DUDEBASE COLLEGES PROGRESS EXAMINATION

Tuesday 18 January 2022 11:30 – 12:30

Computer Science Paper 3 (CST IA)

Answer **one** *question from each Section. Each question is worth the same number of marks.*

Write on **one** *side of the paper only.*

Write your name and the question number at the top of every sheet, and tie your answers into separate bundles (one for each question).

DO NOT TURN OVER THE QUESTION PAPER UNTIL TOLD BY THE INVIGILATOR THAT YOU MAY DO SO

SECTION A

1 Databases

(*a*) In a relational database based on an entity-relationship model, what factors govern whether the schema should be extended to store additional information by adding attributes to an existing entity/relationship, or by adding a new type of entity and a relationship between it and an existing one? [4 marks]

(*b*) A new web-scale service requires a vast database of jokes. A *Joke* consists of a unique ID, a *question* (e.g. "What did the pirate say on his 80th birthday?") and an *answer* (e.g. "Aye matey"). Identify a suitable database technology and provide an implementation to satisfy the use-case of fast lookup by Joke ID. [2 marks]

(*c*) It is now required to store when and where a comedian used a joke. Show how you would extend your database given that the two use-cases are to list every occasion when a particular joke (specified by ID) was used; and to list every joke used by a particular comedian. [4 marks]

(*d*) How would you ensure that your database does not contain duplicate copies of any joke? [1 mark]

(*e*) Further extend your database to store 5-star joke ratings (an integer between 1 and 5 inclusive). Consider the storage space required by your solution in the event that some jokes attract a large number of ratings. [3 marks]

(*f*) Suppose the earlier use-case to retrieve a joke by ID must now include an average rating, a minimum rating and a maximum rating. How would you modify your database (or application), and why? [2 marks]

(*g*) Suppose it became necessary to have to find the highest-ranked joke. How would your approach differ if this were a one-off request, were required very regularly, or if the requirement were modified to find the highest-rated joke within a particular time frame (e.g. the joke with highest ranking in 2017)? [4 marks]

2 Databases

(*a*) Using entity-relationship diagrams and standard notation, depict...

(*i*) A entity *User* with attributes `CRSID`, `Name`, and `TelNum`. `CRSID` is the primary key. [2 marks]

(*ii*) The sub-entities *Student* and *Supervisor*, each keyed by `CRSID` and inheriting the attributes of a *User*. Students have an attribute `TriposName` and supervisors have an attribute `TriposTaught`. [3 marks]

(*iii*) The weak entity *SupervisionArrangement*, depending on the existence of both a student and a supervisor, and having the structured attribute `StartTime` consisting of `year`, `month`, `day`, `hour`, and `minute`. [3 marks]

(*b*) Consider an SQL implementation of the database in part (a). Write SQL queries to find:

(*i*) the number of occasions when someone supervised themselves; [2 marks]

(*ii*) the number of occasions when someone supervised someone with the same name as themselves (but not themselves); [2 marks]

(*iii*) a list of `(Name, Size)` pairs sorted by `Name`, where `Name` is a supervisor's name and `Size` is the number of students attending a supervision. This list should contain one row per supervisor, per `StartTime`. [3 marks]

(*iv*) a list of the average size of supervisions given by each supervisor, sorted by supervisor name; [3 marks]

(*v*) a list of supervision 'grandchildren': `(Name1, Name2)` where supervisor `Name1` supervised some person, `X`, for a course, and `X` supervised `Name2` for the same course. [2 marks]

SECTION B

3 Introduction to Graphics

(a) Recall the two processes involved in digitising an analogue image: sampling and quantising.

(i) What issues might be introduced as a result of each process? [2 marks]

(ii) Assuming in a greyscale image that the pixel intensities are each uniformly distributed between 0 and 1, what are the average error and the average magnitude of the error introduced by 8-bit quantisation? [2 marks]

(iii) Why does taking multiple samples per pixel ameliorate the issues introduced by sampling? [1 mark]

(iv) Describe a method for reducing the undesirable effects of sampling using a Poisson disc. [2 marks]

(b) Consider the vector $(1, 1)$ in Euclidean space.

(i) Give an expression defining the points in a homogeneous co-ordinate system to which this vector maps. What are the advantages of representing the point in such a homogeneous co-ordinate system? [2 marks]

(ii) Construct a matrix that rotates the vector $(1, 1)$ by an angle of 45° about the point (a, b). [2 marks]

(c) Explain the differences between pragmatic and uniform colour spaces, citing examples of each. What is the purpose of the XYZ colour space, and what is the meaning of the primary Y? [3 marks]

(d) Consider a solid 2D triangle that has vertices $(0, 0)$, $(2, 0)$ and $(2, 2)$. The surface is Lambertian and absorbs 20% of blue light, 80% of green light and 10% of red light. A point light source at $(1, 25)$ emits light with intensity 1.0 in each of the R, G and B channels. Calculate the colour to which the pixel at $(1, 1)$ should be set. [6 marks]

4 Introduction to Graphics

(a) When raytracing a 3D scene, why do we use modelling, viewing and projection transformations? [3 marks]

(b) Which transformation(s) would change if...

(i) an object in the scene was moving? [1 mark]

(ii) an object in the scene was changing shape? [1 mark]

(iii) we wanted to model light bending due to extreme gravity of a blackhole? [1 mark]

(c) Suppose your graphics hardware has a z-buffer that supports image sizes up to 2000×2000 pixels in 32-bit colour. You have an integer-valued mathematical function, $f(x, y)$ and, over the (32-bit) integers x and y. Using pseduocode, explain how you could use the z-buffer to find, for every x between -1000 and 999, the value of y that minimises $f(x, y)$. [5 marks]

(d) How does bump mapping cause raytracing to render an object so it appears to have an uneven surface? Why might someone prefer displacement mapping? [5 marks]

(e) Why do we triangularise 3D objects before rasterising them? [4 marks]

END OF PAPER