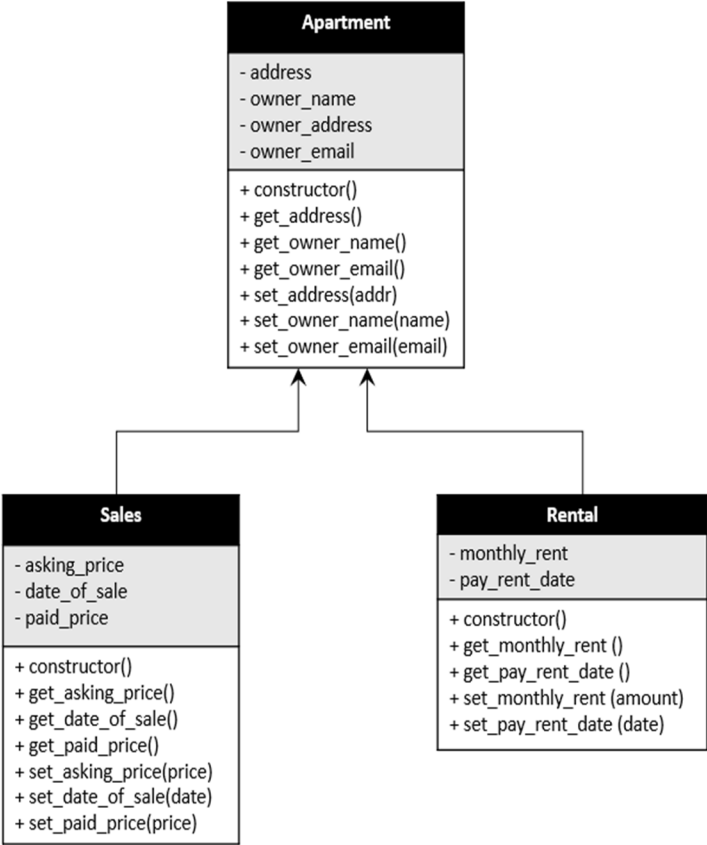# 2022 GCE A Level Theory Paper
Collaborative Learning
(Research, Discuss, Phrase)

Contributor 1: David Tan
Contributor 2: WSM

	Question	Suggested Solution	Comment
1(a)	ER Diagram		
(b)	Table Descriptions	(i) Club (<u>club_name</u>, name, email) (ii) Event (<u>event_no</u>, date, start_time, location) (iii) EventCompetitor (<u>competitor_no</u>, <u>event_no*</u>, competitor_name, club_name*, finish_position, finish_time) (iv) EventClubScore (<u>event_no*</u>, <u>club_name*</u>, score)	Need to store MaxCyclist?
(c)	SQL Query	SELECT club_name, score, name, email FROM Club, EventClubScore WHERE event_no=23 AND EventClubScore.club_name = Club.club_name ORDER BY score DESC;	SELECT Club.club_name, EventClubScore.score, Club.name, Club.email FROM Club INNER JOIN EventClubScore ON EventClubScore.club_name = Club.club_name WHERE EventClubScore.event_no = 23 ORDER BY EventClubScore DESC
(d)	(i) Two aims of normalisation	1. Reduces redundant data so as to optimise storage space. 2. Prevent data inconsistency due to insert, update or delete anomalies.	
	(ii) Two requirements for 1NF table to be 3NF	1. Ensure all non-key attributes are fully functionally dependent on the primary key in each table; hence the	

		tables are in 2NF. 2. For the 2NF tables, ensure that there is no transitive dependency for all the non-key attributes in the tables.	
(e)	Most suitable validation technique for an event score.	Range check : between 0 and 25, inclusive	Validation checks and validation techniques are used interchangeably (referencing from the syllabus and O-level computing textbook v2 and v3) From Examiner's Report: "Some candidates gave alternative validation methods that were not the most suitable within the given context, such as format check or type check."
2(a)	Difference between a class and an object	A class is a template for creating objects in a program, whereas the object is an instance of a class. (Elabration not required) In this question, all the objects created with the Apartment class must contain an address, owner's name and email. Some examples of the objects may be Mr Tan's or Uncle Steven's house which are available for sale or rental.	Does sequencing of question matter?

(b)	Class UML diagram	

Apartment

- address
- owner_name
- owner_address
- owner_email

+ constructor()
+ get_address()
+ get_owner_name()
+ get_owner_email()
+ set_address(addr)
+ set_owner_name(name)
+ set_owner_email(email)

Sales

- asking_price
- date_of_sale
- paid_price

+ constructor()
+ get_asking_price()
+ get_date_of_sale()
+ get_paid_price()
+ set_asking_price(price)
+ set_date_of_sale(date)
+ set_paid_price(price)

Rental

- monthly_rent
- pay_rent_date

+ constructor()
+ get_monthly_rent ()
+ get_pay_rent_date ()
+ set_monthly_rent (amount)
+ set_pay_rent_date (date)

(c)	Explain an advantage of using private properties	Private properties in a class can only be accessed within the class using the accessor and modifiers included. This will prevent unauthorised access or changes by other program code.	
(d)	Explain a benefit of using inheritance	All the properties and methods in the parent class are available for use in the child classes. Hence the software developer need not rewrite all the program code in the child classes again.	

3(a)	State how a message can be made meaningless to anyone other than the intended recipient.	Encrypt the message using the recipient's public key so that only the recipient has the private key to decrypt the message.	
(b)	Explain what sending and receiving devices can do to detect any malicious alteration of a message.	The sender can hash the message with a hashing algorithm like SHA-256 and send both the message and hash value to the receiver. The receiver will hash the received message using the same hashing algorithm to obtain another hash value; compare this hash value with the sender's hash value, the message is not altered if both the hash values are the same. To prevent the man-in-the-middle replacing the message and the hash value from the sender with an altered message and newly generated hash value, the sender can encrypt the original message and the hash value using the receiver's public key so that only the receiver can decrypt it. Alternatively, the sender can encrypt the original message and the hash value using the sender's private key so that the receiver can use the sender's public key to decrypt it. In this way, the receiver will be able to authenticate that the message is indeed from the sender.	Hashing the message can help to detect alteration to the original message, but it will not be able to tell if it is a "malicious" alteration. Qu: "devices" - is it referring to hardware? Digital Signature: <u>Sender Side</u> • 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 <u>Receiver Side</u> • 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
(c)	Describe two situations where authentication is important. For each situation, state how it can be achieved.	1. Authentication is required to ensure only authorised personnel can access the network or the data kept in the server. This can be done by issuing the authorised personnel an userid and password for access. 2. Authentication is important to ensure the true identity	WSM: Similar points 1. A client needs to verify the identity of a server before establishing a connection. 2 way encryption can be used, where the server sends a message encrypted with its private key. If the message can be decrypted using the server's public key, then the identity of the server is verified.

		of a person in situations when personal data needs to be accessed. This can be done by performing a biometric verification with a fingerprint, an iris or retina scan,or facial recognition.	 From ChatGPT: Sure, here are two situations where authentication is important in the computing context and how it can be achieved: User Account Access: Authentication is important to ensure that only authorized users can access their accounts on a computer system. This is typically achieved through the use of a username and password combination. The system will prompt the user to enter their unique username and password, which will then be compared to a database of authorized users to determine if they have access to the system. Network Access: Authentication is important to ensure that only authorized users or devices can access a network, such as a Wi-Fi network or a company's internal network. This is often achieved through the use of a network password or other authentication methods such as digital certificates or biometric authentication. The system will prompt the user or device to enter the correct password or other authentication factor before granting access to the network.
4(a)	State two advantages of using a fixed-capacity array over a linked list to store the ordered items.	1. The elements in the fixed-capacity array can be accessed easily using the index position. 2. No additional memory required for an array, but additional memory is required to store the pointers when using a linked list.	WSM: Additional: In the context, the data items are ordered. By accessing the index, a binary search can be performed on the array, but a linked list is constrained to linear search. ChatGPT: Constant-time access: A fixed-capacity array allows for constant-time access to any element in the array since elements are stored at contiguous memory

			locations. This is not the case with a linked list, where elements may be scattered throughout memory, requiring traversal of the list to access a specific element. Space efficiency: A fixed-capacity array uses a single block of memory to store all elements, whereas a linked list requires additional memory for each element's pointer. This can result in better space efficiency for smaller data sets or when memory is a limited resource.
(b)	State two advantages of using a linked list over a fixed-capacity array to store the ordered items.	1. An element can be easily added in between elements within a linked list by simply reassigning the pointers. 2. A linked list is dynamic size and can keep as many data items as the memory storage is available.	ChatGPT: Dynamic size: A linked list can grow or shrink dynamically as elements are added or removed, whereas a fixed-capacity array has a static size and cannot be resized without creating a new array and copying the elements. Easy insertion and deletion: Insertion and deletion of elements in a linked list can be done in constant time, regardless of the size of the list, by simply updating the appropriate pointers. In contrast, insertion and deletion in a fixed-capacity array may require shifting elements to maintain the ordering of the array, which can be time-consuming and inefficient for large arrays.
(c)	(i) State what line 06 indicates about function Z	Function Z is a recursive function.	
	(ii) State what lines 02 and 03 represent	The represent the terminating case for the recursion function when the value of the pointer of the node is -1, this indicate the node is a final node in the linked list and the function will return 0 which will be added to all the other earlier recursive calls of the function Z.	
	(iii) State the purpose of the function Z	Function Z counts the number of the items in the linked list.	

(d)	Write a pseudo-code function similar to function Z to reverse the order of the data items in the linked list LL.
	
	The Head points to the first item in the current list.
		``
		Function Y(PreviousPointer) RETURNS None
		  Temp = LL(PreviousPointer).Pointer
		  IF PreviousPointer = Head THEN
		    LL(PreviousPointer).Pointer = -1
		    Y(Temp)
		  ELIF Temp = -1 THEN
		    Head = PreviousPointer
		    LL(PreviousPointer).Pointer = PreviousPointer
		    RETURN None
		  ELSE
		    LL(PreviousPointer).Pointer = PreviousPointer
		    Y(Temp)
		  ENDIF
		ENDFUNCTION
		``
		
		Not tested using Python on IDLE
			WSM: This is a difficult qn for students.
			``
			FUNCTION REVERSE(CurrentPointer) RETURNS None
			    IF CurrentPointer = -1 THEN
			        RETURN
			    ENDIF
			    NextPointer = LL(CurrentPointer).Pointer
			    IF CurrentPointer = Head THEN
			        LL(CurrentPointer).Pointer = -1
			        REVERSE(NextPointer)
			    ELSE
			        LL(CurrentPointer).Pointer = Head
			        Head = CurrentPointer
			        REVERSE(NextPointer)
			    ENDIF
			ENDFUNCTION
			
			REVERSE(Head)
			
			Tested to work
			``

(e)	It is common to sort items in a fixed-capacity array. Explain why a merge sort may be faster than a quicksort in this situation.	The time complexity for merge sort is always O(n*log n). But for Quicksort, the best and average case time complexity is O(n*log n) and the worst case is O(n^2). The worst case happens when all the pivots chosen are either the largest or smallest element. Since the array contains an ordered data, if the first or last term is chosen as the pivot, then it would be a worst case scenario for the quicksort. Hence merge sort may be faster than quicksort in this situation.	Sorting in a fixed-capacity array is an in-place sorting but merge sort is not an in-place sorting algorithm. WSM (Additional): Quicksort algorithms can possibly use the first term or last term as a pivot. In this situation, since the data set is ordered, the partitions created by the quicksort will be imbalanced, leading to a worse case scenario of O(n^2).
5(a)	Explain how a mistake in the algorithm is causing the recurring problem.	1. The chosen floor number is stored in the memory. 2. When a user chooses to go to a different floor, the elevator will move to the required floor and release the memory used. 3. But when the user chooses the same floor number as he is, then the program code will end without releasing the used memory. Hence run out of memory	

		eventually and the elevator ceases to operate.	
(b)	Explain how the algorithm should be changed to prevent the problem recurring.	Add the command line to "Release the memory used to store the floor number" if the elevator is at the chosen floor before ending the program, ie. the command line "Finished".	
(c)	Suggest two reasons why the mistake was not identified during testing.	The program code did not run into any error during testing: 1. During testing, it is unlikely that the programmer will enter the elevator and choose the same floor number where he is. Even if he did, the program will finish without running into any error even though the memory used is not released. 2. Only a small amount of memory is used to store the floor number, hence it will exceed the memory capacity after running the program code and selecting the same floor where he is for a large number of times.	WSM: Similar ans In a user acceptance test, the time period of testing can be short (1 to 2 days), and the scenario of memory overflow would not occur. Under real-world conditions over months and years, memory would run out/overflow, leading to an error.
6(a)	(i) Describe the purpose of creating a backup.	1. The purpose of the backup is to create a copy of current data in the event of a primary data failure due to hardware or software failure, data corruption, data loss due to malicious attack (virus or malware), or accidental deletion of data. 2. The backup copies will allow the data to be restored to the latest backup point thereby reducing the amount of data loss and for the business to recover and resume operation quickly from the unplanned event.	
	(ii) Describe the purpose of archiving.	1. Archiving is securely storing away inactive information for long periods of time using a cheaper storage medium. 2a. Some businesses are required by some regulatory organisations to archive the old data for a period of time. 2b. Such archived information may be referred to again in the future when the need arises.	

(b)	Explain why backup copies of the data should be stored off-site.	1. Backup copies stored off-site will not be destroyed together with the primary data at the original site. 2. Backup copies stored off-site will reduce the possibility of data loss due to breaking-in into the physical building or network in the primary site. Point (2) may be a weak answer.	
(c)	Describe the consequences of the business not backing up the data.	1. Loss of current production data. 2. Business will not be able to resume operation quickly without the backup data during an unplanned failure. 3. Company reputation will be affected and may face legal actions which require the company to compensate the customer for the business or financial loss due to the failure. Point (3) may be a weak answer.	WSM: Your point 3 is good/better than mine. I mentioned that a consequence would be to rely on archives to restore part of the data, and extra effort to repopulate the database needs to be made, affecting the business operations.
7(a)	(i) Explain why transmitted data is divided into packets.	1. For efficiency, the network bandwidth can be shared for transmission instead of dedicating the channel for a particular transmission in the case of circuit switching. 2. Re-transmission is only required for the packets which are lost during transmission. 3. Different packets are transmitted via a path that is the most efficient at that instant, hence it improves the overall transmission speed.	WSM: - The data to be transmitted could be too big to fit into one packet. - Sending data as smaller packets prevents network congestion.
	(ii) Explain why the packets are sequentially numbered.	The packets can travel in different paths from the sender to the receiver, hence they need to be sequentially numbered before transmission so that they can be assembled back into the original order at the receiving end. Will this get 2 marks? https://en.wikipedia.org/wiki/Maximum_transmission_unit	WSM: Packets sent over the network can travel in different paths and arrive in a different order. Without sequence numbers, assembling the packets in order of arrival can lead to an incorrect sequence/result.
	(iii) State two items,	Besides the packet number, the packet header	WSM: I only mention source and destination IP

	other than the packet number, that are stored in the packet header.	contains the source address, a destination address, and the protocol used. (Qu: Which layer contain what?)	addresses
(b)	Explain why protocols are required to enable reliable communication over the internet.	1. Communication protocols allow different network devices to communicate with each other. 2. For a successful communication to take place, a set of rules is to define the start and end of the communication, the format and standard for the transmission and how to handle transmission errors.	
(c)	A router is a device that allows the connection of a LAN to the internet. Explain how the router directs arriving data packets to the correct device on the LAN.	1. The arriving data packet's header contains the IP address of the destination. 2. The router uses the subnet mask to determine which subnet the destination belongs to and route the packet to the respective network switch. 3. The network switch will refer to the MAC address table to direct the packet to the correct destination device.	WSM: The router uses port forwarding to check which is the destination device and its MAC address, and converts the IP packet into ethernet frames to send it to the correct device with the correct MAC address.
(d)	A firewall is often placed between a LAN and the internet. Explain how a firewall can provide security to the LAN.	A firewall uses a set of rules to examine the incoming traffic before it enters the LAN. It protects the LAN by performing Packet Filtering - to allow or block the traffic based on the packet's source address, the destination address or the application protocols.	
8(a)	Explain the advantage a hash table search might have over a linear search and a binary search when searching for a specific record. Refer to time complexity in your answer.	Describe how hash table search is done -> The time complexity to perform a search in a hash table is O(1). The time complexity to perform a linear search is O(n). The time complexity to perform a binary search is O(log n) and it is required that the data are sorted before performing the binary search.	
(b)	Explain the meaning of a collision in the context of a hash table search.	To search if a particular data exists in a table, the hash value of the data is first generated. Using the generated hash value as the index, access the element in the table:	WSM: A collision occurs when elements in a hash table share the same hash value. When searching for an item, and another item is found, it does not mean that the search item is not in the table. A collision resolution algorithm is used

		- the data is found in the table if the element in the table at the index is the same as the data; - the data does not exists in the table if it is empty; - there is a collision if the element in the table at the index is not the same as the data. This does not mean that the search item is not in the table because the collision resolution algorithm may have stored the item at another location.	to continue searching for the item at other locations before a conclusion is made.

(c)	Describe one method that can be used to handle the consequence of a collision.	Either one of the following:
		1. Open hashing (Separate Chain)
		  - describe
		
		2. Closed hashing (Linear probing)
		  - describe	

(d)	A hashing algorithm is used to calculate the index of a hash table from a record key. Give three features of an effective hashing algorithm.	Any of the following 3: 1. Produce a hash value corresponding to the indices of the table. 2. Equal probability of generating the indices or the indices generated are spread evenly across all the full range of the indices of the table. 3. Regardless of the length or the data data, the hash function will generate an integer number corresponding to the indices of the table. 4. Use all the parts of the input data to generate the hash value.	TQR: Generate a unique hash value for every unique input Time complexity of O(1)

Diagrams (Please label the question no. and part clearly.)