UM-SJTU JOINT INSTITUTE

Introduction to Algorithms
(VE477)

Homework #5

Prof. Manuel

Xinmiao Yu
518021910792

Nov. 1, 2020

Q1.

1. *Linear partition problem:* A given arrangement S consisting n nonnegative integers $s_1, ..., s_n$ and an integer k, partition S into k ranges so as to minimize the maximum sum over all the ranges.

   Often arises in parallel processing, because of the demand to balance the work done across processors so to minimize the total elapsed run time.

2. No, not a good solution. It will not systematically evaluate all the possibilities. Consider $S = \{3, 3, 3, 2, 5, 5\}$ and $k = 3$. With the this approach we have the average size of partition as $21/3 = 7$, then the set will be divided as $3, 3 \mid\mid 3, 2 \mid\mid 5, 5$ with maximum sum as 10. However, the solution should be $3, 3, 3 \mid\mid 2, 5 \mid\mid 5$ with maximum sum as 9.

3. The problem could be transformed into finding the minimum value of the lager one between 1) cost of the last partition $\sum_{j=i+1}^{n} s_j$ and 2) the cost of the largest partition cost formed to the left of i.

$$M(n, k) = min_{i=1}^{n}\ max(M[i, k-1], \sum_{j=i+1}^{n} s_j)$$

$$\text{With Basis M[1, k]} = s_1, \forall k > 0$$

$$M[n, 1] = \sum_{i=1}^{n} s_i$$

4. If keep $M[i][j]\ \forall i \leq n, j \leq k$, total cell will be $k \cdot n$ in this table. For any $M[n'][k']$, need to find the minimum among n' quantities, each of which is the maximum through table lookup and a sum of at most n' elements. Then fill each cell need $\mathcal{O}(n^2)$. So total need $\mathcal{O}(kn^3) = \mathcal{O}(n^3)$

5. When update each cell, instead of selecting the best of up to n possible points to place the divider, each of which need to sum up to n possible terms, we could store the set of n prefixes sum

$$p[i] = \sum_{k=1}^{i} s_k,\ \text{since p[i] = p[i - 1] + s}_\text{i}$$

Problem 1 continued on next page. . .

6. Dynamic programming approach

Algorithm 1: Linear Partition Problem

Input : arrangement S consisting n nonnegative integers $s_1, ..., s_n$ and an integer k
Output: the cost of the largest range when partition S into k ranges so as to minimize the maximum sum over all the ranges

```
/* compute prefix sum                                   */
p[0] ← 0
for i ← 1 to n do
    p[i] ← p[i − 1] + s_i
end for
/* boundary condition                                   */
for i ← 1 to n do
    M[i, 1] ← p[i]
end for
for j ← 1 to k do
    M[1, j] ← s_1
end for
/* evaluate main recurrence                             */
for i ← 2 to n do
    for j ← 2 to k do
        M[i, j] ← ∞
        for x ← 1 to i - 1 do
            s ← max(M[x, j − 1], p[i] − p[x])
            if M[i,j] > s then
                M[i, j] ← s
                D[i, j] ← x ;                /* used to reconstruct */
            end if
        end for
    end for
end for
return M[n,k]
```

7. First the prefix sum and boundary condition is obviously true. It settle the smallest possible values for each of the arguments of the recurrence. With the evaluation order such that computes the smaller values before the bigger values, it will obtain the right result as long the previous results are true, which must be true as the boundary conditions are true.

8. When update each cell, we do not need to select the best among n possible points to place the divider because of the prefix sum we stored. So for each call, only need linear time. Then the total time complexity would be $\mathcal{O}(kn^2)$.

9. This could be achieved by D, as it record which divider position required to achieve such cost. So to reconstruct the path used to get to the optimal solution, we work backward from

$D[n, k]$ and add a divider at each specified position.

Algorithm 2: Reconstruct

Input : S, D, n, k

Output: S with divider

```
Function Reconstruct(S, D, n, k):
    if k == 1 then
        print the first partition (s_1, s_2, ..., s_n)
    else
        Reconstruct(S, D, D[n,k], k-1)
        print the k-th partition (s_{D[n,k]+1}, ..., s_n)
    end if
end
```

Q2.

As B would produce number $0, 1, 2, 3, 4$ with equal probability $1/5$, we could get number in range $[0, 24]$ by $B * 5 + B$ with equal probability $1/25$. And if drop the number that larger than 23, the probability for generating number in $[0, 23]$ is $1/24$. Then the number in $[0, 7]$ with equal probability could be obtained by $[0, 23]/3 = 1/8$, which returns the integer part of the result.
To extend the generation procedure, the critical part is to generate numbers larger than expected n ($[0, 24]$ previously) with equal probability.

- Denote the original B that produce $[0, 4]$ as B_0
- Denote that produce $[0, 24]$ as B_1 such that $B_1 = 5 * B_0 + B_0 = (5 + 1)$
- $B_2 = 25 * B_0 + B_1$ produce $[0, 124]$ with equal probability as $1/125$
- $B_3 = 125 * B_0 + B_2$ produce $[0, 624]$ with equal probability $1/625$

Then summarize as

$$Range[B_n] = [0, 5^{n+1} - 1], \text{ with P} = \frac{1}{5^{\mathrm{n+1}}}$$

Therefore, we could use the original B, to have any generator B_i we need to produce number in range $[0, 5^{i+1} - 1]$. Restriction on n will be $n \geq 0$.
As in the previous example, the random number in $[0, 7]$ is obtained through $(B_1.output < 24)/3$. Because the range is $[0, 5^2 - 1] = [0, 24]$, which is too large if we just simply keep the number that $B_1.output \leq 7$. So we could apply the same method that find an integer a such that

$$a * n < 5^{i+1} - 1, \quad where\ 5^i - 1 < n \leq 5^{i+1} - 1$$

The the random number is $B_i.output/a$

Algorithm 3: Random Number Generator

Input : nonnegative integer n

Output: a random number in range $[0, n]$

Find i that $5^i - 1 < n \leq 5^{i+1} - 1$
$a \leftarrow 1$
while $(a+1)*(n+1) \leq 5^{i+1} - 1$ **do**
  $a \leftarrow a + 1$
end while
Get the random number generator B_i
$num \leftarrow B_i.output$
while $num > (n+1)*a$ **do**
  $num \leftarrow B_i.output$
end while
return num/a

Q3.

Bellman-ford algorithm

Algorithm 4: Detect negative cycle

Input : weighted graph $G = (V, E)$

Output: whether the graph has negative cycle

Chosen a vertex s randomly
`/* Initialization */`
for *each vertex* $v \in G.V$ **do**
  $v.d \leftarrow \infty$
end for
$s.d \leftarrow 0$
`/* Relax */`
for $i \leftarrow 1$ **to** $|G.V| - 1$ **do**
  for *each edge* $(u, v) \in G.E$ **do**
    if $v.d > u.d + w(u,v)$ **then**
      $v.d \leftarrow u.d + w(u, v)$
    end if
  end for
end for
for *each edge* $(u, v) \in G.E$ **do**
  if $v.d > u.d + w(u,v)$ **then**
    return *True* ; `/* only possible when negative cycle exists */`
  end if
end for
return *False*

Q4.

1. Obviously $1 \leq k \leq n$, otherwise the statement would not be true. To prove, because the hash table has n slots and the probability of n keys to hash to any slot is equal, for each key, it has a probability of $\frac{1}{n}$ to hash to any slot. So the number of keys hash to a same slot follows a binomial distribution with parameter n and p, where $p = \frac{1}{n}$. Then the probability for exactly k keys hash to a same plot is

$$P_k = \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k} \binom{n}{k}.$$

2. The probability of a slot to have k keys is P_k. More than one slots may have k keys, but no slots would have more than k keys. Then, P'_k =the probability of at least one slot has k keys and other slots have no more than k keys, which is **smaller or equal** to the probability that at least one slot has k keys. Because we have n slots, the probability of only one slot has k keys is $\binom{n}{1} P_k = nP_k$, and this probability is **larger or equal** to the probability that at least one slot has k keys. Through this two inequality, $P'_k \leq nP_k$.

3. We have Stirling formula $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, and $1 \leq k \leq n$

$$\begin{aligned}
P_k &= \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k} \frac{n!}{k!(n-k)!} \\
&\approx \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k} k!} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k} \\
&= \sqrt{\frac{n}{n-k}} \left(\frac{n}{n-k} \frac{n-1}{n}\right)^n \left(\frac{n-k}{e} \frac{1}{n} \frac{n}{n-1}\right)^k \frac{1}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^k} \\
&< \sqrt{\frac{n}{2\pi k(n-k)}} \frac{e^{-k}}{\left(\frac{k}{e}\right)^k} \\
&< \frac{e^k}{k^k}
\end{aligned}$$

Q5.

Suppose G is an undirected graph G with weighted edges and the weight of an edge e is decreased where $e \notin T$, $e = (u, v)$.

Algorithm 5: Algorithms in the homework

Input : this file

Output: nice algorithms in the homework

```
Function Reconstruct(this file):
    download file;
    open file;
    compile file;
    while not at end of this document do
        read;
        if understand then
            go to next line;
            current line becomes this one;
        else if want to know more on algorithms in LaTeX then
            refer to algorithm2e documentation
        else
            restart reading from the beginning;
        end if
    end while
    for exercise ← 1 to 7 do
        if algorithm is requested then
            solve the problem;
            A[exercise] ← write the algorithm in LaTeX;
        end if
    end for
    return A
end
```

Q6.