

# Appunti Di Calcolo Parallelo

June 4, 2018

# Contents

<b>1</b>	<b>Grado di parallelismo</b>	<b>3</b>
1.1	Esempi di algoritmi paralleli . . . . .	3
1.1.1	Somma di due vettori . . . . .	3
1.1.2	Somma componenti di un vettore . . . . .	3
1.1.3	Moltiplicazione di matrici . . . . .	4
1.2	Computation DAG . . . . .	5
<b>2</b>	<b>Algoritmi paralleli</b>	<b>9</b>
2.1	Problema dell'ordinamento . . . . .	9
2.2	QUICK-SORT . . . . .	10
2.3	BITONIC SORTING . . . . .	10
2.4	Problema del routing . . . . .	15
2.5	Problema del calcolo dei prefissi . . . . .	18
2.6	Macchine a stati finiti . . . . .	20
2.6.1	Evoluzione di una FSM . . . . .	21
2.6.2	FSM per l'addizione binaria . . . . .	22
2.6.3	Sistemi di funzioni booleane . . . . .	23
2.7	Discrete Fourier Transform . . . . .	23
2.7.1	Algoritmo FFT . . . . .	24
<b>3</b>	<b>Modelli di rappresentazione per reti</b>	<b>25</b>
3.1	Multiprocessore distribuito . . . . .	25

# 1 Grado di parallelismo

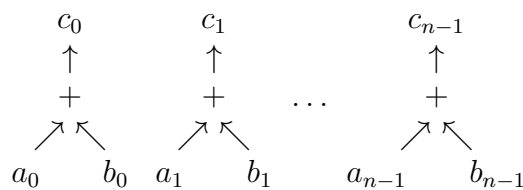
Il grado di parallelismo è il parallelismo potenziale, cioè un valore tale per cui se riscrivo l'algoritmo in modo opportuno posso esplicitare tale livello di parallelismo. Bisogna capire dunque quali operazioni possono essere svolte concorrentemente.

## 1.1 Esempi di algoritmi paralleli

### 1.1.1 Somma di due vettori

$$c = a + b$$

$$a, b, c \in \mathbb{R}^n$$



Tale algoritmo è parallelizzabile se considero  $a, b$  vettori con  $n$  elementi, allora con  $n$  addizionatori posso sommare gli  $n$  elementi ottenendo il vettore  $c$  con una sola operazione per ogni addizionatore.

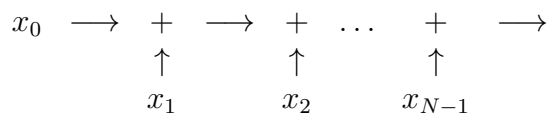
$P_{MAX} = N$  (il livello di parallelismo è  $N$ )

Cammino più lungo per arrivare da un input ad un output: si osserva il grafo che rappresenta le operazioni concorrenti per arrivare all'output e si identificano i cammini critici (nel caso precedente tutti i cammini hanno lunghezza  $L = 1$ ).

### 1.1.2 Somma componenti di un vettore

$$x = (x_0, x_1, \dots, x_{n-1})$$

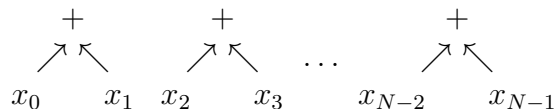
$$S = \sum_{j=0}^{N-1} x_j$$



```

S ← x0
for j ← 1 to n − 1 do
    S ← S + xj
end
    
```

Secondo questo schema l'algoritmo permette l'utilizzo concorrente di un numero di processori pari a  $P_{MAX} = 1$ , con lunghezza del cammino critico  $L = N - 1$ . Tuttavia il problema può essere affrontato in modi diversi, ad esempio la seguente soluzione permette di sfruttare  $P_{MAX} = N/2$  processori, con cammini critici di lunghezza  $L = \log(N)$ .



Il fatto che i due grafi precedenti relativi al problema della somma delle componenti del vettore producano lo stesso risultato è dovuto alla proprietà associativa dell'addizione. Tale proprietà vale per i numeri reali ma si deve considerare non applicabile a numeri floating point.

Ci possiamo fare un'altra domanda: se io utilizzo  $P_{MAX}$  processori riesco ad utilizzarli in modo efficiente? Nel caso del primo problema la risposta è sì. Supponendo che l'utilizzo dei processori abbia un costo per ogni operazione (si pagano anche i cicli inutilizzati se ho deciso di utilizzare tali processori), nel primo caso ogni processore svolge una operazione e il programma termina e dunque non ho uno spreco di risorse. Nel secondo caso invece, solamente al primo stadio sono richiesti tutti gli  $N/2$  processori, tuttavia ad ogni passo successivo la metà di questi rimangono inutilizzati.

Nel secondo caso dunque ho effettivamente accelerato il calcolo ma ho speso un costo eccessivo in quanto i processori inutilizzati potrebbero essere stati sfruttati altrove. Dunque se usiamo  $N/2$  processori otteniamo un tempo migliore ma spendiamo troppo. Se utilizziamo un processore non abbiamo sprechi ma i tempi si allungano. E' interessante allora cercare di capire come variano sprechi e tempi al variare del numero di processori.

Per  $1 \leq P \leq N/2$

Divido il vettore in blocchi di  $N/P$  sottovettori. Faccio poi sommare i blocchi di  $N/P$  ad un singolo processore seguendo lo schema sequenziale (idea di un tipo, poi il prof la deve impostare un po' perchè questa idea potrebbe non essere generalizzabile). Prendo le operazioni del primo stadio e le divido in gruppi da  $P$ , ottenendo il seguente schema:

### 1.1.3 Moltiplicazione di matrici

$C = AB$  con  $A, B$  matrici  $n \times n$

$$c_{ij} = \sum_{h=0}^{n-1} A_{ih} B_{hj}$$

Tale operazione è un insieme di  $n^2$  di prodotti interni, già questo evidenzia un grado  $N^2$  di parallelismo. A sua volta un prodotto interno consiste di  $N$  moltiplicazioni.

Senza entrare troppo nei dettagli, posso ottenere  $L = 1 + \log_2(N)$ . Ogni prodotto interno ha infatti lunghezza del cammino  $\log_2(N)$  e gli  $N^2$  distinti prodotti interni sono

indipendenti l'uno dall'altro e sono eseguibili concorrentemente contribuendo per un fattore 1 sulla lunghezza del cammino per generare ogni output. Tale valore di  $L$  è vero per  $P_{MAX} = N^3$ .

Si può fare a questo punto la stessa analisi di prima:

$$P_{MAX} \cdot T(P_{MAX}) \approx N^3 \cdot \log_2(N)$$

Per ridurre lo spreco si può utilizzare un numero pari a  $P = N^3 / \log_2(N)$  di processori, analiticamente si può suddividere la matrice in blocchi di dimensione  $P$ , ripetendo l'analisi svolta per l'esempio precedente. Questo algoritmo esibisce dunque un altissimo grado di parallelismo.

## 1.2 Computation DAG

Vogliamo ora generalizzare i concetti visti precedentemente.

**Definizione 1.1** (CDAG). Si definisce *Computation DAG* (CDAG) la quaterna

$$\mathcal{C} = (I, V, O, E)$$

dove:

- $I$ , insieme dei nodi di ingresso;
- $V$ , insieme dei nodi di operazione;
- $O$ , insieme dei nodi di uscita;
- $E \subseteq (I + V) \times V$ , insieme degli archi o *dipendenze funzionali*.

e valgono le seguenti relazioni:

- $O \subseteq (I + V)$ ;
- $I \cap V = \emptyset$ ;
- ogni vertice in  $I + V$  è su almeno un cammino  $a \longrightarrow b$  con  $a \in I$  e  $b \in O$ .

Un CDAG modella l'esecuzione di un algoritmo parallelo.

Un *planning* identifica il modo in cui viene gestito il flusso di un programma. Possiamo avere:

- planning *statico*, determinato in fase di compilazione;
- planning *dinamico*, migliore del planning statico e affidato alla macchina in fase di esecuzione (esempio, previsione dei salti condizionati).

Una volta definito un CDAG, bisogna determinare quanto veloce esso può essere eseguito con il numero di risorse fornite.

Ci si chiede inoltre quale sia il minimo numero di processori per poter risolvere il grafo in un determinato numero di passi, potenzialmente il numero minimo. È importante inoltre considerare quale sia un numero ragionevole di processori  $P^*$  per avere un valore di spreco entro il coefficiente 2 ed avvicinarsi il più possibile al numero minimo di passi di risoluzione. Per fare ciò, sono necessarie alcune ipotesi semplificative:

- ogni operazione viene eseguita in un ciclo;
- il risultato di tale operazione è accessibile al ciclo successivo.

In questo caso l'approccio *greedy* è ottimale. Infatti il modo migliore per eseguire le operazioni è la **greedy schedule**, ovvero: eseguire tutto ciò che è possibile eseguire in un dato istante.

**Definizione 1.2** (Greedy Schedule). Si definisce *greedy schedule* del CDAG  $\mathcal{C}$  la  $L$ -upla

$$(V_0, V_1, \dots, V_L)$$

tale che

- $V_0 = I$ ;
- $V_k = \{y \in V - (V_0 + V_1 + \dots + V_{k-1}) : \forall x(x, y) \in E \Rightarrow x \in (V_0 + V_1 + \dots + V_{k-1})\}$

Si inizia cercando di svolgere il calcolo più velocemente possibile: si considera il caso in cui tutti gli input siano disponibili, ci si pone allora il problema di quali operazioni sia possibile svolgere su di essi. Tali operazioni sono rappresentate da  $V_1$ .  $V_2$  è invece l'insieme delle operazioni che possono essere eseguite se gli operandi sono in  $V_0$  e in  $V_1$  ma almeno un operando dev'essere in  $V_1$ , altrimenti  $V_2$  sarebbe stata eseguita prima di  $V_1$ . Questa proprietà vale per ogni  $V_i$  e  $V_{i+1}$ . Arriverà ad un certo punto a  $V_L$  dove il processo termina in quanto non vi sono più operazioni i cui operandi sono in stadi già sistemati.

$$P_1 \rightarrow \sum_{j=0}^L V_j = I + V$$

Siccome il grafo non ha cicli, i cammini sono tutti di lunghezza finita e dunque esiste una lunghezza massima denotata con  $L$ . Oggi si è definita una procedura per scomporre il grafo in pezzi, senza considerare gli input tali pezzi sono  $L$ .

**Proposizione 1.** Se  $(V_0, V_1, \dots, V_L)$  è la *Greedy Schedule* di  $\mathcal{C}$  allora  $L$  è la massima lunghezza di un cammino di  $\mathcal{C}$ .

**Proof:** Sia una quantità  $L$  definita attraverso la Greedy Schedule e si definisce un'altra quantità  $M$  come la lunghezza massima di un cammino di  $\mathcal{C}$ , si verifica che tali grandezze coincidono:  $M \geq L$ . Se io costruisco un cammino di lunghezza  $L$ , dato che  $M$  è il cammino di lunghezza massima allora questa relazione è necessariamente vera. Dalla

Greedy Schedule si ricava che se un nodo appartiene allo stadio  $i$ , esso ha almeno un predecessore nello stadio  $i - 1$ . Di conseguenza esiste dallo stadio  $L$  un cammino di lunghezza  $L$  e termina allo stadio  $V_0$  (per costruzione).

Per dimostrare formalmente la relazione: si costruisce un cammino  $\gamma$  partendo da un vertice  $v_L \in L$  e scegliendo (per  $i = L - 1, \dots, 0$ ),  $v_i \in V_i$  tale che  $(v_i, v_{i+1}) \in E$ . Ovviamente per la Greedy Schedule tale cammino esiste, e  $M \geq \|\gamma\| = L$

$*M \leq L$

Osservazione: se  $[(a, b) \in E, a \in V_i, b \in V_j]$  allora  $j > i$ , cioè due nodi diversi dello stesso cammino devono essere collocati su stadi successivi della Greedy Schedule. Dunque se ho un cammino di lunghezza  $M$  allora devono esistere almeno  $M$  stadi differenti successivi, dunque  $L = M$ . Formalmente se  $\gamma = (u_0, u_1, \dots, u_k)$  è un cammino di  $C$ , allora  $L \geq K$ , scegliendo  $k = M$  ottengo la relazione cercata.

Dunque fra tutti i cammini che arrivano in  $V_i$ , il più lungo ha lunghezza  $i$ . Si sta comunque assumendo l'idealità della macchina, cioè una macchina tale per cui nel momento in cui i valori sono pronti per l'esecuzione dell'operazione successiva, essa viene immediatamente eseguita. Per ogni stadio sono necessari un numero di processori pari alla cardinalità dello stadio stesso, secondo la Greedy Schedule è dunque necessario avere un numero di processori  $P_{MAX} = \max V_j$ , per  $j \leftarrow 1$  a  $L$ . Potrei tuttavia riuscire a percorrere il cammino in tempo  $L$  con un numero di processori  $P < P_{MAX}$ . Sicuramente non è possibile migliorare il tempo  $L$  essendo esso cammino critico per  $P_{MAX}$ . Il problema di determinare il numero minimo di processori per svolgere il cammino critico in tempo  $L$  è NP-Hard. Ci sono problemi dove trovare una soluzione quando esiste un solo tipo di operazione è un problema polinomiale, mentre se vi sono più operazioni tale problema è NP-Hard. A noi interessa calcolare queste quantità  $L$  o  $P_{MAX}$  o ci interessa solo una definizione? Dipende dal contesto, in generale la lunghezza del cammino critico è funzione dell'input e dunque l'analisi dev'essere effettuata in maniera matematica piuttosto che algoritmica. Così come è interessante analizzare il tempo e lo spazio di esecuzione di un algoritmo sequenziale, ci occupiamo di analizzare algoritmi paralleli. Tuttavia questo problema dev'essere trattato anche da algoritmi di compilazione poiché essi devono esporre un certo livello di parallelismo interno al codice per renderlo più efficiente. Il passo successivo che si vuole fare è: quanto efficientemente andremmo ad usare  $P_{MAX}$  processori? Analizzando i  $V_i$  della Greedy Schedule si vede che se ad ogni stadio è presente un numero differente di elementi, allora vi è sempre uno spreco. Come precedentemente affermato, una buona misura del grado di parallelismo si ottiene con:

$$P^* = \frac{|V|}{L}$$

Si dimostrerà che tale numero  $P^*$  rallenterà la Greedy Schedule di un fattore al più 2, ottenendo dunque uno spreco al più pari a 2.

Condizione necessaria e sufficiente affinché un grafo  $C$  sia una schedule per un solo processore è che devono essere rispettate le dipendenze topologiche del grafo, ossia semplicemente se un nodo  $b$  ha un arco entrante da  $a$ , allora l'attività  $b$  dev'essere eseguita dopo  $a$ . La Greedy Schedule non individua un unico ordinamento topologico, però indica le dipendenze tra i nodi. Il numero dei possibili ordinamenti topologici che si possono

ricavare cresce esponenzialmente con il numero di nodi. Per esempio nel caso di due catene indipendenti gli ordinamenti possibili per generare un'unica catena ammissibile che rispetta gli ordinamenti è  $\binom{2n}{n}$ .

Questo valore si ricava dal fatto che per identificare la catena finale è sufficiente indicare quali posizioni sono occupate da nodi  $a$  e quali da nodi  $b$ , poichè l'ordinamento tra i nodi  $a$  e  $b$  è univoco.

A causa del numero esponenziale di differenti ordinamenti che si possono generare il problema di trovare la schedule ottima è computazionalmente difficile. L'ordinamento topologico in sè può essere ricavato in tempo lineare.

Definisco  $P_{MAX} = \max_{i=1}^L |V_i|$  come la cardinalità dello stadio della greedy schedule che ha il maggior numero di nodi.

**Definizione 1.3.**  $T(P)$  è il tempo minimo per eseguire il cdag  $C$  con  $P$  processori.

Nei casi  $P = 1$  e  $P = P_{MAX}$  è facile capire quanto vale mentre per gli altri valori si possono ottenere delle stime superiori o inferiori.

- $T(P) \geq L$
- $T(P) \geq \frac{|V|}{P}$
- $T(P) \geq \max\left(L, \frac{|V|}{P}\right) \geq \frac{1}{2} \left(\frac{|V|}{P} + L\right)$
- $T(P) \leq T_{GREEDY|P} = \sum_{j=1}^L \lceil \frac{|V_j|}{P} \rceil \leq \sum_{j=1}^L \left(\frac{|V_j|}{P} + 1\right) = \frac{|V|}{P} + L$

Non si può essere certi dell'ottimalità della Greedy Schedule, tuttavia essa costituisce un upper bound essendo una schedule valida. Unendo i lower e gli upper bound precedentemente ricavati si ottiene la formula:

$$\frac{1}{2} \left(\frac{|V|}{P} + L\right) \leq T(P) < \frac{|V|}{P} + L$$

E' conveniente definire  $P^* = \frac{|V|}{L}$  (in generale minore di  $P_{MAX}$  anche se ci sono casi in cui  $P^* = P_{MAX}$ ).

- $T_{GREEDY|P}(P^*) < \frac{|V|}{P^*} + L = 2L$
- $T_{GREEDY|P}(P^*)P^* < 2L \frac{|V|}{P^*} = 2|V|$  (numero massimo di operazioni farre con  $P^*$  processori in un tempo  $T_{GREEDY|P}(P^*)$ )

Quest'analisi permette di ridurre gli sprechi massimizzando il guadagno con un numero ragionevole di processori sebbene non permetta di capire in tempo reale (mentre i processori selezionati lavorano) come ridurre gli sprechi.



## 2 Algoritmi paralleli

Nelle ultime lezioni abbiamo definito il modello di calcolo *computation dag* definendo il parallelismo che è possibile individuare in un calcolo. Dobbiamo ora definire la relazione tra un algoritmo e il suo cdag. Un algoritmo è un qualcosa che si può vedere come una scatola nera che in funzione di input restituisce un output. Bisogna tenere presente che un cdag associato ad un algoritmo è dipendente dell'input. Questa non è un'eccezione nell'analisi degli algoritmi poichè anche la più tipica analisi worst-case è basata sull'input.

Gli algoritmi presentati in precedenza hanno una particolarità, una volta fissata la taglia essi generano lo stesso grafo indipendentemente dell'input. Ci sono dunque dei casi significativi di problemi interessanti che hanno questa proprietà, cioè che il grafo di calcolo dipende solo della taglia  $n$  ma non dell'input specifico. Per questi problemi il cdag è particolarmente utile come strumento di analisi. Se ho un algoritmo scritto in maniera sequenziale, l'applicazione del modello del cdag su tale algoritmo mi permette di capirne il livello di parallelismo. Ad oggi il campo degli algoritmi paralleli è molto vasto e tutti i problemi studiati sequenzialmente sono stati ristudiati in modo parallelo.

### 2.1 Problema dell'ordinamento

Andiamo a vedere alcuni algoritmi che conosciamo per analizzarli dal punto di vista del parallelismo. Sotto opportune ipotesi l'ordinamento richiede  $n \log(n)$  confronti e ci sono algoritmi come MERGE-SORT che si avvicinano molto a tale lower bound.

Dal punto di vista sequenziale, il merge si può eseguire in tempo lineare e da luogo a  $n \log(n)$  confronti, ma poichè le due chiamate di SORT lavorano a dati completamente scorrelati e indipendenti, se avessi due processori potrei eseguirle contemporaneamente. Inoltre sviluppando le chiamate ricorsive il parallelismo aumenta ulteriormente. Prima di definire il parallelismo complessivo bisogna osservare il comportamento di merge che comporta problemi in quanto non è molto parallelo. Il MERGE-SORT classico sequenziale ha le seguenti proprietà:

- **Analisi temporale:**  $T(N) = 2T(\frac{N}{2}) + T_M$ .
- **Numero di operazioni:**  $V_S(N) \leq 2V_S(\frac{N}{2}) + V_M(N)$ . Rimane da analizzare il merge e sostituire le funzioni corrette per esso, inoltre si utilizza il segno minore o uguale poichè si prende in considerazione il caso peggiore.
- **Cammino critico:**  $L_S(N) \leq L_S(\frac{N}{2}) + L_M$ . Tale ricorrenza corrisponde al percorrimiento di uno dei rami dell'albero delle ricorsioni per determinarne la lunghezza. Non posso utilizzare l'uguaglianza ma sono costretto usare una maggiorazione per tenere in considerazione tutti i casi in cui il worst-case non avvenga in tutti i blocchi.

*Osservazione:* Supponiamo che un grafo sia suddiviso in più parti, come nel caso di MERGE-SORT dove possiamo distinguere due diverse funzioni, SORT e MERGE, è possibile che i cammini più lunghi della prima metà del grafo si raccordino con cammini più corti della seconda metà del grafo? Quindi nessuno dei cammini è uguale come lunghezza alla somma dei cammini critici. Devo dunque cercare un cammino che sia il

più lungo possibile in tutti i blocchi ma è possibile che tale cammino non si realizzi mai nell'algoritmo.

*Esempio:* Quando tutti i cammini indipendentemente dagli input hanno la stessa lunghezza allora posso ottenere un risultato senza ricorrere alla maggiorazione.

Tornando al MERGE-SORT, se usiamo il MERGE "tradizionale",  $V_M(N) \leq N$ ,  $V_S(1) = 0 \Rightarrow V(N) \leq N \log_2(N)$ . Cosa si può dire invece dell' $L_N$ ? Ogni operazione dipende da un'operazione precedente e dunque per nel cdag dovremmo aggiungere dipendenze aggiuntive rispetto a quelle funzionali, dipendenze di controllo e di indirizzo. Allora  $L_M(N) \leq N$ ,  $L_S(1) = 0 \Rightarrow L_S(N) = \Theta(N)$ . Dunque il MERGE-SORT con MERGE classico ha un parallelismo:

$$P_S^* = \frac{V_S(N)}{L_S(N)} \leq \frac{O(N \log(N))}{\Omega(N)} \leq O(\log(N))$$

## 2.2 QUICK-SORT

C'è una differenza sostanziale tra i due schemi MERGE-SORT e QUICK-SORT, le taglie di  $x_0$  e  $x_1$  dopo la fase di pivoting sono molto probabilmente differenti, al punto che nel caso peggiore riesco a ridurre la taglia del problema ad ogni iterazione di 1, portando il tempo nel worst-case a  $O(N^2)$ , con profondità dell'albero della ricorsione pari a  $N$ . Il QUICK-SORT è uno dei casi in cui l'analisi al caso peggiore non è molto affidabile, poichè il worst-case è molto peggiore del caso medio ma è estremamente raro. Tale algoritmo infatti risulta tipicamente più veloce del MERGE-SORT nel caso medio, anche se analiticamente questa differenza non si vede. Con una probabilità  $\geq 50\%$  in due passi ho almeno dimezzato la taglia del problema, e dunque anche se la profondità non è costante, la profondità massima sarà  $k \log(N)$  per una qualche costante  $k$ . Per aumentare la probabilità di bilanciamento della suddivisione si possono prendere 3 elementi e scegliere quello intermedio come pivot, ottenendo così una riduzione della profondità dell'albero facendo qualche confronto in più.

$$\begin{aligned}\tilde{V}_Q(N) &= \Theta(N \log_2(N)) \\ L(N) &\leq O(\log(N)) + \alpha L(\beta N) = O(\log^2(N)), \quad \beta > \frac{1}{2} \\ \tilde{P}_S^* &= \Theta\left(\frac{N}{\log_2(N)}\right)\end{aligned}$$

Quegli elementi che rendevano il QUICK-SORT sequenziale più veloce del MERGE-SORT tendono ad attenuarsi nel calcolo parallelo. Tale algoritmo è comunque spesso utilizzato in modo parallelo ma esistono altri algoritmi di sorting che sfruttano altre proprietà per ottenere un parallelismo migliore.

## 2.3 BITONIC SORTING

Questo algoritmo è una forma di MERGE-SORT proposto da Kenneth E. Batcher nel 1964 e pubblicato nel 1968. In quegli anni si progettavano algoritmi booleani per realizzare algoritmi cablati e si stimava che circa un quarto del tempo macchina veniva

utilizzato per operazioni di ordinamento. Si cercava dunque di realizzare un circuito cablato per eseguire ordinamenti con l'idea che prima o poi tale circuito sarebbe diventato funzionale sui calcolatori. In realtà questa idea non ha mai visto la luce, i processori odierni non possiedono unità funzionali per svolgere l'ordinamento perché probabilmente non ha la stessa importanza che aveva ai tempi e inoltre perché il circuito ha una taglia fissa e ridotta. Non è facile, tenendo presente questo, realizzare algoritmi veloci che utilizzano come operazione primitiva il confronto tra più elementi (ipotizzando la presenza di un circuito in grado di fare questo) invece che il singolo confronto.

*Osservazione:* I processori odierni:

$\tau = 10 \text{ ps} = 10^{-11} \text{ s}$  (tempo di commutazione di un transistor odierno)

$c = 3 \cdot 10^8 \text{ ms}^{-1}$

$c\tau = 3 \cdot 10^8 \cdot 10^{-11} = 3 \text{ mm}$

Più o meno un chip ha un lato di  $3 \text{ cm}$ , dunque i transistor commutano fino a 10 porte logiche nel tempo che il segnale va da una parte all'altra del chip. Perché non si fanno i processori più grandi? Per una questione di difetti, essi infatti contengono con probabilità significativa dei difetti, probabilità dovuta alle tecniche di produzione distribuita su ogni unità di area. Dunque più grande è il chip, più è probabile che sia difettoso, e la probabilità che non vi siano difetti decresce esponenzialmente in funzione dell'area. Oggi come oggi, quando viene progettato e realizzato un nuovo chip, solamente un 5% riesce senza difetti e non viene buttato. A partire da questo risultato i processi produttivi vengono raffinati e ottimizzati per tale specifico chip per ottenere una percentuale di yield migliore. Dove manca il parallelismo di MERGE-SORT è nell'operazione di MERGE, bisogna dunque trovare un modo più efficiente di svolgere tale operazione. Supponiamo di avere due sequenze  $A$  e  $B$ , non necessariamente della stessa lunghezza, e di voler fare il MERGE. Prendo  $B$ , la rovescio e la concateno con  $A$ , in questo modo utilizzo due processori, uno che parte dai massimi (più o meno dal centro della stringa), e uno dai minimi e li ordina.

Questa idea può essere riproposta per un numero di processori  $> 2$ ? Intuitivamente posso suddividere ogni stringa a metà e su ognuna di questa metà applico due processori come visto precedentemente ma questa operazione non è cenetta. Devo trovare una suddivisione delle due sequenze in problemi tali per cui valgono le proprietà:

- $\max(A_1, B_1) < \min(A_2, B_2)$
- $|A_1| + |B_1| \cong |A_2| + |B_2|$

IDEA: Si cerca l'elemento mediano della sequenza  $A$ , tale operazione si può svolgere in tempo costante e si cerca tale elemento, o un elemento con valore immediatamente successivo o precedente nella sequenza  $B$ . Con questo metodo la stringa  $B$  è molto sbilanciata e l'obiettivo dell'equipartizione non è risolto. Si cerca dunque di migliorare il bilanciamento delle sequenze rendendo più efficiente questa idea invece di riformulare un'altra idea da capo.

Prendiamo allora sia la mediana di  $A$  che la mediana di  $B$ . Naturalmente bisogna osservare che la prima proprietà non è soddisfatta. Se  $m_A > m_B$  lo split dev'essere modificato spostando l'indice in  $A$  verso sinistra e l'indice in  $B$  verso destra, se  $m_A < m_B$  l'operazione è analoga ma contraria. Di quanto vanno spostati questi indici? Sfrutto la

ricerca binaria per trovare gli indici adatti. Se  $A$  e  $B$  hanno la stessa lunghezza il risultato sarà bilanciato. Tuttavia non essendo  $A$  e  $B$  della stessa lunghezza, la ricerca binaria procede dividendo a metà solamente la stringa più corta e spostando l'indice sulla più lunga della stessa quantità.

Tuttavia questa ricerca binaria ha un problema, è un algoritmo sequenziale e inoltre il risultato de confronto ci dice dove andare a prendere i dati, considerazione dipendente dell'input. Questo causa problemi visto che Batchner aveva in mente di realizzare un circuito. Infatti gli algoritmi che più si apprestano ad essere realizzati con circuiti cablati sono quelli che presentano lo stesso cdag per ogni input. L'idea di Batchner per eliminare quest'incertezza sull'esito dei confronti e la seguente: invece di svolgere una ricerca binaria intelligente si svolgono una serie di confronti tutti contemporaneamente. Nel momento in cui il valore trovato in  $A$  è maggiore dell'elemento in  $B$  ho individuato, effettuo lo scambio tra  $A$  e  $B$  di tutti quegli elementi di  $A$  che sono maggiori degli elementi di  $B$ .

In questo modo ottengo automaticamente, senza ulteriore elaborazione, due sequenze  $A$  e  $B$  in cui  $A_i > B_j$  per ogni  $i, j$ . Per effettuare questi confronti posso sfruttare l'idea iniziale di Batchner, se io concateno  $A$  con  $B$  rovesciato ottengo una sequenza molto semplice e di confronti come illustrato in figura.

A questo punto posso risolvere i sottoproblemi ricorsivamente.

Se si va a provare si nota un problema da superare. Chiamiamo *saliscendi* una sequenza che comincia con una porzione crescente e diventa poi decrescente. Quando si procede ricorsivamente su una sequenza di questo tipo non ho la garanzia che i due sottoproblemi che ricavo siano anch'esse sequenze saliscendi. Ma il buon Batchner scoprì che quando l'input è un saliscendi, la sequenza viene spezzata in una rotazione di saliscendi.

**Definizione 2.1.** L'operazione *Left Cyclic Shift* si definisce come:

$$SHIFT_j(x_0, x_1, \dots, x_{N-1}) \hat{=} (x_j, \dots, x_{N-1}, x_0, \dots, x_{j-1})$$

**Definizione 2.2.** Una sequenza  $(x_0, \dots, x_{N-1})$  si dice *bitonica* se esiste una sua rotazione che è in forma saliscendi,  $j \in \{0, \dots, N-1\}$  tale che ponendo  $(y_0, y_1, \dots, y_{N-1}) = SHIFT_j(x_0, \dots, x_{N-1})$  si ha che esiste un  $i \in \{0, \dots, N-1\}$  tale che  $y_0 \leq y_1 \leq \dots \leq y_i \geq y_{i+1} \geq y_{i+2} \geq \dots \geq y_{N-1}$ .

**Definizione 2.3.** Sia  $N$  pari, sia  $\underline{x} = (x_0, \dots, x_{N-1})$ , COMPARISON-EXCHANGE è l'operazione che confronta e scambia direttamente due elementi della sequenza. Si tratta di un'operazione primitiva in molti algoritmi di sorting.

Si nota che tutti gli elementi di  $L_{\underline{x}} \leq U_{\underline{x}}$ . Se effettuo lo shift su una sequenza bitonica allora gli elementi che saranno confrontati dal comparatore saranno gli stessi, sebbene in posizione diverse. Cioè elementi che nella figura precedente erano confrontati dal comparatore saranno ancora accoppiati se la sequenza viene shiftata. L'idea è dunque di applicare ricorsivamente le comparazioni sulle due nuove sequenze  $L_{\underline{x}}$  e  $U_{\underline{x}}$  ottenute dal precedente comparatore. Tale rete è il BITONIC-MERGE con input di 8 elementi.

**Proposizione 2.** Se  $\underline{x}$  è bitonica, allora:

- $\max L_{\underline{X}} \leq \min U_{\underline{X}}$
- $\max L_{\underline{X}}$  e  $\min U_{\underline{X}}$  sono sequenze bitoniche

*Osservazione.* Il principio di induzione matematica dice che se  $P(0)$  vale e inoltre  $\forall n, P(n) \Rightarrow P(n+1)$  allora  $\forall n, P(n)$  vale. Tale principio è un assioma. Se  $Q(n) \Rightarrow P(n)$ , ovvero  $Q(n)$  è affermazione più generale di  $P(n)$ , allora dimostrare che  $Q$  è vera per ogni  $n$  mi dimostra che, transitivamente, anche  $P$  è vera per ogni  $n$  (non vale il viceversa). In certi casi può valere la pena puntare a dimostrare l'ipotesi  $Q(n)$  piuttosto che  $P(n)$ , rafforzando così l'ipotesi induttiva. Nel caso in esame:

- $P(n) = \text{"tutte le sequenze saliscendi sono ordinabili"};$
- $Q(n) = \text{"tutte le sequenze bitoniche sono ordinabili"}.$

Numero di confronti che l'algoritmo svolge (caso  $N$  potenza di 2) è:  $C_{BM}(N) = \frac{N}{2} \log_2 N$ . Il cammino critico dell'algoritmo è:  $L_{BM}(N) = \log_2(N)$ . Il parallelismo è dunque  $P^* = \frac{N}{2}$ , un valore ottimale. Tuttavia il problema è che  $C_{BM}$  è elevato per essere solamente relativo alla componente MERGE dell'algoritmo, c'è dunque un costo da pagare per ottenere questo livello di parallelismo.

**Proposizione 3** (Floyd). In un cdag per il merging il numero di confronti è almeno  $\frac{1}{4}N \log_2 N$ .

D'algoritmo ha una struttura statica, questo significa che il grafo che rappresenta i confronti (attraverso il quale si può realizzare l'algoritmo cablato) non dipende dall'input. Se tale grafo invece varia in funzione dell'input allora si dice che l'algoritmo ha una struttura dinamica, come ad esempio per l'algoritmo di ricerca binaria, dove primo confronto è indipendente dall'input, tuttavia il secondo e i successivi dipendono dal valore dell'input.

**Proposizione 4** (Bilardi-Nicolau '86). ADAPTIVE-BITONIC-MERGE ha un numero di confronti  $C_{ABM} = O(N)$  ed un cammino critico  $L_{ABM} = O(\log_2 N)$ .

Per quanto riguarda il SORT, invece, se vado a contare il numero dei confronti:

$$C_{BS}(N) = \frac{N \log_2 N (\log_2 N + 1)}{2}$$

$$L_{BS}(N) = L_{BM}(2) + \dots + L_{BM}(N) = \frac{\log_2 N (1 + \log_2 N)}{2}$$

**Proposizione 5** (AKS '83). Si è trovata una rete AKS che è in grado di svolgere il SORTING con:

$$C_{AKS}(N) = O(N \log_2 N)$$

$$L_{AKS}(N) = O(\log_2 N)$$

Tuttavia la notazione  $O$  grande nasconde una costante di fronte a  $N \log N$  che è stato stimato essere oltre il migliaio, dunque per tutti gli scopi pratici il BITONIC-SORTING è migliore.

Mettendo assieme le due componenti dell'algoritmo che abbiamo visto possiamo osservare com'è strutturata la rete per fare il BITONIC—SORT:

**Definizione 2.4.** Sia  $B(N)$  il numero di permutazioni bitoniche di  $\{0, 1, \dots, N - 1\}$ .

$$B(1) = 1$$

$$B(2) = 2 = 2!$$

$$B(3) = 6 = 3!$$

$$B(4) \leq 4!$$

Vediamo allora se riusciamo a catalogare tutti i saliscendi e poi vediamo le loro rotazioni. Quest'idea è buona ma bisogna considerare che alcuni elementi vengono conteggiati più volte. Si cerca allora di bloccare il maggior numero di elementi possibili: se la sequenza è  $\{0, \dots, N - 1\}$ , un suo saliscendi comincerà con l'elemento 0 e la fase di salita terminerà sul numero  $N - 1$ .

Se io stabilisco quali elementi sono presenti nella porzione di sequenza tra l'elemento 0 e l'elemento  $N - 1$  (nell'insieme  $S$ ) allora ho identificato univocamente tale stringa, poichè so che tali elementi si configureranno necessariamente in ordine crescente e i rimanenti elementi si ordineranno in maniera decrescente successivamente all'elemento  $N - 1$ . Allora:

$$B(N) = N \cdot 2^{N-2}$$

Potrei avere che due rotazioni in  $S_1$  e in  $S_2$  mi danno la stessa sequenza, dunque considerando che io posso sommare le cardinalità di due insiemi solo se la loro intersezione è vuota, allora devo dimostrare questo fatto. L'unico modo di ottenere due rotazioni uguali è di ruotare le sequenze lo stesso numero di volte, altrimenti il risultato avrebbe l'elemento 0 in posizioni diverse. Tuttavia se le due sequenze sono state ruotate lo stesso numero di volte, allora è banale che le due sequenze devono essere identiche per poter avere la rotazione di indice  $i$  identica. E' dunque allora lecito sommare le cardinalità.

*Osservazione.* Le sequenze bitoniche sono in numero inferiore alle permutazioni totali, che sono invece  $N!$ . La differenza tra la rappresentazione binaria di queste due grandezze è di questo tipo:

$$\begin{aligned}\log_2 N! &= N \log_2 N - O(N) \\ \log_2 B(N) &= \log_2 N + (N - 2)\end{aligned}$$

In realtà il BITONIC-MERGE è in grado di ordinare molte più sequenze di quelle bitoniche.

**Proposizione 6.** Il BITONIC-MERGE( $N$ ) ordina  $2^{C_{BM}(N)}$  permutazioni distinte, dove  $C_{BM}(N) = \frac{1}{2}N \log_2 N$ .

**Proof:** La dimostrazione viene lasciata per esercizio al lettore.

Supponiamo di prendere la rete del BITONIC-MERGE, ci si chiede: tale rete ordina o no tutte le permutazioni possibili?

**Proposizione 7** (0-1 Principle). Se una rete di comparator-exchanger ordina tutte le sequenze in  $\{0, 1\}^N$ , allora ordina tutte le sequenze.

## 2.4 Problema del routing

Il *problema dell'instradamento* o *routing* è un problema che prima ancora di presentarsi nel calcolo si è presentato nell'ambito della telefonia. Esso può essere suddiviso in vari sottoproblemi, tra cui una distinzione molto nota è:

- **routing off-line**, il sistema che deve instradare i pacchetti conosce in anticipo sorgenti e destinazioni dei pacchetti e può dunque svolgere un lavoro di pianificazione;
- **routing on-line**, tutto è determinato in tempo reale, destinazioni, sorgenti e contenuto dei messaggi.

L'algoritmo che studieremo principalmente è nato nel campo della telefonia dalle ricerche condotte negli anni 60 da Benes, che lavorava nei Bell Labs. Supponiamo di avere una situazione in cui ho  $N$  terminali di ingresso e di uscita. Il problema può essere così schematizzato:

Supponiamo il problema con le seguenti restrizioni (permutation routing):

- $\text{Term}(i)$  ha un pacchetto per  $\text{Term}(\pi(i))$ ;
- $\pi$  è una permutazione.

Ogni terminale in ingresso ha uno e un solo pacchetto da inviare ad un terminale di uscita (questo è un caso specifico del problema generale studiato da Benes). Se la centralina me la devo costruire io, come la costruisco? Benes prese come mattone elementare per costruire questa rete di routing l'interruttore o switch binario, un dispositivo a due ingressi, due uscite e uno stato binario  $S$ :

- Se  $S = 0$  allora l'input in  $a$  esce in  $c$  e l'input in  $b$  esce in  $d$ .
- Se  $S = 1$  allora l'input in  $a$  esce in  $d$  e l'input in  $b$  esce in  $c$ .

Gli output sono dunque rappresentati dalle equazioni:

$$\begin{cases} c = Sb + (1 - S)a \\ d = Sa + (1 - S)b \end{cases}$$

L'idea di Benes è dunque di costruire la rete di permutazione con tali switch.

*Osservazione.* Si noti la somiglianza tra i comparator-exchanger (vedi capitolo bitonic sort) e gli switch. I primi possono essere visti come degli switch in grado di autoconfigurarsi dinamicamente in base all'input. Un modo per fare una rete di permutazioni è quello di mettere gli interruttori disposti secondo una rete che svolge l'ordinamento.

*Osservazione.* Una rete per l'ordinamento può essere adattata per eseguire il permutation routing on-line, tant'è vero che in molti router vi sono pezzi della rete di Batcher.

Si desidera ottimizzare il numero di commutazioni, il numero degli switch e la lunghezza dei cammini critici. Il problema off—line è più facile del problema on-line e c'è dunque

qualche speranza in più di trovare soluzioni più efficienti per tale problema. Ma quanto può essere piccolo il numero di interruttori affinché si possano produrre tutte le permutazioni in output? Una rete in grado di eseguire il sorting può essere adattata per fare il routing, ma non è vero il viceversa. I due problemi sono collegati e la riduzione si effettua dunque dal routing al sorting, non viceversa. In questo caso il problema più difficile è l'ordinamento e, di conseguenza, l'upper bound dell'algoritmo di Batcher varrà anche per il problema del routing. Non si può invece affermare il viceversa. Se si possiede una rete di interruttori che può generare tutte le permutazioni, essa automaticamente diventa un codice che assegna ad ogni permutazione una o più stringhe binarie (nel caso la stessa si possa ottenere con configurazioni diverse di interruttori). Si può dunque a questo punto osservare che il numero di configurazioni possibili è  $N!$ .

**Proposizione 8.** Una rete con  $S$  interruttori ha  $2^S$  configurazioni. Se la rete può realizzare tutte le permutazioni, allora:

$$2^S \geq N! \Rightarrow S \geq \log_2 N! = N \log_2 N - O(N)$$

*Osservazione.* Quel limite inferiore al numero di interruttori non può essere stretto, non è cioè possibile verificare tale relazione con un'uguaglianza nei casi  $N \neq 2$  in quanto  $N!$  contiene fattori diversi da 2. Si vedrà tuttavia che la rete di Benes si avvicina molto a questo lower bound. Problema del cammino critico: si osserva che la rete di Benes ha un cammino critico dell'ordine di  $\log N$ . Si prende un terminale; quando il cammino che origina da tale terminale attraversa un interruttore, il pacchetto che lo percorre si può trovare in due punti distinti della rete. Dopo aver attraversato due interruttori il pacchetto può essere in quattro distinti posti della rete. Se si guardano i punti raggiungibili da un terminale di ingresso, se si attraversano  $k$  interruttori, il numero di terminali d'uscita che è possibile raggiungere è al più  $2^k$ . In una rete di permutazioni,  $2^k \geq N$  (poiché è necessario poter raggiungere tutte le possibili permutazioni dell'input).

$$\begin{aligned} k &\geq \log_2 N \\ L_{SN}(N) &\geq \log_2 N \end{aligned}$$

Una rete di Benes è ricorsivamente definita come uno stadio di comparatori tra gli ingressi in due sottoreti di  $N/2$  elementi ed uno stadio di comparatori tra le uscite. Dove  $BN(2)$  coincide con il caso base, ossia lo switch binario. Si prenda ora come esempio una rete di Benes con 8 input ( $BN(8)$ ):

Si osserva che la prima metà della rete, detta  $\Omega(N)$  è isomorfa a  $BM(N)$ , ossia BITONIC-MERGE.

$$\begin{cases} S_{BN}(N) = 2S_{BN}(N/2) + N \\ S_{BN}(2) = 1 \end{cases} \Rightarrow S_{BN}(N) = N \log_2 N - \frac{N}{2}$$

La lunghezza del cammino critico è:

$$\begin{cases} L_{BN}(N) = L_{BN}(\frac{N}{2} + 2) \\ L_{BN}(2) = 1 \end{cases} \Rightarrow L_{BN}(N) = 2 \log_2 N - 1$$



Da  $L$  si può ricavare  $S$ , osservando che ad ogni stadio vi sono  $N/2$  interruttori, cioè  $S_{BN}(N) = S_{BN}(N) \frac{N}{2}$ .

$$BN(N) \approx \Omega(N)\Omega^{-1}(N)$$

Se, invece,  $\Omega(N)\Omega(N)$  sia una rete di permutazione non si sa dimostrare, sebbene le simulazioni mostrano che la risposta potrebbe essere positiva, esso è tuttora un problema aperto.

**Proposizione 9.**  $\forall N = 2^d$ ,  $BN(N)$  realizza  $N!$  permutazioni distinte.

**Proof:** Per induzione. Caso base:  $d = 1$ ,  $N = 2$ .

Passo induttivo:

- ipotesi induttiva:  $BN(\frac{N}{2})$  realizza tutte le  $\frac{N}{2}!$  permutazioni
- tesi: (con esempio)

Il primo interruttore può essere configurato a piacere, in questo caso come passante. Con questa configurazione il valore 7 dev'essere fatto uscire necessariamente nel 4° pin e quindi l'uscita corrispondente dev'essere incrociante. Si deve poi proseguire controllando quali input sono necessari alle varie uscite e di conseguenza che configurazione devono avere gli interruttori. Giunto al più n. 5 ricade nella scelta del primo pin che ho già compiuto.

- L'algoritmo configura il primo e l'ultimo stadio
- Assegna permutazioni alle due sottorete  $BN(\frac{N}{2})$

Il tempo di configurazione è il seguente:

$$T_{CONF}(N) = 2T_{CONF}\left(\frac{N}{2}\right) + O(N) = O(N \log N)$$

Questo risultato non può essere migliorato poichè, considerando tempo costante per la configurazione di ogni interruttore, essi sono  $n \log n$  ed è necessario configurarli tutti. Si osserva che l'algoritmo di Benes è difficilmente parallelizzabile:

$$\begin{aligned} SW_{PERM}(N) &\geq \log_2(N!) = N \log_2 N - (\log_2 e)N + o(N) \\ SW_{BN}(N) &= N \log_2 N - \frac{1}{2}N \end{aligned}$$

Togliendo il primo switch (ovvero settandolo come passante) otterremmo:

$$SW_{BN}(N) = N \log_2 N - 0.5(N) - (1 + 2 + 2^2 + \dots + 2^{\log_2 N - 2})$$

dove

$$\begin{aligned} (1 + 2 + 2^2 + \dots + 2^{\log_2 N - 2}) &= \frac{1 - 2^{\log_2 N - 1}}{1 - 2} = \frac{N}{2} - 1 \\ SW_{BN}(N) &= N \log N - N + 1 \end{aligned}$$

$\forall i, \pi(i) = i$  (permutazione identica),  $BN(N)$  realizza  $\geq 2^{N/2(\log_2 N - 1)}$

## 2.5 Problema del calcolo dei prefissi

**Definizione 2.5.** Data una sequenza di  $n$  numeri, vogliamo calcolare la somma di ciascun prefisso di lunghezza  $k$ .

Il calcolo dei prefissi è un problema assolutamente banale se affrontato in modo sequenziale e, prima dell'avvento del calcolo parallelo, è stato trascurato.

*Osservazione.* Dal punto di vista analitico la somma è l'integrale definito mentre la somma di prefissi è l'integrale indefinito.

Per semplificare le cose assumeremo che l'universo degli operandi sia definito.

Consideriamo il semigrupp  $S = \langle A, \bullet \rangle$ , dove  $A$  è un insieme, mentre un'operazione binaria su  $A$  è tale che  $A \times A \rightarrow A$ , è associativa, cioè  $\forall a, b, c \in A, (a \bullet b) \bullet c = a \bullet (b \bullet c)$ .

Esempi:

- $\langle \mathbb{Z}, + \rangle, \langle \mathbb{Z}, \cdot \rangle$  sono semigrupp
- $\langle B, \vee \rangle, \langle B, \wedge \rangle, \langle B, \oplus \rangle$  dove  $B = \text{boolean}$

Un'operazione che ci risulterà molto utile è la composizione di funzioni.

**Proposizione 10** (Cayley). Ogni gruppo finito è isomorfo ad un gruppo di permutazioni. Le permutazioni sono funzioni invertibili che si possono comporre tra loro e formano un gruppo. Equivalentemente ogni semigrupp finito è isomorfo ad un gruppo di funzioni dove l'operazione è la composizione di funzioni.

**Definizione 2.6** (Problema del prefisso). INPUT:  $(x_0, x_1, \dots, x_{N-1} \in A^N$   
OUTPUT:  $(y_0, y_1, \dots, y_{N-1}) \in A^N$  dove  $y_j = x_0 x_1 \dots x_j$ , con  $j = 0, 1, \dots, N-1$

In maniera equivalente posso anche dire che  $y_0 = x_0$  e  $y_j = y_{j-1} x_j$  (un prefisso è uguale al prefisso precedente dopo aver applicato  $\bullet$  per l' $x$  attuale).

In un grafo rappresentante un cdag che tratta operazioni non commutative bisogna specificare l'ordine degli operandi all'ingresso di un'operazione. Se si vuole fare sequenzialmente l'operazione per un prefisso  $y$ , è necessario trattare anche i prefissi di lunghezza inferiore che sono risultati parziali. La situazione tuttavia cambia con il calcolo parallelo. Una prima soluzione parallela per il problema è la seguente:

Questo algoritmo è un ottimo algoritmo parallelo, se ho un numero di processori pari a  $N/2$  posso eseguire l'algoritmo come evidenziato nella greedy schedule anche se il parallelismo reale è inferiore. Sul cammino evidenziato dal grafico si vede che solo  $y_1, y_3$  e  $y_7$  sono validi prefissi, poichè sono gli unici che sono influenzati da  $x_0$ . Ci si pone la domanda, quanto posso tenere piccolo il cammino critico se voglio avere tutti i valori intermedi come output? Bisogna trovare un cdag per il calcolo di tutti gli  $y$  con profondità logaritmica. Se l'unico obiettivo è il cammino critico posso andare alla

definizione e per ogni prodotto fare un albero a sè, se il numero non è una potenza di due non verrà un albero bilanciato ma non è un gran problema, la profondità sarà  $\lfloor \log n \rfloor$ . Ottengo in questo modo un grafo di questo tipo con:

$$|V| = 1 + 2 + \dots + N - 1 = \frac{N(N-1)}{2} = \binom{N}{2}$$

$$L = \log_2 N$$

Questo algoritmo mostra che sebbene si sprechi sul piano del lavoro aggiuntivo non c'è motivo per pensare che non sia possibile affrontare in modo parallelo questo problema. Una domanda che ci possiamo fare è: se si hanno a disposizione tutti i valori interni di quell'albero, invece di andare per gli altri  $y$  a ripetere il calcolo da capo, è possibile comporre in qualche modo i risultati parziali di  $y_7$  per ottenere i prefissi precedenti? Supponiamo di volere  $y_4$ , si evidenzia dal grafo che  $y_3$  è disponibile e moltiplicandolo per  $x_4$  si ottiene  $y_4$ . Per calcolare  $x_6$ , si nota che si sta praticamente raccogliendo tutti gli elementi che stanno alla sinistra del cammino di  $x_7$ . Percorrendo tale cammino si incontrano un certo numero di alberi sinistri che originano da esso i quali, con un'opportuna composizione, producono la soluzione per  $x_6$ .

Ogni prefisso  $y_j$  si può ottenere in questo modo e la lunghezza del cammino critico è al più logaritmica. Abbiamo constatato dunque che mentre nel grafo relativo all'esecuzione sequenziale i risultati intermedi erano sempre disponibili, in quello parallelo no ma è possibile ottenerli in modo relativamente semplice. Tale grafo ha dunque costo:

$$|V| \leq N \log_2 N$$

$$L \leq 2 \log_2 N$$

$$L \leq \log_2 N + \log_2 \log_2 N$$

Se mentre si costruisce l'albero in parallelo si accumulano dei risultati parziali in parallelo all'alberone è possibile ottenere un cammino critico:

$$L \leq \log_2 N$$

Non è introdotto dunque nessun ritardo globale rispetto al calcolo senza considerare i parziali.

Ci si rende tuttavia conto che il numero di operazioni da  $N$  dell'algoritmo sequenziale passa a  $N \log N$  e ci si chiede se questo fattore  $\log N$  possa essere tutto o in parte rieparmiato.

*Osservazione.* La descrizione dello stato di una macchina a stati finiti può essere associato al problema del calcolo dei prefissi. E' dunque interessante costruire un sistema al quale dando un segnale di input di durata potenzialmente infinita produce costantemente un output. Una prima rappresentazione grafica di questo sistema è la seguente:

Si osserva che se si aggiunge una linea ideale (tratteggiata) si ottiene un albero totalmente simmetrico a quello iniziale. Questo albero è stato chiamato in letteratura *Twisted Reflected Tree* (TRT) (Bilardi, Preparata). Questa rappresentazione non costituisce una greedy schedule, bensì permette di trattare segnali infiniti mettendo un buffer tra uno stadio e quello successivo.

$$|V| = (N - 1) + (N - 1) + N = O(N)$$

$$L \leq 2 \log_2 N$$

Il circuito si dice *pipelinable*. Osserviamo cosa significa il simbolo:

Il circuito, privo di questo elemento, è in grado di calcolare tutti i prefissi di un singolo fronte di input. Se vogliamo invece calcolarli per un segnale infinito questo elemento fa l'accumulo dei prefissi del fronte. Il fronte immediatamente successivo dovrà acquisire, per calcolare i suoi prefissi, tutti quelli del fronte precedente. Questo avviene tramite il buffer che funge anche da radice del secondo albero ed è in grado di temporizzare correttamente il segnale. Questa tecnica ci permette di trattare anche segnali molto lunghi. Questo elemento circuitale è dunque in grado di svolgere il prodotto tra semigrupp, esegue cioè ciclicamente il prodotto tra il nuovo fronte e quelli precedenti, che poi propaga a tutte le foglie del secondo albero, così da rendere validi tutti i nuovi prefissi calcolati. Il grafo precedente lo possiamo dunque vedere con o senza questo elemento, ed esso rappresenterebbe rispettivamente il cdag di un algoritmo per il calcolo dei prefissi ed un sistema dinamico in grado di calcolare prefissi con segnali finiti e infiniti.

## 2.6 Macchine a stati finiti

La macchina a stati finiti o *Finite State Machine* (FSM) è il modello più realistico di un sistema dinamico che è possibile realizzare in ingegneria.

La macchina è così definita:

- $M = (\Sigma, Q, \Gamma, \delta, \eta, q_0)$
- $u(t) \in \Sigma$
- $q(t) \in Q$ , si osserva che  $q(t)$  è la funzione che determina l'output e ha la proprietà di avere memoria
- $z(t) \in \Gamma$
- $\delta : Q \times \Sigma \rightarrow Q$
- $\eta : Q \times \Sigma \rightarrow \Gamma$

$$q(t+1) = \delta(q(t), u(t))$$

$$z(t) = \eta(q(t), u(t))$$

Esistono differenze tra le definizioni delle macchine a stati finiti. Si distinguono infatti:

- FSM di Melay,  $z(t) = \eta(q(t), n(t))$
- FSM di Moore  $z(t) = \eta(q(t))$

E' possibile dimostrare che per ognuna FSM di Melay ne esiste una equivalente di Moore. La funzione  $z(t)$  nel DFA (Deterministic Finite-State Automaton) vista nel corso di informatica teorica è:

$$\eta(q) = \begin{cases} 1 & q \in F \\ 0 & q \notin F \end{cases}$$

La DFA è dunque un caso particolare di FSM.

### 2.6.1 Evoluzione di una FSM

INPUT:  $q_0 = (u_0, u_1, \dots, u_{N-1}) \in \Sigma^N$

OUTPUT:  $(q_1, \dots, q_N) = (q(1), \dots, q(N)) \in Q^N$

s.t.  $q(0) = q_0, q(t) = \delta(q(t), u_t)$

Questo è un problema tipico di teoria dei sistemi, ad esempio nei circuiti elettrici l'analisi viene fatta con segnali in forma analitica. Questo tipo di analisi è differente: si basa sull'analisi dello stato di un sistema che ha in input un segnale rappresentato come sequenza di campioni. Si cerca ora dunque un algoritmo di tipo parallelo che produce l'output della macchina a partire dall'input, poiché se, basandosi sulla definizione, si applica un singolo cambiamento ad ogni istante temporale, si ottiene un cammino critico  $N \cdot k$ . Per velocizzare l'operazione si utilizza il ragionamento fatto per il calcolo dei prefissi.

Se un elemento particolare  $a \in \Sigma$ , è possibile definire  $\delta_a(q) = \delta(q, a)$ . Riscrittura della formula della transizione di stato in funzione della sola variabile  $q$ . Caso  $N=2 =$

$$\delta_a : Q \rightarrow Q$$

Riscrittura della formula della transizione di stato in funzione della sola variabile  $q$ .

Caso  $N = 2, \Rightarrow u_0, u_1$

$$\begin{aligned} q_1 &= \delta(q_0, u_0) = \delta_{u_0}(q_0) \\ q_2 &= \delta(q_1, u_1) = \delta_{u_1}(q_1) = \delta_{u_1}(\delta_{u_0}(q_0)) = (\delta_{u_0} \circ \delta_{u_1})(q_0) \end{aligned}$$

dove

$$\begin{aligned} f(g(x)) &= (g \circ f)(x) \\ S &= \langle A, \circ \rangle \end{aligned}$$

Con  $A = Q^Q$  e la funzione  $f : Q \rightarrow Q$

$$q_{t+1} = (\delta_{u_0} \circ \delta_{u_1} \circ \dots \circ \delta_{u_t})(q_0) \quad t = 1, \dots, N$$

### 2.6.2 FSM per l'addizione binaria

La macchina prende in input due sequenze di bit dove  $a_0$  e  $b_0$  sono le cifre meno significative. Gli insiemi della macchina:

- $\Sigma = \{00, 01, 10, 11\} = \{0, 1\}^2$
- $Q = \{0, 1\}$
- $\Gamma = \{0, 1\}$
- $\delta(q, u) = \delta(q, (a, b)) = TH_2(q, a, b) = qa + qb + ab$ , dove  $TH_2$  è la funzione soglia che è uguale a 1 se  $k$  o più elementi sono = 1
- $\eta(q, u) = \eta(q, (a, b)) = q \oplus a \oplus b$

Si osserva ora il semigruppato  $M$ :

$$M = (\Sigma, Q, \Gamma, \delta, \eta)$$

Esso si ottiene dal gruppo di generatori  $\delta_u : u \in \Sigma$ .  $M$  è il semigruppato generato da  $\delta_u : u \in \Sigma$ .

$q$	$\delta_{00}$	$\delta_{01} = \delta_{10}$	$\delta_{11}$
0	0	0	1
1	0	1	1

Se nel sommare si arriva in una colonnina in cui entrambi in numeri sono entrambi zero allora sicuramente quella colonna non genera riporto.

Se si comincia a moltiplicare (comporre) fra loro questi generatori si ottengono dei generatori:

$f$	$g$	$\delta_{00}$	$\delta_{01} = \delta_{10}$	$\delta_{11}$
$\delta_{00}$		$\delta_{00}$	$\delta_{00}$	$\delta_{11}$
$\delta_{01} = \delta_{10}$		$\delta_{00}$	$\delta_{01}$	$\delta_{11}$
$\delta_{11}$		$\delta_{00}$	$\delta_{11}$	$\delta_{11}$

Se questi risultati tra generatori avessero dato risultati al di fuori di quelli sopra illustrati la tabella sarebbe stata estesa. L'estensione della tabella pitagorica può continuare finché vengono aggiunti elementi di un nuovo tipo.

- $\delta_{00} \equiv R$ , **carry reset**
- $\delta_{01} \equiv P$ , **carry propagate**
- $\delta_{11} \equiv S$ , **carry set**

Il semigruppone viene chiamato *semigruppone dei riporti*, o semigruppone carry. 00 è un carry reset perchè che ci sia o meno un riporto che viene da destra, muovendosi da destra verso sinistra nel modo più naturale di svolgere le operazioni, da lì in poi non c'è riporto in nessun caso, quindi anche se non si conosce cosa ci sia alla destra di una posizione 00 si è comunque certi che quella posizione non dà riporto. É dunque possibile cominciare a lavorare a sinistra di quella posizione senza sapere quello che c'è a destra, tali punti permettono di decomporre il problema. Dal punto di vista statistico questa sequenza si ripete quindi è possibile parallelizzare idealmente l'operazione di addizione. Questo in pratica non avviene perchè è necessario possedere una procedura deterministica. Paradossalmente il propagate è quello che rende difficile la parallelizzazione dell'addizione. Si dice paradossalmente poichè è proprio la funzione identità a causare il problema.

$$\forall X \in R, S, P \quad XR = R, \quad XS = S, \quad XP = X$$

Ora, in possesso di questo semigruppone, l'impianto matematico è completo ma è fondamentale ora sintetizzare l'algoritmo con porte logiche booleane e il primo passo è quello di scegliere la rappresentazione. Siccome si hanno tre elementi ed è necessario utilizzare una lunghezza fissa, servono 2 bit per rappresentarli. Come scegliere questa rappresentazione? Innanzitutto è fondamentale passare da  $u$  a  $\delta_u$  per fare questo è possibile rappresentare  $\delta_u$  direttamente con i bit di  $u$ . Cioè 00 per  $\delta_{00}$  o  $R$  e 11 per  $\delta_{11}$  o  $S$ .  $P$  è invece problematico; la cosa più semplice che si può fare è rappresentare l'identità in due modi diversi, cioè sia 10 che 01 sono rappresentazioni legali per  $\delta_{01}$ . Non è vietato rappresentare un elemento del semigruppone in due modi diversi. Si è fatta una scelta semplice e banale, ora bisogna considerare le conseguenze e realizzare dei moltiplicatori di semigruppone. Cos'è un moltiplicatore carry-save? Se invece di rappresentare i numeri binari con il numero minimo di cifre possibili essi si rappresentano con una certa ridondanza, cioè ad ogni posizione binaria invece di associare un bit se ne associano due, oltre a sprecare un buon numero di bit, l'operazione di somma non ha effetti a propagazione illimitata. Cioè per sapere quanto vale l'operazione di somma in una determinata posizione si dovrebbe analizzare un numero costante di posizioni dell'input. Questa operazione ha un cammino critico anche più piccolo del sommatore che si sta cercando di realizzare. L'idea è allora di usare questo formato per costruire un albero della somma in cui ogni sottoinsieme ha complessità costante e il totale è logaritmico. Il risultato in forma carry-safe si converte nella forma "normale" con una semplice operazione di somma. Questo è un esempio in cui una rappresentazione ridondante apparentemente inefficiente facilita invece l'operazione.

### 2.6.3 Sistesie di funzioni booleane

Si scrive la tabella di verità e usando le mappe di Karnaugh si ricavano delle funzioni booleane semplificate:

Si noti che le  $D_i$  sono le **don't care condition**. Nel casi centrali del  $P$  non è importante come i bit vengono disposti, tuttavia è fondamentale che siano uno il complementare dell'altro. Allora  $\bar{D}_i$  è il complementare di  $D_i$ .

## 2.7 Discrete Fourier Transform

Si immagina di avere un anello  $A(+, \cdot)$ , con  $\omega$  radice principale di ordine  $N$ , se

- $\omega^N = 1$
- $\omega^1, \omega^2, \dots, \omega^{N-1} \neq 1$
- $\sum_{j=0}^{N-1} \omega^{ij} = \begin{cases} N & i = 0 \\ 0 & i \neq 0 \end{cases}$

### 2.7.1 Algoritmo FFT

L'algoritmo per il calcolo della DFT proposto da Cooley-Tukey nel 1965, ma anticipato in un manoscritto di Gauss riportato alla luce solo negli anni '80, è una strategia Divide&Conquer per il calcolo della trasformata e richiede che il numero  $N$  sia fattorizzabile ( $N = pq$ ). Se  $N$  è un numero primo questo algoritmo non dà nessun vantaggio.

- Si dispone l'input in una matrice  $p \times q$  row major.
- Si esegue la trasformata delle colonne  $F(w^q)$ .
- Si moltiplica per il *twiddle factor* (la casella di indice  $i, j$  va moltiplicata per  $\omega^{ij}$ ).
- Si esegue la trasformata delle righe.

L'algoritmo è di interesse nel caso generale di  $N = pq$  ma per comodità sarà trattato il caso  $N = 2^d$ . Le potenze di due sono i numeri più fattorizzabili, ma proprio per questo possiamo andare a scegliere  $p$  e  $q$  in tanti modi diversi. Il caso più tipico è quello  $p = 2$  e  $q = N/2$  e il suo simmetrico  $p = N/2, q = 2$ . Si osserva come esempio il caso  $N = 8$ :

$$F_2(\omega^{N/2}) = F_2(-1) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ a-b \end{bmatrix}$$

La permutazione eseguita dalla prima fase del circuito FFT è detta UNSHUFFLE poichè ripristina l'ordinamento degli input. Nelle prossime lezioni sarà presentata una rete detta shuffle-exchanger basata sulla stessa operazione. Questo tipo di rete oggi non va per la maggiore ma era considerata fino agli anni 80 la rete ideale per connettere tra loro più processori. Si cerca di capire cosa succede nella rete di unshuffle osservando gli input in base 2:

$$h \longrightarrow UNSHUFFLE(h)$$

$$\begin{array}{llll} 0 & 0 & 0 & \longrightarrow 0 & 0 & 0 \\ 0 & 0 & 1 & \longrightarrow 1 & 0 & 0 \\ 0 & 1 & 0 & \longrightarrow 0 & 0 & 1 \\ 0 & 1 & 1 & \longrightarrow 1 & 0 & 1 \\ 1 & 0 & 0 & \longrightarrow 0 & 1 & 0 \\ 1 & 0 & 1 & \longrightarrow 1 & 1 & 0 \\ 1 & 1 & 0 & \longrightarrow 0 & 1 & 1 \\ 1 & 1 & 1 & \longrightarrow 1 & 1 & 1 \end{array}$$



Si nota che quest'operazione coincide con una rotazione ciclica verso destra di una posizione. Come ci si convince che questa regola non funziona solo nel caso  $N = 8$ ? In pratica abbiamo preso gli elementi in posizione pari da mettere sopra e quelli dispari da mettere sotto. Essa infine è una bit-permutation. In che modo l'unshuffle è una bit permutation? Si è partiti da una permutazione di numeri e se si guarda la loro rappresentazione binaria si osservano tutte le stringhe di lunghezza  $\log_2 N$ . Non tutte le permutazioni hanno la caratteristica di poter essere descritte dalla permutazione di singoli bit. Le permutazioni che sono bit-permutation sono in effetti rare. Le permutazioni sui bit sono  $\frac{(\log_2 N)!}{N!}$

### 3 Modelli di rappresentazione per reti

Il modello cdag finora analizzato prende in considerazione solo il numero di processori (risorse) a disposizione ed è dunque un modello semplificato e idealizzato. Esso non considera in effetti aspetti importanti come l'architettura dei processori e il tempo impiegato per le comunicazioni, in particolare tra risorse diverse. Il modello finora utilizzato è un buon modo per entrare nell'ottica del calcolo parallelo sebbene semplificato.

Nei modelli dove il costo delle comunicazioni è significativo, questo generalmente fa sì che si debba stare molto più attenti a come mappare l'algoritmo sul sistema di calcolo o anche cambiare l'ordine di algoritmi diversi. Di fatto la maggior parte degli algoritmi ottimizzati per le comunicazioni non sono stati pensati in quel modo, bensì con un modello idealizzato e poi adattati. È meglio dunque pensare all'algoritmo cercando di estrarne il massimo del parallelismo possibile e poi realizzarne un'implementazione ottimizzata. Ci sono tuttavia dei casi in cui nel momento in cui si introduce il vincolo delle comunicazioni ci si trova costretti a rivedere totalmente il progetto iniziale. Si parlerà ora di qualche modello di calcolo un po' più realistico. Vi sono molti modelli distinti poiché essi perseguono spesso obiettivi diversi:

- avere modelli molto accurati
- avere modelli molto portabili, cioè catturare i vincoli essenziali in modo tale che essi valgano su più macchine possibili, analogamente a quello che si fa nella progettazione degli algoritmi sequenziali ad alto livello

Ci si concentra ora su un modello dove ogni nodo ha risorse di calcolo e risorse di memoria.

#### 3.1 Multiprocessore distribuito