BCS Higher Education Qualification

Professional Graduate Diploma

October 2023

EXAMINERS' REPORT TEMPLATE

Advanced Database Management Systems

Questions Report:

	All markers of section A are required to add comments here:
A1	
	Approximately 30% of candidates attempted this question of which just over half achieved a pass grade in the question.
	This question was poorly answered, with candidates seeming to know the fundamental ideas related to object orientation and how it compares to database modelling approaches. However, answers generally lacked depth of understanding. In part a) as well as part b) the vagueness of understanding of object oriented concepts was most obvious, with many students failing to identify the relation between an entity and tables/relations and the respective OO concepts, as well as a lack of understanding of inheritance, overwriting, encapsulation and how these transpire in database modelling. The substantial question on ORM often had very short answers that simply missed many of the aspects that needed considering for describing how data between objects and database is mapped. Part c) was generally well understood, but some candidates returned too many or too few values as well as not always identifying the correct columns to be returned.
A2	
	This question was attempted by 75% of candidates, with approximately 50% of these passing the question.
	Most candidates could describe the key concepts, uses of and approaches in data warehousing. Many would have scored higher by providing some detail beyond the keywords and ensuring that they did not only select one or two of the possible four or five aspects that were of relevance. Also, most parts were asking for examples and candidates omitted providing these. The prior observations are relevant to all the subquestions in this question, but in the final part on metadata many failed to express the essence that metadata is data about data.
A3	
	Approximately half of the candidates answered this question, with approximately 40% of those passing the question.
	This question was poorly answered, with candidates often not answering subquestions or doing so very superficially. The technical results were generally correct (i.e. the result of running queries), but the explanations were imprecise. For part a) many students explained what each line of the code would do but did not explain the overall function of the trigger. Part d) was the most problematic – it required discussion of a number of tuning methods and many answers did not manage to explain clearly what the mentioned techniques were or how they help improve performance.

B4	
	This was the most popular question on the paper attempted by around 90% of candidates. Overall performance was one of the best on the paper with around 64% achieving a pass mark. There was a good spread of marks with a top mark of 25 achieved.
Part a)	This question had three subparts, related to data security, all of which attracted a good set of marks overall. Most candidates could explain concepts of SQL injection and described how attacks occur. Examples of SQL code that result in an injection attack were expected and were rewarded with higher marks. This also applies to the remedial action to prevent attacks. The role of a database administrator (DBA) in handling a database was familiar to most candidates, as were the data masking techniques. Examples of each technique were expected to show how these techniques were used in practice.
Part b)	This question had three subparts related to distributed databases, specifically data fragmentation. This part was generally well answered and showed a sound understanding of this topic. The final subpart included an example to apply data fragmentation. Given that the benefits of data fragmentation were well understood by candidates such as local autonomy, it was surprising that many candidates incorrectly produced two vertical fragments that included duplicated columns that of the original table. Each fragment should autonomously reflect the two views: academic and finance.
Part c)	This part covered conformity of input into a database. A significant number of candidates seemed to have difficulty understanding the requirements of this question. Many answers concentrated on data integrity constraints applied within the database using SQL DDL (such as Entity Referential Integrity) rather than covering the shape of data. This means: When data is inserted; is it of the right type, length and format? An example was expected such as the date of birth needs to be valid; such as in the past have a suitable shape (DD-MM-YYYY). This is validated when data is inserted into the database and is usually achieved by suitable masks on data entry fields.
B5	
	This was a fairly popular question, attempted by around 67% of candidates. Overall performance was one of the highest on the paper, having a pass rate of around 63%. There was a good spread of marks. This question covered concurrency control in a multi-access RDBMS.
Part a)	This part had two subparts covering concepts related to ACID support for transactions. Most candidates were familiar with the ACID acronym and gave well-reasoned answers to the conflict of maintaining ACID properties when there are transactions that run for a long time before being able to commit. It was good to see that many candidates provided a suitable example of a long-running transaction.
Part b)	This part had two subparts where candidates were required to explain the effects of different concurrency control scenarios. Two further subparts related to row level clocking and conservative 2 phase locking, which were fairly well answered. The first scenario depicted two concurrent transactions that place locks on the same data item. Many candidates didn't consider the type of lock applied, whether it was an exclusive lock or a shared lock. If the transactions simply read the data item, then

	the logic is that conflicts only apply when no other transaction is trying to update that data item. The second scenario included a schedule of two concurrent transactions that had no locks applied to data items that interleaved read and update operations on the same data item. Many candidates had difficulty expanding the schedule by applying the two different concurrency control strategies (pessimistic vs optimistic). Candidates were generally fairly familiar with the pessimistic approach as this followed a strict locking strategy to guarantee consistency. An optimistic locking strategy based on versioning was unfamiliar to many candidates and as a result, scored fewer marks than for the pessimistic locking strategy. The final outcome is that in both cases (optimistic or pessimistic), one of the transactions would either have to abort or roll back to prevent a lost update.
Part c)	This part covered concurrency control in a distributed database where consistency between two databases was managed by coordinating the throughput of transactions to each database. Knowledge of the 2 phase commit protocol was required in order to answer this question. A small number of candidates were sufficiently familiar with the protocol and managed to conclude that coordinator C would send an 'abort' transaction to both participating databases.