1 Loop Invariant and Iterative Algorithm

Consider the following algorithm that takes as input a list of integers $(A[1], \ldots, A[n])$ and an integer $1 \leq k \leq n$ and outputs the array of sums of windows of length k, $S[I] = \sum_{i=I}^{i=I+k-1} A[i]$ for $1 \leq I \leq n-k+1$. For example if the input is: $(53, 10, 4, 7, 12, 18, 72, 71)$ and $k = 3$, then the output should be $S[1..6] = 67, 21, 23, 37, 102, 161$.

procedure `SumWindows`$(A[1], \ldots, A[n]; k)$: $1 \leq k \leq n$.)

1. $S = A[1]$
2. For $J = 2$ to k do: $S = S + A[J]$
3. $S[1] = S$
4. For $I = 2$ to $n - k + 1$ do:
5. $\quad S = S - A[I-1] + A[I+k-1]$
6. $\quad S[I] = S$.
7. Return $S[1..n-k+1]$.

Consider the following loop invariant:

For $1 \leq j \leq n-k+1$, *after the iteration of the second loop when* $I = j$, $S = \sum_{i=j}^{i=j+k-1} A[i]$

Fill in the blanks of the proof that the loop invariant is correct and proves the correctness of the algorithm.

(a) (3 points) **Base Case:** At $j = 1$,before the loop starts, state why the loop invariant is true :

Before the loop starts, $j = 1$, the loop in line 2 sets S to $\sum_{i=1}^{i=k} A[i] = \sum_{i=1}^{i=1+k-1} A[i]$, so the invariant is true for $j = 1$.

Inductive Hypothesis: Suppose that for some $1 \le j \le n-k$, after the iteration when $I = j$, $S = \sum_{i=j}^{i=j+k-1} A[i]$.

Inductive Step: Show that after the next iteration. $S = \sum_{i=j+1}^{i=j+k} A[i]$.

Let S_0 be the value of S before the loop when $I = j+1$, and S_1 its value after the loop. Then in line 5 of the algorithm, S_1 is set to $S_0 - A[(j+1)-1] + A[j+1+k-1] = S_0 - A[j] + A[j+k]$. By the inductive hypothesis, $S_0 = \sum_{i=j}^{i=j+k-1} A[i] = A[j] + \sum_{i=j+1}^{i=j+k-1} A[i]$. So $S_1 = (\sum_{i=j+1}^{i=j+k-1} A[i]) + A[j+k] = \sum_{i=j+1}^{i=j+k} A[i]$, and the invariant is true after the next iteration.

(b) Use the loop invariant to show that all values of the output array $S[I]$ are correct.

In line S, at the end of the loop when $I = j$, we set $S[j] = S$. By the invariant, this means we set $S[j] = \sum_{i=j}^{i=j+k-1} A[i]$ for each $1 \le j \le n-k+1$, which is the definition of the problem.

2 Recursive Algorithm

Consider the following *recursive* algorithm that solves the same problem as above, i.e. it takes as input a list of integers $(A[1], \ldots, A[n])$ and a $1 \le k \le n$ and returns the array $S[1..n-k+1]$ where $S[I] = \sum_{i=I}^{i=I+k-1} A[i]$.

procedure `WindowSumRec`$(A[1\ldots, n], k)$

1. **if** $n = 1$ return the array $(A[1])$.
2. **if** $n = k$:
3. $\quad S[1] = WindowSumRec(A[1..n-1], k-1)$
4. $\quad$ Return the array $(S[1] + A[n])$.
5. $S[1..n-k] = WindowSumRec(A[1..n-1], k)$
6. $S[n-k+1] = S[n] - A[n-k-1] + A[n]$ **return** $S[1..n-k+1]$

(a) Write the recursion for the runtime $T(n)$ for this algorithm assuming arithmetic takes constant time.

If $n = 1$, we return in constant time in line 1. If $n > 1$, we either make one recursive call to an instance of size $n-1$ in line 3 or in line 5. The other lines take constant time. Thus, $T(n) = T(n-1) + c$ for $n \ge 2$,$T(1) = c$ for some c.

(b) Solve the recursion in the previous part to give the order of the runtime $T(n)$ for this algorithm assuming arithmetic takes constant time.

If $T(n) = T(n-1) + c$, for $n \ge 2$, $T(n) = c + T(n-1) = 2c + T(n-2) = ...kc + T(n-k)$. Using this with $k = n-1$, $T(n) = (n-1)c + T(1) = nc \in \Theta(n)$ time.

3 Counting

(15 points total) A license plate consists of 7 characters (each character can be any of the 10 digits or the 26 letters of the alphabet.)
For each problem, you can leave your answer in terms of exponentials, factorials, or binomial

(a) If the first three characters must be numbers and the last four characters must be letters then how many different license plates are there?

By the product rule: $(10)^3(26)^4$

(b) If the first three characters must be different numbers in strictly increasing order and the last four characters must be different letters in alphabetical order then how many different license plates are there?

Each subset of three numbers and each subset of four letters determine such a plate uniquely: $\binom{10}{3} * \binom{26}{4}$

(c) If all 7 characters must be digits in non-decreasing order, then how many different license plates are there?

We need to use stars and bars, with a star representing reusing the same number, starting at 0, and a bar representing moving on to the next number. We have 10 digits, so nine bars, and seven places for *'s, so the total is $\binom{}{169}$

(d) Suppose that all 7 characters must be letters. There is a very bad 4-letter word (all letters are distinct, let's say the bad word is "MATH") that you cannot have on the license plate. How many license plates avoid this 4-letter word as a consecutive substring?

We subtract the combinations including the bad word from the total. There are four places where the bad word could start, and we can't have the same word in multiple places, since all the letters are distinct. So the total is $(26)^7 - 4(26)^3$.

(e) Suppose that any character can be a digit or a letter but each license plate must have at least one digit and at least one character.

We use inclusion-exclusion, but note that no plate can have both no letters and no numbers, so the last term is 0. $(36)^7 - (26)^7 - (10)^7$

Scratch work only. This page will not be graded.

4 Graphs

(10 points) Suppose you are renting your apartment out for D days and you are using Airbnb to find people to rent it. Airbnb finds n people to rent it and gives you the start and end days of each person. You do not want your apartment to be empty on any day that you are gone. You cannot rent to two people at the same time and checkout is in the morning and checkin is in the afternoon so you could have a guest arrive on the same day that the previous guest left. Design a graph that models this problem and express what graph problem you must solve in order to decide if it is possible to rent your apartment with no vacant days.

(a) (2 points) Specify the vertices for your graph

There are two correct approaches that students found. One is to make the vertices the days you are gone, and the other to make the vertices the prospective renters.

(b) (2 points) Specify the edges for your graph

In the first version, for each renter, we put a directed edge from the renters arrrival date to the renters checkout date. In the second version, for each pair of renters, we put a directed edge from the first to the second if the checkout date for the first equals the arrival date of the second.

(c) (2 points) Draw your graph for the following example where you want to rent your apartment continuously from day 1 to day 10 and the renters are interested in the following intervals:

$$(1,5),(2,3),(8,10),(3,8),(1,2),(6,8),(4,10)$$

Omitted.

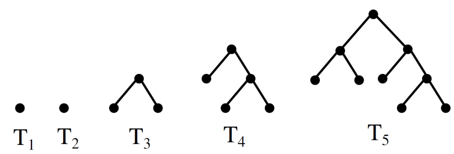
(d) (4 points) What graph problem would solve this problem and what graph algorithm would

In both versions, we are trying to determine the existence of appropriate paths. For the first, we want a path from the day you leave to the day you return. For the second, we want to find a path from some guest arriving on day 1 to some guest checking out on day D. You can use any version of graph search to solve these problems; for the first, we search from vertex 1, and see if day D is reachable. For the second, we search from each starting vertex arriving at 1, and see if any checking out on day D is reachable. (Alternatively, you can add a dummy start vertex with edges to each guest arriving on day 1).

Scratch work only. This page will not be graded.

5 Trees and Recursive counting

(9 points) Let T_1 be the rooted tree consisting of a single vertex and let $T_2 = T_1$.
For $n \geq 3$, let T_n be the rooted tree whose left subtree is T_{n-2} and whose right subtree is T_{n-1}.
The first few trees T_n for $n \leq 5$ is given in the figure below.



Give a recurrence for each of the following. No explanation needed.

(a) $E(n)$, the number of edges in T_n

$E(n) = E(n-1) + E(n-2) + 2$ for $n \geq 3$, $E(1) = 0, E(2) = 0$

Base case(s):

(b) $V(n)$, the number of vertices in T_n

$V(n) = V(n-1) + V(n-2) + 1$ for $n \geq 3$, $V(1) = 1, V(2) = 1$

Base case(s):

(c) $L(n)$, the number of leaves in T_n

$L(n) = L(n-1) + L(n-2)$ for $n \geq 3$, $L(1) = 1, L(2) = 1$

Base case(s):

6 Expected Value

(10 points) Given a sequence of n integers: $(A[1], \ldots, A[n])$ such that each integer is in the range $(1, \ldots, n)$, a fixed point is any entry of the sequence where $A[i] = i$.

(a) If each entry of the sequence is selected independently from the range $(1, \ldots, n)$, compute the expected number of fixed points.

For example: $(2, 2, 3, 5, 1)$ then $A[2] = 2, A[3] = 3$ so this has 2 fixed points.

Each element maps to a random value so has a $1/n$ chance of being a fixed point. Since there are n elements, by additivity of expectations, the expected value is $n * 1/n = 1$.

(b) A sequence of this sort is called a permutation if all values $(1, \ldots, n)$ each appear exactly once. If A is a permutation selected uniformly at random, compute the expected number of fixed points.

For example: $(4, 3, 2, 5, 1)$ has no fixed points.

Although there are now dependencies between the events, the observations above are the same: Each individual element maps to a random value so has a $1/n$ chance of being a fixed point. Since there are n elements, by additivity of expectations, the expected value is $n * 1/n = 1$.

Scratch work only. This page will not be graded.

7 Probability

(12 points) In the game of craps, you roll 2 fair six-sided dice.

(a) What is the probability that after rolling two dice, the sum of the two dice is 7 or 11? There are six ways of geting a 7, 2 of 11. So the total is $8/36 = 2/9$

(b) What is the probability that after rolling two dice, the sum of the two dice is 7 or eleven if you know that at least one of the dice is a six?

There are 11 combinations where at least one is a 6. Two sum to 7, two sum to 11. So the conditional probability is 4/11.

(c) What is the probability that after rolling two dice, the sum of the two dice is 7 or eleven if you know that at least one of the dice is greater than 3?

If no die is greater than 3, it is impossible to sum to 7 or 11. The probability of not rolling larger than 3 either time is 1/4. $Prob[7or11|atleastone > 3] = Prob[7or11andatleastonedie > 3]/Prob[Atleastonedie > 3] = Prob[7or11]/Prob[Atleastonedie > 3] = 2/9/(3/4) = 8/27$

(d) Suppose you roll the pair of dice over and over until you get a 7 or a 10. What is the probability that you roll a 10 before you roll a 7?

It doesn't matter how many times you rolled previously. The last roll is either a 7 or 10, so the probability that it is a 10 is the conditional probability of rolling 10 given that you rolled 7 or 10. Since there is a $3/36 = 1/12$ chance of rolling 10, and a $1/6$ chance of rolling 7, this is $1/12/(1/12 + +1/6) = 1/3$.

Scratch work only. This page will not be graded.

8 Multiple Choices

(18 points) No justification necessary, please write the *letter* clearly in the box to the right of the problem.

1. True or false: If $n \in \Omega(f(n))$ then $\sqrt{n} \in \Omega(f(n))$.

   (A) True (B) False

   B 1.

2. True or false: If $\lim_{n\to\infty} f(n)/g(n) = 1000$ then $f(n) \in O(g(n))$

   (A) True (B) False

   A 2.

3. True or false: $n^2 \log(n) \in \Theta(n \log(n^2))$.

   (A) True (B) False 3.

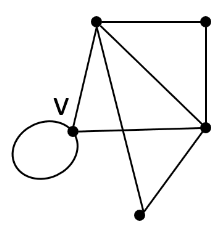
   B

4. True or false: If $T(n) = T(n-1) + T(n-2) + O(1)$ and $T(1) = c$ then $T(n) \in O(2^n)$.

   (A) True (B) False 4.

   A

5. The degree of vertex **v** in the graph below is



   (A) 1 (B) 2 (C) 3 (D) 4 (E) None of the above 5.

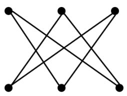
   D: 4. Loops count as degree 2.

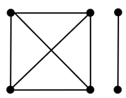
6. Suppose an undirected graph has eight vertices: two vertices of degree 3, three vertices of degree 4, one vertex of degree 2, and two isolated vertices. The number of edges in the graph is

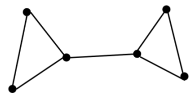
   (A) 9 (B) 15 (C) 20 (D) 10 (E) 30 6.

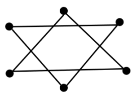
Scratch work only. This page will not be graded.

7. Which of the following graphs has an Eulerian circuit? (note: the vertices are the bold dark nodes, there are not necessarily vertices at the intersection of edges.


(A)


(B)


(C)


(D)

(E) None of the above

7.

Only graph A. Graphs B and D are disconnected, and C has two odd degree vertics, so it has an Eulerian path, but not a circuit.

8. The number of simple, undirected graphs with 3 (labelled) vertices is

$(A)\ 2^3$ $(B)\ 2^8$ $(C)\ 2^9$ $(D)\ 2^{27}$ (E) None of the above

8.

There are three edge positions, which can either be present or absent so A.

9. Consider the following algorithm

procedure *PrintSomething* (n: a power of 2)

```
for i = 1 to n
  j = n
  while j > 1
      print (i, j)
      j = j - 1
```

The runtime of this algorithm is

$(A)\ \Theta(\log n)$ $(B)\ \Theta(n)$ $(C)\ \Theta(n \log(n))$ $(D)\ \Theta(n^2)$ (E) None of the above

D. 9.

10. Suppose $T(n) = 4T(n/2) + n^2$ then the big-O class for $T(n)$ given by the Master Theorem (MT) will be

$(A)\ O(n)$ $(B)\ O(n^2)$ $(C)\ O(n^2 \log(n))$ $(D)\ O(n \log n)$ (E) None of the above

C.
10.

11. Suppose $T(n) = 7T(n/2) + n^2$ then the big-O class for $T(n)$ given by the Master Theorem (MT) will be

$(A)\ O(n)$ $(B)\ O(n^2)$ $(C)\ O(n^2 \log(n))$ $(D)\ O(n \log n)$ (E) None of the above

$\boxed{\text{$a = 7, b = 2, d = 2$. Since $7 > 2^2$, the master theorem gives $T(n) \in O(n^{log_2 7})$., so E.}}$

11.

Scratch work only. This page will not be graded.

12. Given a set of n^2 objects, under a **fixed-length** encoding scheme, what is the minimum number of bits required to encode each element in this set?

(A) 2^n (B) $\binom{n}{2}$ (C) 4^n (D) $\lceil 2\log(n) \rceil$ (E) none of the above

12.

D

13. Let T_n be the set of ternary strings (strings with 0, 1, and 2 bits) of length n that never have the same symbol appearing twice in a row. For example, 01212010 is such a string of length $n = 8$. Under a **fixed-length** encoding scheme, what is the minimum number of bits required to encode each element in T?

(A) n (B) $n+1$ (C) 3^n (D) $\lceil n\log(3) \rceil$ (E) none of the above

13.

The first number takes 2 bits, then each other takes 1. So B.

14. Suppose that A and B are events with probabilities $P(A) = 2/3$ and $P(B) = 2/3$. What is the largest $P(A \cup B)$ can be?

(A) 0 (B) $\frac{1}{3}$ (C) $\frac{4}{9}$ (D) $\frac{2}{3}$ (E) 1 (F) $\frac{7}{6}$ (G) not enough information

14.

E. No event can be greater than 1 probability.

15. Suppose that A and B are events with probabilities $P(A) = 2/3$ and $P(B) = 2/3$. What is the smallest $P(A \cap B)$ can be?

(A) 0 (B) $\frac{1}{3}$ (C) $\frac{4}{9}$ (D) $\frac{2}{3}$ (E) 1 (F) $\frac{7}{6}$ (G) not enough information

B.
15.

16. Suppose that events E and F are disjoint with $P(E) \neq 0$ and $P(F) \neq 0$. Which of the following can never be true:

(A) $P(E) = P(F)$ (B) $P(E) + P(F) < 1$ (C) $P(E|F) = P(E)$

(D) $P(E \cup F) = P(E) + P(F)$ (E) none of the above

16.

D cannot be true, since disjoint nonzero events are always dependent.

17. Suppose that events E and F are independent and $P(E) \neq 0$ and $P(F) \neq 0$. Then which of the following is false:

(A) $P(E|F) = P(F|E)$ (B) $P(E|F) = \frac{P(E \cap F)}{P(F)}$ (C) $P(E \cap F) = P(E)P(F)$

17.

The answer we were looking for was A, but logicians also forced us to accept E.

18. Suppose that X and Y are random variables (not necessarily independent.) Which of the following is false?

(A) $E(X+Y) = E(X) + E(Y)$ (B) $V(X) = E(X^2) - E(X)^2$ (C) $\sigma(X) \geq V(X)$

(D) $E(XY) = E(X)E(Y)$ (E) more than one of the above

18.

C and D are usually false, so the correct answer is E.

END