CSC373 Worksheet 3 Solution

July 30, 2020

1. Using the following formula

$$M[i,j] = \begin{cases} 0 & \text{if } i = j \\ \min_{i \leq k \leq j} M[i,k] + M[k+1,j] + p_{i-1}p_k p_j & \text{if } i < j \end{cases} \quad (1)$$

we have

C	1	2	3	4	5	6
1	0	350	770	612	1212	1422
2	x	0	840	462	1662	1362
3	x	x	0	252	1092	1098
4	x	x	x	0	1440	936
5	x	x	x	x	0	720
6	x	x	x	x	x	0

And an optimal parenthesization is

My Work:

$(A_1A_2)(A_3A_4)(A_5A_6)$

Correct Solution:

Using the following formula

$$M[i,j] = \begin{cases} 0 & \text{if } i = j \\ \min_{i \leq k \leq j} M[i,k] + M[k+1,j] + p_{i-1}p_k p_j & \text{if } i < j \end{cases} \quad (2)$$

we have

m (i/j)	1	2	3	4	5	6
1	0	350	770	612	1212	1422
2	x	0	840	462	1662	1362
3	x	x	0	252	1092	1098
4	x	x	x	0	1440	936
5	x	x	x	x	0	720
6	x	x	x	x	x	0

s (i/j)	1	2	3	4	5	6
1	0	1	2	1	4	4
2	x	0	2	2	4	4
3	x	x	0	3	4	4
4	x	x	x	0	4	4
5	x	x	x	x	0	5
5	x	x	x	x	x	0

And an optimal parenthesization is

My Work:

$(A_1(A_2(A_3A_4)))(A_5A_6)$

Notes:

- Sequence of Dimensions

  The sequence of dimensions $< p_0 = 5, p_1 = 10, p_2 = 3, p_2 = 12, p_3 = 5, p_4 = 50, p_5 = 6 >$ means there are 6 matrices with dimensions $p_{i-1} \times p_i$

  - $A_1 \to 5 \times 10$
  - $A_2 \to 10 \times 3$
  - $A_3 \to 3 \times 12$
  - $A_4 \to 12 \times 5$
  - $A_5 \to 5 \times 50$
  - $A_6 \to 50 \times 6$
- Dynamic Programming
  - Is applied to optimization problems
  - Applies when the subproblems overlap
  - Uses the following sequence of steps
    1. Characterize the structure of an optimal solution

2. Recursively define the value of an optimal solution
3. Construct an optimal solution from computed information

- Matrix-chain Multiplication
  - Is an optimization problem solved using dynamic programming
  - Goal is to find matrix parenthesis with fewest number of operations

  Example:

  Given chain of matrices $< A, B, C >$, it's fully parenthesized product is:

  * $(AB)C$ needs $(10 \times 30 \times 5) + (10 \times 5 \times 60) = 1500 + 3000 = 4500$ operations
  * $A(BC)$ needs $(30 \times 5 \times 60) + (10 \times 30 \times 60) = 27000$ operations

  Thus, $(AB)C$ performs more efficiently than $A(BC)$.

  - Is stated as: given a chain $< A_1, A_2, ..., A_n >$ of n matrices, where for $i = 1, 2, ..., n$ matrix A_i has dimension $p_{i-1} \times p_i$, fully parenthesize the product $A_1 A_2 ... A_n$ in a way that minimizes the number of scalar multiplications.
  - Steps
    1. **Check is the problem has Optimal Substructure**

       Let us adopt the notation $A_{i...j}$ where $i \leq j$, for the matrix that results from evaluating the product $A_i A_{i+1} ... A_j$.

       Assume the solution has the following parentheses:

       $$(A_{i...k})(A_{k+1...j})$$

       If there is a better way to multiply $(A_{i...k})$, then we would have a more optimal solution.

       This would be a contradiction, as we already stated that we have the optimal solution for $A_{i...j}$.

       Therefore, this problem has optimal substructure.

    2. **Find the Recursive Solution**

       Let $M[i, j]$ be the cost of multiplying matrices from A_i to A_j

       We want to find out at which $'k'$ returns the fewest number of multiplications, or the minimum number of M.

The recursive formula for the cost of multiplying from A_i to A_j is

$$M[i,j] = \begin{cases} 0 & \text{if } i = j \\ \min_{i \le k \le j} M[i,k] + M[k+1,j] + p_{i-1}p_kp_j & \text{if } i < j \end{cases} \tag{3}$$

3. **Computing the Estimated Cost**

   * Steps
     1) Fill the table for $i = j$
     2) Fill the table for $i < j$ with a spread of 1
     3) Repeat 2 with the increased value of spread

   Example:

   Given
   $< A_1, A_2, A_3, A_4, A_5 >$

   where

   * $A_1 \rightarrow 4 \times 10$
   * $A_2 \rightarrow 10 \times 3$
   * $A_3 \rightarrow 3 \times 12$
   * $A_4 \rightarrow 12 \times 20$
   * $A_5 \rightarrow 20 \times 7$

   we have:

   1) Fill the table for $i = j$

   1) i = j

i \ j	1	2	3	4	5
1	0				
2	x	0			
3	x	x	0		
4	x	x	x	0	
5	x	x	x	x	0

$$M[i,j] = \begin{cases} \boxed{0 \qquad\qquad \text{if } i = j} \\ \min_{i \le k \le j} M[i,k] + M[k+1,j] + p_{i-1}p_kp_j & \text{if } i < j \end{cases}$$

   2) Fill the table for $i < j$ with a spread of 1

2) (i = 1, j = 2), (i = 2, j = 3), (i = 3, j = 4), (i = 4, j = 5)

i \ j	1	2	3	4	5
1	0	120			
2	x	0	360		
3	x	x	0	720	
4	x	x	x	0	1680
5	x	x	x	x	0

since

* $i = 1, j = 2$

$$\begin{aligned} M[1,2] &= \min_{1\le k\le 2}(M[1,1] + M[1,2] + p_{i-1}p_kp_j) && (4)\\ &= \min_{1\le k\le 2}(0 + 0 + p_0p_1p_2) && (5)\\ &= \min_{1\le k\le 2}(0 + 0 + 4\cdot 10\cdot 3) && (6)\\ &= 120 && (7)\end{aligned}$$

where $p_0 = 3$ is from the dimension 3×10 of A_1, $p_k = 10$ is from the dimension of 3×10 of A_1.

* $i = 2, j = 3$

$$\begin{aligned} M[2,3] &= \min_{2\le k\le 3}(M[2,2] + M[3,3] + p_{i-1}p_kp_j) && (8)\\ &= \min_{2\le k\le 3}(0 + 0 + p_1p_2p_3) && (9)\\ &= \min_{2\le k\le 3}(0 + 0 + 10\cdot 3\cdot 12) && (10)\\ &= 360 && (11)\end{aligned}$$

* $i = 3, j = 4$

$$\begin{aligned} M[3,4] &= \min_{3\le k\le 4}(M[3,3] + M[4,4] + p_{i-1}p_kp_j) && (12)\\ &= \min_{3\le k\le 4}(0 + 0 + p_2p_3p_4) && (13)\\ &= \min_{3\le k\le 4}(0 + 0 + 3\cdot 12\cdot 20) && (14)\\ &= 720 && (15)\end{aligned}$$

* $i = 4, j = 5$

$$M[4,5] = \min_{4 \le k \le 5} (M[4,4] + M[5,5] + p_{i-1}p_kp_j) \tag{16}$$
$$= \min_{4 \le k \le 5} (0 + 0 + p_3p_4p_5) \tag{17}$$
$$= \min_{4 \le k \le 5} (0 + 0 + 12 \cdot 20 \cdot 7) \tag{18}$$
$$= 1680 \tag{19}$$

3) Repeat 2 with the increased value of spread

2) (i = 1, j = 2), (i = 2, j = 3), (i = 3, j = 4), (i = 4, j = 5)

i \ j	1	2	3	4	5
1	0	120	264	1080	1344
2	x	0	360	1320	1350
3	x	x	0	720	1140
4	x	x	x	0	1680
5	x	x	x	x	0

* $i = 1, j = 3$

$\underline{k = 1}$

$$M[1,3] = M[1,1] + M[2,3] + p_{i-1}p_kp_j \tag{20}$$
$$= 0 + 360 + p_0p_1p_3 \tag{21}$$
$$= 0 + 360 + 4 \cdot 10 \cdot 12 \tag{22}$$
$$= 0 + 360 + 480 \tag{23}$$
$$= 840 \tag{24}$$

$\underline{k = 2}$

$$M[1,3] = M[1,2] + M[3,3] + p_{i-1}p_kp_j \tag{25}$$
$$= 120 + 0 + p_0p_2p_3 \tag{26}$$
$$= 120 + 0 + 4 \cdot 10 \cdot 12 \tag{27}$$
$$= 120 + 0 + 144 \tag{28}$$
$$= 264 \tag{29}$$

Thus, $\min_{1 \le k \le 3} M[1,3] = 264$.

* $i = 2, j = 4$

  $\underline{k = 2}$

$$\begin{aligned} M[2,4] &= M[2,2] + M[3,4] + p_{i-1}p_kp_j && (30)\\ &= 0 + 720 + p_1p_2p_4 && (31)\\ &= 0 + 720 + 10 \cdot 3 \cdot 20 && (32)\\ &= 0 + 720 + 600 && (33)\\ &= 1320 && (34) \end{aligned}$$

  $\underline{k = 3}$

$$\begin{aligned} M[2,4] &= M[2,2] + M[3,4] + p_{i-1}p_kp_j && (35)\\ &= 360 + 0 + p_1p_3p_4 && (36)\\ &= 360 + 0 + 10 \cdot 12 \cdot 20 && (37)\\ &= 360 + 0 + 2400 && (38)\\ &= 2760 && (39) \end{aligned}$$

  Thus, $\min_{2 \leq k \leq 4} M[2,4] = 1320$.

* $i = 3, j = 5$

  $\underline{k = 3}$

$$\begin{aligned} M[3,5] &= M[3,3] + M[3,5] + p_{i-1}p_kp_j && (40)\\ &= 0 + 1680 + p_2p_3p_5 && (41)\\ &= 0 + 1680 + 3 \cdot 12 \cdot 7 && (42)\\ &= 0 + 1680 + 252 && (43)\\ &= 1932 && (44) \end{aligned}$$

  $\underline{k = 4}$

$$\begin{aligned} M[3,5] &= M[3,4] + M[5,5] + p_{i-1}p_kp_j && (45)\\ &= 720 + 0 + p_2p_4p_5 && (46)\\ &= 720 + 0 + 3 \cdot 20 \cdot 7 && (47)\\ &= 720 + 420 && (48)\\ &= 1140 && (49) \end{aligned}$$

  Thus, $\min_{3 \leq k \leq 5} M[3,5] = 1140$.

* $i = 2, j = 5$

  $\underline{k = 2}$

$$\begin{aligned}
M[2,5] &= M[2,2] + M[3,5] + p_{i-1}p_kp_j && (50)\\
&= 0 + 1140 + p_1p_2p_5 && (51)\\
&= 0 + 1140 + 10 \cdot 3 \cdot 7 && (52)\\
&= 0 + 1140 + 210 && (53)\\
&= 1350 && (54)
\end{aligned}$$

  $\underline{k = 3}$

$$\begin{aligned}
M[2,5] &= M[2,3] + M[4,5] + p_{i-1}p_kp_j && (55)\\
&= 360 + 1680 + p_1p_3p_5 && (56)\\
&= 2040 + 10 \cdot 12 \cdot 7 && (57)\\
&= 2040 + 840 && (58)\\
&= 2880 && (59)
\end{aligned}$$

  $\underline{k = 4}$

$$\begin{aligned}
M[2,5] &= M[2,4] + M[5,5] + p_{i-1}p_kp_j && (60)\\
&= 1320 + p_1p_3p_5 && (61)\\
&= 1320 + 10 \cdot 20 \cdot 7 && (62)\\
&= 1320 + 1400 && (63)\\
&= 2720 && (64)
\end{aligned}$$

  Thus, $\min_{2 \leq k \leq 5} M[2,5] = 1350$.

* $i = 1, j = 5$

  $\underline{k = 1}$

$$\begin{aligned}
M[1,5] &= M[1,1] + M[3,5] + p_{i-1}p_kp_j && (65)\\
&= 0 + 1350 + p_0p_1p_5 && (66)\\
&= 0 + 1350 + 4 \cdot 10 \cdot 7 && (67)\\
&= 0 + 1350 + 280 && (68)\\
&= 1630 && (69)
\end{aligned}$$

<u>$k = 2$</u>

$$\begin{aligned}
M[1,5] &= M[1,2] + M[3,5] + p_{i-1}p_kp_j && (70)\\
&= 120 + 1140 + p_0p_2p_5 && (71)\\
&= 120 + 1140 + 4 \cdot 3 \cdot 7 && (72)\\
&= 1260 + 84 && (73)\\
&= 1344 && (74)
\end{aligned}$$

<u>$k = 3$</u>

$$\begin{aligned}
M[1,5] &= M[1,3] + M[4,5] + p_{i-1}p_kp_j && (75)\\
&= 264 + 1680 + p_0p_3p_5 && (76)\\
&= 264 + 1680 + 4 \cdot 12 \cdot 7 && (77)\\
&= 1944 + 336 && (78)\\
&= 2280 && (79)
\end{aligned}$$

<u>$k = 4$</u>

$$\begin{aligned}
M[1,5] &= M[1,4] + M[5,5] + p_{i-1}p_kp_j && (80)\\
&= 1080 + 0 + p_0p_4p_5 && (81)\\
&= 1080 + 4 \cdot 20 \cdot 7 && (82)\\
&= 1080 + 560 && (83)\\
&= 1640 && (84)
\end{aligned}$$

Thus, $\min_{1 \le k \le 5} M[1,5] = 1344$.

4. **Constructing the Optimal Solution (Needs revision)**

3)

i \ j	1	2	3	4	5
1	0	$\boxed{120}$	264	1080	1344
2	x	0	360	1320	1350
3	x	x	0	$\boxed{720}$	1140
4	x	x	x	0	1680
5	x	x	x	x	0

$$(A_1A_2)((A_3A_4)A_5)$$

So, the optimal solution is $(A_1A_2)((A_3A_4)A_5)$

References:

1) CSBreakdown, Chain Multiplication - Dynamic Programming, link
2) University of Maryland, CMSC351 - Fall 2014 Homework # 4, link

2.
```
procedure MATRIX-CHAIN-MULTIPLY(A,s,i,j)
    if i == j then
        return A[i]
    end if

    if i < j
        a = MATRIX-CHAIN-MULTIPLY(A,s,i,s[i,j])
        b = MATRIX-CHAIN-MULTIPLY(A,s,s[i,j] + 1,j)

        return MATRIX-CHAIN-MULTIPLY(a,b)

    end if

end procedure
```

Example:

- *MATRIX-CHAIN-ORDER* computes the table s containing optimal costs
- s table consists of k value at which $m[i, j]$ is minimum!!

  $A_{1..s[1,n]}A_{s[1,n]+1..n}$

- Table of optimal costs m is used with table s to construct solution to matrix-chain multiplication problem

3. First, we need to determine the total number of times $m[i, j]$ is referred in the innermost loop.

   We know from the header that the loop runs from $k = i$ to $k = j - 1$.

   Using this fact, we can write the innermost loop has $j - i = l - 1$ iterations.

   Since $m[i.j]$ is referred twice, the total number of $m[i, j]$ referred in the loop is:

$$(l-1)2 \tag{1}$$

Second, we need to determine the total number of times $m[i, j]$ is referred in the intermediate loop

We know from the header that the loop runs from $i = 1$ to $i = n - l + 1$.

Using this fact, we can write the intermediate loop runs $n - l + 1$ iterations.

Since each iteration referrs $m[i, j]$ $(l - 1)2$ many times, the total number of times $m[i, j]$ is referred in the itnermediate loop is:

$$(n - l + 1)(l - 1)2 \tag{2}$$

Finally, we need to determine the total number of times $m[i, j]$ is referenced in the outermost loop.

We know from the header that the loop runs from $l = 2$ to n.

Since each iteration referrs $m[i, j]$ $(n - l + 1)(l - 1)2$ many times, the total number of times $m[i, j]$ is referred in the outermost loop is:

$$\begin{aligned}
\sum_{l=2}^{n}(n-l+1)(l-1)2 &= 2\sum_{l'=1}^{n-1}(n-l')(l') && (3)\\
&= 2\Big[\sum_{l'=1}^{n-1} nl' - (l')^2\Big] && (4)\\
&= 2\Big[n\sum_{l'=1}^{n-1} l' - \sum_{l'=1}^{n-1}(l')^2\Big] && (5)\\
&= 2\Big[n\sum_{l'=1}^{n-1} l' - \sum_{l'=1}^{n-1}(l')^2\Big] && (6)\\
&= 2\Big[n\sum_{l'=0}^{n-1} l' - \sum_{l'=0}^{n-1}(l')^2\Big] && (7)\\
&= 2\Big[\frac{(n-1)n^2}{2} - \frac{(n-1)n(2n-1)}{6}\Big] && (8)\\
&= 2\Big[\frac{n^3-n^2}{2} - \frac{(n-1)n(2n-1)}{6}\Big] && (9)\\
&= 2\Big[\frac{n^3-n^2}{2} - \frac{(n^2-n)(2n-1)}{6}\Big] && (10)\\
&= 2\Big[\frac{n^3-n^2}{2} - \frac{(3n^3-3n^2+n)}{6}\Big] && (11)\\
&= 2\Big[\frac{3n^3-3n^2}{6} - \frac{(3n^3-3n^2+n)}{6}\Big] && (12)\\
&= 2\Big[\frac{n^3-n}{6}\Big] && (13)\\
&= \frac{n^3-n}{3} && (14)
\end{aligned}$$

Notes:

- Hint from equation $A.3$

  $\sum\limits_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$
- I feel the need to gain more insight regarding the question's 'other table entries in a call of MATRIX-CHAIN-ORDER'. What does 'other table entries' mean?

  Answer:

The other entries :)

MATRIX-CHAIN-ORDER(p)

$n = p.length - 1$

let $m[1\,..\,n, 1\,..\,n]$ and $s[1\,..\,n-1, 2\,..\,n]$ be new tables

for $i = 1$ **to** n

    $m[i,i] = 0$

for $l = 2$ **to** n // l is the chain length

    for $i = 1$ **to** $n - l + 1$

        $j = i + l - 1$

        $m[i,j] = \infty$

        for $k = i$ **to** $j - 1$

            $q = m[i,k] + m[k+1,j] + p_{i-1}p_k p_j$

            if $q < m[i,j]$

                $m[i,j] = q$

                $s[i,j] = k$

return m and s

- In the problem '$m[i,j]$ is referenced' refers to $m[i,j]$ used in assignments $q = m[i,j]$...

MATRIX-CHAIN-ORDER(p)

$n = p.length - 1$

let $m[1\,..\,n, 1\,..\,n]$ and $s[1\,..\,n-1, 2\,..\,n]$ be new tables

for $i = 1$ **to** n

    $m[i,i] = 0$

for $l = 2$ **to** n // l is the chain length

    for $i = 1$ **to** $n - l + 1$

        $j = i + l - 1$

        $m[i,j] = \infty$

        for $k = i$ **to** $j - 1$

            $q = \boxed{m[i,k]} + \boxed{m[k+1,j]} + p_{i-1}p_k p_j$

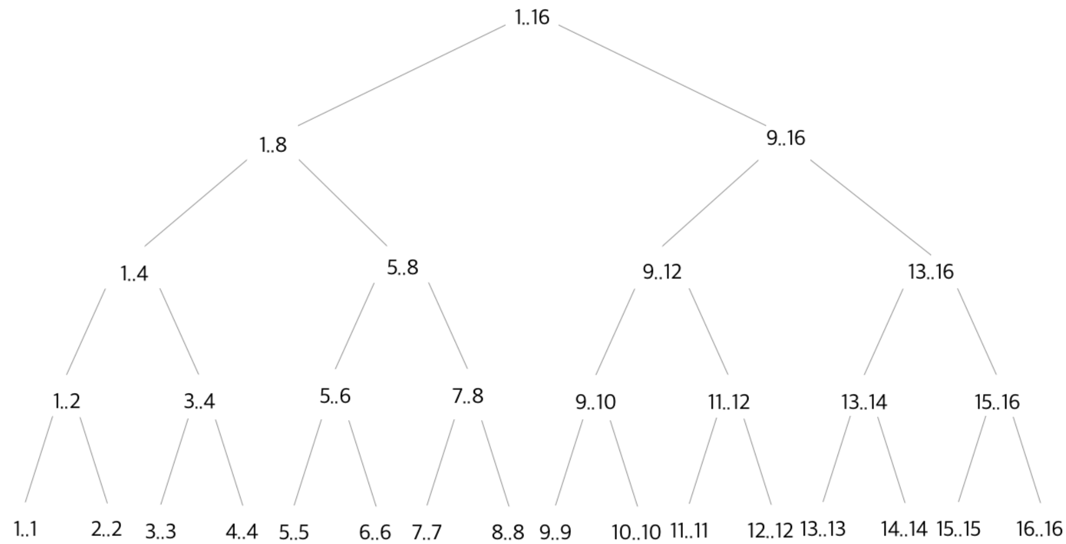
            if $q < m[i,j]$

                $m[i,j] = q$

                $s[i,j] = k$

return m and s

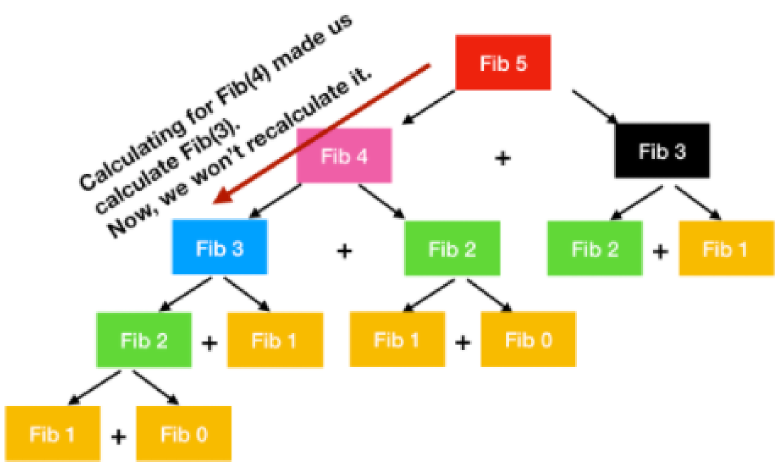
4. **Solution**



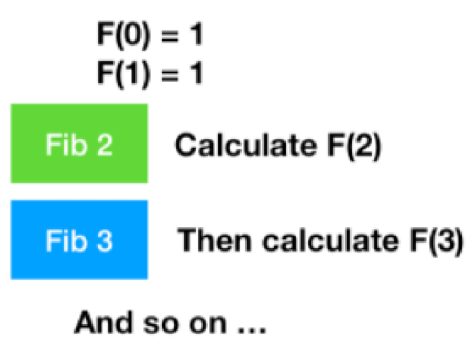
Memoization fails to speed up the algorithm because it lacks overlapping subproblems.

Notes:

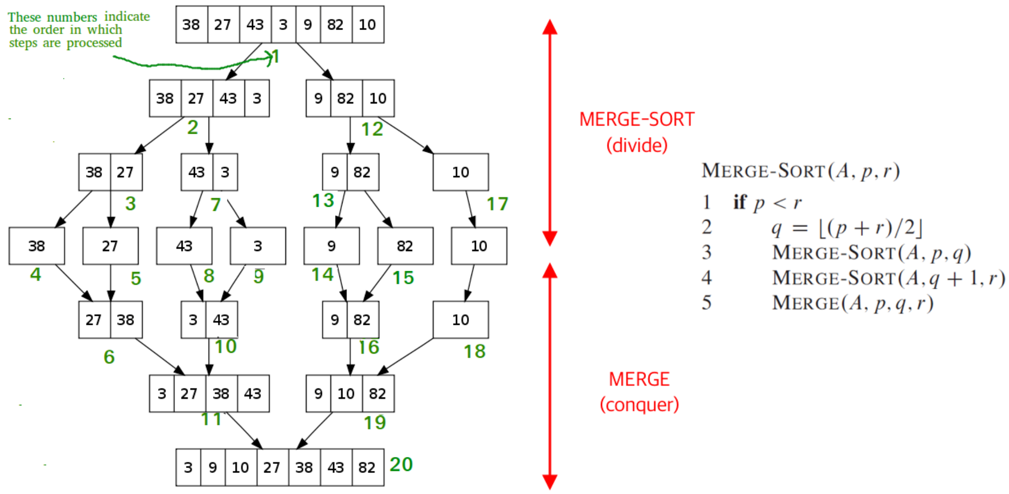
- Elements of Dynamic Programming
  - Optimal Substructure
    - Is the first step to solving dynamic programming problem
    - Exists if an optimal solution to the problem contains within it optimal solution to subproblems
  - Overlapping Subproblems
    - Is the second step to solving dynamic programming
    - Exists when an algorithm revisits the same problem repeatedly
  - Memoization
    - Maintains an entry in a table for the solution to each subproblem
    - Ensures that a method doesn't run for the same inputs more than once
- Top-Down Dynamic Programming
  - Uses recursion
  - Is preferred
    - When all sub-solutions need to be solved
    - Because it's easier



- Bottom-Up Dynamic Programming
  - Uses for-loop
  - Is preferred
    - When all sub-solutions need to be solved
    - Because it is sometimes faster (No recursive call and no unnecessary Random Memory access)
  - Is preferred when not all sub-solutions need to be computed



- Merge Sort
  - How it works
    1. Find the middle point to divide the array into two halves
    2. Call mergeSort for first half
    3. Call mergeSort for second half
    4. Merge two halves in sorted order



References

1)

5. Yes. This problem does exhibit optimal substructure

Proof. Assume that the optimal structure of $A_iA_{i+1}..A_j$ exists. That is, there exists some $k \in \mathbb{N} - \{0\}$ such that the following parenthesis $(A_{i..k})(A_{k+1..j})$ produces maximum operation cost.

I need to show that the substructures $A_{i..k}$ and $A_{k+1..j}$ are also optimal.

I will do so in parts.

Part 1 (Proving optimal substructure of $A_{i..k}$)

Assume for the sake of contradiction that the substructure $A_{i..k}$ is not optimal.

Then, we can write that there exists some $k' \in \mathbb{N} - \{0\}$ such that $(A_{i...k'})(A_{k'+1..k})$ has larger operation cost than $(A_{i..k})$.

Then, we can write that $((A_{i...k'})(A_{k'+1..k}))(A_{k+1..j})$ has larger operation costs than $(A_{i..k})(A_{k+1..j})$, which contradicts the original assumption that $(A_{i..k})(A_{k+1..j})$ has maximum operation cost.

Thus, $(A_{i..k})$ must be optimal.

Part 2 (Proving optimal substructure of $A_{k+1..j}$)

This proof is nearly verbatim as part 1, where the only difference is using $A_{k+1..j}$ instead of $A_{i..j}$. □

Notes:

- A problem has **optimal substructure** if an optimal solution can be constructed from optimal solutions of its subproblems. [3]
- Showing optimal subtructure for the original Matrix-Chain Multiplication problem

> Assume that the optimal structure of $A_i A_{i+1}..A_j$ exists. That is, there exists some $k \in \mathbb{N} - \{0\}$ such that the following parenthesis $(A_{i..k})(A_{k+1..j})$ produces minimum operation cost.
>
> I need to show that the substructures $A_{i..k}$ and $A_{k+1..j}$ are also optimal.
>
> I will do so in parts.
>
> **Part 1 (Proving optimal substructure of $A_{i..k}$)**
>
> Assume for the sake of contradiction that the substructure $A_{i..k}$ is not optimal.
>
> Then, we can write that there exists some $k' \in \mathbb{N} - \{0\}$ such that $(A_{i...k'})(A_{k'+1..k})$ has smaller operation cost than $(A_{i..k})$.
>
> Then, we can write that $((A_{i...k'})(A_{k'+1..k}))(A_{k+1..j})$ has smaller operation costs than $(A_{i..k})(A_{k+1..j})$, which contradicts the original assumption that $(A_{i..k})(A_{k+1..j})$ has minimum operation cost.
>
> Thus, $(A_{i..k})$ must be optimal.
>
> **Part 2 (Proving optimal substructure of $A_{k+1..j}$)**
>
> This proof is nearly verbatim as part 1, where the only difference is using $A_{k+1..j}$ instead of $A_{i..j}$.

References:

1) CSBreakdown, Chain Multiplication - Dynamic Programming, link
2) CodeScope, Dynamic Programming, link
3) Wikipedia, Optimal Substructure, link

6. Let $< 2, 10, 20, 5 >$.

Then, we see that $p_0 p_1 p_3$ is minimum, and by Professor Capulet's claim, splitting at $k = 1$ should result in minimum-cost matrix multiplication.

But we have

m	1	2	3
1	0	400	600
2	x	0	1000
3	x	x	0

s	1	2	3
1	0	1	2
2	x	0	2
3	x	x	0

And the minimum-cost matrix multiplication occurs when $A_{i..j}$ is splitted at $k = 2$.

Thus, Professor Capulet's claim is false.

Notes:

- I need to find the matrices $A_i, ..., A_j$ where the total operating cost of Professor Capulet's method is bigger than the properly parenthesized solution
- I feel the need for clarafication regarding the phrase 'always choosing the matrix A_k at which to split the product...' Is k in A_k the same in any matrix multiplications $A_i A_{i+1} ... A_j$?

  Answer:

  No. k in A_k is the value that makes $p_{i-1}p_k p_j$ minimum.

7. **Notes:**

- Longest Common Sequence