Assignment No. 1 for Ch. 02 (Spring 2024)

1. Consider the relation PAY in Figure below. Suppose there is one application that access PAY. It is using 'TITLE' in PAY included in its simple predicates.
   1) Find all possible simple predicates for 'TITLE' in PAY.
   2) Derive the primary horizontal fragmentation of PAY using the simple predicates in 1).
   3) Using the fragmentation of PAY in 2), perform a derived horizontal fragmentation for EMP.
   4) Using the fragmentation of EMP in 3), perform a derived horizontal fragmentation for ASG.

EMP

ENO	ENAME	TITLE
E1	J. Doe	Elect. Eng
E2	M. Smith	Syst. Anal.
E3	A. Lee	Mech. Eng.
E4	J. Miller	Programmer
E5	B. Casey	Syst. Anal.
E6	L. Chu	Elect. Eng.
E7	R. Davis	Mech. Eng.
E8	J. Jones	Syst. Anal.

ASG

ENO	PNO	RESP	DUR
E1	P1	Manager	12
E2	P1	Analyst	24
E2	P2	Analyst	6
E3	P3	Consultant	10
E3	P4	Engineer	48
E4	P2	Programmer	18
E5	P2	Manager	24
E6	P4	Manager	48
E7	P3	Engineer	36
E8	P3	Manager	40

PROJ

PNO	PNAME	BUDGET
P1	Instrumentation	150000
P2	Database Develop.	135000
P3	CAD/CAM	250000
P4	Maintenance	310000

PAY

TITLE	SAL
Elect. Eng.	40000
Syst. Anal.	34000
Mech. Eng.	27000
Programmer	24000

2. Consider relations ASG and PROJ in Figure above. ASG and PROJ are horizontally fragmented as follows:

$$ASG_1 = \sigma_{PNO = \text{“P1”}}(ASG)$$

$$ASG_2 = \sigma_{PNO = \text{“P2”}}(ASG)$$

$$ASG_3 = \sigma_{PNO = \text{“P3”}}(ASG)$$

$$ASG_4 = \sigma_{PNO = \text{“P4”}}(ASG)$$

$$PROJ_1 = \sigma_{BUDGET \geq 200000}(PROJ)$$

$$PROJ_2 = \sigma_{BUDGET < 200000}(PROJ)$$

1) Draw the join graph of ASG $\ltimes_{PNO}$ PROJ. Is the graph simple or partitioned?

2) If it is partitioned, modify the fragmentation of ASG so that the join graph of ASG $\ltimes_{PNO}$ PROJ is simple, and provide new fragmentation.

3) Draw the join graph(s) based on the new fragmentation.

3. Let Q = {q1, q2, q3, q4, q5} be a set of queries, A = {A1, A2, A3, A4, A5} be a set of attributes, and S = {S1, S2, S3} be a set of sites. The matrix of Figure (a) below describes the attribute usage values, and the matrix of Figure (b) gives the application access frequencies. Assume that $ref_s(q_k) = 1$ for all q_k and S_i.
   1) Derive an AA (Attribute Affinity) matrix (Provide all the details how to get it).
   2) Derive a CA (Clustered Affinity) matrix using the bond energy algorithm (Provide all the details how to get it).
   3) Using the partitioning algorithm ($Z = CTQ*CBQ - COQ^2$), partition the attributes in the above CA to two partitions.

	A1	A2	A3	A4	A5
Q1	1	0	0	1	0
Q2	1	1	1	0	0
Q3	0	0	0	1	1
Q4	1	0	1	0	0
Q5	0	1	0	1	1

(a)

	S1	S2	S3
Q1	20	30	0
Q2	10	0	15
Q3	0	45	15
Q4	0	20	0
Q5	0	25	0

(b)