

# CSCI 5870: DATA STRUCTURES AND ALGORITHMS

## HOMEWORK 2

INSTRUCTOR: DR. BOB KRAMER

RABIN THAPA

### Part I. 10 points each.

#### 1. [BvG 3.1]

**Show that every 2-tree with  $n$  internal nodes has  $n + 1$  external nodes.**

*Proof:* (using Induction)

Let  $T_i$  denote the set of 2-tree(s) with  $i$  nodes where  $i$  is an odd integer ( $\because$  starting from a trivial (single node) 2-tree, the order of next "big" tree in the sequence increase by 2). Let  $I(T)$  denote the number of internal nodes and  $E(T)$  denote the number of external nodes of a 2-tree  $T$ .

We need to show  $E(T) = I(T) + 1$  for any 2-tree.

Base Case ( $i = 1$ ): Let  $T \in T_1$ .  $T$ , which is only the root node is an external node. So,  $I(T) = 0$  and  $E(T) = 1$ , hence  $E(T) = I(T) + 1$ .

Inductive Step: Let's assume  $E(T) = I(T) + 1$  is true for any 2-tree in  $T_i$  for  $i = k$ . We need to show the condition holds for  $T_{k+2}$ .

Let  $T \in T_{k+2}$ . Let's consider a tree  $T'$ , a subgraph of  $T$  formed by removing both the leaves ( $v_1$  and  $v_2$ ) from a same parent node  $v \in V(T)$  ( $v$  located at height  $h - 1$  where  $h$  is the height of  $T$ ). Here,  $T' \in T_k$  which means, by hypotheses,  $E(T') = I(T') + 1$ . Now, let's add those leaves ( $v_1$  and  $v_2$ ) back to  $v$  to form  $T' + v_1 + v_2 = T$ . So,  $v_1, v_2$  are external nodes and  $v$  is now an internal node. This gives  $I(T) = I(T') + 1$  and  $E(T) = E(T') + 2 - 1 = E(T') + 1 = (I(T') + 1) + 1 = I(T) + 1$ . Thus, the condition still holds.

Hence, by induction it is true that every 2-tree with  $n$  internal nodes has  $n + 1$  external nodes. ■

#### 2. [BvG 3.3]

**Show that the external path length  $epl$  in a 2-tree with  $m$  external nodes satisfies  $epl \leq \frac{1}{2}(m^2 + m - 2)$ . Conclude that  $epl \leq \frac{1}{2}n(n + 3)$  for a 2-tree with  $n$  internal nodes.**

*Proof:* (using Induction)

We need to show external path length  $epl$  in a 2-tree with  $m$  external nodes satisfies  $epl \leq \frac{1}{2}(m^2 + m - 2)$ . Let  $T_i$  denote the set of 2-tree(s) with  $i$  nodes where  $i$  is an odd integer.

Base Case ( $m = 1$ ): Let  $T \in T_1$ . Here  $epl = 0 \leq \frac{1}{2}(1 + 1 - 2) = 0$ . So, the condition holds.

Inductive Step: Let's assume the condition is true for any 2-trees with  $m$  external nodes i.e. Such 2-trees have  $m - 1$  internal nodes, so they belong to  $T_{2m-1}$ . This means for a 2-tree  $T \in T_{2m-1}$ ,  $epl \leq \frac{1}{2}(m^2 + m - 2)$ . Now we need to show that the condition is true for any 2-tree in  $T_{2m+1}$ .

Let  $T \in T_{2m+1}$ ;  $T$  has  $m + 1$  external nodes and  $m$  internal nodes. Let's consider a tree  $T'$ , a subgraph of  $T$  formed by removing both the leaves ( $v_1$  and  $v_2$ ) from a same parent node  $v \in V(T)$  ( $v$  located at height  $h - 1$  where  $h$  is the height of  $T$ ). Here,  $T' \in T_{2m-1}$  which means  $epl_{T'} \leq \frac{1}{2}(m^2 + m - 2)$ . Now, let's add those leaves ( $v_1$  and  $v_2$ ) back to  $v$  to form  $T' + v_1 + v_2 = T$ . So,  $v_1, v_2$  are external nodes, each of which can be at a maximum distance of  $(m + 1) - 1 = m$  and  $v$  is now an internal node in  $T$ .

This gives  $epl_T \leq \frac{1}{2}(m^2 + m - 2) + (\text{max increase in } epl \text{ upon addition of } v_1 \text{ and } v_2) = \frac{1}{2}(m^2 + m - 2) + [(m + m) - (m - 1)] = \frac{1}{2}(m^2 + m - 2) + m + 1 = \frac{1}{2}((m^2 + 2m + 1) + (m + 1) - 2) = \frac{1}{2}((m + 1)^2 + (m + 1) - 2)$ . Hence, the condition holds for any  $T \in T_{2m+1}$  with  $m + 1$  external nodes.

Thus, it is proved that  $epl$  in a 2-tree with  $m$  external nodes satisfies  $epl \leq \frac{1}{2}(m^2 + m - 2)$ . ■

### 3. [4.5-1]

Use the master method to give tight asymptotic bounds for the following recurrences.

a.  $T(n) = 2T(n/4) + 1$

Here,  $n^{\log_b a} = n^{\log_4 2} = n^{0.5}$ . For  $\epsilon = 0.5$ ,  $f(n) = 1 \in O(n^{n^{0.5-\epsilon}}) = O(1)$ , so  $T(n) \in \Theta(n^{0.5}) = \Theta(\sqrt{n})$  [Master Theorem, case 1].

b.  $T(n) = 2T(n/4) + \sqrt{n}$

Here,  $n^{\log_b a} = n^{\log_4 2} = n^{0.5}$ . Now,  $f(n) = \sqrt{n} \in \Theta(n^{n^{0.5}}) = \Theta(n^{\sqrt{n}})$ , so  $T(n) \in \Theta(n^{0.5} \lg n) = \Theta(\sqrt{n} \lg n)$  [Master Theorem, case 2].

c.  $T(n) = 2T(n/4) + n$

Here,  $n^{\log_b a} = n^{\log_4 2} = n^{0.5}$ . For  $\epsilon = 0.5$ ,  $f(n) = n \in \Omega(n^{n^{0.5+\epsilon}}) = \Omega(n)$  and  $2f(n/4) = n/2 \leq cn = cf(n)$  for  $c = 0.9 < 1$  and all  $n \geq 0$ . So,  $T(n) \in \Theta(f(n)) = \Theta(n)$  [Master Theorem, case 3].

d.  $T(n) = 2T(n/4) + n^2$

Here,  $n^{\log_b a} = n^{\log_4 2} = n^{0.5}$ . For  $\epsilon = 1.5$ ,  $f(n) = n^2 \in \Omega(n^{n^{0.5+\epsilon}}) = \Omega(n^2)$  and  $2f(n/4) = n^2/8 \leq cn^2 = cf(n)$  for  $c = 0.25 < 1$  and all  $n \geq 0$ . So,  $T(n) \in \Theta(f(n)) = \Theta(n^2)$  [Master Theorem, case 3].

### 4. [BvG 3.6 (modified)]

In this exercise, all integers are considered to be nonnegative, for simplicity. A divisor of an integer  $k$  is any integer  $d \neq 0$  such that  $k/d$  has no remainder. A *common divisor* for a set of integers is an integer that is a divisor for each integer in the set. Euclid's algorithm for finding the greatest common divisor (GCD) of two nonnegative integers,  $m$  and  $n$ , can be written as follows:

```

1: procedure GCD(int  $m$ , int  $n$ )
2:   if  $n = 0$  then
3:      $\text{answer} \leftarrow m$ 
4:   else if  $m < n$  then
5:      $\text{answer} \leftarrow \text{GCD}(n, m)$ 
6:   else
7:      $r \leftarrow m - n \cdot \lfloor \frac{m}{n} \rfloor$   $\triangleright r$  is the remainder of  $\frac{m}{n}$ 
8:      $\text{answer} \leftarrow \text{GCD}(n, r)$ 
9:   end if
10:  return  $\text{answer}$ 
11: end procedure

```

The preconditions for  $\text{GCD}(m, n)$  are that  $m \geq 0, n \geq 0$  and  $m + n > 0$ . Prove the following using induction.

a. If the preconditions of  $\text{GCD}(m, n)$  are satisfied, then the value that the function returns is *some* common divisor of  $m$  and  $n$ .

b. If the preconditions of  $\text{GCD}(m, n)$  are satisfied, then the value that the function returns is the greatest common divisor of  $m$  and  $n$ .

*Hints:* If  $d$  is a divisor of  $k$ , how can you rewrite  $k$  in terms of  $d$ ? How do you show that two sets are equal?

a. *Proof:* (using Strong Induction) Let  $m$  and  $n$  be integers such that  $m \geq 0, n \geq 0$  and  $m + n > 0$  and  $P(n) := \text{GCD}(m, n)$  is some common divisor of  $m$  and  $n$ .

Base Case ( $n = 0$ ): We need to show  $P(n)$  is true for  $n = 0$  and some  $m > 0$ . Here,  $\text{GCD}(m, 0) = m$  (by line 2-3). Since  $m|m$  and  $m|0$ ,  $\text{GCD}(m, 0)|m, n$ . So,  $P(0)$  is true.

Inductive Step: Let  $P(n)$  be true for all  $n < k$  for some  $k \in \mathbb{Z}^+$ . We need to prove  $P(k)$  is true.

Now, let  $m = kq + r$  where  $q = \lfloor \frac{m}{k} \rfloor$  and  $r < k$ . So (by line 4-8),  $\text{GCD}(m, k) = \text{GCD}(k, r)$ . Since  $r < k$ ,  $P(r)$  holds true, so  $\text{GCD}(k, r)|k, r \Rightarrow \text{GCD}(k, r)|k, kq + r$  but  $kq + r = m$ . So,  $\text{GCD}(k, r) = \text{GCD}(m, k)|m, k$ . This shows  $P(k)$  is true.

Hence, if the preconditions of  $\text{GCD}(m, n)$  are satisfied, then the value that the function returns is *some* common divisor of  $m$  and  $n$ . ■

b. *Proof:* (using Strong Induction)

From Part (a),  $\text{GCD}(m, n)|m, n$ . Now we need to show if  $p|m, n$  then  $p|\text{GCD}(m, n)$  where  $p \in \mathbb{Z}^+$ .

Base Case ( $n=0$ ):  $\text{GCD}(m, 0) = m$ . If  $p|m, 0$  then  $p|\text{GCD}(m, 0), 0 \Rightarrow p|\text{GCD}(m, 0)$ . So the base case holds true.

Inductive Step: If  $p|m, n$  for all  $n < k \in \mathbb{Z}^+$ ,  $p|\text{GCD}(m, n)$ . Assume this is true. Now, we need to show if  $p|m, k$  then  $p|\text{GCD}(m, k)$ .

Here,  $p|m, k \Rightarrow m = c_1p, k = c_2p$  for some integers  $c_1$  and  $c_2$ . Now like in Part (a),  $m = kq + r \Rightarrow r = m - kq = c_1p - c_2pq = p(c_1 - c_2q) \Rightarrow p|r$ . So by hypothesis,  $p|k, r \Rightarrow p|\text{GCD}(k, r) = \text{GCD}(m, k)$  since  $r < k$ .

Hence, if  $p|m, n$  then  $p|\text{GCD}(m, n) \Rightarrow p \leq \text{GCD}(m, n)$  since  $\text{GCD}(m, n), p > 0$ . Thus, if the preconditions of  $\text{GCD}(m, n)$  are satisfied, then the value that the function returns is the *greatest* common divisor of  $m$  and  $n$ . ■

**5. [Problem 4-3 (modified)]**

Give an asymptotically tight bound for  $T(n)$  in each of the following recurrences.

**a.  $T(n) = 4T(n/3) + n \lg n$** 

Here,  $n^{\log_b a} = n^{\log_3 4} \approx n^{1.26}$ . For  $\epsilon = 0.26$ ,  $f(n) = n \lg n \in O(n^{1.26-\epsilon}) = O(n)$ . So,  $T(n) \in \Theta(n^{\log_3 4})$  [by Master Theorem, case 1].

**b.  $T(n) = 3T(n/3) + n/\lg n$** 

Here,  $n^{\log_b a} = n^{\log_3 3} = n$ . And,  $f(n) = n/\lg n \in \Theta(n^{\log_b a}) = \Theta(n)$ . So,  $T(n) \in \Theta(n^{\log_b a} \lg n) = \Theta(n \lg n)$  [by Master Theorem, case 2].

**c.  $T(n) = 4T(n/2) + n^2 \sqrt{n}$** 

Here,  $n^{\log_b a} = n^{\log_2 4} = n^2$ . For  $\epsilon = 0.5$ ,  $f(n) = n^2 \sqrt{n} = n^{2.5} \in \Omega(n^{2+\epsilon})$  and  $4f(n/2) = n^2 \sqrt{n}/\sqrt{2} \leq cn^2 \sqrt{n} = cf(n)$  for  $c = 0.9$  and all  $n > 0$ . So,  $T(n) \in \Theta(f(n)) = \Theta(n^2 \sqrt{n})$  [by Master Theorem, case 3].

**d.  $T(n) = 3T(n/3 - 2) + n/2$** 

For large values of  $n$ , its asymptotic behaviour is same as that of  $T(n) = 3T(n/3) + n/2$  where  $n^{\log_b a} = n^{\log_3 3} = n$ . And  $f(n) = n/2 \in \Theta(n^{\log_b a}) = \Theta(n)$ . So,  $T(n) \in \Theta(n^{\log_b a} \lg n) = \Theta(n \lg n)$  [by Master Theorem, case 2].

**e.  $T(n) = 2T(n/2) + n/\lg n$** 

Here,  $n^{\log_b a} = n^{\log_2 2} = n$ . And,  $f(n) = n/\lg n \in \Theta(n^{\log_b a}) = \Theta(n)$ . So,  $T(n) \in \Theta(n^{\log_b a} \lg n) = \Theta(n \lg n)$  [by Master Theorem, case 2].

**f.  $T(n) = T(n/2) + T(n/4) + T(n/8) + n$** 

Let's use substitution method to solve this recurrence. We make guess that  $T(n) \in \Theta(n)$ . This method requires us to prove  $c_1 n \leq T(n) \leq c_2 n$  for  $n \geq n_0$  and  $c_1, c_2 > 0$ . So, we start by assuming the bounds hold for  $n/2$ :

$$\left(\frac{7c_1 + 8}{16}\right)n = c_1 \left(\frac{n}{4} + \frac{n}{8} + \frac{n}{16} + \frac{n}{2}\right) \leq T(n) \leq c_2 \left(\frac{n}{4} + \frac{n}{8} + \frac{n}{16} + \frac{n}{2}\right) = \left(\frac{7c_2 + 8}{16}\right)n$$

Now,

$$c_1 \leq \frac{7c_1 + 8}{16} \Rightarrow c_1 \leq \frac{8}{9} \text{ and } \frac{7c_2 + 8}{16} \leq c_2 \Rightarrow c_2 \geq \frac{8}{9}.$$

So, the statement,  $c_1 n \leq T(n) \leq c_2 n$ , holds for any  $c_1$  and  $c_2$  :  $c_1 \leq \frac{8}{9}$  and  $c_2 \geq \frac{8}{9}$ . Hence,  $T(n) \in \Theta(n)$ .

## Part I. 25 points each.

## 6. [Problem 4-4]

This problem develops properties of the Fibonacci numbers, which are defined by recurrence (3.22). We shall use the technique of generating functions to solve the Fibonacci recurrence.

Define the generating function (or formal power series)  $\mathcal{F}$  as

$$\begin{aligned}\mathcal{F}(z) &= \sum_{i=0}^{\infty} F_i z^i \\ &= 0 + z + z^2 + 2z^3 + 3z^4 + 5z^5 + 8z^6 + 13z^7 + 21z^8 + \dots\end{aligned}$$

where  $F_i$  is the  $i^{\text{th}}$  Fibonacci number.

**a** Show that  $\mathcal{F}(z) = z + z\mathcal{F}(z) + z^2\mathcal{F}(z)$ .

**b** Show that

$$\begin{aligned}\mathcal{F}(z) &= \frac{z}{1 - z - z^2} \\ &= \frac{z}{(1 - \phi z)(1 - \hat{\phi} z)} \\ &= \frac{1}{\sqrt{5}} \left( \frac{1}{1 - \phi z} - \frac{1}{1 - \hat{\phi} z} \right)\end{aligned}$$

where

$$\phi = \frac{1 + \sqrt{5}}{2} \approx 1.61803\dots$$

and

$$\hat{\phi} = \frac{1 - \sqrt{5}}{2} \approx -0.61803\dots$$

**c** Show that

$$\mathcal{F}(z) = \sum_{i=0}^{\infty} \frac{1}{\sqrt{5}} (\phi^i - \hat{\phi}^i) z^i$$

**d** Use part (c) to prove that  $F_i = \phi^i / \sqrt{5}$  for  $i \geq 0$ , rounded to the nearest integer. *Hint:* Observe that  $|\hat{\phi}| < 1$ .

a. *Proof:*

$$\begin{aligned}
z + \mathcal{F}(z) + z^2 \mathcal{F}(z) &= z + z \sum_{i=0}^{\infty} \mathcal{F}_i z^i + z^2 \sum_{i=0}^{\infty} \mathcal{F}_i z^i \\
&= z + \mathcal{F}_0 z + \sum_{i=1}^{\infty} \mathcal{F}_i z^{i+1} + \sum_{i=0}^{\infty} \mathcal{F}_i z^{i+2} \\
&= z + 0 + \sum_{i=2}^{\infty} \mathcal{F}_{i-1} z^i + \sum_{i=2}^{\infty} \mathcal{F}_{i-2} z^i \\
&= 0 + z + \sum_{i=2}^{\infty} (\mathcal{F}_{i-1} + \mathcal{F}_{i-2}) z^i \text{ [since each sum converges]} \\
&= \mathcal{F}_0 z^0 + \mathcal{F}_1 z^1 + \sum_{i=2}^{\infty} \mathcal{F}_i z^i = \sum_{i=0}^{\infty} \mathcal{F}_i z^i = \mathcal{F}(z).
\end{aligned}$$

Hence,  $\mathcal{F}(z) = z + \mathcal{F}(z) + z^2 \mathcal{F}(z)$ . ■

b. *Proof:*

$$\begin{aligned}
\text{From a, we get } \mathcal{F}(z) &= z + \mathcal{F}(z) + z^2 \mathcal{F}(z) \\
&\Rightarrow (1 - z - z^2) \mathcal{F}(z) = z \\
&\Rightarrow \mathcal{F}(z) = \frac{z}{1 - z - z^2}.
\end{aligned}$$

Now let's consider

$$\begin{aligned}
\frac{1}{\sqrt{5}} \left( \frac{1}{(1 - \phi z)} - \frac{1}{(1 - \hat{\phi} z)} \right) &= \frac{1}{\sqrt{5}} \left( \frac{1 - \hat{\phi} z - 1 + \phi z}{(1 - \phi z)(1 - \hat{\phi} z)} \right) \\
&= \frac{1}{\sqrt{5}} \left( \frac{(\phi - \hat{\phi}) z}{(1 - \phi z)(1 - \hat{\phi} z)} \right) \\
&= \frac{1}{\sqrt{5}} \left( \frac{\sqrt{5} z}{(1 - \phi z)(1 - \hat{\phi} z)} \right) [\because \phi - \hat{\phi} = \sqrt{5}] \\
&= \frac{z}{(1 - \phi z)(1 - \hat{\phi} z)}.
\end{aligned}$$

Then, using the fact that  $\phi + \hat{\phi} = 1$  and  $\phi \hat{\phi} = -1$ , we get

$$\begin{aligned}
\frac{z}{(1 - \phi z)(1 - \hat{\phi} z)} &= \frac{z}{1 - (\phi + \hat{\phi}) z + \phi \hat{\phi} z} \\
&= \frac{z}{(1 - \phi z)(1 - \hat{\phi} z)} \\
&= \frac{z}{1 - z - z^2}.
\end{aligned}$$

Combining all the results, we get

$$\mathcal{F}(z) = \frac{z}{1 - z - z^2} = \frac{z}{(1 - \phi z)(1 - \hat{\phi} z)} = \frac{1}{\sqrt{5}} \left( \frac{1}{(1 - \phi z)} - \frac{1}{(1 - \hat{\phi} z)} \right).$$

■

c. *Proof:* Using the result from b,

$$\begin{aligned} \mathcal{F}(z) &= \frac{1}{\sqrt{5}} \left( \frac{1}{(1 - \phi z)} - \frac{1}{(1 - \hat{\phi} z)} \right) = \frac{1}{\sqrt{5}} \left( \sum_{i=0}^{\infty} (\phi z)^i - \sum_{i=0}^{\infty} (\hat{\phi} z)^i \right) \\ &= \frac{1}{\sqrt{5}} \sum_{i=0}^{\infty} (\phi z)^i - (\hat{\phi} z)^i \\ &= \frac{1}{\sqrt{5}} \sum_{i=0}^{\infty} (\phi^i - \hat{\phi}^i) z^i. \end{aligned}$$

■

d. *Proof:*

$$\mathcal{F}(z) = \frac{1}{\sqrt{5}} \sum_{i=0}^{\infty} (\phi^i - \hat{\phi}^i) z^i \text{ where } \frac{1}{\sqrt{5}} (\phi^i - \hat{\phi}^i) \text{ is the coefficient of } z^i. \text{ But by definition } \mathcal{F}(z) = \sum_{i=0}^{\infty} \mathcal{F}_i z^i,$$

$\mathcal{F}_i$  is the coefficient of  $z^i$ . So,  $\mathcal{F}_i = \frac{1}{\sqrt{5}} (\phi^i - \hat{\phi}^i)$ . Now, let's consider  $\phi^i / \sqrt{5}$  rounded to nearest integer

$$\begin{aligned} \text{which is given by } \left\lceil \frac{\phi^i}{\sqrt{5}} - \frac{1}{2} \right\rceil &= \left\lceil \left( \frac{\phi^i}{\sqrt{5}} - \frac{\hat{\phi}^i}{\sqrt{5}} \right) + \left( \frac{\hat{\phi}^i}{\sqrt{5}} - \frac{1}{2} \right) \right\rceil \\ &= \left\lceil \mathcal{F}_i + \left( \frac{\hat{\phi}^i}{\sqrt{5}} - \frac{1}{2} \right) \right\rceil = \mathcal{F}_i + \left\lceil \frac{\hat{\phi}^i}{\sqrt{5}} - \frac{1}{2} \right\rceil = \mathcal{F}_i + 0 = \mathcal{F}_i \end{aligned}$$

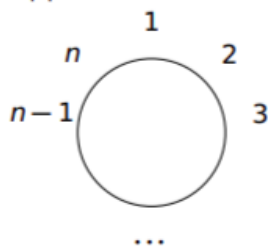
$$\text{since } |\hat{\phi}| < 1 \Rightarrow \frac{|\hat{\phi}^i|}{\sqrt{5}} < \frac{|\hat{\phi}|}{\sqrt{5}} < \frac{1}{2} \Rightarrow \frac{1}{2} < \left( \frac{\hat{\phi}^i}{\sqrt{5}} - \frac{1}{2} \right) < 0 \Rightarrow \left\lceil \frac{\hat{\phi}^i}{\sqrt{5}} - \frac{1}{2} \right\rceil = 0.$$

Hence,  $\mathcal{F}_i = \hat{\phi}^i / \sqrt{5}$  for  $i \geq 0$ , rounded to the nearest integer.

■

## 6. [The Josephus Problem]

Suppose we have  $n$  items arranged in a circle as shown:



We proceed around the circle, removing every other item (item 2 is the first to be removed) until one item remains. For example, if  $n = 10$ , the sequence of removed items is 2, 4, 6, 8, 10, 3, 7, 1, 9, with item 5 being the last one left. The *survivor's number*,  $J(n)$ , is the number of the last remaining item from a set of  $n$  items (thus,  $J(10) = 5$ ).

- Suppose there are  $2n$  items. After  $n$  items have been removed, there are  $n$  items remaining. What are the numbers of the items that remain? How do those numbers relate to the numbering used for an initial set of  $n$ ? Use this information to express  $J(2n)$  in terms of  $J(n)$ .
- Suppose there are now  $2n + 1$  items. After  $n + 1$  items have been removed, there are  $n$  items remaining. What are the numbers of the items that remain? How do those numbers relate to the numbering used for an initial set of  $n$ ? Use this information to express  $J(2n + 1)$  in terms of  $J(n)$ .
- Combine parts **a** and **b**, along with the base case  $J(1) = 1$ , to form a recurrence relation for  $J(n)$ .
- Use the recurrence relation to create a small table of  $J(n)$  values (you shouldn't need more than 20 to see the pattern). Use this table to find a closed form (*i.e.*, non-recursive) for  $J(n)$ . *Hint*: Express  $n$  as  $n = 2^m + l$ , where  $2^m \leq n < 2^{m+1}$ .
- Prove that your closed form solution is correct. *Hint*: Use induction on  $m$ .

*Solution:* Here,  $\mathcal{J}(n)$  is the number of the last remaining item from a set of  $n$  items. Let us define start-point as the item from which we proceed around the circle. The initial start-point in this problem is always at 1.

a. Let  $\mathcal{J}(n) = k$ , where  $1 \leq k \leq n$ . Now in a set of  $2n$  items, when  $n$  of them are removed, the  $n$  odd-numbered items are remaining, with new start-point at 1. This case is "isomorphic" to finding  $\mathcal{J}(n)$ , in which  $k^{\text{th}}$  item is the last one left. The  $k^{\text{th}}$  number in a list of first  $n$  odd numbers is  $2k - 1$ . So,  $\mathcal{J}(2n) = 2k - 1 = 2\mathcal{J}(n) - 1$ .

b. Now in a set of  $2n + 1$  items, when  $n + 1$  of them are removed, the  $n$  odd-numbered items are remaining, with new start-point at 3. Again, this case is "isomorphic" to finding  $\mathcal{J}(n)$ , in which  $k^{\text{th}}$  item is the last one left. In a list of  $n$  increasing odd numbers starting at 3, the  $k^{\text{th}}$  number is  $3 + 2(k - 1) = 2k + 1$ . So,  $\mathcal{J}(2n + 1) = 2k + 1 = 2\mathcal{J}(n) + 1$ .



c. Combining the result of a and b with  $\mathcal{J}(1) = 1$ , we get

$$\mathcal{J}(n) = \begin{cases} 1 & \text{if } n = 1 \\ 2\mathcal{J}(k) - 1, & \text{if } n = 2k \text{ for } k \in \mathbb{Z}^+ \\ 2\mathcal{J}(k) + 1, & \text{if } n = 2k + 1 \text{ for } k \in \mathbb{Z}^+ \end{cases}$$

d.

$n$	$\mathcal{J}(n)$
$1 = 2^0 + 0$	1
$2 = 2^1 + 0$	1
$3 = 2^1 + 1$	3
$4 = 2^2 + 0$	1
$5 = 2^2 + 1$	3
$6 = 2^2 + 2$	5
$7 = 2^2 + 3$	7
$8 = 2^3 + 0$	1
$9 = 2^3 + 1$	3
$10 = 2^3 + 2$	5
$11 = 2^3 + 3$	7
$12 = 2^3 + 4$	9
$13 = 2^3 + 5$	11
$14 = 2^3 + 6$	13
$15 = 2^3 + 7$	15
$16 = 2^4 + 0$	1
$17 = 2^4 + 1$	3
$18 = 2^4 + 2$	5
$19 = 2^4 + 3$	7
$20 = 2^4 + 4$	9

From the table above, it can be conjectured that  $\mathcal{J}(n) = 2l + 1$  where  $n = 2^m + l$ ,  $0 \leq l < 2^m$ .

e. To prove:  $P(n)$  : If  $n = 2^m + l$ ,  $0 \leq l < 2^m$ ,  $\mathcal{J}(n) = 2l + 1$ .

*Proof:* (using Strong Induction)

Base case:  $P(1) = 1$  and  $P(2) = 1$  are true (from the table).

Induction step: Let's assume  $P(k)$  is true for all  $k \leq n, k \in \mathbb{Z}^+$ . We need to prove  $P(n+1)$  is true, which we shall divide into two cases.

Case 1: When  $n+1$  is even,  $\mathcal{J}(n+1) = 2\mathcal{J}(\frac{n+1}{2}) - 1$ . Since  $\frac{n+1}{2} \leq n$ , by hypothesis

$$\mathcal{J}(\frac{n+1}{2}) = 2l_1 + 1 \text{ where } \frac{n+1}{2} = 2^{m_1} + l_1, 0 \leq l_1 < 2^{m_1}, \text{ and } l_1, m_1 \in \mathbb{N}.$$

So,  $\mathcal{J}(n+1) = 2(2l_1 + 1) - 1 = 2(2l_1) + 1$  where  $n+1 = 2^{m_1+1} + 2l_1$ ,  $0 \leq l < 2^{m_1+1}$ .

Setting  $2l_1 = l$  and  $m_1 + 1 = m$ , we get

$$\mathcal{J}(n+1) = 2l + 1 \text{ where } n+1 = 2^m + l, 0 \leq l < 2^m.$$

Case 2: When  $n + 1$  is odd,  $\mathcal{J}(n + 1) = 2\mathcal{J}(\frac{n}{2}) + 1$ . Since  $\frac{n}{2} \leq n$ , by hypothesis

$$\mathcal{J}(\frac{n}{2}) = 2l_1 + 1 \text{ where } \frac{n}{2} = 2^{m_1} + l_1, \ 0 \leq l_1 < 2^{m_1}, \text{ and } l_1, m_1 \in \mathbb{N}.$$

So,  $\mathcal{J}(n + 1) = 2(2l_1 + 1) + 1$  where  $n + 1 = 2^{m_1+1} + (2l_1 + 1)$ ,  $1 \leq 2l_1 + 1 < 2^{m_1+1} + 1$ .

Setting  $2l_1 + 1 = l$  and  $m_1 + 1 = m$ , we get

$$\mathcal{J}(n + 1) = 2l + 1 \text{ where } n + 1 = 2^m + l, \ 1 \leq l < 2^m + 1.$$

[Note: I think the range of  $l$  in the last line is not correct; it should have come out to be  $0 \leq l < 2^m$ .]

■