TD 5: Data Exchange

Exercise 1.

Given the schema mapping: $\Sigma : E(x, y) \rightarrow \exists z, H(x, z) \wedge H(z, y)$ with the source instance $I = E(a, b)$.

- Are the following J_i solutions? Justify

Q1.1. $J_1 = \{H(a,b), H(b,b)\}$

Q1.2. $J_2 = \{H(a,a), H(a,b)\}$

Q1.3. $J_3 = \{H(a,X), H(X,b)\}$

Q1.4. $J_4 = \{H(a,X), H(X,b), H(a,Y), H(Y,b)\}$

Q1.5. $J_5 = \{H(a,X), H(X,b), H(Y,Y)\}$

Solution:

J_1 is a solution with $(z = b)$

J_2 is a solution with $(z = a)$

J_3 is a solution with $(z = X)$

J_4 is a solution with $(z = X)$ and $H(a,Y), H(Y,b) \in J_4$

J_5 is a solution with $(z = X)$ and $H(Y,Y) \in J_5$

- Specify for each J_i whether it is a universal solution or not, and justify.

Solution:

J_1 is not a universal solution because for each homomorphism h we have h(a) = a. Therefore, there is no homomorphism from J_1 to J_2.

J_2 is not a solution for the same reason. It is impossible to find a homomorphism from J_2 to J_1.

J_3 is a universal solution. For each solution J, there is z such that H(a,z) and H(a,b) are in J. So we construct the homomorphism h(X) = a.

J_4 is a universal solution (we can choose h(X) = h(Y)).

J_5 is not a universal solution. There is no homomorphism from J_5 to J_1.

Exercise 2.

- Create the dependency graph for the following schema mapping and specify if/why the sets of tgds are weakly acyclic.

2.1.

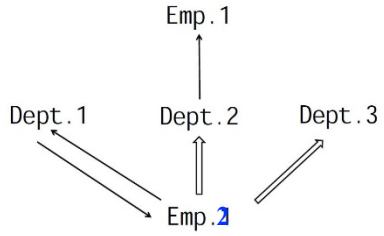
$$\Sigma_{st} = \{DeptEmp(d,n,e) \rightarrow \exists M\,(Dept(d,M,n) \wedge Emp(e,d))\}$$
$$\Sigma_t = \{Dept(d,m,e) \rightarrow Emp(m,d), Emp(e,d) \rightarrow \exists M \exists N\, Dept(d,M,N)\}$$

2.2.

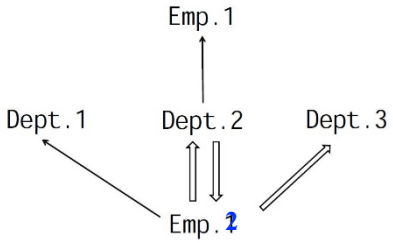
$$\Sigma_{st} = \{DeptEmp(d,n,e) \rightarrow \exists M\,(Dept(d,M,n) \wedge Emp(e,d))\}$$
$$\Sigma_t = \{Dept(d,m,e) \rightarrow \exists D\, Emp(m,D), Emp(e,d) \rightarrow \exists M \exists N\, Dept(d,M,N)\}$$

Solution:

- The graph for 2.1 is weakly acyclic



- The graph for 2.1 is not weakly acyclic



Exercise 3.

Let us consider the following source instance:

NYSE

name	symbol
Google	GOOG
Yahoo!	YHOO

Public-Company

name	city
Apple	Cup
Adobe	SJ

Public-Grant

company	investigator	amount
Apple	Mike B.	25,000
Adobe	Anne C.	50,000

NSF-Grantee

id	name	symbol
23	Yahoo!	YHOO
25	Adobe	ADBE

NSF-Grant

company	amount
23	18,000
25	50,000

and the constraints $\Sigma_{st}=\{m_1,m_2,m_3,m_4\}$ and $\Sigma_t=\{t_1,e_1\}$ such that:

$m_1: \forall s,n,NYSE(s,n) \rightarrow \exists I\, Company(I,n,s)$

m_2: ∀ n, c , a, pi, PublicCompany (n, c) ∧ PublicGrant (n,i,a) →∃ I ,∃ S Company (I , n, S) ∧Grant (a, I)

$m_3: \forall i,n,s,NFSGrantee(i,n,s) \rightarrow Company(i,n,s)$

$m_4: \forall a,c,NFSGrant(c,a) \rightarrow Grant(a,c)$

$t_1: \forall a, c, Grant(a, c) \rightarrow \exists N, \exists S\, Company(c, N, S)$

$e_1: \forall n, n', i, i', s, Company(i, n, s) \wedge Company(i', n', s) \rightarrow (n = n') \wedge (i = i')$

- Specify if the following instances are universal solutions? not universal? Or other.

J_1

Company

id	name	symbol
N1	Google	GOOG
N2	Yahoo	YHOO
I1	Apple	S1
I2	Adobe	S2
23	Yahoo!	YHOO
25	Adobe	ADBE

Grant

amount	company
25,000	I1
50,000	I2
18,000	23
50,000	25

J_2

Company

id	name	symbol
N1	Google	GOOG
I1	Apple	S1
I2	Adobe	S2
23	Yahoo!	YHOO
25	Adobe	ADBE

Grant

amount	company
25,000	I1
50,000	I2
18,000	23
50,000	25

J_3

Company

id	name	symbol
N1	Google	GOOG
I1	Apple	NULL
23	Yahoo!	YHOO
25	Adobe	ADBE

Grant

amount	company
25,000	I1
18,000	23
50,000	25

J_4

Company

id	name	symbol
N1	Google	GOOG
I1	Apple	NULL
23	Yahoo!	YHOO
25	Adobe	ADBE

Grant

amount	company
25,000	I1
18,000	I2
50,000	25
80,000	N1

Solution:

The minimal universal solution is:

$Company =$

id	name	symbol
$I1$	Apple	$S1$
$I2$	Google	GOOG
23	Yahoo!	YHOO
25	Adobe	ADBE

$Grant =$

amount	company
18000	23
50000	25
25000	$I1$

J_0

Solution:

J_1 is not a solution because e_1 is not satisfied for s = YHOO.

J_2 is a universal solution because we can define a homomorphism h from J_2 to J_0 with h(I2)=25, h(S2)=ABDE, h (I1)=I1 and h(N1)=I2.

J_3 is a solution because it satisfies all the constraints. On the other hand, for any homomorphism h we have h(NULL)=NULL, so there is no homomorphism from J_3 to J_0. So J_3 is not a universal solution.

J_4 is not a solution because it does not satisfy all the constraints (m4 is not satisfied).

.