HWA CHONG INSTITUTION
C2 PRELIMINARY EXAMINATION 2022

COMPUTING
Higher 2

13 Sept 2022   **Paper 1 (9569 / 01)**   **1400 -- 1700 hrs**

_____

READ THESE INSTRUCTIONS FIRST

An answer booklet will be provided with this question paper. You should follow the instructions on the front cover of the answer booklet. If you need additional answer paper ask the invigilator for a continuation booklet.

Answer *ALL* questions.

Approved calculators are allowed.

The number of marks is given in brackets [ ] at the end of each question or part question.

The total number of marks for this paper is **100**.

_____

This document consists of **8** printed pages.

1 An organization would like to develop a system using object-oriented approach to manage the information of the employees.

The system computes:
- the monthly salary of full-time employees
- the monthly overtime allowance of the full-time employees based on the hourly rate
- the monthly salary of daily-rated employees based on daily rate.

(a) Draw a class diagram, which exhibits the following:
- Suitable classes with appropriate properties and methods
- Inheritance
- Polymorphism   [6]

(b) Explain how your design in **(a)** demonstrates code reuse.   [2]

(c) Explain the term **polymorphism** and how it is applied in the design in **(a)**.   [2]
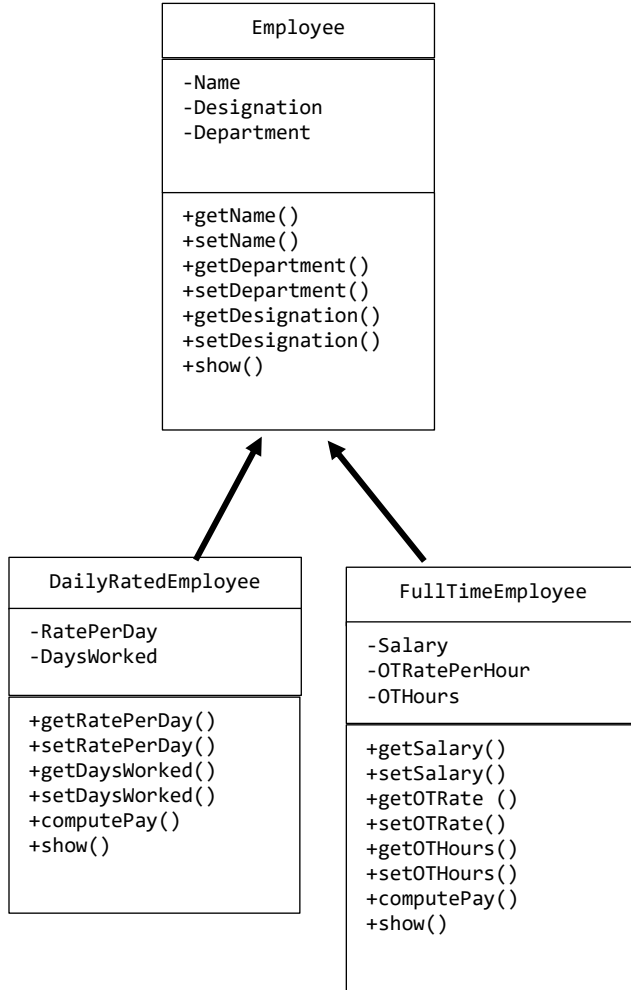
Due to temporary increase in the amount of work, the organization decided to allow daily-rated employees to receive overtime allowance based on a hourly rate.

(d) Suggest and explain the changes to the class diagram to enable the computation of daily-rated employees overtime allowance.   [3]

Solution:

```
                    ┌─────────────────────────┐
                    │        Employee         │
                    ├─────────────────────────┤
                    │ -Name                   │
                    │ -Designation            │
                    │ -Department             │
                    │                         │
                    ├─────────────────────────┤
                    │ +getName()              │
                    │ +setName()              │
                    │ +getDepartment()        │
                    │ +setDepartment()        │
                    │ +getDesignation()       │
                    │ +setDesignation()       │
                    │ +show()                 │
                    │                         │
                    └─────────────────────────┘
```

```
┌──────────────────────────┐      ┌──────────────────────────┐
│   DailyRatedEmployee     │      │     FullTimeEmployee     │
├──────────────────────────┤      ├──────────────────────────┤
│ -RatePerDay              │      │ -Salary                  │
│ -DaysWorked              │      │ -OTRatePerHour           │
│                          │      │ -OTHours                 │
├──────────────────────────┤      ├──────────────────────────┤
│ +getRatePerDay()         │      │ +getSalary()             │
│ +setRatePerDay()         │      │ +setSalary()             │
│ +getDaysWorked()         │      │ +getOTRate ()            │
│ +setDaysWorked()         │      │ +setOTRate()             │
│ +computePay()            │      │ +getOTHours()            │
│ +show()                  │      │ +setOTHours()            │
│                          │      │ +computePay()            │
└──────────────────────────┘      │ +show()                  │
                                  │                          │
                                  └──────────────────────────┘
```

(b)   The sub-classes (DailyRatedEmployee and FullTimeEmployee) **inherit all the attributes and methods** of the Employee class. The methods **getName, SetName, getDepartment, setDepartment, getDesignation, setDesignation are inherited without changes to the implementation and the no coding are required, hence code reused is achieved.**

(c) Polymorphism refers to **an object's ability to take different forms**. It allows subclasses to have methods with the same name as methods in their superclasses. It gives the ability for a program to call the correct method depending on the type of object that is used to call it.
The method *show()* in both the subclasses **overrides** the superclass *show()*. If the subclass object (DailyRatedEmployee) is used to call *show()*, then the subclass's version of the method is invoked. If the superclass object (Employee) is used to call *show()*, then the superclass method will be invoked.

(d)
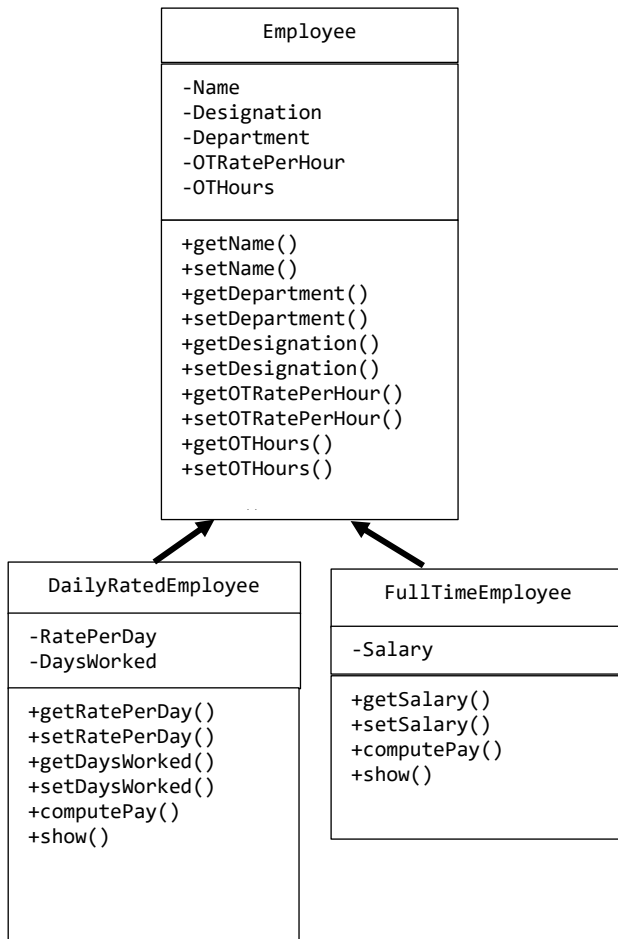
2 Mrs Wang is tested positive for flu on day 1 of the month. In order to protect her baby from the virus, she designs her own quarantine rule based on the government policy. For the first seven days, she must fully quarantine at home. After day 7, she performs the diagnostic test in the morning every day. If she has symptoms or tests positive in the morning, she will still go to work for the day but isolate herself at home. Only when she shows no symptoms and tests negative in the morning, she will be fully back to normal life.

(a) Create a decision table showing all the possible outcomes and results. [4]

(b) Simplify your decision table by removing redundancies. [2]

(c) The simplified table produced in part (b) is used as a design for program code. Write, **in pseudocode**, a program to display Mrs Wang's quarantine result for a particular day. [2]

Solution:

(a)

Day > 7	T	T	T	T	F	F	F	F
Symptom	T	F	F	T	T	F	F	T
Test positive	F	T	F	T	F	T	F	T
Fully Quarantine at home					Y	Y	Y	Y
Work + isolate herself at home	Y	Y		Y				
Back to normal life			Y					

(b)

Day > 7	T	T	T	F
Symptom	T	F	F	-
Test positive	-	T	F	-
Fully Quarantine at home				Y
Work + isolate herself at home	Y	Y		
Back to normal life			Y	

(c)
```
DECLARE Day: INTEGER
DECLARE Symptom: BOOLEAN // Symptom = True means sick
symptoms
DECLARE Test: BOOLEAN // Test = True means she tests
positive
IF Day < = 7
THEN
      OUTPUT 'Fully Quarantine at home'
ELSE
      IF Symptom = True OR Test = True
      THEN
          Output 'Work + isolate herself at home'
      ELSE
          Output 'Back to normal life'
      ENDIF
ENDIF
```

3 **(a)** Describe **two** reasons for layering in the TCP/IP model. [2]

(b) Describe the use of switch and router in the TCP/IP model. [4]

(c) Describe how data is transmitted in a packet switching network. Give **two** advantages of packet switching over circuit switching network. [3]

(d) Describe how a digital signature can be used to give confidence that a received message has not been altered. [6]

(e) Explain the purpose of data verification and give one example. [2]

(f) An SMTP server provides email services to the client.
  (i) Describe **one** disadvantage of client-server network. [1]

  (ii) Name **one** malware and describe how this malware compromise computer systems. State **one** protection scheme that can protect the server from this malware. [2]

  (iii) Describe **two** codes of conduct for the IT staff who builds and maintains the server. [2]

Solution:
(a)
Reason for layering: Simplifies the network model; Enables programmers to specialize in a particular layer of the model; Provides design modularity; Allows for standardized interfaces to be produced by networking vendors
(b)
At the **Network** layer, router uses **IP address** to send data packets to the designated **host in the network**
At the **Data Link** layer, switch uses **MAC address** to send data frame to the designated **machine in the local network**
(c)
Data is **separated** into packets and each packet **independently** find **the best route** to the receiver.

Advantage:

Different packets may travel in **different** route and thus **more efficient**, **saves bandwidth** and **avoids congestion**. It is also more **secured** since it became much harder to attack **all the routes** instead of one route in circuit switching network.

(d)

Sender Side
- The sender uses a hash algorithm to **create a hashed version** of the message
- The sender uses **its private key** to **encrypt the hash** to the digital signature
- Both the **message** (encrypted or not) and the **digital signature** are sent to the receiver

Receiver Side
- The receiver uses the **sender's public key** to **decrypt** the digital signature back to the sender's version of hash
- The receiver uses the **same** hash algorithm to create a **new hash** from the received message
- If the two hashes **match**, it means the data is not altered and is sent by the known sender

(e) Data verification ensures the input data matches the original resource.
Example: enter password twice, proofread before submitting forms

(f)
(i) **If server is down, whole network is down.** So the server is more vulnerable to cyberattack;
Centralized server is **more expensive** to build up, and **requires professional to maintain**.
(ii)
Virus: attach itself to another program and remains dormant until it is executed by the user. It replicates to infect other computers in the network
Worm: **self-replicating** to infect other computers in the network which aims to **damage** computer systems or gain unauthorized access to them.
Trojan Horse: appears as an legitimate program. Once it gains access into the computer, it ruans malicious code to cause damage to the computer
Ransomware: locks the computer or encrypt the data and force the user to pay a ransom to get their computer back
Adware: unwanted software which displays advertisement on the computer
Spyware: hidden program that secretly collects information and transmits information to attackers without the user's knowledge
Protection Scheme: firewall, anti-software, Intrusion Detection/Protection Scheme
(iii)
ensure the servers are up-to-date and equipped with protection schemes; keep the confidentiality and the safely of the emails from the clients; constantly upgrade their knowledge and skills

4 The first two items of a sequence are 2 and 3 respectively. From the third item onwards, each number is the multiplication of its two preceding items. A function, SEQ, uses recursion, to generate the n-th item of the sequence.

```
01   def SEQ(n):
02       if n < 3:
03           return 2
04       else:
05           return SEQ(n - 1) * SEQ(n - 2)
```

(a) State what is meant by recursion. [1]

(b) Explain how a stack is used when a recursive call is made. [4]

(c) State the line number that indicate function SEQ is recursive. [1]

(d) Modify line 03 to make the function work. [1]

Solution:

(a)   recursion means a function that is defined in terms of itself; a function that calls a simpler version of itself
(b)   Each time when the function SEQ is called, an activation record is generated including its **parameter n, return value and return address**. This activation record is **pushed** onto the run-time stack. When a recursive version of SEQ terminates, its activation record is **popped** from the stack. The **top** activation record in the stack is always the procedure currently executed.
(c)   line 05
(d)   return n + 1

5 A country has a number of cross-country running clubs. A database is to be created storing data about the races and runners.
- Each runner belongs to one club.
- Each race has a date and the distance.
- Each club has name and contact telephone number.
- The clubs have agreed to stage one race only on any date.

This table shows the data about the races and runners.

Runner ID	Runner Name	Club ID	Club Name	Club Tel	Race Date	Race Distance
1	Smith	1	Rede2run	65442781	20-03-2022	15
2	Jones	2	Run4Life	62568892	20-03-2022	15
3	Roberts	1	Rede2run	65442781	20-03-2022	15
1	Smith	1	Rede2run	65442781	15-02-2022	18
4	Harris	3	Run4Fun	64519768	15-02-2022	18
5	Charles	4	FastWalker	64432091	18-04-2022	20
4	Harris	3	Run4Fun	64519768	18-04-2022	20
5	Charles	4	FastWalker	64432091	15-02-2022	18
6	Patel	5	Walk2Fast	62789431	20-03-2022	15
7	Harrison	5	Walk2Fast	62789431	15-02-2022	18

(a) Explain whether the above table is in first normal form (1NF). [1]

The following is an attempt to reduce data redundancy:

Table: `Race`

Race Date	Race Distance
20-03-2022	15
15-02-2022	18
18-04-2022	20

Table: `RunnerClub`

Runner ID	Runner Name	Club ID	Club Name	Club Tel	Race Date
1	Smith	1	Rede2run	65442781	20-03-2022
2	Jones	2	Run4Life	62568892	20-03-2022
3	Roberts	1	Rede2run	65442781	20-03-2022
1	Smith	1	Rede2run	65442781	15-02-2022
4	Harris	3	Run4Fun	64519768	15-02-2022
5	Charles	4	FastWalker	64432091	18-04-2022
4	Harris	3	Run4Fun	64519768	18-04-2022
5	Charles	4	FastWalker	64432091	15-02-2022
6	Patel	5	Walk2Fast	62789431	20-03-2022
7	Harrison	5	Walk2Fast	62789431	15-02-2022

(b) Explain why table `RunnerClub` is not in third normal form (3NF). [2]

A relational database is to be used. Based on the information given above, design the database that consists of a number of tables.

(c) Draw the Entity-Relationship (E-R) diagram to show the tables in third normal form (3NF) and their relationships between them. [4]

A table description can be expressed as:

```
TableName (Attribute1, Attribute2, Attribute3, ...)
```

The primary key is indicated by underlining one or more attributes. Foreign keys are indicated by using a dashed underline.

The `RACE` table has been identified with the table descriptions:

```
Race(RaceDate, RaceDistance)
```

(d) Using the information given, write table descriptions for the other tables you identified in part **(c)**. [4]

Records have been created for all the runners registered for the race on 20-03-2022

(e) Write an SQL query to output the runner's name, club's name of all the runners entered for this race, in descending order of runner's name, based on the table descriptions for the tables in part **(d)**. [5]

To promote more races, the agreement between the clubs to stage one race only on any date was cancelled.

(f) Explain whether your existing table design is able to support the change and state the changes required to the design in part **(d)**. [3]

In order to notify the runners when there are new races, information such as contact numbers, addresses need to be stored in the database.

(g) State **two** actions that needs to be done regarding the collection, disclosure and use of this data under the Personal Data Protection Act. [2]

Solution:

(a) Yes. All cells in the table has atomic values

(b) Not in 3NF. There are transitive dependencies in `RunnerClub` table.
`ClubName, ClubTel` is transitive dependent on `RunnerID`.

(c)
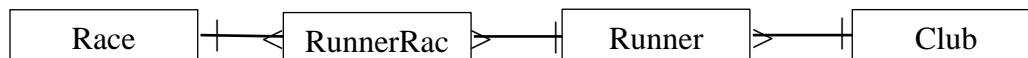
(d)

Race(<u>RaceDate</u>, RaceDistance)
Runner(<u>RunnerID</u>, RunnerName, ClubID)
Club(<u>ClubID</u>, ClubName, ClubTel)
RunnerRace(<u>RunnerID</u>, <u>RaceDate</u>)

(e)

SELECT Runner.RunnerName, Club.ClubName
FROM RunnerRace INNER JOIN Runner ON RunnerRace.RunnerID = Runner.RunnerID INNER JOIN Club ON Runner.ClubID = Club.ClubID
WHERE RunnerRace.RaceDate = '20-03-2022'
ORDER BY Runner.RunnerName DESC

(f) Existing table design supports one race per day hence it is possible to **use the column RaceDate as an unique identifier for the record in the table**. With the cancellation of

the agreement, it would means that there could be **one or more races on the same day**, this will means that the **column RaceDate is not able to uniquely identify the record**.
  Two tables need to be changed.:

      a. Race(<u>RaceID</u>, RaceDate, Race Distance)
      b. RunnerRace(<u>RunnerID</u>, <u>RaceID</u>)

(g) Any 2 – describe the actions to be done not the obligation name

    1. Under Consent Obligation -Only collect, use or disclose personal data when an individual has given his/her consent.

    2. Under Purpose Limitation Obligation - An organisation may collect, use or disclose personal data about an individual for the purposes that a reasonable person would consider appropriate in the circumstances and for which the individual has given consent.

    3. Under Notification Obligation - Notify individuals of the purposes for which your organisation is intending to collect, use or disclose their personal data on or before such collection, use or disclosure of personal data.

    4. Under Access and Correction Obligation - Upon request, the personal data of an individual and information about the ways in which his or her personal data may have been used or disclosed in the past year should be provided. Organisations are also required to correct any error or omission in an individual's personal data upon his or her request.

    5. Under Accuracy Obligation - Make reasonable effort to ensure that personal data collected by or on behalf of your organisation is accurate and complete, if it is likely to be used to make a decision that affects the individual, or if it is likely to be disclosed to another organisation.

    6. Under Protection Obligation - Make security arrangements to protect the personal data that your organisation possesses or controls to prevent unauthorised access, collection, use, disclosure or similar risks.

    7. Under Retention Limitation Obligation - Cease retention of personal data or remove the means by which the personal data can be associated with particular individuals when it is no longer necessary for any business or legal purpose.

    8. Under Transfer Limitation Obligation -Transfer personal data to another country only according to the requirements prescribed under the regulations, to ensure that the standard of protection provided to the personal data so transferred will be comparable to the protection under the PDPA.

    9. Under Accountability Obligation -Make information about your data protection policies, practices and complaints process available on request.

6 Following an outbreak of catpox, the National Infectious Disease Center (NCID) keeps a record of all the infection cases. The records are stored in a list sorted by their date and hour of diagnosis by a doctor.

(a) For contact tracing purposes, NCID wants to identify a patient who was diagnosed on a certain date and time. Explain whether a linear search or a binary search would be more suitable. [2]

As the number of infections increase, it was determined that storing the records in a sorted list is inefficient for contact tracing as the search algorithm is still taking too long. It was suggested to use a hash table with the date and hour of diagnosis as the hash key. For example, someone who was diagnosed at 3:15pm on 13 September 2022 would be assigned the key 13092215 where 130922 is the date and 15 is the hour of diagnosis.

(b) Explain what a hash key is and state the worst case scenario time complexity of a hash table search. [2]

(c) Give **one** advantage and **one** disadvantage to the method provided in finding the hash key. [2]

(d) State **two** methods to resolve collisions in a hash table. Explain which method would be better suited in the context of contact tracing. [3]

Solution:

(a)	A binary search algorithm should be used instead as the list is sorted chronologically. It would be a more efficient method as the time complexity is O(log n) compared to O(n) if a linear search was used.
(b)	A hash key is used to find a location in a Hash Table to search for or store a data. The worst case scenario time complexity for hash table search is O(n).
(c)	This method of finding a hash key is simple. (advantage) However, as most diagnosis would happen in the day time, during the working hours of clinics and hospitals, most of the data would be clustered within a certain range of the hash table, leading to higher chances of collision. (disadvantage)
(d)	2 possible methods are chaining and linear probing. Chaining is the better strategy. Chaining is a better option as it would allow better tracking of trends. Linear Probing would also have a huge disadvantage of having a record stored at a time that is not related to the person who was diagnosed. For example, if 24 people were diagnosed in the same hour, the 24th person would be stored in a record that is 1 full day away. Linear Probing is also not as good as linear probing may result in skipping certain hash keys, leading to a breakage in the "links" or breakage of the "contact tracing", which may be better tracked if chaining was used.

7 A common way to represent negative values using a binary string is the "one's complement" method. To convert a number from its positive to negative (and vice-versa), the binary string needs to be inverted. To invert a binary string, every 0 bit is inverted to become 1 while a 1 bit is inverted to 0. For example, since 5 is 0101 in binary, inverting 0101 will give 1010 which would represent −5.

(a) Using pseudocode, write a function `inv(bin_str)` which inverts the 4-bit binary string `bin_str`. [2]

(b) State the 4-bit binary string that represents −2. [1]

A function `add(num1,num2)` calculates the sum of the values represented by the 4-bit binary strings `num1` and `num2`. The algorithm is given as follows:

```
01   def add(num1,num2):
02       ret = "
03       carry = 0
04       for i in range(3,-1,-1):
05           x = int(num1[i])
06           y = int(num2[i])
07           value = x + y + carry
08           if value <= 1:
09               carry = 0
10           else:
11               carry = 1
12               value -= 2
13           ret = str(value) + ret
14       return ret
```
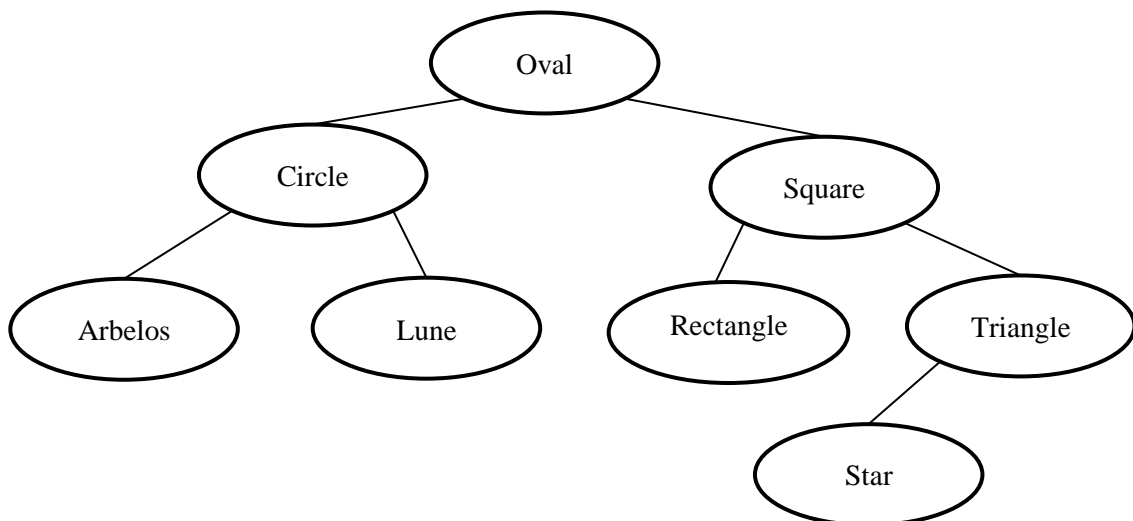
(c) Explain the purpose of lines 07 to 12. [3]

(d) The 4-bit binary strings `bit2` and `bit4` which have numerical values of 2 and 4. Hence, using the functions defined in the question and variables `bit2` and `bit4`, write down the statement to calculate 2 − 4. [2]

(e) Using a suitable example, explain how a miscalculation may occur for the function `add()`. [2]

Solution:

a	`1101`
| b | ```
FUNCTION INV(BIN_STR) RETURNS STRING
   RET_STR ← '0000'
   FOR INDEX ← 0 TO 3
     IF BIN_STR[INDEX] = '0'
       RET_STR[INDEX] = '1'
     ELSE-IF BIN_STR[INDEX] = '1'
       RET_STR[INDEX] = '0'
   RETURN RET_STR
ENDFUNCTION
``` |
c	Line 7 is the line adding the bits from num1 and num2 and any carry-over value. Lines 8 determines that if the value is at most 1, then there would be no need to carry over. (line 9) Otherwise, in lines 10-12, the value is at least 2 and cannot be representing in 1 bit, so a carry bit is added and the value reduces by 2.
d	`add(bit4,inv(bit2))`
e	`4 + 4` should give 8. However, 8 can not defined 4bit by one's complement. Hence, 8 is considered to be too big, and it 'overflows' to return a smaller/negative value. (in this case, it returns `1000` which has a value of `-7`)

8 A data analyst is developing a database of shapes so that different shapes can be recognised using artificial intelligence. Eight different shapes are stored using a binary search tree as shown below.

(a) State the output if a preorder traversal algorithm is used to display the tree. [1]

(b) Draw the resulting tree after the following 3 steps:
- insert 'Geoid'
- delete 'Arbelos'
- delete 'Square' [3]

The binary search tree can be implemented using object-oriented programming. Each node comprises three attributes: a `left` pointer, the `data` and a `right` pointer. `None` indicates there are no further nodes in a particular direction.

(c) Write, **in pseudocode**, a recursive procedure that takes the `root` node and outputs the result of an inorder traversal. [2]

(d) Write, **in pseudocode**, a recursive procedure that takes the `root` node, the `target` data and return `True` if `target` data is found in the tree or `False` otherwise. [4]

Solution:

a	Oval, Circle, Arbelos, Lune, Square, Rectangle, Triangle, Star

b	Oval Circle Rectangle Lune Triangle Geoid Star Alternative: Oval Circle Star Lune Rectangle Triangle Geoid
| c | ```
FUNCTION InOrder(node)
    IF node <> None
        InOrder(node.left)
        OUTPUT node.value
        InOrder(node.right)
    ENDIF
ENDFUNCTION
``` |
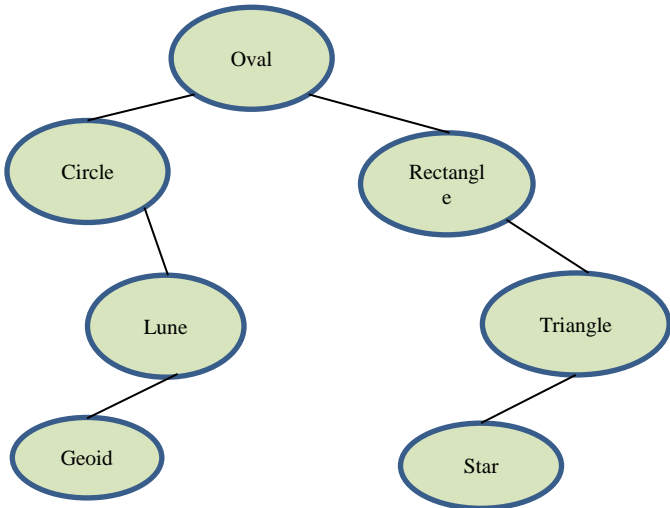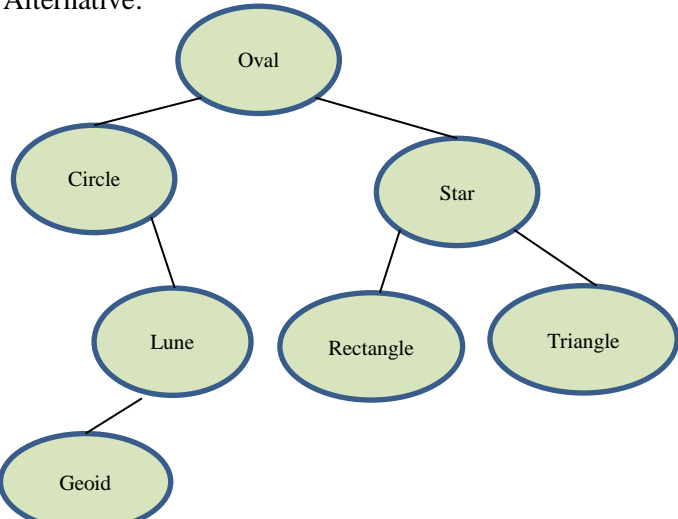| d | ```
FUNCTION Search(tree, target)
    IF tree = None
        RETURN False
    ELSE-IF target < tree.data
        RETURN Search(tree.left, target)
    ELSE-IF target > tree.data
        RETURN Search(tree.right, target)
    ELSE:
        RETURN True
ENDFUNCTION
``` |