Assignment 2

Please make sure that you always use notations consistent with lecture notes. Different notations will not be accepted.

The deadline for assignment 2 is: 5pm Fri 17th, November

Question 1 (12 marks)

Consider a relation R (A, B, C, D, E, G, H, I, J) and its FD set F = {A -> EI, BC -> DEG, CEH -> GJ, D -> CJ, DHJ ->AB}

Regarding the following questions. Give and justify your answers if the question is specified.

1) Check if BD→G. Justify your answer. (1 mark)
2) Find all the candidate keys for R. (2 mark)
3) Determine the highest normal form of R with respect to F. Justify your answer. (2 marks)
4) Find a minimal cover F_m for F. (2 marks)
5) Regarding F, does the decomposition R1 = {ABDE}, R2 = {CEHJ}, R3 = {AGIJ} of R satisfy the lossless join property? Please justify your answer. (2 marks)
6) Provide a step-by-step lossless decomposition of R into BCNF normal form. (3 marks)

Question 2 (6 marks)

Consider the following query:
P1, P2, P3, P4, P3, P1, P5, P2, P3, P6, P7, P5, P2, P1.
(The user is trying to read page 1 from disk, then page 2, page 3, …)
Assume there are 4 buffers in the buffer pool.

1) Sketch the process of how blocks are replaced in the First In First Out (FIFO) policy. (2 marks)
2) Sketch the process of how blocks are replaced in the Least Recently Used (LRU) policy. (2 marks)
3) Between FIFO and LRU policies, which one performs better in the given query? Why? (2 marks)

Question 3 (8 marks)

Consider the schedule below. Here, R(*) and W(*) stand for 'Read' and 'Write', respectively. T1, T2, T3 and T4 represent four transactions and t_i represents a time slot.

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}	t_{14}	t_{15}	t_{16}
T1	R(A)					W(C)						R(A)			W(A)	
T2				R(B)			R(C)	W(B)		W(A)						
T3		R(B)	R(D)								W(B)		W(D)			
T4					R(C)				W(C)					R(D)		W(D)

Each transaction begins at the time slot of its first operation and commits right after its last operation (same time slot).

Regarding the following questions. Give and justify your answers.

1) Assume a checkpoint is made between t_4 and t_5, what should be done to the four transactions when the crash happens between t_{11} and t_{12}. (2 marks)

2) Is the transaction schedule conflict serializable? Give the full precedence graph to justify your answer. (2 marks)

3) Construct a schedule (which is different from above) of these four transactions which **causes** deadlock when using two-phase locking protocol. You should clearly indicate all the locks and the corresponding unlocks in your schedule. If no such schedule exists, explain why. (4 marks)

Assignment Submission

- Students must submit an electronic copy of their answers to the above questions to the course website in Moodle.
- Only **.doc** or **.pdf** file is accepted. The file name should be ass2_**studentID**.doc or ass2_**studentID**.pdf (e.g., **ass2_z5100000.doc** or **ass2_z5100000.pdf**).

Note:

1. If you have problems relating to your submission, please write to the course email at junhua.zhang@unsw.edu.au or yiheng.hu@unsw.edu.au.
2. All submissions will be checked for plagiarism. The university regards plagiarism as a form of academic misconduct and has very strict rules. Not knowing the rules *will not be considered a valid excuse when you are caught.*
   a. For UNSW policies, penalties, and information to help avoid plagiarism, please see: https://student.unsw.edu.au/plagiarism.
   b. For guidelines in the online ELISE tutorials for all new UNSW students: https://subjectguides.library.unsw.edu.au/elise/plagiarism.

Late Submission Penalty

- 5% of the max assessment mark will be deducted for each day late (1 second late is considered as one day late), up to 5 days (5*24 hours).
- Submissions that are more than five days late will not be marked.