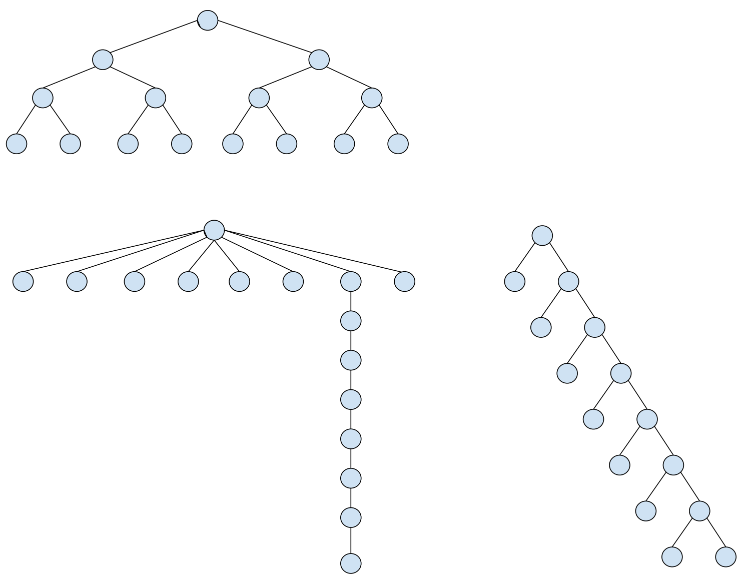
Esercizi

Domanda 1. Per i tre DAG di esecuzione mostrati in figura, determinare:

- la lunghezza del cammino critico;
- il massimo possibile speedup assumendo un numero infinito di processori;
- il minimo numero di processori necessari per ottenere il massimo possibile speedup;
- il massimo possibile speedup se il numero di processori è limitato a 2, 4 o 8, e uno schedule dell'esecuzione che permette di raggiungere tale speedup.



Nell'analisi si assuma che n è il numero totale di nodi del DAG e si considerino gli archi orientati dall'alto al basso.

Soluzione. I tre DAG sono rispettivamente: un albero binario completo con $n = 15$ nodi; una star in cui la radice ha $n/2 = 8$ figli, uno dei quali è la testa di una catena di $n/2 = 8$ nodi; un albero sbilanciato con $n = 15$ nodi tale che ogni figlio sinistro è una foglia. Il numero di livelli nei tre grafi è rispettivamente $\lceil \log_2 n \rceil$, $\frac{n}{2} + 1$, e $\lceil n/2 \rceil$.

1. *Albero binario completo* ($n = 15$)

    - Se consideriamo il numero dei nodi, il cammino critico ha lunghezza 4 (oppure 3, se consideriamo il numero di archi). Generalizzando, il cammino critico in un albero binario completo contiene $\lceil \log_2 n \rceil$ nodi.
      Notate che considerare archi o nodi è del tutto equivalente a livello asintotico, poiché a ogni nodo corrispondono $O(1)$ istruzioni. Concettualmente in questo scenario sembra comunque preferibile (ed è quello che farò nella soluzione dell'esercizio) ragionare sul numero di nodi: ogni nodo nel cammino critico deve infatti essere "eseguito" sequenzialmente.

- Work e span sono rispettivamente $T_1 = n$ e $T_\infty = \lceil \log_2 n \rceil$[1]. Il massimo possibile speedup, corrispondente a un numero infinito di processori, è quindi $\frac{n}{\lceil \log_2 n \rceil}$. Nell'esempio illustrato, lo speedup massimo è 15/4.
- Il minimo numero di processori che consente di ottenere il massimo possibile speedup è $8 = \lceil \frac{n}{2} \rceil$. Infatti i $\lceil \frac{n}{2} \rceil$ nodi dell'ultimo livello dell'albero (il più ampio) devono poter essere eseguiti in parallelo per garantire speedup massimo. Si ha quindi $T_8 = T_\infty = 4$.
- Come già osservato al punto precedente, 8 processori consentono di ottenere massimo speedup. Con 4 processori i nodi dell'ultimo livello dovranno essere eseguiti in due "ondate", e quindi $T_4 = 5$ e lo speedup con 4 processori sarà $S_4 = 15/5$. Con due processori i nodi del penultimo livello dovranno essere eseguiti in due "ondate" e i nodi dell'ultimo livello dovranno essere eseguiti in quattro "ondate", quindi $T_2 = 8$ e $S_2 = 15/8$.

2. *Star + catena* ($n = 16$)

- Il cammino critico contiene $9 = \frac{n}{2} + 1$ nodi.
- Work e span sono rispettivamente $T_1 = n$ e $T_\infty = \frac{n}{2} + 1$. Il massimo possibile speedup è quindi $\frac{n}{\frac{n}{2}+1} = \Theta(1)$. Nell'esempio illustrato, lo speedup massimo è 16/9.
- In questo DAG, anche se il secondo livello dell'albero contiene $n/2$ nodi, non è necessario eseguire tutti questi nodi in parallelo. Per ottenere il massimo possibile speedup 2 processori saranno infatti sufficienti: dopo aver eseguito la radice, basta che lo scheduler esegua il primo nodo del secondo livello dell'albero in parallelo con la testa della catena, poi il secondo nodo del secondo livello dell'albero in parallelo con il secondo nodo della catena, poi il terzo nodo del secondo livello dell'albero in parallelo con il terzo nodo della catena, etc. Poiché il numero di nodi sul secondo livello dell'albero è uguale a quello nella catena, questa strategia comporterà 9 "ondate" di scheduling, che corrisponde esattamente alla lunghezza del cammino critico, da cui $T_2 = T_\infty = 9$.
- Se il numero di processori è limitato a 2, è possibile ottenere speedup massimo, come discusso al punto precedente. A maggior ragione, questo risulta vero se i processori sono 4 oppure 8: pur potendo eseguire in parallelo più nodi tra quelli sul secondo livello dell'albero, non sarà però possibile utilizzare in modo efficace durante l'intera esecuzione i processori in eccesso per via delle dipendenze sul cammino critico.

3. *Albero sbilanciato* ($n = 15$)

- Il cammino critico contiene $8 = \lceil n/2 \rceil$ nodi.
- Work e span sono rispettivamente $T_1 = n$ e $T_\infty = \lceil n/2 \rceil$. Il massimo possibile speedup è quindi $\frac{n}{\lceil n/2 \rceil} = \Theta(1)$. Nell'esempio illustrato, lo speedup massimo è 15/8.
- Ogni livello dell'albero contiene al più due nodi, quindi 2 processori saranno sufficienti per ottenere il massimo speedup (con 2 processori è possibile realizzare 8 "ondate" di scheduling, quindi $T_2 = T_\infty = 8$).
- Se il numero di processori è limitato a 2, è possibile ottenere speedup massimo, come discusso al punto precedente. A maggior ragione, questo risulta vero se i processori sono 4 oppure 8: non sarà infatti possibile utilizzare efficacemente i processori in eccesso per via delle dipendenze sul cammino critico.

Non è necessario usare formule particolari per risolvere questo esercizio. Basta utilizzare la definizione di speedup su P processori. In particolare, non sappiamo come applicare la legge di Amdahl (e non dovete cercare di farlo), poiché dai DAG non è chiaro come distinguere la frazione sequenziale e quella parallela dell'esecuzione.

[1] Sarebbe più corretto scrivere $T_1 = \Theta(n)$ e e $T_\infty = \Theta(\log_2 n)$, poiché a ogni nodo del DAG corrispondono $\Theta(1)$ istruzioni. Per alleggerire la notazione, ometteremo la notazione Θ immaginando che ogni nodo del DAG corrisponda ad una ed una sola istruzione. La validità dei risultati non è comunque compromessa a livello asintotico da questa semplificazione.

Domanda 2. Si consideri una routine f di un programma $\mathcal{P}$ e si assuma che f richieda il 60% del tempo di esecuzione totale di $\mathcal{P}$ su un dato input X. Se volessimo dimezzare il tempo richiesto dal programma $\mathcal{P}$ su input X, assumendo che f sia perfettamente parallelizzabile, quanti processori dovremmo usare? E se la premessa fosse stata che f richiede il 40% invece del 60% del tempo totale del programma $\mathcal{P}$?

Soluzione. Dimezzare il tempo richiesto dal programma $\mathcal{P}$ equivale a dire che si vuole ottenere uno speedup $S = 2$. In base alla legge di Amdahl, lo speedup è definito come

$$S = \frac{1}{s + \frac{1-s}{P}}$$

dove P è il numero di processori e s la frazione sequenziale di codice. Nell'esercizio, $s = 40\%$, poiché f è la parte parallelizzabile e consuma il 60% del tempo. Quindi, per ottenere speedup 2, basta impostare:

$$2 = \frac{1}{0.4 + \frac{0.6}{P}} \qquad \text{da cui} \qquad 0.4 + \frac{0.6}{P} = \frac{1}{2}$$

Si ha quindi $\frac{0.6}{P} = 0.1$ da cui $P = \frac{0.6}{0.1} = 6$. Per ottenere speedup 2 servono quindi 6 processori.

Nella seconda domanda, la frazione sequenziale e quella parallela si invertono: la parte sequenziale prende il 60% del tempo di esecuzione. Impostiamo quindi:

$$2 = \frac{1}{0.6 + \frac{0.4}{P}} \qquad \text{da cui} \qquad 0.6 + \frac{0.4}{P} = \frac{1}{2}$$

Si ha quindi $\frac{0.6}{P} = -0.1$, che è impossibile per $P \geq 0$. Nel secondo caso è quindi impossibile ottenere speedup 2. Si poteva anche rispondere alla seconda domanda osservando che il massimo speedup, ottenibile con un numero infinito di processori, è $\frac{1}{s} = \frac{1}{0.6} = \frac{10}{6} = 1.\bar{6} < 2$.

Domanda 3. Un programma ha efficienza 89% con una frazione seriale del 2%. Approssimativamente, quanti processori sono usati in base alla legge di Amdahl? (Si veda la lezione 2 per la definizione di efficienza).

Soluzione. L'efficienza è definita come lo speedup S diviso per il numero di processori P:

$$E = \frac{S}{P}$$

Lo speedup è dato dalla legge di Amdahl, e quindi:

$$S = \frac{1}{s + \frac{1-s}{P}} \qquad \text{da cui} \qquad E = \frac{1}{P \cdot (s + \frac{1-s}{P})} = \frac{1}{Ps + (1-s)}$$

Nell'esercizio, conosciamo $E = 89\%$ e $s = 2\%$. Quindi:

$$\frac{89}{100} = \frac{1}{\frac{2P}{100} + \frac{98}{100}} = \frac{100}{2P + 98} \qquad \text{da cui} \qquad 2P = \frac{10000}{89} - 98$$

Risulta quindi $P \approx 7$.

Domanda 4. Dato uno speedup di 20 su 32 processori, qual è la frazione seriale secondo la legge di Amdahl? E secondo quella di Gustafson?

Soluzione. Ricordiamo la legge di Amdahl:

$$S = \frac{1}{s + \frac{1-s}{P}} \qquad \text{da cui} \qquad 20 = \frac{1}{s + \frac{1-s}{32}}$$

ovvero

$$s + \frac{1}{32} - \frac{s}{32} = \frac{1}{20} \qquad \text{da cui} \qquad \frac{31}{32}s = \frac{1}{20} - \frac{1}{32} = \frac{8-5}{160} = \frac{3}{160}$$

da cui $s \approx 0.02 = 2\%$.

Ricordiamo ora la legge di Gustafson:

$$S_P = P - \alpha(P-1) \qquad \text{da cui} \qquad 20 = 32 - 31\alpha$$

ovvero $31\alpha = 12$ e quindi $\alpha \approx 38\%$.