(t_u) = p_s(t_s) = \mu e^{-\mu ts}$$

Nel caso *ii*), invece, occorrerà attendere dapprima l'arrivo di un nuovo cliente e quindi l'espletamento del servizio. La variabile aleatoria t_u sarà dunque data dalla somma delle due variabili aleatorie indipendenti t_a e t_s:

$$p_u(t_u) = p_a(t_a) * p_s(t_s) = \lambda e^{-\lambda ta} * \mu e^{-\mu ts}$$

A questo punto, poiché il caso *i*) si presenta con probabilità ρ e il caso *ii*) con probabilità $(1 - \rho)$, applicando il teorema della probabilità totale ed effettuando la trasformata di Laplace, si ottiene:

$$P_u(s) = \rho \frac{\mu}{s + \mu} + (1 - \rho) \frac{\lambda}{(s + \lambda)} \frac{\mu}{(s + \mu)}$$

da cui, essendo $\rho = \lambda / \mu$, con facili passaggi si ha la tesi:

$$P_u(s) = \frac{\lambda}{(s + \lambda)}$$

–

2. Reti aperte

Una *rete di code aperta* è un insieme di M stazioni interconnesse da un sistema di trasporto, in cui possono avvenire *ingressi* di clienti nel sistema, *uscite* dal sistema, o *trasferimenti* di un cliente da una stazione a un'altra. In altri termini, allorché una stazione termina un servizio, o il cliente esce dal sistema, oppure rimane al suo interno dirigendosi verso un'altra stazione. Analogamente, l'arrivo di un cliente a una stazione può avvenire dall'esterno del sistema oppure da un'altra stazione.

La scelta della stazione ove il cliente si va a porre può essere sia deterministica che stocastica. Un esempio del primo caso si ha quando, ad esempio, un pezzo deve visitare una successione prefissata di macchine, ognuna delle quali è preposta a una particolare operazione (taglio, lavaggio, misura…). Quindi, con certezza, dopo avere subìto una certa operazione, il pezzo dovrà subirne una certa altra. Il secondo caso si ha invece, ad esempio, allorché il successivo instradamento dipende dall'esito di un'operazione (tipicamente di collaudo o misura): qualora il pezzo risultasse difettoso, deve essere rispedito ad una precedente stazione per una rilavorazione. In genere è possibile stimare con una certa accuratezza la probabilità con cui questo accade.

Una prima considerazione è un'immediata conseguenza dei teoremi visti nell'introduzione. Si considerino alcune stazioni *M/M/s* tali che le loro uscite vanno ad alimentare una nuova stazione (Fig.1). Per il Teorema 3, ciascuno dei “flussi” di clienti è un processo poissoniano. L’unione di questi flussi di clienti darà luogo ancora a un processo poissoniano, di parametro pari alla somma di quelli dei singoli processi (Teorema 2). Il tempo d'interarrivo alla nuova stazione è dunque distribuito esponenzialmente (Teorema 1).

Analogamente, si consideri un processo poissoniano di parametro λ che si scinde in vari "stream", nel senso che ogni cliente del processo ha probabilità p_i di essere instradato nell'i-esimo stream (ovviamente, $\Sigma_i\, p_i = 1$): si può facilmente mostrare che ciascuno stream è ancora un processo poissoniano di parametro $p_i\lambda$.

Esempio 1. Si consideri il sistema in Fig.1. Sei macchine (1–6) processano lotti di pezzi, che vengono poi inviati a una stazione di collaudo (C). Il 97% dei lotti sono spediti alla stazione di imballaggio (P), il rimanente 3% è difettoso e rinviato in lavorazione. Ogni macchina ha una propria coda. I tempi di interarrivo alle macchine e tutti i tempi di servizio sono esponenziali; ogni macchina processa in media 10 lotti al giorno. Vogliamo determinare i parametri dei processi di arrivo alle stazioni C e P.

Le uscite delle 6 macchine si fondono in un unico stream, in ingresso alla stazione di collaudo. Per il Teorema 3, tale processo è poissoniano, di parametro λ_C=60 lotti/giorno. Il processo in uscita da C è ancora poissoniano, e si suddivide in due stream: quello in ingresso all'imballaggio ha parametro λ_P = (0.97) 60 = 58.2 lotti/giorno.

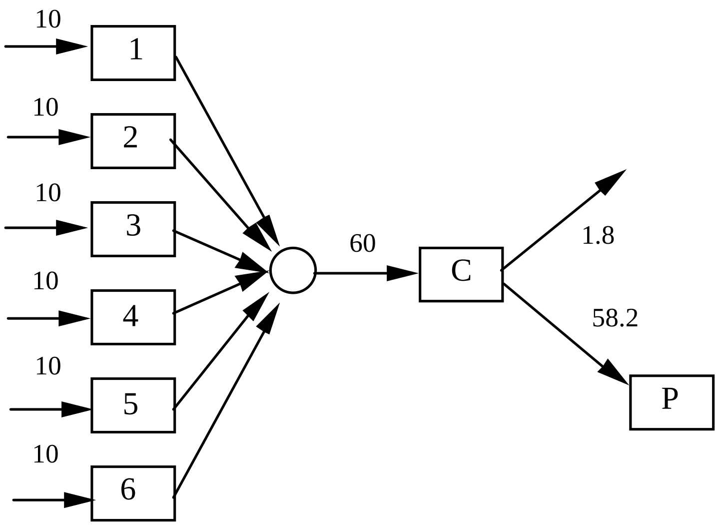

Figura 1. Unione e separazione di processi poissoniani.

2.1 Un sistema tandem

Prima di arrivare al teorema fondamentale sulle reti di code aperte, vediamo un semplice sistema, rappresentante il reparto assiematura di un sistema produttivo (Fig.2). In tale reparto, i componenti sono prelevati da un magazzino e posti in scatole (*kit*), dopodiché sono avviati all'assiematura (A), e successivamente al collaudo e imballaggio (I). Supponiamo che il processo degli arrivi alla stazione di assiematura (A) sia poissoniano. Sia A che I hanno tempi di lavorazione distribuiti esponenzialmente. Ciascuna stazione è costituita da un solo servente (ossia, $s_A = s_I$ = 1), ed è dotata in ingresso di un buffer di dimensioni illimitate.

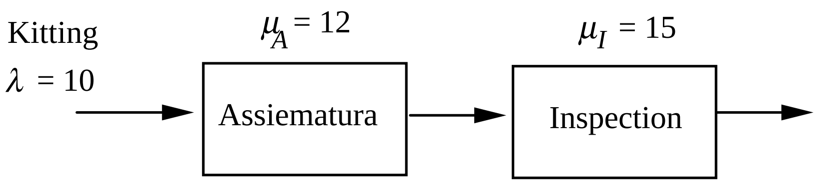

Figura 2. Un semplice sistema seriale.

Questo semplice sistema è dunque un sistema *seriale*, nel senso che l'uscita da una stazione costituisce l'ingresso nella successiva (tali sistemi si chiamano anche, secondo la terminologia delle reti di code, *sistemi tandem*). Siccome il processo degli arrivi a A è poissoniano, per il Teorema 3 lo è pure quello a I.

Lo *stato* di un sistema di servizio di tipo *M/M/s* può essere descritto come il numero di clienti presenti nel sistema a un certo istante. Com'è noto, se vale la condizione $\lambda/s\mu < 1$, esiste una *distribuzione stazionaria di probabilità*, ossia è possibile esprimere la probabilità P_i che nel sistema, a un generico istante, siano presenti esattamente i clienti. Tale probabilità è data da:

$$P_i = \frac{1}{i!} \left(\frac{\lambda}{\mu}\right)^i P_0 \qquad \text{per } i=s$$

$$P_i = \frac{1}{s!\, s^{i-s}} \left(\frac{\lambda}{\mu}\right)^i P_0 \qquad \text{per } i=s$$

dove P_0 è la probabilità che il sistema sia vuoto. Si ricorda che nel caso s=1, $P_0 = 1-\rho$. Passando ora a considerare un sistema, come il nostro, costituito da due stazioni *M/M/s*, appare naturale considerare come stato del sistema *il numero di clienti in ciascuna stazione*, ossia il vettore (n_A, n_I), ove n_A rappresenta il numero di clienti nella stazione di assiematura e n_I nella stazione di collaudo/imballaggio. Associamo allora a tale stato una *probabilità stazionaria* $P(n_A, n_I)$.

Vogliamo ora scrivere, per il nostro sistema, l'equazione di equilibrio ingresso/uscita per ciascuno stato. Da questo sistema di equazioni ricaveremo quindi i valori delle probabilità $P(n_A,n_I)$. Ricordando che ciascuna stazione costituisce un particolare processo di nascita e morte, possiamo prendere in considerazione, per ciascun n_j, la probabilità di passare, in un intervallo di tempo sufficientemente piccolo dt, in uno dei due stati "adiacenti", ovvero n_j+1 e n_j-1, dal momento che la probabilità di avere almeno due nascite o due morti è un infinitesimo di ordine superiore rispetto a dt. Ricordiamo invece che la probabilità di avere una nascita (morte) è proporzionale a dt. Similmente, se dt è sufficientemente piccolo, anche la probabilità di avere due nascite o due morti in due diverse stazioni del sistema può essere trascurata, e conseguentemente gli unici stati adiacenti a (n_1,n_2) che andremo a considerare sono quelli che differiscono da esso per *un solo cliente*, nel senso che verrà ora chiarito.

Supponiamo, per cominciare, di trovarci nello stato (0,0). Vogliamo esprimere la probabilità di *entrare* in tale stato nell'intervallo dt. L'unico stato dal quale è possibile entrare in (0,0) è evidentemente (0,1), e la transizione in (0,0) si ha se l'unico cliente presente in I esce dal sistema. Se ci troviamo in (0,1), ciò avviene con probabilità μ_I dt. Dunque, la probabilità che, in un intervallo di tempo dt, avvenga un ingresso nello stato (0,0), è pari a $\mu_I P(0,1) dt$. Esprimiamo ora la probabilità di *uscire* da (0,0). Ciò può avvenire solo grazie all'ingresso di un cliente nel sistema (con il che lo stato diventa (1,0)). La probabilità che ciò avvenga è $\lambda_A P(0,0) dt$. Perché il sistema ammetta una distribuzione di equilibrio, il valore atteso del numero di volte che si osserva un ingresso in qualunque stato nell'unità di tempo deve uguagliare il valore atteso del numero di volte che si osserva un'uscita da quello stato (altrimenti tale stato sarebbe assorbente, o, al contrario, transiente); è possibile allora scrivere l'equazione di equilibrio:

$$\lambda_A P(0,0) = \mu_I P(0,1)$$

Consideriamo ora la situazione che si ha quando vi sono n_A pezzi nella stazione di assiematura, e nessuno in quella di collaudo/imballaggio, ossia gli stati del tipo $(n_A,0)$. Stavolta, si può entrare in tale stato o perché vi erano n_A pezzi in A, uno in I e quest'ultimo ha terminato il suo servizio, oppure perché vi erano $n_A–1$ pezzi in A (e nessuno in I), e un nuovo pezzo è entrato in A. Da $(n_A,0)$, invece, si può uscire o per un nuovo arrivo nel sistema (e quindi in A), o perché un pezzo ha terminato la fase di assiematura e passa in I. Corrispondentemente, si ottiene la seguente equazione:

$$(\lambda_A + \mu_A) P(n_A,0) = \lambda_A P(n_A–1,0) + \mu_I P(n_A,1) \qquad \text{per ogni } n_A>0$$

ragionando analogamente per gli stati del tipo $(0,n_I)$:

$$(\lambda_A + \mu_I) P(0,n_I) = \mu_A P(1,n_I–1) + \mu_I P(0,n_I+1) \qquad \text{per ogni } n_I>0.$$

Per il generico stato (n_A,n_I), occorre considerare tutti e tre i tipi di eventi che possono verificarsi, ossia arrivi a A, trasferimenti da A a I, partenze da I. In definitiva si ottiene:

$$(\lambda_A+\mu_A+\mu_I) P(n_A,n_I) = \lambda_A P(n_A–1,n_I) + \mu_A P(n_A+1,n_I–1) + \mu_I P(n_A,n_I+1)$$
$$\text{per ogni } n_A>0, n_I>0.$$

Come si vede, per ogni stato (n_A,n_I) si ha un'equazione lineare. Per risolvere tale sistema di equazioni basta imporre (come per una singola stazione modellata come processo di nascita e morte), che la somma di tutte le probabilità sia pari a 1. Ciò posto, è possibile verificare facilmente per sostituzione che la soluzione a tale sistema è data da

$$P(n_A,n_I) = (1-\rho_A)\ \ \rho_A{}^{n_A}(1-\rho_I)\ \ \rho_I{}^{n_I}$$

laddove $\rho_A = \lambda_A / \mu_A$ e $\rho_I = \lambda_I / \mu_I$.

Ricordiamo che, per un sistema *M/M/*1 di parametri λ_A e μ_A, la probabilità di avere i clienti nel sistema, nell'ipotesi che $\lambda_A < \mu_A$, è data da $P_i = (1-\rho)\rho^i$. Osserviamo allora che la probabilità di avere n_A clienti nella prima stazione e n_I nella seconda, altri non è che il *prodotto* delle probabilità relative alle due stazioni, considerate come se fossero due sistemi *M/M/*1 *indipendenti*. Dal punto di vista probabilistico, quindi, le grandezze caratteristiche (N,W, W_q, L) di ciascuna stazione sono le stesse che si avrebbero analizzando ciascuna stazione separatamente. La struttura di questo tipo di soluzione prende il nome di *soluzione in forma prodotto*, e costituisce una caratteristica chiaramente desiderabile di una rete di code: infatti, il calcolo delle grandezze caratteristiche relative all'intera rete di M stazioni è ricondotto al calcolo di M grandezze, ciascuna relativa a una singola stazione. Si osservi che il valore di queste grandezze non dipende dall'ordine in cui sono visitate le varie stazioni.

<u>Esempio 2</u>. Con riferimento al sistema di Fig.2, si determini il valore atteso del numero di clienti nel sistema, e il tempo medio di attraversamento del sistema.
La frequenza media degli arrivi per ambedue le stazioni è $\lambda = 10$ kit/ora. Si noti che infatti essa non può essere diversa, dal momento che tutti i job che entrano in A passano poi in I, senza perdite né aggiunte. Le capacità di servizio siano invece $\mu_A=12$ kit/ora e $\mu_I = 15$ kit/ora. Anzitutto, osserviamo che $\lambda < \mu_A$ e $\lambda < \mu_I$; dunque ambedue le stazioni ammettono, separatamente, una distribuzione stazionaria di probabilità. Si ha: $\rho_A= 10/12$ e $\rho_I= 10/15$. Il valore atteso del numero di clienti nel sistema è dato evidentemente dalla somma del valore atteso del numero di clienti in A più quelli in I. Quindi: $W_A = 1/(\mu_A - \lambda_A) = 0.5$ ore e $W_I = 1/(\mu_I - \lambda_I) = 0.2$ ore, da cui $N_A= \lambda W_A = 5$ kit e $N_I= \lambda W_I = 2$ kit. Dunque, il valore atteso del numero di kit nel sistema è pari a 7. Anche il tempo di attraversamento è determinabile con estrema facilità, osservando che ogni kit visita ciascuna stazione esattamente una volta. Si ha perciò $W = W_A + W_I = 0.7$ ore.

_

2.2 Reti di code aperte generiche

L'esempio visto nel precedente paragrafo era un caso particolare di rete di code, nella fattispecie costituita da sue sole stazioni, in serie. Arriveremo però a mostrare che la possibilità di scrivere in forma prodotto la probabilità con cui il sistema si trova in un certo stato vale, con opportune generalizzazioni, per una classe molto più ampia di reti.

Per ora, supporremo che esista *un solo tipo di pezzo* che viene lavorato nel sistema (questa assunzione verrà in seguito rimossa). La classe di reti di code cui facciamo riferimento nel seguito presenta le seguenti caratteristiche:

i) Gli arrivi *dall'esterno* a ciascuna stazione j sono poissoniani con frequenza media λ_j;
ii) I tempi di servizio in ciascun servente della stazione j sono esponenziali di parametro μ_j;
iii) la disciplina di servizio è FCFS (primo arrivato, primo servito);
iv) un cliente che esce dalla stazione j ha probabilità p_{jk} di essere instradato alla stazione k e probabilità p_{j0} di uscire dal sistema ($p_{j0} = 1 - \sum_{k=1}^{M} p_{jk}$);
v) ogni stazione ha un buffer d'ingresso illimitato.

Si osservi che la classe di reti di code individuata da queste caratteristiche è abbastanza generale. In particolare, arrivi dall'esterno possono giungere a qualunque stazione, come pure clienti possono uscire dal sistema da qualunque stazione, e infine ci possono essere "ricircoli" di materiale tra diverse stazioni, come ad esempio avviene tipicamente allorché un pezzo che risulta difettoso al collaudo viene rispedito in lavorazione. Va sottolineato però che tutti i pezzi che arrivano a ciascun centro di lavorazione vengono trattati allo stesso modo — con tempo medio di lavorazione $1/\mu_j$ — indipendentemente dal fatto che essi visitino la stazione j per la prima volta o no. Inoltre, la validità modellistica di questa classe di reti di code è fortemente legata all'assunzione di tempi di servizio e arrivi esponenziali. Non di rado, e soprattutto negli FMS, i tempi di lavorazione possono considerarsi con ottima approssimazione deterministici. Ciò richiederebbe, per costruire la rete di code, l'uso di "mattoni" differenti dai sistemi *M/M/s*, e questo complicherebbe molto la trattazione: peraltro, studi sono stati effettuati proprio per definire delle opportune reti caratterizzate da distribuzioni esponenziali "equivalenti" a reti con distribuzioni più generali (Yao e Buzacott 1986). I risultati ottenuti sono stati molto buoni e comunque, a meno di casi molto particolari, la maggiore precisione che si avrebbe adottando modelli più aderenti alla realtà non è tale da compensare lo sforzo computazionale aggiuntivo. Un altro punto essenziale della nostra analisi riguarda la possibilità di esprimere effettivamente in forma probabilistica l'instradamento dei pezzi da una stazione all'altra: ciò, come vedremo, non è in genere una grave limitazione, e spesso i valori p_{jk} provengono dall'aggregazione di dati relativi a diversi tipi di pezzi presenti nel sistema.

2.3 Equazione di equilibrio

Vogliamo ora scrivere per una generica rete aperta (sotto le condizioni *i* – *v*) la stessa equazione di equilibrio che abbiamo scritto nel caso più particolare delle reti tandem. Sia $P(n_1, n_2,\ldots, n_M)$ la probabilità stazionaria che il sistema sia nello stato $\boldsymbol{n}$ + $(n_1, n_2,\ldots, n_M)$. Indichiamo ora con $\mathbf{n}_{jk}$ lo stato $(n_1,n_2,\ldots,n_j+1,\ldots,n_k-1,\ldots,n_M)$, vale a dire lo stato che differisce da $\boldsymbol{n}$ per il fatto di avere un pezzo in più in *j* e un pezzo in meno in *k*. Inoltre, $\mathbf{n}_{0k}$ e $\boldsymbol{n}_{j0}$ indicheranno rispettivamente gli stati $(n_1,n_2,\ldots,n_k-1,\ldots,n_M)$ e $(n_1,n_2,\ldots,n_j+1,\ldots,n_M)$. Infine, $r_j(\boldsymbol{n})$ è la frequenza dei servizi nella stazione *j* quando ci troviamo nello stato $\boldsymbol{n}=(n_1,n_2,\ldots,n_M)$. Tale valore è ovviamente dato da $\min(n_j,s_j)\,\mu_j$.

Se il sistema si trova già nello stato $\boldsymbol{n}$, esso ne può uscire per due motivi: o perché un nuovo pezzo arriva dall'esterno a una delle stazioni, o perché una stazione (tra quelle con almeno un cliente) termina un servizio, e dunque un pezzo lascia una stazione (diretto verso un'altra stazione o fuori del sistema). La probabilità che il sistema esca dallo stato $\boldsymbol{n}$ (nell'unità di tempo) è dunque data da

$$\sum_{j=1}^{M} \lambda_j P(\boldsymbol{n}) + \sum_{j\,:n_j=1} r_j(\boldsymbol{n})P(\boldsymbol{n})$$

Passiamo invece a considerare la probabilità di osservare un ingresso nello stato $\boldsymbol{n}$ a partire da un altro stato. Questo può accadere per tre motivi diversi: *i*) un nuovo cliente entra nel sistema dall'esterno, *ii*) un cliente esce dal sistema, oppure *iii*) un cliente esce da una stazione M_i e si trasferisce su M_j. Il contributo di un nuovo arrivo alla probabilità di entrare nello stato $\boldsymbol{n}$ è dato da:

$$\sum_{j\,:n_j=1} \lambda_j P(\boldsymbol{n}_{0j})$$

Nel secondo caso, ricordando che p_{j0} indica la probabilità che un pezzo in uscita da M_j esca dal sistema, il contributo è dato da:

$$\sum_{j=1}^{M} r_j(\boldsymbol{n}_{j0})\, p_{j0}\, P(\boldsymbol{n}_{j0})$$

e infine il termine associato al trasferimento di un pezzo da M_j a M_i è

$$\sum_{j=1}^{M} \sum_{i\,:n_i=1} p_{ji}\; r_j(\boldsymbol{n}_{ji})P(\boldsymbol{n}_{ji})$$

In definitiva, l'equazione di equilibrio per un generico stato è dunque:

$$\sum_{j=1}^{M} \lambda_j P(\boldsymbol{n}) + \sum_{j:n_j=1} r_j(\boldsymbol{n})P(\boldsymbol{n}) =$$

$$\sum_{j:n_j=1} \lambda_j P(\boldsymbol{n}_{0j}) + \sum_{j=1}^{M} p_{j0}\, r_j(\boldsymbol{n}_{j0})\, P(\boldsymbol{n}_{j0}) + \sum_{j=1}^{M} \sum_{i:n_i=1} p_{ji}\, r_j(\boldsymbol{n}_{ji})P(\boldsymbol{n}_{ji}) \qquad (1)$$

Vogliamo giungere a un'espressione in forma chiusa dei valori $P(n_1,n_2,\ldots,n_M)$ che risolvono tutte le (infinite) equazioni di equilibrio. Come vedremo, nelle ipotesi *i–v*, queste espressioni sono particolarmente semplici e si decompongono nel prodotto di termini relativi alle singole stazioni di servizio (soluzione in forma prodotto).

2.4 Frequenza effettiva degli arrivi

In una generica rete di code aperta, un pezzo può visitare una stazione j anche diverse volte, prima di uscire dal sistema. Anzi, visto che gli instradamenti sono caratterizzati in modo probabilistico, in generale (a differenza dell'esempio del sistema tandem) diversi pezzi visitano le stazioni un numero differente di volte. La stazione j, quindi, in generale "vede" arrivare sia alcuni pezzi nuovi e sia altri che già l'avevano visitata in precedenza. Di conseguenza, il valore della *frequenza media effettiva degli arrivi a* j può essere superiore al valore λ_j, che tiene conto solo degli arrivi dall'esterno. La frequenza media effettiva tiene invece conto anche dei clienti che arrivano a j da un'altra stazione. Se indichiamo con λ_j' tale frequenza media effettiva, essa è chiaramente data da:

$$\lambda_j' = \lambda_j + \sum_{k=1}^{M} \lambda_k' p_{kj} \qquad (2)$$

infatti, se a k arrivano in media λ_k' clienti, siccome una frazione pari a p_{kj} viene convogliata verso j, ecco che $\lambda_k' p_{kj}$ rappresenta proprio il contributo agli arrivi in j da parte della stazione k. Siccome le λ_j sono quantità note, le (2) costituiscono un sistema lineare di M equazioni in altrettante incognite, non omogeneo (se almeno un $\lambda_j > 0$, come supponiamo). Si noti che in generale p_{jj}?0.

Una condizione necessaria affinché la rete ammetta una soluzione finita è che la matrice dei coefficienti delle (2) sia non singolare. Tale condizione è senz'altro verificata in tutti i casi di interesse pratico: da un'analisi un po' più accurata della struttura delle (2), infatti, si può vedere che, se la matrice risultasse singolare, esisterebbe un sottoinsieme di stazioni nel quale i pezzi continuerebbero a ricircolare, senza possibilità di uscita — e dunque non si potrebbe, in questo caso, raggiungere una distribuzione stazionaria di probabilità.

Il fatto che la rete sia topologicamente "ben posta" (il che dipende unicamente dai valori p_{jk}) non è però sufficiente a garantire che possa esistere una distribuzione stazionaria di probabilità. Il valore di qualche λ_j', infatti, ancorché finito, potrebbe rivelarsi troppo elevato per la capacità della stazione j, espressa, come sappiamo, dalla quantità μ_j. Come vedremo tra poco, in perfetta analogia con i sistemi *M/M/s*, la condizione sotto la quale una rete di code ammette una distribuzione stazionaria di probabilità (nelle ipotesi *i*–*v*), è che la capacità produttiva di ciascuna stazione sia strettamente maggiore della frequenza effettiva degli arrivi a quella stazione.

2.5 Il teorema di Jackson

In questo paragrafo vedremo un'espressione in forma chiusa per risolvere le equazioni di equilibrio (1). A questo scopo, è necessario premettere il seguente risultato, che diamo senza dimostrazione:

Lemma 1. *Se esiste una distribuzione stazionaria di probabilità, allora per ogni stato* $\boldsymbol{n}$ + $(n_1,n_2,\ldots,n_M)$ *e per ogni macchina* j *tale che* n_j=1, *è soddisfatta la seguente uguaglianza:*

$$\lambda_j' \, P(\boldsymbol{n}_{0j}) = r_j(\boldsymbol{n}) \, P(\boldsymbol{n}) \tag{3}$$

Si noti che il Lemma 1 estende una proprietà che è ben nota nel caso dei singoli processi di nascita e morte. Infatti, le probabilità stazionarie $\{P_i\}$ di un processo di nascita e morte obbediscono alla nota relazione

$$\mu_i \; P_i = \lambda_{i-1} \, P_{i-1}$$

che può interpretarsi dicendo che, in condizioni di equilibrio, la probabilità di osservare, in una finestra temporale sufficientemente piccola, una transizione dallo stato i allo stato $i-1$ è uguale a quella di osservare una transizione dallo stato $i-1$ allo stato i. La (3) sancisce che ciò vale anche se il processo di nascita e morte che stiamo osservando è inserito in una rete di processi. In altre parole, l'interazione tra le varie stazioni può interamente riassumersi attraverso le frequenze effettive degli arrivi λ_i', note le quali l'evoluzione di una singola stazione è la stessa, sia che i pezzi gli pervengano dall'esterno, sia da altre stazioni. In definitiva, nel caso delle reti aperte, è possibile trattare separatamente le evoluzioni delle singole stazioni, con ovvi benefici dal punto di vista computazionale. A questo punto non deve quindi stupire il risultato espresso dal seguente teorema:

Teorema 4 (Jackson 1957). *Data una rete di code con M stazioni, per le quali valgono le ipotesi i– v, ponendo*

$$f_j^A(n_j) = \begin{cases} \dfrac{1}{n!}\left(\dfrac{\lambda'_j}{\mu_j}\right)^n f_j^A(0) & per\ n \le s_j \\ \dfrac{1}{s_j!\, s_j^{n-s_j}}\left(\dfrac{\lambda'_j}{\mu_j}\right)^n f_j^A(0) & per\ n \ge s_j \end{cases}$$

e $f_j^A(0)$ *è tale che, per ogni* j *si ha* $\sum_{n=0}^{8} f_j^A(n) = 1$*, e se per ogni stazione* j *vale la*

$$\lambda_j' < s_j\, \mu_j \tag{5}$$

allora esiste una distribuzione stazionaria di probabilità, e la probabilità che il sistema si trovi nello stato $(n_1,n_2,\dots,n_M)$ *è data da:*

$$P(n_1,n_2,\dots,n_M) = \prod_{j=1}^{M} f_j^A(n_j) \tag{6}$$

DIM. — Per semplicità, dimostreremo il risultato nel caso particolare in cui tutte le stazioni hanno un solo servente. In questo caso, se $n_j = 1$, $r_j(\boldsymbol{n}) = \mu_j$. Il coefficiente di utilizzazione della stazione j–esima, ρ_j, è pari al rapporto λ_j'/μ_j e infine $f_j^A(n_j) = \rho_j^{n_j} f_j^A(0)$. Inoltre, la (3) diventa

$$P(\boldsymbol{n}) = \rho_j\, P(\boldsymbol{n}_{0j})$$

Applicando ripetutamente n_j volte questa relazione, si ottiene

$$P(\boldsymbol{n}) = \rho_j^{n_j}\, P(n_1,n_2,\dots,n_{j-1},0,n_{j+1},\dots,n_M)$$

e iterando il procedimento per ogni stazione, si ha

$$P(n_1,n_2,\dots,n_M) = \prod_{j=1}^{M} \rho_j^{n_j}\, P(0,0,\dots,0)$$

per avere una soluzione completa delle (1), rimane solo da calcolare $P(0,0,\dots,0)$. A questo scopo, sfruttiamo la condizione che la somma delle probabilità di tutti gli stati deve essere pari a 1. Si ha dunque:

$$P(0,0,\ldots,0)=\frac{1}{\sum_{n}\prod_{j=1}^{M}\rho_j^{n_j}}$$

dove la sommatoria è estesa a tutti i possibili stati della rete. A questo punto è evidente che perché la soluzione esista, il denominatore deve assumere valore finito (in tal caso si può dimostrare anche che la soluzione è unica). In tal caso, è possibile allora invertire i simboli di produttoria e sommatoria, e scrivere

$$P(0,0,\ldots,0)=\frac{1}{\sum_{n}\prod_{j=1}^{\infty}\rho_j^{n_j}}=\prod_{j=1}^{M}\left(1-\rho_j\right)$$

da cui si ottiene quindi

$$P\left(n_1,n_2,\ldots,n_M\right)=\prod_{j=1}^{M}\rho_j^{n_j}\left(1-\rho_j\right)$$

vale a dire, ponendo $f_j^A(0) = (1 - \rho_j)$, la tesi.

L'utilità del teorema di Jackson risiede principalmente nel fatto che i termini $f_j^A(n_j)$ che figurano nell'espressione della probabilità $P(\boldsymbol{n})$ sono esattamente gli stessi che esprimono la probabilità di avere n_j clienti nella stazione M_j, presa singolarmente, quando il processo degli arrivi a questa stazione è descritto da una distribuzione esponenziale di parametro λ_j', e con capacità di servizio μ_j. La (6) costituisce cioè quella che (come già si era anticipato) va sotto il nome di *soluzione in forma prodotto*, in quanto la probabilità che il sistema si trovi in un certo stato $(n_1,n_2,\ldots,n_M)$ è esattamente pari al prodotto delle probabilità che le singole stazioni si trovino nei rispettivi stati n_j, *indipendentemente da ciò che avviene nelle altre stazioni*. Dunque, sotto le ipotesi del Teorema 4, è possibile effettuare lo studio delle reti di code aperte secondo una procedura di massima che possiamo riassumere nei seguenti tre punti:

1. Determinare le frequenze effettive degli arrivi alle stazioni λ_j' tramite le (2);
2. Analizzare ogni stazione indipendentemente e verificare che le (5) valgono;
3. Aggregare opportunamente i risultati per avere indicazioni sul sistema complessivo.

Essendoci finora soffermati sui primi due passi, vediamo a questo punto in che modo vanno aggregati i dati relativi alle singole stazioni. Il fatto che le stazioni si comportino come sistemi di servizio indipendenti facilita grandemente il problema di calcolare i valori dei parametri di maggiore interesse relativi al funzionamento globale del sistema. Tali parametri sono, innanzitutto, il valore atteso del numero di clienti nel sistema (N) e il valore atteso del tempo di permanenza nel sistema (W), ma notevole interesse può anche rivestire il tempo trascorso in media da un pezzo in una particolare stazione j — che, come ora vedremo, è in generale diverso da W_j.

Il valore N può essere banalmente calcolato sommando i valori N_j relativi alle singole stazioni. A questo punto, per avere il tempo medio di attraversamento del sistema basta applicare la legge di Little $W=N/\lambda$, laddove

$$\lambda = \sum_{j=1}^{M} \lambda_j$$

Supponiamo invece di voler conoscere il valore atteso del tempo trascorso in media da un pezzo nella stazione j. A tale scopo, osserviamo che $W_j = 1/(\mu_j - \lambda_j')$ rappresenta il valore atteso del tempo che un pezzo trascorre in *j ogni volta che deve essere processato* da j. Tale valore coincide con quello che stiamo cercando solo se il pezzo visita la stazione j una e una sola volta. In generale, tuttavia, un pezzo visita il centro j un numero diverso di volte. Indichiamo allora con v_j il valore atteso del numero di visite di un pezzo alla stazione j. Tale grandezza prende il nome di *visit count* (*numero di visite*). Si noti infatti che, analizzando le singole stazioni separatamente, non sappiamo se i λ_j' arrivi per unità di tempo alla stazione sono dovuti a pezzi sempre diversi (che giungono dall'esterno o da altre stazioni), oppure anche a pezzi che hanno già visitato la stazione in esame. Noti i valori v_j, il tempo trascorso nella stazione j dal generico pezzo è pari a $v_j W_j$.

Il calcolo dei visit count è estremamente semplice. Una frazione λ_j/λ dei pezzi provenienti dall'esterno visitano j come prima macchina, e dunque tale è la probabilità che la prima visita di un pezzo sia sulla stazione j. Inoltre, una frazione p_{ij} dei pezzi che visitano la stazione i vengono instradati a j e dunque, con un ragionamento del tutto analogo a quello svolto per le (2), si ha

$$v_j = \frac{\lambda_j}{\sum_{j=1}^{M} \lambda_j} + \sum_{i=1}^{M} p_{ij} v_i = \frac{\lambda_j}{\lambda} + \sum_{i=1}^{M} p_{ij} v_i$$

la cui soluzione è data dalle:

$$v_j = \frac{\lambda_j'}{\lambda} \tag{7}$$

Osserviamo infine che W può anche calcolarsi come:

$$\sum_{j=1}^{M} v_j W_j$$

<u>Esempio 3</u>. Riprendendo l'Esempio 2, supponiamo ora che il 10% dei pezzi collaudati risultino difettosi, e debbano essere quindi rinviati alla stazione di assiematura. Vogliamo vedere come viene influenzato il tempo di attraversamento. Innanzitutto determiniamo le frequenze effettive degli arrivi λ_A' e λ_I': Dalle (2):

$$\lambda_A' = \lambda_A + p_{IA}\,\lambda_I'$$
$$\lambda_I' = \lambda_A'$$

da cui, essendo p_{IA} =0.1 e λ_A=10 arrivi/ora,

$$\lambda_I' = \lambda_A' = \lambda\,/\,(1 - p_{IA}) = 11.11 \text{ arrivi/ora.}$$

A questo punto, si ottiene $\rho_A = \lambda_A' / \mu_A = 0.9259$ e $\rho_I = \lambda_I' / \mu_I = 0.7407$. Siccome ambedue i valori sono inferiori a 1, il sistema ammette una distribuzione stazionaria di probabilità e dunque si può applicare il teorema di Jackson. Dalle formule valide per i sistemi *M/M/1* si ottiene N_A=12.5 e N_I=2.85, da cui N=15.35 e W=1.53 ore. Il valore atteso del numero di visite alla stazione di assiematura e a quella di collaudo sono uguali, e pari a $v_A = v_I = 1.11$.

2.6 Throughput e work-in-process

Per un semplice sistema *M/M/1*, la relazione che intercorre tra il valore atteso del numero di clienti nel sistema (N) e il throughput (λ) assume l'espressione

$$N = \frac{\lambda}{\mu - \lambda}$$

da cui si osserva che, ad aumenti del work-in-process oltre un certo valore, corrispondono aumenti del throughput sempre meno significativi, tendendo asintoticamente al valore μ, che, come sappiamo, rappresenta infatti la capacità produttiva del sistema (Fig.3).

Per una generica rete di code aperta, il calcolo della relazione tra λ e N è in generale più laborioso, ma dà luogo, qualitativamente, allo stesso andamento. Il valore asintotico raggiungibile con un work-in-process infinitamente alto è limitato, evidentemente, dalla stazione di servizio meno efficiente di tutto il sistema, spesso chiamata *collo di bottiglia*. Tale valore μ_0, che dunque rappresenta la capacità produttiva dell'intera rete, è chiaramente dato da

$$\mu_0 = \min_j \left\{ \frac{s_j \, \mu_j}{v_j} \right\}$$

E' interessante osservare che lo stesso può dirsi per una rete chiusa, allorché però v_j va opportunamente definito, come vedremo più avanti.

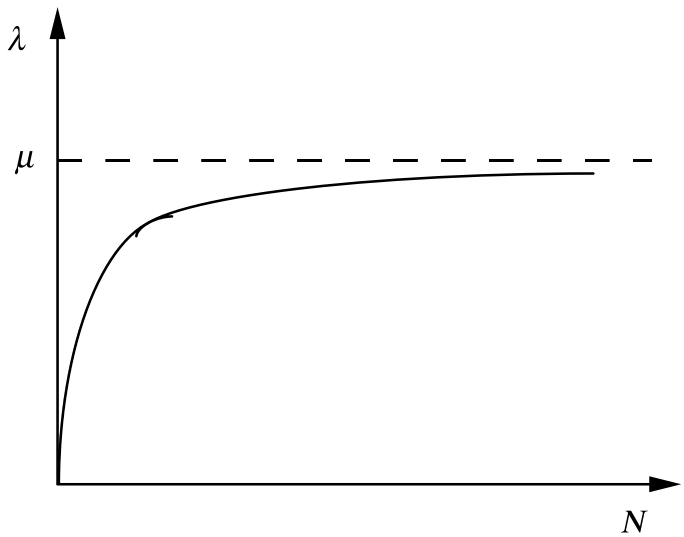

Figura 3. Relazione tra throughput λ e work-in-process N.

2.7 Diversi tipi di pezzi

Finora l'analisi è stata condotta con riferimento a un solo tipo di pezzo, che può eventualmente (con diverse probabilità) seguire diversi instradamenti, ma per tutti i pezzi che entrano una medesima stazione il tempo di servizio subìto ha la stessa distribuzione probabilistica. Vogliamo ora vedere come si può condurre il calcolo delle stesse grandezze analizzate finora, con riferimento però al caso in cui pezzi *di diverso tipo* coesistono nel sistema. Continueremo a supporre che, per ogni tipo di pezzo, i tempi di servizio siano distribuiti esponenzialmente, ma stavolta il valore atteso del tempo di servizio è diverso da tipo a tipo.

Una tecnica semplice per ricondursi al caso precedente, è quella di definire una sorta di *pezzo composito*, come media pesata dei pezzi che costituiscono il mix produttivo, analizzare poi le stazioni indipendentemente e infine aggregare i dati relativi a stazioni diverse, facendo però stavolta attenzione se si vogliono calcolare grandezze relative a pezzi di un certo tipo.

Si consideri dunque la situazione in cui nel sistema sono presenti pezzi di K tipi diversi, e sia r_{kj} la frazione complessiva di arrivi di tipo k alla stazione j. Si noti che r_{kj} tiene conto del fatto che uno stesso pezzo di tipo k può visitare j più volte. λ_{kj} è la frequenza degli arrivi dall'esterno di pezzi di tipo k alla stazione j. La frequenza media degli arrivi di tipo k all'intero sistema è indicata con λ^k ed è ovviamente pari a

$$\lambda^k = \sum_{j=1}^{M} \lambda_{kj}$$

Siano λ_{kj}' e μ_{kj} rispettivamente la frequenza effettiva degli arrivi di pezzi di tipo k alla stazione j e l'inverso del valore atteso del tempo di lavorazione della stazione j su un pezzo di tipo k. I valori λ_{kj}' si possono determinare semplicemente usando le (2) separatamente per ogni tipo di pezzo. Vogliamo ora determinare i valori di λ_j' e μ_j per ogni stazione. I primi si ottengono semplicemente come

$$\lambda_j' = \sum_{k=1}^{K} \lambda_{kj}'$$

mentre per i secondi abbiamo

$$\frac{1}{\mu_j} = \sum_{k=1}^{K} \frac{r_{kj}}{\mu_{kj}}$$

A questo punto, dopo aver verificato, al solito, che $\lambda_j' < s_j\, \mu_j$, potremmo calcolare, con le tecniche già viste, il valore di W per l'intero sistema — e quindi gli altri parametri. Siccome però sono presenti nel sistema pezzi di tipo diverso, può avere interesse calcolare il valore W_{kj}, che esprime il valore atteso del tempo di attraversamento per un cliente di tipo k in j, ogni volta che richiede un servizio. Stavolta va tenuto conto esplicitamente del fatto che ciascun k ha un tempo di servizio distribuito secondo il parametro μ_{kj}. Ogni volta che un pezzo di tipo k entra in j, trascorre mediamente nella sola coda (come tutti gli altri tipi di pezzi) un tempo pari a W_{qj}, mentre il *proprio* tempo medio di processamento è $1/\mu_{kj}$. Ricordando l'espressione di W_j in funzione di λ_j' e μ_j si ha allora che, ogni volta che un pezzo di tipo k visita la stazione j, esso trascorre in j un tempo di permanenza in media pari a:

$$W_{kj} = \frac{1}{\mu_j - \lambda'_j} - \frac{1}{\mu_j} + \frac{1}{\mu_{kj}}$$

Quindi, separatamente per ogni k e per ogni j, va calcolato il visit count v_{kj} (valore atteso del numero di visite di ciascun pezzo di tipo k a j). A tale scopo, si osservi che la frequenza effettiva degli arrivi di tipo k alla stazione j è data da $r_{kj}\lambda_j'$ e dunque, dalle (7), $v_{kj} = r_{kj}\lambda_j' / \lambda^k$. In definitiva, il valore atteso del tempo di attraversamento del sistema di un pezzo di tipo k è dato da:

$$W^k = \sum_{j=1}^{M} v_{kj} W_{kj}.$$

effettuando una somma pesata dei W^k, secondo le proporzioni relative λ^k/λ, ovviamente si riottiene W (si osservi che $v_j = \sum_{k=1}^{K} (\lambda^k/\lambda) v_{kj}$).

3. Reti chiuse

Nelle reti di code aperte non vi è teoricamente limite alla quantità di pezzi presenti, a un dato istante, nel sistema. I pezzi arrivano dall'esterno e vanno ad affollare i buffer delle varie stazioni, che abbiamo supposto in grado di ospitare una quantità comunque elevata di pezzi. Di fatto, molto spesso, questa non è la situazione che si verifica in pratica. Un valore elevato di work-in-process implica costi elevati: costi legati al capitale immobilizzato, allo spazio fisico occupato dai pezzi, alle strutture di movimentazione, ai problemi di supervisione che un sistema a elevato traffico pone. D'altro canto, l'approccio tradizionale alla massimizzazione della produttività consiste nell'utilizzare i serventi il più possibile, ossia nel minimizzare le attese dei centri di lavorazione; ma per fare in modo che ogni centro di lavorazione abbia sempre lavoro pronto da eseguire, ecco che i magazzini di semilavorati vengono a riempirsi. Solo negli anni '80, con l'avvento e la diffusione delle tecniche giapponesi di controllo della produzione (just-in-time) si è ribaltato il punto di vista, mirando piuttosto a tenere basso il work-in-process, e a minimizzare le attese dei pezzi, anziché delle macchine.

Appare di estremo interesse disporre di strumenti analitici per la valutazione delle prestazioni di una rete di code caratterizzata dal fatto che il numero di clienti in esso è costante nel tempo, ossia in cui il valore del work-in-process è costante ed è un dato del problema. Chiaramente, lo studio delle prestazioni del sistema al variare del work-in-process potrà suggerire fino a che punto aumentare il numero di clienti nel sistema è ancora utile ai fini della produttività, e quando invece l'aumento dei costi non è bilanciato da un corrispondente miglioramento del throughput.

Diciamo subito che, anche se nel modello analitico i "clienti" continuano a fluire nel sistema da una stazione di servizio a un'altra senza mai uscirne, nella realtà avvengono, ovviamente, ingressi e uscite dal sistema. In particolare, dal punto di vista operativo, non appena un pezzo esce dal sistema viene subito rimpiazzato da un nuovo pezzo entrante. Dal punto di vista modellistico, una delle stazioni di servizio rappresenterà l'ingresso e l'uscita di pezzi dal sistema.

Un caso in cui il numero di pezzi è costante si ha allorché il numero dei pallet è limitato e quindi l'ingresso di un nuovo pezzo può avvenire solo quando un pallet occupato si rende disponibile (per l'uscita di un pezzo finito). In sistemi di questo tipo è evidente l'utilità di un modello che aiuti a dimensionare opportunamente il numero di pallet.

3.1 Stati del sistema

Nel seguito, N indicherà il numero complessivo di pezzi nel sistema, sottolineando che questo non è più il valore atteso di una variabile aleatoria bensì un parametro di controllo. Come già fatto per le reti aperte, ci limiteremo qui a riportare e commentare alcuni risultati relativi ai casi più semplici, in cui valgono cioè ipotesi sostanzialmente uguali a quelle del teorema di Jackson. Supporremo dunque, in tutto il capitolo, che ciascuna stazione sia un sistema di servizio a s_j serventi, ciascuno con tempo di servizio esponenzialmente distribuito di parametro μ_j. Un pezzo uscente dalla stazione j ha probabilità p_{jk} di essere instradato alla stazione k. I buffer in ingresso a ciascuna stazione hanno capacità sufficientemente alta (ossia, possiamo supporre senza perdita di generalità, pari a N).

La maggiore complicazione concettuale delle reti chiuse risiede nel fatto che, chiaramente, non si può più sperare di ottenere una soluzione del problema in cui il comportamento delle diverse stazioni di lavoro sia completamente disaccoppiato (come espresso invece dal teorema di Jackson). Prendiamo ad esempio un sistema costituito da 2 stazioni, e si consideri la probabilità che nella prima stazione vi siano n_1 pezzi: dal momento che $n_1+n_2=N$, tale probabilità *coincide* con quella che al secondo centro i pezzi siano $N-n_1$, ed è quindi tutt'altro che indipendente da essa!

Cionondimeno, anche per le reti chiuse si può parlare di soluzione in forma prodotto, benché, per il motivo sopra citato, esisterà sempre un coefficiente che tiene conto della distribuzione complessiva di pezzi tra stazioni e che quindi tiene "legate" le evoluzioni delle singole stazioni. Nel seguito, per semplicità faremo riferimento a un solo tipo di pezzo presente nel sistema, ma quanto detto potrebbe facilmente estendersi al caso di più tipi di pezzi, passando attraverso la definizione di un pezzo composito, come già fatto per le reti aperte.

Il numero di stati di una rete chiusa in cui circolano N clienti è evidentemente finito. Esso è pari a tutti i modi possibili di distribuire N oggetti identici tra M contenitori, ovvero il numero delle M-partizioni di N elementi. Tale valore è dato da (Kleinrock 1975)

$$\binom{M+N-1}{M-1}$$

Esempio 4. In Fig.4 è rappresentata una rete con M=3 stazioni. Supponendo che N=4, dalla formula precedente abbiamo che il sistema ammette 6! / 2! 4! = 15 stati. Essi sono: (4,0,0); (3,1,0); (3,0,1); (2,2,0); (2,1,1); (2,0,2); (1,3,0); (1,2,1); (1,1,2); (1,0,3); (0,4,0); (0,3,1); (0,2,2); (0,1,3); (0,0,4).

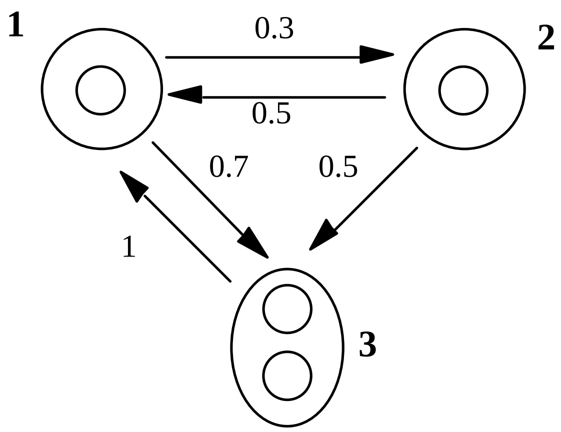

Figura 4. Una rete di code chiusa.

Una conseguenza importante del fatto che il numero di clienti è limitato è che, ovviamente, non esistono problemi di stabilità per una rete di code chiusa: di conseguenza, nelle stesse ipotesi *i*–*v* già viste nello scorso capitolo, una distribuzione stazionaria di probabilità esiste sempre.

3.2 Visit count

Come già per le reti aperte, anche ora il problema è quello di calcolare le probabilità (stazionarie) di ciascuno stato $\boldsymbol{n}$ + $(n_1,n_2,\ldots,n_M)$. Prima di fare ciò, è necessario definire, per ogni stazione j, cosa intendiamo adesso per valore atteso del numero di volte che ciascun pezzo visiterà la stazione j (che chiameremo ancora *visit count*).

Infatti, essendo la rete chiusa, occorrerà specificare *convenzionalmente* l'ingresso e l'uscita di un pezzo dal sistema.

Ragionando come nel caso di reti aperte, possiamo riscrivere le stesse equazioni (2), tenendo conto del fatto che stavolta non abbiamo arrivi dall'esterno, e quindi λ_j=0 per tutti i j. Si ottiene:

$$v_j = \sum_{k=1}^{M} p_{kj} v_k \tag{8}$$

A differenza delle (2), la matrice dei coefficienti delle (8) *è* singolare, dal momento che la somma degli elementi di ciascuna colonna dà 0 (non ci sono uscite dal sistema). Siccome anche il vettore dei termini noti è nullo, (almeno) una equazione è ridondante e dunque il sistema ammette infinite soluzioni, ovvero è possibile fissare *arbitrariamente* il valore di (almeno) una variabile e le altre risultano fissate di conseguenza — nel seguito, supporremo che il rango della matrice dei coefficienti sia M–1 e dunque *esattamente una* variabile possa essere fissata arbitrariamente. La quantità v_j che viene fissata assume dunque il significato di una variabile di riferimento, in quanto tutti gli altri valori v_i saranno *relativi* a tale v_j. In effetti, non poteva essere altrimenti, dal momento che, essendo la rete chiusa, non esiste un riferimento esterno assoluto, ma sarà invece solo possibile calcolare che, ad esempio, il numero medio di volte che il centro j viene visitato è il doppio del numero medio di volte che viene visitato un altro centro h. D'altro canto, se è possibile fissare uno di questi v_j pari al numero di volte che *realmente* un pezzo visita una particolare stazione, i valori v_j assumono un significato fisico ben preciso. A tale scopo, conviene riferirsi a quella stazione di servizio avente il significato di stazione di ingresso/uscita (sia essa la prima stazione), in quanto siamo *certi* che tale stazione (fittizia o meno) verrà visitata una e una sola volta da ciascun pezzo *fisico* entrante nel sistema. Fissando allora v_1=1, i valori v_j che risulteranno determinati di conseguenza indicheranno esattamente il valore atteso del numero di volte che la stazione j è visitata da ciascun pezzo.[1] Quindi, nel seguito potremo ancora pensare i v_j come visit

---

[1]Vale la pena di sottolineare che questa scelta non è comunque obbligatoria. Il fatto di optare per la stazione di carico e scarico è dovuto al fatto che generalmente, costruendo una rete di code chiusa per rappresentare un sistema reale, occorrerà sempre introdurre tale stazione e, allo stesso tempo, il numero di volte che essa viene visitata non dipende dalle probabilità di instradamento p_{jk} come accade, in generale, per le altre stazioni. Tuttavia, se ad esempio esiste un centro j che viene visitato con certezza esattamente q volte da ciascun pezzo, si può porre n_j=q.

count nello stesso senso visto per le reti aperte; tuttavia è importante ricordare che essi rappresentano, in generale, solo un indice relativo.

3.3 Equazione di equilibrio

Un parametro estremamente importante associato a ciascuna stazione in una rete di code chiusa è la quantità di tempo che, mediamente, un pezzo trascorre in servizio in un centro j. Tale quantità viene indicata con x_j ed è ovviamente legata a μ_j e al numero di volte che il centro j viene visitato: si ha cioè

$$x_j = \frac{v_j}{\mu_j} \tag{9}$$

Come già visto per le reti aperte, $r_j(\boldsymbol{n})$ e $P(\boldsymbol{n})$ indicano rispettivamente la frequenza dei servizi nella stazione j quando ci si trova nello stato $\boldsymbol{n}$ e la probabilità stazionaria che il sistema si trovi nello stato $\boldsymbol{n}$.

A questo punto, per ciascuno stato della rete, è possibile scrivere l'equazione di equilibrio ingresso-uscita, come già visto per le reti aperte. Infatti, considerando lo stato $\boldsymbol{n} = (n_1, n_2, \ldots, n_M)$, il valore atteso del numero di volte che si *esce* da tale stato nell'unità di tempo è dato da:

$$P(\boldsymbol{n}) \sum_{j=1}^{M} r_j(\boldsymbol{n})$$

Valutando il valore atteso del numero di volte che si *entra* nello stato $\boldsymbol{n}$ nell'unità di tempo, occorrerà considerare tutti gli stati "adiacenti" a $\boldsymbol{n}$, vale a dire appunto quelli che differiscono da $\boldsymbol{n}$ per un solo trasferimento di pezzo. Ricordando che un pezzo uscente da j ha probabilità p_{jk} di andare a k, si ha:

$$\sum_{k:n_k=1} \sum_{j=1}^{M} p_{jk} \, P(\boldsymbol{n}_{jk}) \, r_j(\boldsymbol{n}_{jk})$$

si noti che la sommatoria è estesa solo a quei centri k tali che, in $\boldsymbol{n}$, hanno almeno un pezzo. Si ottiene dunque, per lo stato $\boldsymbol{n}$:

$$P(\boldsymbol{n}) \sum_{j=1}^{M} r_j(\boldsymbol{n}) = \sum_{k:n_k=1} \sum_{j=1}^{M} p_{jk} \, P(\boldsymbol{n}_{jk}) \, r_j(\boldsymbol{n}_{jk}) \tag{10}$$

Esempio 5. Per il sistema dell'Esempio 4, si supponga che s_1=s_2=1, mentre s_3=2. Le probabilità di transizione sono riportate in Fig.4. Scriviamo l'equazione di equilibrio per lo stato (1,1,2):

$$P(1,1,2)\ (\mu_1+\mu_2+2\mu_3) = P(2,1,1)\ 0.7\ \mu_1 + P(2,0,2)\ 0.3\ \mu_1 + P(0,2,2)\ 0.5\ \mu_2 + P(1,2,1)\ 0.5\ \mu_2 + P(0,1,3)\ 2\mu_3$$

si noti infatti che r_3(0,1,3)=2 e che lo stato (1,0,3) non compare a secondo membro in quanto un pezzo uscente da M_3 viene sempre instradato verso M_1. _

A questo punto, possiamo enunciare il teorema che costituisce il corrispettivo del teorema di Jackson per reti di code chiuse (Gordon e Newell 1967):

Teorema 5. *Il sistema di equazioni (10) è risolto dalle seguenti espressioni:*

$$P(n) = \frac{1}{G(N)} \prod_{j=1}^{M} f_j(n_j) \tag{11}$$

dove

$$f_j(n_j) = \begin{cases} \dfrac{x_j^{n_j}}{n_j!} & per \quad n_j \le s_j \\ \dfrac{x_j^{n_j}}{s_j! s_j^{n_j - s_j}} & per \quad n_j \ge s_j \end{cases}$$

mentre

$$G(N) = \sum_{n} \prod_{j=1}^{M} f_j(n_j) \tag{12}$$

G(N) è una costante di normalizzazione *tale che la somma delle probabilità di tutti gli stati sia 1.*

DIM. — Anche qui, per semplicità, ci limitiamo a riportare la dimostrazione per il caso in cui tutti i centri siano monoservente, con il che $f_j(n_j) = x_j^{n_j}$ per tutti i j = 1,…, M. A differenza del teorema di Jackson, mostriamo che la tesi vale andando direttamente a sostituire l’espressione (11) nell’equazione di equilibrio (10) per un generico stato $\boldsymbol{n}$. Osserviamo che il primo membro della (10) può scriversi

$$\frac{1}{G(N)} \left(\sum_{i:n_i \ge 0} \mu_i \right) \prod_{j=1}^{M} x_j^{n_j} \tag{13}$$

mentre il secondo membro può scriversi

$$\frac{1}{G(N)}\sum_{k:n_k>0}\sum_{j=1}^{M} p_{jk}\mu_j\left(x_1^{n_1}x_2^{n_2}\dots x_j^{n_j+1}\dots x_k^{n_k-1}\dots x_M^{n_M}\right)=$$

$$=\frac{1}{G(N)}\sum_{k:n_k>0}\sum_{j=1}^{M} p_{jk}\mu_j\frac{x_j}{x_k}\prod_{j=1}^{M} x_j^{n_j}=$$

dalla definizione di x_j questa diventa

$$=\frac{1}{G(N)}\sum_{k:n_k>0}\mu_k\sum_{j=1}^{M} p_{jk}\frac{v_j}{v_k}\prod_{j=1}^{M} x_j^{n_j}$$

da cui, ricordando le (8), si vede che questa quantità coincide con la (13) e si ha dunque la tesi.

_

Peraltro, la soluzione data dalle (11)-(12) è unica (Gordon e Newell 1967). E' abbastanza evidente la somiglianza formale delle (11) con le corrispondenti espressioni viste per le reti aperte (6). In particolare, nell'espressione di $P(\boldsymbol{n})$ compare ancora il prodotto di M funzioni, ciascuna dipendente solo dal valore n_j dei clienti nella stazione j. L'interazione tra stazioni viene "riassunta" unicamente nel coefficiente di normalizzazione $G(N)$. Come dovrebbe risultare chiaro dalla dimostrazione del Teorema 5, il coefficiente $G(N)$ che compare in tutti i $P(n_1,n_2,\dots,n_M)$ non serve a garantire il soddisfacimento delle equazioni di equilibrio, bensì a fare in modo che la somma di tutte le probabilità sia 1. In questo senso, i prodotti delle $f_j(n_j)$ possono considerarsi delle "probabilità non normalizzate", dal momento che, divise per $G(N)$, danno appunto la probabilità di $\boldsymbol{n}$.

Si noti che $f_j(n_j)$ ha le dimensioni di un tempo alla n_j-esima potenza; siccome $P(\boldsymbol{n})$ è ovviamente un numero, si ha che la costante di normalizzazione ha le dimensioni di un tempo alla N-esima potenza.

3.4 Calcolo di G(N)

La soluzione del problema di analisi di una rete di code chiusa, nelle ipotesi viste, richiede dunque il calcolo della costante di normalizzazione $G(N)$. Dalla (12), osserviamo che $G(N)$ è la somma dei prodotti delle $f_j(n_j)$, per tutti gli stati $\boldsymbol{n}$ caratterizzati dal fatto di avere esattamente N pezzi, ossia per tutti gli stati possibili del sistema. Abbiamo cioè un prodotto per ognuno dei modi di assegnare N pezzi a M macchine. Come già osservato, tale numero, benché finito, può comunque risultare alquanto elevato, e dunque il calcolo di $G(N)$ per enumerazione totale degli stati è in generale improponibile, in quanto il tempo di calcolo cresce esponenzialmente in N.

E' possibile però calcolare $G(N)$ in modo più efficiente con un procedimento di programmazione dinamica, secondo quanto segue. Sia $G(j,n)$ la somma dei termini di prodotto relativi a tutti i modi possibili di assegnare n pezzi alle *prime j* stazioni, con $n=N$ e $j=M$. Quello che vogliamo determinare è $G(M,N)=G(N)$. Anzitutto abbiamo che $G(j,0)=1$ per ogni $j = 1,\ldots, M$ e $G(1,n)=f_1(n)$ per ogni $n = 1,\ldots, N$. In generale, possiamo esprimere $G(j,n)$ ragionando come segue: se nella stazione j-esima vi sono k pezzi, la somma dei termini di prodotto si ottiene moltiplicando $f_j(k)$ per il termine relativo a tutti i modi possibili di assegnare gli altri $n–k$ pezzi alle rimanenti $j–1$ macchine, ossia

$$f_j(k)G(j–1,n–k)$$

quindi, sommando per tutti i possibili k, otteniamo tutti i termini relativi ai modi di distribuire n pezzi tra le stazioni $1,2,\ldots,j$:

$$G(j,n) = \sum_{k=0}^{n} f_j(k)G(j-1,n-k) \tag{14}$$

la (14) dà un'espressione che consente il calcolo di $G(N)$ in $O(MN^2)$: infatti, ogni $G(j,n)$ può essere calcolato in $O(N)$, e chiaramente la matrice $G(j,n)$ consiste di $O(MN)$ termini che vanno calcolati per giungere a $G(N)$. Si noti che mentre nel calcolo di $G(j,n)$ è essenziale specificare la numerazione delle stazioni, il valore $G(M,N)$ è indipendente da esso.

Esiste un modo ancora più conveniente di calcolare $G(N)$, valido però solo se $s_j=1$ per tutti i j. Si consideri il coefficiente $G(M,N)$. Come visto, esso è la somma di tutti i termini di prodotto corrispondenti ai modi di assegnare gli N pezzi tra le M macchine. Ora, dividiamo in due gruppi questi addendi: nel primo collochiamo quelli corrispondenti a stati con $n_M = 0$, nel secondo quelli relativi a stati con $n_M=1$. Chiaramente, nei termini del primo gruppo non comparirà mai x_M, in quanto è elevato alla potenza 0, mentre x_M comparirà in tutti i termini del secondo gruppo. Osserviamo allora che i termini del primo gruppo — dal momento che la stazione M non è mai utilizzata — corrispondono a tutti i modi possibili di distribuire N pezzi tra le prime M-1 stazioni, e questo è pari a $G(M–1,N)$. Consideriamo ora i termini del secondo gruppo: tra essi, è possibile mettere in evidenza x_M; ciò che rimane, avendo tolto un pezzo, corrisponde a tutti i modi possibili di assegnare $N–1$ pezzi tra le M stazioni, e quindi in definitiva si ottiene:

$$G(M,N) = G(M-1,N) + x_M\, G(M,N-1) \qquad (15)$$

ricorsivamente è dunque possibile calcolare ciascun $G(j,n)$ come somma di due termini, e perciò $G(M,N)$ è ottenibile in $O(MN)$.

<u>Esempio 6</u>. Si consideri un sistema uguale a quello di Fig.4, tranne per il fatto che supponiamo s_j=1 per tutte e tre le stazioni, con il che $f_j(n_j) = x_j^{n_j}$. Supponiamo adesso N=3. Il numero totale degli stati è dunque pari a 10. Dalla definizione di G(3,3) si ha:

$$G(3,3) = x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3 + x_1 x_2 x_3 + x_1^2 x_3 + x_2^2 x_3 + x_1 x_3^2 + x_2 x_3^2 + x_3^3 =$$

$$= x_1^3 + x_1^2 x_2 + x_1 x_2^2 + x_2^3 + x_3\,(x_1 x_2 + x_1^2 + x_2^2 + x_1 x_3 + x_2 x_3 + x_3^2) =$$

$$G(2,3) + x_3\, G(3,2).$$

_

3.5 Calcolo delle prestazioni del sistema

Abbiamo già avuto modo di osservare che nelle reti di code chiuse N è una quantità nota, e non va dunque calcolata, come avveniva nelle reti aperte. Viceversa, una grandezza di estremo interesse che va calcolata è il throughput, che invece nelle reti aperte, in condizioni di stabilità, è pari alla frequenza complessiva degli arrivi al sistema. Richiamiamo l'attenzione sul fatto che il throughput, come del resto gli altri indici di valutazione che vedremo, sono noti una volta definiti i valori v_j, e dunque, da un punto di vista formale, vanno comunque sempre pensati come quantità relative.

Non vedremo qui una dimostrazione generale dell'espressione del throughput, per la quale si rimanda a Solberg (1981). Vogliamo però derivarla, pur non approfondendo tutti gli aspetti formali, nel caso in cui tutte le stazioni abbiano un solo servente. Consideriamo l'ultima stazione, ossia l'M-sima. La probabilità che essa sia vuota è pari alla somma delle probabilità di tutti gli stati con $n_M = 0$: dalle definizioni sopra introdotte, tale probabilità è pari a $G(M-1,N)$, diviso per la costante di normalizzazione $G(M,N)$. Quindi, la probabilità che la stazione M sia attiva, è pari a:

$$1 - \frac{G(M-1;N)}{G(M;N)}$$

ma dalla (15), osserviamo che questa quantità è pari a

$$x_M \, \frac{G(M;N-1)}{G(M;N)}$$

avendo supposto le stazioni monoservente, questa quantità esprime la frazione di tempo che M opera su qualche pezzo. Siccome ogni pezzo trascorre mediamente un tempo x_M sul centro M, ecco che il numero X di pezzi prodotti nell'unità di tempo è pari a:

$$X = \frac{G(M;N-1)}{G(M;N)}$$

Sottolineiamo che, benché l'espressione di X sia stata derivata solo nel caso monoservente, essa è di validità geenrale. Il calcolo del throughput X mostra come non solo il valore $G(M,N)$, ma anche i valori intermedi ottenuti durante il calcolo della tabella $G(j,n)$, siano di estremo interesse. Noto X, è possibile calcolare altre grandezze importanti. Innanzitutto, si può mostrare che il valore atteso del numero di serventi attivi alla stazione j è dato da

$$B_j = x_j\, X$$

il che è intuitivo, dal momento che $x_j\, X$ rappresenta il tempo che complessivamente la stazione j è attiva per lavorare i pezzi che vengono prodotti nell'unità di tempo. Inoltre, è anche possibile esprimere la frequenza media degli arrivi alla stazione j. Dal momento che vengono prodotti X pezzi per unità di tempo, e che ciascuno di questi pezzi arriva in media v_j volte alla stazione j, si ha:

$$\lambda_j = X\, v_j$$

L'uso del simbolo λ_j non dovrebbe trarre in inganno il lettore, tentato forse, a questo punto, di interpretare il fenomeno degli arrivi alla stazione j come un processo poissoniano di parametro λ_j. Attenzione: il processo degli arrivi a una stazione non è poissoniano — basta riflettere, ancora una volta, sul fatto che, essendo la rete chiusa, gli arrivi a una stazione sono tanto meno probabili quanti più clienti sono presenti nella stazione stessa.

Una quantità che può essere interessante calcolare è la probabilità che nella stazione M vi siano esattamente k pezzi. Essa è evidentemente pari alla somma delle probabilità di tutti gli stati caratterizzati dal fatto di avere $n_M = k$, e complessivamente $(N-k)$ pezzi tra tutte le altre $M-1$ stazioni. Ricordando il significato delle $G(j,n)$, possiamo scrivere:

$$Pr\,(n_M=k) = f_M(k)\, \frac{G(M-1;N-k)}{G(M;N)} \tag{16}$$

Si noti che se si fosse voluta conoscere la probabilità di avere k pezzi in un'altra stazione j, avremmo dovuto rinumerare le stazioni ponendo j come ultima stazione. La conoscenza di queste ultime probabilità rende possibile il calcolo del valore atteso del numero di pezzi in ciascuna stazione.

Se tutti i centri hanno un solo servente, si può usare un procedimento alternativo per il calcolo della probabilità che in ogni stazione vi sia un numero k di pezzi. Calcoliamo dapprima la probabilità che nella generica stazione j vi siano *almeno* k pezzi. A tale scopo, sommiamo le probabilità di tutti gli stati in cui k pezzi sono allocati alla macchina j, e i rimanenti $N - k$ sono distribuiti, in tutti i modi possibili, tra *tutte* le stazioni (compresa, nuovamente, la stazione j). Ovvero:

$$Pr\ (n_j = k) = x_j^k \frac{G(M;N-k)}{G(M;N)} \qquad (17)$$

a questo punto $Pr\ (n_j = k)$ può essere calcolata leggendo direttamente i valori dalla tabella $G(j, n)$. Infatti, basta osservare che

$$Pr\ (n_j = k) = Pr\ (n_j = k) - Pr\ (n_j = k + 1)$$

e applicare quindi la (17).

<u>Esempio 7</u>. Si consideri nuovamente il sistema dell'Esempio 5, con N=4 pezzi. Supponiamo che un'operazione alla stazione 1 rappresenti l'uscita di un pezzo dal sistema e l'ingresso di un nuovo pezzo; dunque v_1=1. Dalle (8), si ha v_2= 0.3 e v_3 = 0.85. A questo punto è necessario conoscere i valori μ_j alle varie stazioni, che sono: μ_1=2 pezzi/ora, μ_2=μ_3=1 pezzo/ora. Da questi valori possiamo calcolare x_1=0.5 ore; x_2=0.3 ore; x_3=0.85 ore. Possiamo allora, usando ricorsivamente le (14), riempire la seguente tabella:

n $\quad j$	1	2	3
0	1	1	1
1	0.5	0.8	1.65
2	0.25	0.49	1.5312
3	0.125	0.272	1.131
4	0.0625	0.1441	0.7403

si noti che $f_1(n) = x_1^n$ e $f_2(n) = x_2^n$, mentre $f_3(n)$ è pari a $x_3^n\ /n!$ se n=2 e a $x_3^n\ /2^{n-1}$ se n>2. Noti tutti gli elementi della tabella, possiamo calcolare il throughput:

$$X = \frac{G(3;3)}{G(3;4)} = \frac{1.131}{0.7403} = 1.528 \text{ pezzi/ora}$$

ci si può allora chiedere se l'eliminazione di un pallet, con conseguente diminuzione della complessità e dei costi di movimentazione, si tradurrebbe in un calo eccessivo di throughput. Per conoscere il valore di throughput con N=3, basta calcolare il rapporto

$$X' = \frac{G(3;2)}{G(3;3)} = \frac{1.5312}{1.131} = 1.354 \text{ pezzi/ora}$$

che può rappresentare o meno un calo contenuto rispetto a X. Tornando al caso N=4, vogliamo conoscere il valore atteso del numero di serventi utilizzati in ogni stazione (per i centri 1 e 2 questo coincide con la probabilità che il servente sia attivo). Avendo X, tale calcolo è immediato, e dà: $B_1=x_1X = 0.764$; $B_2=x_2X = 0.458$; $B_3=x_3X = 1.3$. Infine, vogliamo conoscere il valore atteso del numero di pezzi nella terza stazione. Dalle (16), abbiamo:

$$Pr\,(n_3=0) = f_3(0)\,\frac{G(2;4)}{G(3;4)} = 0.195$$

$$Pr\,(n_3=1) = f_3(1)\,\frac{G(2;3)}{G(3;4)} = 0.312$$

$$Pr\,(n_3=2) = f_3(2)\,\frac{G(2;2)}{G(3;4)} = 0.239$$

$$Pr\,(n_3=3) = f_3(3)\,\frac{G(2;1)}{G(3;4)} = 0.166$$

$$Pr\,(n_3=4) = f_3(4)\,\frac{G(2;0)}{G(3;4)} = 0.088$$

e quindi il valore atteso del numero di clienti nella terza stazione è pari a:

$$N_3 = \sum_{i=0}^{4} i\,Pr\,(n_3=i) = 0.312 + 2\,(0.239) + 3\,(0.166) + 4\,(0.088) = 1.64 \text{ pezzi.}$$

Per conoscere il valore atteso del tempo di attraversamento della terza stazione calcoliamo prima $\lambda_3 = v_3X = 0.85\ 1.528 = 1.3$ pezzi/ora. Si ha quindi, dalla legge di Little (che, lo ricordiamo, è valida sotto ipotesi del tutto generali): $W_3 = N_3 / \lambda_3 = 1.64 / 1.3 = 1.261$ ore.

Conclusioni

In queste poche pagine si è cercato di fornire alcune nozioni introduttive per lo studio delle reti di code, con particolare riferimento al loro uso nella rappresentazione dei sistemi manufatturieri. Va sempre tenuto presente che i risultati illustrati, ancorché abastanza potenti, sono comunque limitati, sostanzialmente, dalla validità delle ipotesi sotto le quali si può trovare una soluzione in forma di prodotto, sia nelle reti aperte che chiuse. Alcune di tali ipotesi riguardano la descrizione del fenomeno degli arrivi (nelle reti aperte) e dei servizi, che si suppongono esponenziali; altre sono per così dire "strutturali", in quanto suppongono che la disciplina di servizio sia rigidamente FIFO, e che gli instradamenti siano noti probabilisticamente. Inoltre, un pezzo che torni su un centro sul quale è già stato si comporta come un pezzo che utilizza quel centro per la prima volta. Non abbiamo parlato né di code con priorità, né di situazioni con tempi di servizio diversi (ad esempio deterministici). E' però rassicurante il fatto che è spesso possibile costruire un modello poissoniano "equivalente" a uno generale senza commettere errori eccessivi nella valutazione delle grandezze in gioco, e questa è la principale motivazione per affrontare l'introduzione allo studio delle reti di code nei modi visti in queste pagine.

Riferimenti

Askin, R.G., Standridge, C.R., 1993, *Modeling and analysis of manufacturing systems*, Wiley and Sons, New York.

De Julio, S., La Bella, A., 1975, *Elementi di teoria delle file d'attesa*, dispense a uso esclusivo degli studenti della facoltà di Ingegneria, Roma.

Gordon, W.J., G.F. Newell, 1967, Closed queueing systems with exponential servers, *Operations Research*, 15, 254–265.

Jackson, J.R., 1957, Networks of waiting lines, *Operations Research*, 5, 518–521.

Kleinrock, L., 1975, *Queueing systems*, Wiley and Sons, New York.

Solberg, J.J., 1981, Capacity planning with a stochastic flow model, *AIIE Transactions*, 13, 116–122.

Yao, D.D., Buzacott, J.A., 1986, The exponentialization approach to flexible manufacturing systems models with general processing times, *European Journal of Operational Research*, 24, 410–416.