CS300 Homework #2

TA: Yeonghun Kim
neutrinoant@kaist.ac.kr

Total 100 points Due: 2020-10-05 18:00:00 KST

- Write in English or Korean.
- Make sure your writing is readable before upload it (especially for scanned image).

1. Matrix Multiplication (50 pts)

Let A and B be $n \times n$ matrices where n is an exact power of 2. Let $C = A \cdot B$. Suppose that we partition each of A, B and C into four $n/2 \times n/2$ matrices.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

$$C = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

Suppose we create following matrices:

$$\begin{aligned}
S_1 &= A_{21} + A_{22} & P_1 &= S_2S_4 & T_1 &= P_1 + P_2 \\
S_2 &= S_1 - A_{11} & P_2 &= A_{11}B_{11} & T_2 &= T_1 + P_4 \\
S_3 &= B_{12} - B_{11} & P_3 &= A_{12}B_{21} & & \\
S_4 &= B_{22} - S_3 & P_4 &= (A_{11} - A_{21})(B_{22} - B_{12}) & & \\
& & P_5 &= S_1S_3 & & \\
& & P_6 &= (A_{12} - S_2)B_{22} & & \\
& & P_7 &= A_{22}(S_4 - B_{21}) & &
\end{aligned}$$

(a) (14 pts) Show that $C = \begin{pmatrix} P_2 + P_3 & T_1 + P_5 + P_6 \\ T_2 - P_7 & T_2 + P_5 \end{pmatrix}$

Answer) Let $C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$, then

$P_2 + P_3 = A_{11}B_{11} + A_{12}B_{21} = C_{11}$

$$\begin{aligned}
T_1 + P_5 + P_6 &= P_1 + P_2 + P_5 + P_6 = S_2S_4 + A_{11}B_{11} + S_1S_3 + (A_{12} - S_2)B_{22} \\
&= (S_1 - A_{11})(B_{22} - S_3) + A_{11}B_{11} + S_1S_3 + A_{12}B_{22} - B_{22}S_2 \\
&= S_1B_{22} - S_1S_3 - A_{11}B_{22} + A_{11}S_3 + A_{11}B_{11} + S_1S_3 + A_{12}B_{22} - S_1B_{22} + A_{11}B_{22} \\
&= A_{11}S_3 + A_{11}B_{11} + A_{12}B_{22} \\
&= A_{11}(B_{12} - B_{11}) + A_{11}B_{11} + A_{12}B_{22} \\
&= A_{11}B_{12} + A_{12}B_{22} = C_{12}
\end{aligned}$$

$$\begin{aligned}
T_2 - P_7 &= T_1 + P_4 - P_7 = P_1 + P_2 + P_4 - P_7 \\
&= S_2S_4 + A_{11}B_{11} + A_{11}B_{22} - A_{11}B_{12} - A_{21}B_{22} + A_{21}B_{12} - A_{22}(B_{22} - S_3 - B_{21}) \\
&= (A_{21} + A_{22} - A_{11})(B_{22} - B_{12} + B_{11}) + A_{11}B_{22} - A_{11}B_{12} - A_{21}B_{22} \\
&\quad + A_{21}B_{12} - A_{22}B_{22} + A_{22}B_{12} - A_{22}B_{11} + A_{22}B_{21} \\
&= A_{21}B_{11} + A_{22}B_{21} = C_{21}
\end{aligned}$$

$$\begin{aligned}
T_2 + P_5 &= T_1 + P_4 + S_1S_3 = S_2S_4 + P_2 + P_4 + S_1S_3 \\
&= (A_{21} + A_{22} - A_{11})(B_{22} - B_{12} + B_{11}) + A_{11}B_{11} \\
&\quad + (A_{11} - A_{21})(B_{22} - B_{12}) + (A_{21} + A_{22})(B_{12} - B_{11}) \\
&= A_{21}B_{12} + A_{22}B_{22} = C_{22}
\end{aligned}$$

(b) (10 pts) Use the above method to compute the matrix product $\begin{pmatrix} 4 & 2 \\ 7 & 5 \end{pmatrix}\begin{pmatrix} 6 & 8 \\ 1 & 3 \end{pmatrix}$

Answer)

$S_1 = 12, S_2 = 8, S_3 = 2, S_4 = 1$

$P_1 = 8, P_2 = 24, P_3 = 2, P_4 = 15, P_5 = 24, P_6 = -18, P_7 = 0$

$T_1 = 32, T_2 = 47$

$$C = \begin{pmatrix} P_2 + P_3 & T_1 + P_5 + P_6 \\ T_2 - P_7 & T_2 + P_5 \end{pmatrix} = \begin{pmatrix} 24 + 2 & 32 + 24 - 18 \\ 47 - 0 & 47 + 24 \end{pmatrix} = \begin{pmatrix} 26 & 38 \\ 47 & 71 \end{pmatrix}$$

※ If you calculate C without using calculated values of $S_1 \sim S_4, P_1 \sim P_7, T_1 \sim T_2$, get 6pts.

(c) (6 pts) Is this way more efficient that Strassen's method? Explain in terms of the numbers of $n/2 \times n/2$ matrix multiplications and additions(subtractions).

Answer)

Strassen's method : 7 multiplication and 18 addition(subtractions).

Problem's method : 7 multiplication and 15 addition(subtractions).

Problem's method is more efficient than Strassen's.

If the numbers of operations(mul, add) you calculated are wrong or if you argue that efficiency is same since numbers of multiplications are both 7, get 4pts. (The difference would be little, but strictly speaking, you cannot say the number of additions does not have any effect.)

(d) (6 pts) Constuct a recurrence equation for this method.

Answer) $T(n) = 7T(n/2) + \Theta(n^2)$

(e) (6 pts) Calculate the asymptotic running time of this method in Θ-notation.

Answer) $T(n) = \Theta(n^{\lg 7})$

(f) (8 pts) Even if n is not an exact power of 2, you can use the above method by finding the smallest power of 2 such that $n \leq 2^m$, and padding your matrices to $2^m \times 2^m$. Use it to prove that the asymptotic running time of this method is the same as the answer to (e) for any positive integer n.

Answer) After zero-padding, you have a $2^m \times 2^m$ matrix with m satisfying $2^{m-1} < n \leq 2^m$. Thus $T(n) = T(2^m) = \Theta(2^{m \lg 7})$ by using the result of (e). Note that $n \leq 2^m < 2n$ and then $n^{\lg 7} \leq 2^{m \lg 7} < 7n^{\lg 7}$, which gives $T(n) = \Theta(2^{m \lg 7}) = \Theta(n^{\lg 7})$ by the definition of Θ notation.

※ Each proof of big-O and Ω gives 4pts, or minimum 2pts if there is a logical problem.

2. Randomized Quicksort (50 pts)

Recall RANDOMIZED-QUICKSORT in Chapter 7.3 in CLRS.

(a) (7 pts) Explain the Randomized Quicksort and why we should use it.

Answer) Randomized Quicksort picks the pivot randomly. In this case, running time is independent of the input order, no specific input elicits the worst-case behavior.

※ Grading: What is Randomized Quicksort (2pts) + Why we should use it (5pts).

(b) (1 pts) Argue that, given an array of size n, the probability that any particular element is chosen as the pivot is $1/n$. Use this to define indicator random variables

$$X_i = \begin{cases} 1 & \text{if } i \text{ th smallest element is chosen as the pivot;} \\ 0 & \text{otherwise} \end{cases}$$

which is slightly different from the lecture slides. What is $E[X_i]$?

Answer) $E(n) = \frac{1}{n}$

(c) (6 pts) Let $T(n)$ be a random variable denoting the running time of quicksort on an array of size n. Show that

$$E[T(n)] = E\left[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))\right]. \tag{1}$$

Answer) We can apply linearity of expectation over all of the events X_i. Suppose we have a particular X_i be true, then we will have one of the subarrays be length $i-1$, and the other be $n-i$, and will of course still need linear time to run the partition procedure.

(d) (12 pts) Show that we can rewrite equation (1) as

$$E[T(n)] = \frac{2}{n}\sum_{q=2}^{n-1} E[T(q)] + \Theta(n). \tag{2}$$

Answer)

$$\begin{aligned}
E\left[\sum_{q=1}^{n} X_q(T(q-1) + T(n-q) + \Theta(n))\right] &= \sum_{q=1}^{n} E[X_q(T(q-1) + T(n-q) + \Theta(n))] \\
&= \sum_{q=1}^{n} E[X_q] \cdot E[T(q-1) + T(n-q) + \Theta(n)] \\
&= \frac{1}{n}\sum_{q=1}^{n} E[T(q-1) + T(n-q) + \Theta(n)] \\
&= \frac{1}{n}\left(\sum_{q=1}^{n} E[T(q-1)] + \sum_{q=1}^{n} E[T(n-q)]\right) + \Theta(n) \\
&= \frac{2}{n}\sum_{q=1}^{n} E[T(q-1)] + \Theta(n) \\
&= \frac{2}{n}\sum_{q=2}^{n-1} E[T(q)] + \Theta(n)
\end{aligned}$$

(e) (10 pts) Show that

$$\sum_{k=2}^{n-1} k \lg k \leq \frac{1}{2}n^2 \lg n - \frac{1}{8}n^2. \tag{3}$$

where $n \geq 3$. (Hint: Split the summation into two parts, one for $k = 2, 3, ..., \lceil n/2 \rceil - 1$ and the other for $k = \lceil n/2 \rceil, ..., n-1$)

Answer 1)

$$\begin{aligned}
\sum_{k=2}^{n-1} k\lg k &= \sum_{k=2}^{\lceil \frac{n}{2}\rceil -1} k\lg k + \sum_{k=\lceil \frac{n}{2}\rceil}^{n-1} k\lg k \\
&\le \sum_{k=2}^{\lceil \frac{n}{2}\rceil -1} k\lg \frac{n}{2} + \sum_{k=\lceil \frac{n}{2}\rceil}^{n-1} k\lg n \\
&= \lg n \sum_{k=2}^{\lceil \frac{n}{2}\rceil -1} k - \lg 2 \sum_{k=2}^{\lceil \frac{n}{2}\rceil -1} k + \lg n \sum_{k=\lceil \frac{n}{2}\rceil}^{n-1} k \\
&= \lg n \sum_{k=2}^{n-1} k - \sum_{k=2}^{\lceil \frac{n}{2}\rceil -1} k \\
&= \lg n \frac{(n-2)(n+1)}{2} - \frac{(\lceil \frac{n}{2}\rceil - 2)(\lceil \frac{n}{2}\rceil + 1)}{2} \\
&\le \lg n \frac{(n-2)(n+1)}{2} - \frac{(\frac{n}{2} - 2)(\frac{n}{2} + 1)}{2} \\
&= \lg n \frac{n^2-n-2}{2} - (\frac{n^2}{8} - \frac{n}{4} - 1) \\
&= (\frac{1}{2}n^2\lg n - \frac{1}{8}n^2) - (\frac{1}{2}n\lg n - \frac{1}{4}n + \lg n - 1) \\
&\le \frac{1}{2}n^2\lg n - \frac{1}{8}n^2 \quad (n \ge 3)
\end{aligned}$$

※ If you miss the ceiling functions, get 6pts.

※ Similarly, using expression like $\sum_{k=1}^{n/2} k$ get 6pts. It's not a valid expression.

Answer 2)

If we let $f(k) = k\lg k$ treated as a continuous function, then $f'(k) = \lg k + 1$. Note now that the summation written out is the left-hand approximation of the integral of $f(k)$ from 2 to n with step size 1. By integration by parts, the anti-derivative of $k\lg k$ is

$$\frac{1}{\ln 2}\left(\frac{k^2}{2}\ln k - \frac{k^2}{4}\right)$$

So plugging in the bounds and subtracting, we get

$$\frac{n^2\lg n}{2} - \frac{n^2}{4\ln 2} - 1$$

Since f has a positive derivative over the entire interval that the integral is being evaluated over, the left-hand rule provides a underapproximation of the integral, so we have that

$$\sum_{k=2}^{n-1} k \lg k \leq \frac{n^2 \lg n}{2} - \frac{n^2}{4 \ln 2} - 1 \leq \frac{n^2 \lg n}{2} - \frac{n^2}{8}$$

where the last inequality uses the fact that $\ln 2 < 2$.

(f) (14 pts) Using the bound from equation (3), show that the recurrence in equation (2) has the solution $E[T(n)] = O(n \lg n)$.
(Hint: Show, by substitution, that $E[T(n)] \leq an \lg n$ for sufficiently large n and for some positive constant a.)

Answer) By substitution method, prove $E[T(n)] \leq an \lg n$ for positive constant a.

$$\begin{aligned} E[T(n)] &= \frac{2}{n} \sum_{q=2}^{n-1} E[T(q)] + \Theta(n) \\ &\leq \frac{2}{n} \sum_{q=2}^{n-1} aq \lg q + \Theta(n) \\ &\leq \frac{2a}{n} \left(\frac{1}{2} n^2 \lg n - \frac{1}{8} n^2 \right) + \Theta(n) \\ &= an \lg n - \left(\frac{1}{4} an - \Theta(n) \right) \end{aligned}$$

If a is chosen large enough so that $\frac{1}{4}an$ dominates $\Theta(n)$, then $E[T(n)] \leq an \lg n$ is satisfied.

※ If you don't describe the condition that $\frac{1}{4}an$ dominates $\Theta(n)$ (or a similar condition), get 12pts. $an \lg n - (\frac{1}{4}an - \Theta(n)) \leq an \lg n$ cannot be satisfied without the condition.
※ Also get 7pts if you give wrong proof especially on the substitution method. Be careful not to assume the upper bound you want to proof as like "$\leq n \lg n + \Theta(n)$".

※ lg means logarithm with base 2.