

# CSC3170 Introduction to Database Systems

## (Spring 2024) Assignment 2 –

### Non-SQL PART

Please answer all the questions below and submit your answer to blackboard

#### Question 1 (16 marks)

Consider a relation  $R(A, B, C, D, E, G, H, I, J)$  and its FD set  $F = \{AB \rightarrow DE, C \rightarrow GH, E \rightarrow BCD, D \rightarrow CI, H \rightarrow G, EH \rightarrow I\}$ .

- 1) Check if  $E \rightarrow G \in F^+$ . Justify your answer. (2 mark)
- 2) List all the candidate keys for  $R$ . (2 marks)
- 3) How many super keys can be found for  $R$ ? Compute the total number of super keys and list 5 of them. (2 marks)
- 4) Find a minimal cover  $F_m$  for  $F$ . (2 marks)
- 5) Determine the highest normal form of  $R$  with respect to  $F$ . Justify your answer. (2 marks)
- 6) Regarding  $F$ , is the decomposition  $R_1 = \{ABCDE\}$ ,  $R_2 = \{CGH\}$ ,  $R_3 = \{EIJ\}$  of  $R$  dependency-preserving? Please justify your answer. (2 marks)
- 7) Regarding  $F$ , is the decomposition  $R_1 = \{ABCDE\}$ ,  $R_2 = \{CGH\}$ ,  $R_3 = \{EIJ\}$  of  $R$  lossless-join? Please justify your answer. (2 marks)
- 8) Decompose it into a collection of BCNF relations if it is not in BCNF. Make sure your decomposition is lossless-joined and briefly justify your answers. (2 marks)

## Question 2 (8 marks)

Consider the schedule below. Here,  $R(*)$  and  $W(*)$  stand for 'Read' and 'Write', respectively.  $T_1$ ,  $T_2$ ,  $T_3$  and  $T_4$  represent four transactions and  $t_i$  represents a time slot.

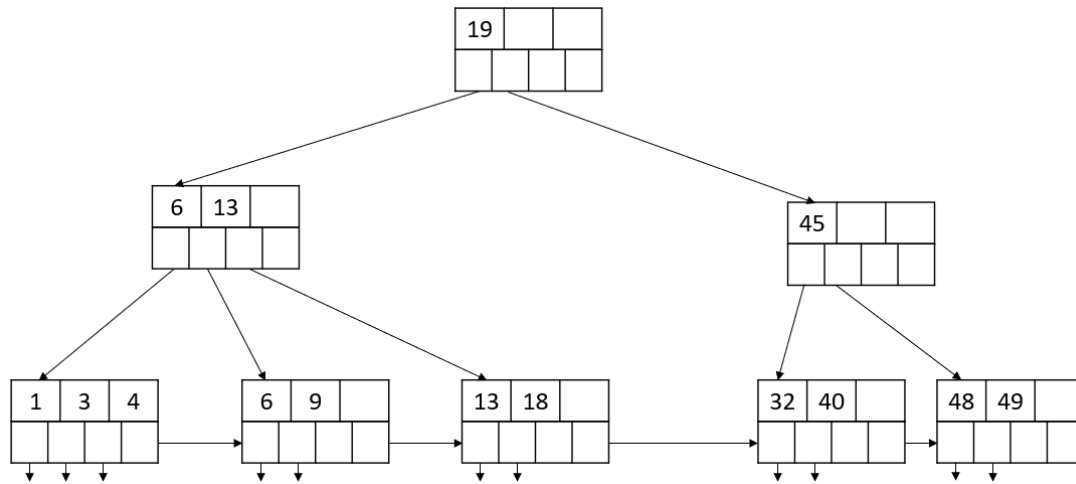
Time	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	$t_7$	$t_8$	$t_9$	$t_{10}$	$t_{11}$	$t_{12}$
$T_1$	R(B)					R(A)	W(B)				W(A)	
$T_2$			R(A)		R(A)				R(B)	W(B)		W(A)
$T_3$								R(B)				W(B)
$T_4$		R(A)		W(A)						R(B)	W(B)	

Each transaction begins at the time slot of its first Read and commits right after its last Write (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  $t_6$  and  $t_7$ . (2 marks)
- 2) Is the transaction schedule conflict serializable? Give the precedence graph to justify your answer. (2 marks)
- 3) Give a serial schedule of these four transactions. (2 marks)
- 4) Construct a schedule (which is different from above) of these four transactions which causes deadlock when using two-phase locking protocol. If no such schedule exists, explain why. (2 marks)

### Question 3 (6 marks)



- 1) There are currently 11 records in this tree. How many additional records could be added to this tree without changing its height (give the maximum possible number)? (3 marks)
- 2) Show the B+ tree after deleting the data entry with key 49 from the original tree. (3 marks)