Bases de Données Relationnelles
TD2
SI3

November 7, 2022

1 Formalisation en algèbre relationnelle de requêtes exprimées en langue naturelle

On considère le schéma de base de données suivant:

- employe(Nom, Prenom, DateDeNaissance, Adresse, NumeroSecuriteSociale, Salaire, NumeroDepartement, Superieur)
- departement(NomDepartement, NumeroDepartement, Directeur)
- projet(NomProjet, NumeroProjet, Lieu, NumeroDepartement)
- travaille(NumeroSecuriteSociale, NumeroProjet,Heures)

L'attribut Superieur d'un employé contient le numéro de sécurité sociale du supérieur direct de l'employé. Tout employé est rattaché à un département et travaille sur un nombre quelconque de projets. Chaque projet est rattaché à un département. Un employé peut travailler sur un projet qui n'est pas rattaché au même département que lui.

Pour chaque relation, on a souligné le ou les attributs qui constituent une clé de la relation.

Exprimer (lorsque c'est possible) en algèbre relationnelle les requêtes suivantes (i.e., construire la formule algébrique qui définit chacune d'elles).

1. Date de naissance et adresse de Juliette Rochat

$\Pi_{DateDeNaissance,Adresse}\ [(\sigma_{Prenom='Juliette' \wedge Nom='Rochat'}(employe)]$

---

2. Noms et adresses des employés rattachés au département "Recherche"

$\Pi_{Nom,Adresse}\ [employe \bowtie \sigma_{NomDepartement='Recherche'}(departement)]$

---

3. Noms des projets sur lesquels travaillent Jean Muller ou Annie Grandjean

---

$\Pi_{NomProjet}\ [projet \bowtie travaille \bowtie$
$\Pi_{NumeroSecuriteSociale}(\sigma_{Nom='Muller' \ \wedge \ Prenom='Jean'}(employe) \ \cup \sigma_{Nom='Grandjean' \ \wedge \ Prenom='Annie'}(employe))]$

La jointure des trois relations se fait sur les attributs NumeroProjet et NumeroSecuriteSociale. Attention, avant la jointure il faut faire une projection sur employe comme dans la solution donnée ou sur projet pour ne pas faire de jointure sur l'attribut NumeroDepartement, auquel cas on n'aurait que les projets du même département que Annie ou Jean.

---

4. Noms des projets sur lesquels travaillent à la fois Jean Muller et Annie Grandjean

---

Attention, on ne peut pas faire l'intersection sur les deux numéros de sécurité sociale (comme on a fait l'union à la question précédente), on doit faire l'intersection des numéros de projets.

Commençons par rechercher les numéros de projets communs:

$NPC = (\Pi_{NumeroProjet}(\Pi_{NumeroSecuriteSociale}(\sigma_{Nom='Muller' \wedge Prenom='Jean'}(employe)) \bowtie travaille))$
$\cap (\Pi_{NumeroProjet}(\Pi_{NumeroSecuriteSociale}(\sigma_{Nom='Grandjean' \wedge Prenom='Annie'}(employe)) \bowtie travaille))$
La réponse est donc $\Pi_{NomProjet}\ [projet \bowtie NPC]$

---

5. Noms et prénoms des employés qui ne travaillent sur aucun projet

---

$\Pi_{Nom,Prenom}[employe - (employe \bowtie \Pi_{NumeroSecuriteSocial}(travaille))]$
Comme NumeroSecuriteSociale est un attribut de employe, le résultat de la jointure est du même schéma que employe et donc on peut faire une différence entre employe et cette table.

---

6. Noms des employés qui ne travaillent sur aucun des projets localisés à Sophia Antipolis

---

On recherche d'abord les employés qui travaillent sur au moins un projet localisé à Sophia Antipolis
$Sophia = \Pi_{Nom,NumeroSecuriteSociale}(travaille \bowtie employe \bowtie$
$\Pi_{NumeroProjet}(\sigma_{Lieu="SophiaAntipolis"}(projet))$
$Resultat = \Pi_{Nom}[\Pi_{Nom,NumeroSecuriteSociale}(employe) - Sophia]$

---

7. Noms des employés qui ne travaillent que sur des projets localisés à Sophia Antipolis

---

On recherche d'abord les employés qui travaillent sur au moins un projet localisé ailleurs qu'à Sophia Antipolis
$Ailleurs = \Pi_{Nom,NumeroSecuriteSociale}(travaille \bowtie employe \bowtie$
$\Pi_{NumeroProjet}(\sigma_{Lieu \neq "SophiaAntipolis"}(projet))$
$Sophia = \Pi_{Nom,NumeroSecuriteSociale}(travaille \bowtie employe \bowtie$
$\Pi_{NumeroProjet}(\sigma_{Lieu="SophiaAntipolis"}(projet))$
$Resultat = \Pi_{Nom}[Sophia - Ailleurs]$
Dans cette solution, on a exclu les employés ne travaillant sur aucun projet.
Il est important de garder les Numéros de Securité Sociale jusqu'à la dernière projection, à cause des possibles homonymes.

---

8. Noms et prénoms des employés dont le supérieur est Juliette Rochat

---

$\Pi_{Nom,Prenom}\ [employe \bowtie$
$\delta_{NumeroSecuriteSocial \leftarrow Superieur}(\Pi_{NumeroSecuriteSociale}(\sigma_{Nom='Rochat'\ \wedge\ Prenom='Juliette'}(employe)))]$

---

9. Numéro des projets qui ont au moins un participant dans chaque département

---

On commence par calculer le complémentaire : les projets P qui n'ont pas de participant dans chaque département. Pour cela on construit d'abord l'ensemble des couples possibles (NumeroProjet,NumeroDepartement) avec le produit cartésien. Puis on soustrait les couples présents dans la base (i.e., pour lesquels au moins un employé du département travaille dans le projet). Le résultat est donc l'ensemble des couples tels que le département ne participe pas au projet. En projetant sur NumeroProjet on obtient donc l'ensemble des projets dans lesquels au moins un département ne participe pas, i.e. qui n'ont pas de participant dans chaque département. Finalement il faut soustraire de l'ensemble de tous les projets ces projets-là pour obtenir l'ensemble des projets qui ont au moins un participant dans chaque département.
$P = (\Pi_{NumeroProjet}(projet)\ \times\ (\Pi_{NumeroDepartement}(departement))-$
$\Pi_{NumeroProjet,NumeroDepartement}(travaille \bowtie employe)$
$Resultat = \Pi_{NumeroProjet}(projet) - \Pi_{NumeroProjet}(P)$

Autre solution: les projets qui ont au moins un participant dans chaque département sont ceux auxquels participent tous les départements. On construit la table reliant les projets aux départements participant et on la divise par la table des départements.

$ProDept = \Pi_{NumeroProjet,NumeroDepartement}(\Pi_{NumeroSecuriteSociale,NumeroProjet}(travaille) \bowtie$
$\Pi_{NumeroSecuriteSociale,NumeroDepartement}(employe))$

$Resultat = ProDept \div \Pi_{NumeroDepartement}(departement)$

---

2 Traduction en langue naturelle de requêtes formalisées en algèbre relationnelle

1. $\Pi_{Nom,Prenom}(\sigma_{Superieur=X \wedge Salaire>Y}(employe \bowtie$
$\delta_{NumeroSecuriteSocial \leftarrow X, Salaire \leftarrow Y}(\Pi_{NumeroSecuriteSocial,Salaire}(employe))))$

Nom et Prénom des employés dont le salaire est supérieur à celui de leur Supérieur immédiat.

---

2. $projet - \Pi_{NomProjet,NumeroProjet,Lieu,NumeroDepartement}(employe \bowtie projet \bowtie travaille)$

Projets sur lesquels ne travaille aucun employé rattaché au département de rattachement du projet. La jointure se fait sur NumeroSecuriteSociale, NumeroProjet mais aussi sur NumeroDepartement. La table résultat relie donc les employés et les seuls projets de leur département de rattachement dans lesquels ils travaillent.

---

3 Requêtes en algèbre relationnelle, retour sur le schéma du TD1

On considère le schéma de base de données suivant:

- marque(<u>IdM</u>, NomM, Classe, Pays, IdProp)
- societe(<u>IdS</u> ,Nom, Pays, <u>Site</u>)
- enreg(<u>NumE</u>, IdM, <u>Pays</u>, DateE, IdDeposant)
- vente(<u>NumV</u>, IdM, DateV, Pays, IdVend, IdAch)

Pour chaque relation, on a souligné le ou les attributs qui constituent un identifiant unique de la relation.

Contrairement au TD1, il n'y a pas d'autres contraintes sur le schéma.

Exprimer (lorsque c'est possible) en algèbre relationnelle les requêtes suivantes (i.e., construire les formules algébriques qui les définissent).

1. Les noms et pays des sociétés possédant au moins une marque.

$$\Pi_{Nom,Pays}\left[societe \ \bowtie \ (\delta_{IdProp \leftarrow IdS}\ (\Pi_{IdProp}\ (marque)))\right]$$

---

2. Les noms et sites des sociétés possédant au moins une marque dans la classe 24.

$$\Pi_{Nom,Site}\left[societe \ \bowtie \ (\delta_{IdProp \leftarrow IdS}\ (\Pi_{IdProp}(\sigma_{Classe=24}\ (marque))))\right]$$

---

3. Les noms de marques homonymes utilisés pour au moins deux marques toutes les deux françaises, toutes les deux enregistrées, mais dans deux classes différentes.

On calcule d'abord la table `A(NomM,Classe)` qui contient tous les couples (nomM, Classe) des marques françaises enregistrées, puis on fait une auto-jointure sur cette table.

$A = \Pi_{NomM,Classe}(enreg \bowtie \Pi_{IdM,NomM,Classe}(\sigma_{Pays=Fr}[marque]))$
$B = \delta_{NomM \leftarrow N, Classe \leftarrow C}(A)$
Ensemble recherché: $\Pi_{NomM}(\sigma_{N=NomM, Classe \neq C}(A \bowtie B))$

ou
$A = \Pi_{NomM,Classe}(enreg \bowtie \Pi_{IdM,NomM,Classe}(\sigma_{Pays=Fr}[marque]))$
$B = \delta_{Classe \leftarrow C}(A)$
Ensemble recherché: $\Pi_{NomM}(\sigma_{Classe \neq C}(A \bowtie B))$

---

4. Les identifiants des marques enregistrées dans tous les pays. On supposera que dans chaque pays il y a au moins une marque enregistrée.

$A = \Pi_{IdM,Pays}(enreg) \div \Pi_{Pays}(enreg)$

---

5. Le nom des marques et le nom et pays de leur propriétaire pour les marques enregistrées avant le 29 janvier 95.

$$\Pi_{NomM,Nom,Pays}\ [\delta_{IdS \leftarrow IdProp}\ (societe) \bowtie [\Pi_{IdM}\ (\sigma_{DateE \leq 950129}\ (enreg)) \bowtie \Pi_{IdM,NomM,IdProp}\ (marque)]]$$

---

6. Les noms et pays des sociétés dont toutes les marques qu'elles possèdent sont dans la classe 14.

Remarque : Une société qui ne possède aucune marque doit apparaitre dans les réponses.

$$\Pi_{Nom,Pays}\ [(\delta_{IdS \leftarrow IdProp}\ societe) \bowtie [\delta_{IDS \leftarrow IdProp}(\Pi_{IdS}\ societe) \setminus (\Pi_{IdProp}\ (\sigma_{Classe \neq 14}\ marque))]]$$

---

7. Est-ce que toutes les marques ont été enregistrées ?

Oui si et seulement si $\Pi_{IdM}\ marque \setminus \Pi_{IdM}\ enreg = \emptyset$

---

8. Les noms, sites et pays des propriétaires qui ont déposé eux-mêmes toutes les marques qu'ils possèdent et qui ont été enregistrées. Un propriétaire est une société qui possède au moins une marque.

On recherche d'abord l'ensemble des marques enregistrées avec leurs propriétaires et leurs déposants:
$A = \Pi_{IdM,IdProp}(marque) \bowtie \Pi_{IdM,IdDeposant}(enreg)$

Puis on calcule l'ensemble des propriétaires qui ont déposé au moins une de leur(s) marque(s):
$ProprioDeposant = \Pi_{IdProp}(\sigma_{IdProp=IdDeposant}(A))$

Puis on calcule l'ensemble des propriétaires qui n'ont pas déposé au moins une de leur(s) marque(s):
$ProprioNonDeposant = \Pi_{IdProp}(\sigma_{IdProp \neq IdDeposant}(A))$

On calcule alors la différence entre ces deux ensembles pour avoir les seuls propriétaires ayant déposé *toutes* leurs marques et on joint le résultat avec la relation societe pour récupérer les nom, site et pays de ces propriétaires:
$\Pi_{Nom,Site,Pays}(\delta_{IdS \leftarrow IdProp}(societe) \bowtie (ProprioDeposant \setminus ProprioNonDeposant))$

---

9. Les noms des sociétés n'ayant vendu aucune des marques qu'elles possèdent.

On calcule l'ensemble de tous les propriétaires et on lui soustrait l'ensemble des propriétaires qui ont vendu une des marques qu'ils possèdent:

$$A = (\Pi_{IdProp}\ (marque))$$
$$\setminus\ (\Pi_{IdProp}\ (marque\ \bowtie\ \delta_{IdVend \leftarrow IdProp, Pays \leftarrow PaysV}\ (vente))$$

Ensemble recherché: $\Pi_{Nom}(\ societe\ \bowtie\ \delta_{IdProp \leftarrow IdS}\ (A))$

---

10. L'avant-dernier propriétaire, s'il existe, de la marque "Chanel" enregistrée en France dans la classe 14.

La difficulté vient du fait que le même propriétaire peut avoir vendu la même marque au propriétaire actuel. Il faut donc raisonner avec les tuples `<NumV,IdVend>`. L'avant dernier propriétaire est le dernier vendeur. On va donc successivement :

(a) Rechercher le propriétaire actuel de la marque Chanel enregistrée en France dans la classe 14:
$$P = \Pi_{IdM,IdProp}\ [\sigma_{NomM=channel, Classe=14}\ (marque)$$
$$\bowtie\ \Pi_{IdM}\ (\sigma_{Pays=Fr}(enreg))]$$

(b) Rechercher l'ensemble des vendeurs qui ont vendu cette marque à ce propriétaire :
$$V = \Pi_{NumV,IdVend}\ [\delta_{IdAch \leftarrow IdProp}\ (P)\ \bowtie\ vente]$$

(c) Identifier le dernier vendeur:
$$ADP = V\ \setminus\ \Pi_{NumV,IdVend}\ [\sigma_{NumV < NumnV'}\ (V\ \bowtie$$
$$\delta_{NumV \leftarrow NumnV', IdVend \leftarrow IdVend'}\ (V))]$$

---

4 Relations sur des expressions de l'algèbre relationnelle

1. Soit $A \subseteq R$, et soient r et s deux relations sur R. Quelles sont les relations d'inclusion ou d'égalité entre les expressions suivantes :

   $\Pi_A(r \cap s)$ et $\Pi_A(r) \cap \Pi_A(s)$
   $\Pi_A(r \cup s)$ et $\Pi_A(r) \cup \Pi_A(s)$
   $\Pi_A(r \setminus s)$ et $\Pi_A(r) \setminus \Pi_A(s)$

On a:

$\Pi_A(r \cap s) \quad \subseteq \quad \Pi_A(r) \cap \Pi_A(s)$ car les tuples qui diffèrent uniquement par les attributs de $R \setminus A$ sont conservés si on effectue la projection avant l'intersection.

Par exemple si

R={Nom,Prenom} A ={Nom},

r={(Dupont, Louis),(Durant,Marcel)},

s={(Durand, Michel),(Durant,Marie)},

l'intersection de r et s est vide, tandis que l'intersection des projections contient le Nom Durant.

$\Pi_A(r \cup s) \quad = \quad \Pi_A(r) \cup \Pi_A(s)$ car contrairement à l'intersection, l'union n'élimine pas de tuples.

$\Pi_A(r \setminus s) \quad \supseteq \quad \Pi_A(r) \setminus \Pi_A(s)$ car les tuples qui diffèrent uniquement par les attributs de $R \setminus A$ sont éliminés si on effectue la projection avant la différence.

---

2. Exprimez $r \cap s$ en fonction de $r \bowtie s$

Si r et s sont deux instances d'un même schéma, alors $r \cap s = r \bowtie s$, sinon l'intersection n'est pas définie, alors que la jointure l'est.

---

3. Soient $r(R)$ et $s(S)$ deux instances de relations. Quelles sont les relations d'inclusion existant entre r, s, $r \bowtie s$, $\Pi_R(r \bowtie s)$, $\Pi_S(r \bowtie s)$?

$\Pi_R(r \bowtie s) \subseteq r$. Ce n'est pas une égalité car $\Pi_R(r \bowtie s)$ élimine les tuples de r qui ne sont joignables avec aucun tuple de s.

De même $\Pi_S(r \bowtie s) \subseteq s$

On a finalement $r \bowtie s = \Pi_R(r \bowtie s) \bowtie \Pi_S(r \bowtie s)$
