

---

# Examiners' commentary

## 2018–2019

---

### CO2209 Database systems – Zone B

#### General remarks

This year's examination did not present any surprises for those who completed the coursework and revised from previous years' examinations. A few candidates actively, and hopefully many more passively, took advantage of the course discussion board on the VLE to present proposed solutions to past examinations – both the VLE tutor and fellow students commented on these solutions at length. This was a particularly pertinent method for revision, because the same kinds of problems that proved challenging in past papers, also proved challenging in this year's paper, although perhaps not so challenging for those who followed the discussions on the VLE.

A small number of candidates presumed on the goodwill, and the eyesight, of the examiners, by submitting a script covered with barely-legible writing. One paper's answers were written in letters no larger, by measurement, than three millimetres.

For those reading this report in the hope of doing well in the next examination, it cannot be emphasized enough that close attention to past papers and past examination commentaries, and at least an acquaintance with the course discussion board, plus completion of the coursework assignments, is the key to success.

#### Comments on specific questions

##### Question 1

This question covered Primary Key choice. A table was presented with its proper Primary Key given, and candidates were asked to extrapolate the consequences of choosing a different set of attributes for the Primary Key. Too 'narrow' a choice of Primary Key will not allow some valid tuples to be added, too 'wide' a choice will let in invalid data. Some candidates interpreted the question to mean what would happen if the relation as presented to them, with data already in it, suddenly had its Primary Key changed, rather than what would happen if the original relation had had this key and was presented with the data shown. Allowance was made for this interpretation. Some candidates over-used the phrase 'the uniqueness property would be lost'. In examinations, it is best to illustrate phrases like this with an example, or to spell out your answer with further information – just using a phrase like 'the uniqueness property would be lost' raises the suspicion that this is just a memorized phrase, offered in the hopes that it will 'fit' the question.

Candidates attempting this question also had some definitions to write. Telling the examiner that 'Attribute Integrity' is about the 'integrity of Attributes' is unlikely to gain many – or any – marks. There are only a small number of basic concepts in relational database theory, probably no more than two dozen, at a stretch. Some of them are likely to appear in any examination on the subject. It's not difficult to learn them, letter-perfect. Examiners will naturally give some latitude when marking an essay. But a sloppy, careless definition – using 'key' when 'attribute' is meant, for example – simply suggests that the candidate could not be bothered.

## Question 2

This question began with a 5-mark question about how a database system might use the statistics it collects regarding usage of the database. Although most candidates were able to suggest a plausible use, it was clear that understanding here was thin, suggesting that next year's coursework should orient students towards some of the physical aspects of database usage. The rest of the question looked at concurrent access, the problems this can cause, the possible solutions, and the possible problems that can arise from these solutions. It was generally adequately answered, although the last part of the question gave a number of candidates a problem: it was clear that they really didn't know what a 'logical unit of work' meant. Nor could every candidate illustrate the various circumstances in which a transaction faces a system failure before it has been completed and written to permanent storage.

## Question 3

Here candidates encountered a range of subjects, beginning with 'Entities and Relationships'. A simple E/R diagram showing a one-to-many relationship with optional participation on one side was required. Some candidates got the cardinality indicator the wrong way around. Still on the subject of E/R diagrams, candidates needed to give an illustration of a 'Chasm Trap' situation and show how it might be resolved. This was often done very badly, it must be said. The impression given was that revision had taken place, but only in terms of memorizing a diagram without understanding its meaning. So many wrong variants of a 'Chasm Trap' were presented which were not a Chasm Trap. Some Fan Traps were presented which were Fan Traps, but these are not the same as a Chasm Trap. For a discussion of Fan and Chasm Traps, see Section 5.3.3 in Volume 1 of the subject guide.

Part C. of the question presented some terms to be defined, which was generally done well, although the usual errors also appeared: the word 'key' used where 'attribute' was meant, and the reuse of the term to be defined in the definition itself: it conveys no information to say that a 'determinant' 'determines' something – it should be shown what 'determination' actually means.

Parts D. and E. of the question, about database architecture and the difference between security and integrity, were not inherently difficult and most candidates did well on them. A few defined 'security' rather narrowly, presenting it as only securing the database by restricting what its own authorized users can see, while security also involves its protection from any unwanted would-be accessor.

## Question 4

This question was the usual 'SQL question', and candidates did well here. As usual, a significant number could not handle the two questions requiring set operations or their sub-query equivalents. The seventh query required a sub-query to a given salary to the maximum salary, but many candidates constructed a query that simply displayed the maximum salary alongside the first name in the table. The part of the question dealing with VIEWS did not present a problem for the great majority of candidates, although some of them seemed to think that the only advantage of a VIEW was as a security measure, and many believed that a VIEW created a new, separate relation.

## Question 5

This question probed candidates' understanding of the concept of functional dependency. Many candidates wasted time and paper by elaborately deriving

their final set of relations from the original unnormalized relation using Heath's Rule. The worthy Mr Heath is not necessary here – simply following the rule, 'let every Functional Dependency be a candidate key' will result in the desired set of normalized tables. Anyone reading this in preparation for a future examination must make sure that they really understand what an unnormalized table 'anomaly' is. 'Anomaly' is not a synonym for 'mistake'. As a way of making sure that a proposed 'anomaly' is indeed one, ask this question: is this a problem that could occur with an unnormalized table, but could NOT occur with a normalized table? If it COULD occur with a normalized table, then it is not an anomaly.

## **Conclusion**

These examinations are not intended to be arcane, or tests of candidates' abilities to engage in subtle decoding of the examiner's hidden intent. Systematic revision, undertaken over several weeks, not several days before the date of the examination, combined with use of the course discussion board, will come close to guaranteeing success.