Assignment No. 3 for Ch. 4.2 and Ch. 5 (Spring 2024)

- Example database
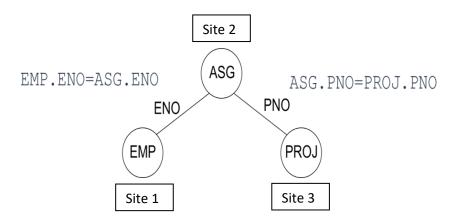
EMP

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

1. (Ch. 04) Consider the join graph below and the following information: size(EMP) = 300, size(ASG) = 400, size(PROJ) = 500, size(EMP ⋈ ASG) = 500, and size(ASG ⋈ PROJ) = 400. Describe an optimal join program that minimizes response time (consider only communication) without using semi-join.

2. Consider the join graph above in 1, and give a program (possibly not optimal) that reduces each relation fully by semijoins. The final query result needs to be in Site 2.

3. (Ch. 04) Consider the following query, the join graph and the distribution of the relations and/or fragmentations to the sites depicted below. We assume all records have same size. size(EMP ⋈ ASG) = 3000, and size(ASG ⋈ PROJ) = 2000. The goal is to minimize communication cost. The final query result needs to be in Site 2.

> SELECT ENAME
> FROM EMP, ASG, PROJ
> WHERE PROJ.PNO = ASG.PNO AND
>   EMP.ENO = ASG.ENO

Relations	Site 1	Site 2	Site 3	Total
EMP	2000	2000	2000	6000
ASG		3000		3000
PROJ			2000	2000
Total	2000	5000	4000	11000

1) Which join do we need to perform first, either EMP ⋈ ASG or ASG ⋈ PROJ? Why?
2) Explain how to perform the first join in terms of which relation needs to be sent to which site. You do not need to consider semijoin.
3) Explain how to perform the second join which is using the result of the first join in terms of which relation(s) need(s) to be sent to which site(s). You do not need to consider semijoin.

4. (Ch. 05) For the following histories,

H1 = {W2(x), W1(x), R3(x), R1(x), W2(y), R3(y), R2(x), R3(z)}

H2 = {R3(y), R3(z), W2(y), R2(z), W1(x), R3(x), W2(x), R1(x)}

H3 = {W2(x), R3(z), W2(y), R1(x), R3(x), R2(z), R3(y), W1(x)}

H4 = {R2(z), W2(x), W2(y), W1(x), R1(x), R3(x), R3(y), R3(z))}

1) Provide all possible pair of conflicting operations.
2) Which histories are conflict equivalent? You need to provide a reason (or reasons) for your answer.

5. (Ch. 05) Which of the above histories are serializable? You need to provide a reason (or reasons) for your answer using serialization graph.