Bases de données

L2 sciences et technologies, mention informatique

la structure du modèle relationnel

ou : quel modèle mathématique pour représenter ceci ?

films	titre	réalisateur	année
	starwars	lucas	1977
	nikita	besson	1990
	locataires	ki-duk	2005
	dune	lynch	1984

realisateurs	nom	nationalité
	lucas	américaine
	lynch	américaine
	besson	française
	ki-duk	coréenne

patrick.marcel@univ-tours.fr
http://celene.univ-tours.fr/course/view.php?id=3131

Cine nu cunoaște lema, nu cunoaște teorema

vieux proverbe roumain

Cine nu cunoaște lema, nu cunoaște teorema

vieux proverbe roumain

celui qui ne connait pas le lemme
ne connait pas le théorème

Cine nu cunoaște lema, nu cunoaște teorema

vieux proverbe roumain

celui qui ne connait pas le lemme
ne connait pas le théorème

c'est quoi un lemme?

étudiant de L2 ayant souhaité conserver l'anonymat

plan de la séance

1. rappels de mathématiques élémentaires
2. structure basique du modèle relationnel
3. différentes approches

rappels mathématiques

ensemble

un ensemble est une collection d'éléments

ensemble

un ensemble est une collection d'éléments

on note :

- $x \in S$ l'appartenance de l'élément x à l'ensemble S
- $\emptyset$ l'ensemble vide
- $|S|$ la cardinalité de l'ensemble S
- $\mathcal{P}(S)$ ou 2^S l'ensemble des parties de l'ensemble S

ensemble

quelles sont les opérations que nous connaissons sur les ensembles ?

ensemble

quelles sont les opérations que nous connaissons sur les ensembles ?

- $\cup$
- $\cap$
- $-$
- $\times$

ensemble

quels sont les comparateurs ensemblistes que nous connaissons ?

ensemble

quels sont les comparateurs ensemblistes que nous connaissons ?

- $\subset$
- $\not\subset$
- $=$
- $\neq$

partition

soit S un ensemble non vide et I un intervalle

une *partition* de S est une famille d'ensembles $\{S_i | i \in I\}$ telle que :

1. $\cup_{i \in I} S_i = S$
2. $S_i \cap S_j = \emptyset$ pour $i \neq j$
3. $S_i \neq \emptyset$ pour tout $i \in I$

exemple

$$S = \{a,b,c,d\}$$

exemple

$S = \{a,b,c,d\}$

$S_1 = \{a,c\}$ et $S_2 = \{b,d\}$ forment une partition de S

exemple

$S = \{a,b,c,d\}$

$S_1 = \{a,c\}$ et $S_2 = \{b,d\}$ forment une partition de S

$S_1 = \{a,b\}$ et $S_2 = \{b,c,d\}$ ne forment pas une partition de S

relation

une *relation binaire* R sur un ensemble S est un sous-ensemble de $S \times S$

relation

une *relation binaire* R sur un ensemble S est un sous-ensemble de $S \times S$

donc une relation est un ensemble

relation

une *relation binaire* R sur un ensemble S est un sous-ensemble de $S \times S$

donc une relation est un ensemble

on note $R(x,y)$ pour indiquer que $(x,y) \in R$

on appelle (x,y) couple ou doublet ou *2-uplet*

relation

une *relation n-aire* R sur S est définie de manière analogue comme un sous-ensemble de S^n

relation

une *relation n-aire* R sur S est définie de manière analogue comme un sous-ensemble de S^n

on note $R(x_1, \ldots, x_n)$ pour indiquer que $(x_1, \ldots, x_n) \in R$

on appelle $(x_1, \ldots, x_n)$ n-uplet ou *tuple*

exemple de relations binaires

$<$ est une relation sur $\mathbb{N}$, $(1{,}2) \in <$

exemple de relations binaires

$<$ est une relation sur $\mathbb{N}$, $(1,2) \in <$

si S est un ensemble, $\subset$ est une relation sur $\mathcal{P}(\mathcal{S})$
$(\{1\},\{1,2\}) \in \subset$

exemple de relation n-aire

soit l'ensemble de constantes $S =$

$$\{\text{lucas, lynch, dune, 1977, starwars, 1984}\}$$

exemple de relation n-aire

soit l'ensemble de constantes $S =$

$$\{\text{lucas, lynch, dune, 1977, starwars, 1984}\}$$

soit la relation *films* sur S

(lucas, starwars, 1977) $\in$ films
(lynch, dune, 1984) $\in$ films

fermeture transitive

la *fermeture transitive* d'une relation binaire R est une relation R^+ telle que :

1. si $(x,y) \in R$ alors $(x,y) \in R^+$

fermeture transitive

la *fermeture transitive* d'une relation binaire R est une relation R^+ telle que :

1. si $(x,y) \in R$ alors $(x,y) \in R^+$
2. si $(x,y) \in R^+$ et $(y,z) \in R^+$ alors $(x,z) \in R^+$

fermeture transitive

la *fermeture transitive* d'une relation binaire R est une relation R^+ telle que :

1. si $(x,y) \in R$ alors $(x,y) \in R^+$
2. si $(x,y) \in R^+$ et $(y,z) \in R^+$ alors $(x,z) \in R^+$
3. R^+ ne contient rien qui ne satisfait pas les conditions 1 et 2

exemple

$$pere = \{(adam,cain),(adam,abel),(cain,enoch),(enoch,irad)\}$$

exemple

$$pere = \{(adam,cain),(adam,abel),(cain,enoch),(enoch,irad)\}$$

$$pere^{+} = \{(adam,cain),(adam,abel),(cain,enoch),(enoch,irad),\\ (adam,enoch),(adam,irad),(cain,irad)\}$$

la structure

définitions formelles

dans ce qui suit, sauf indication contraire, les ensembles utilisés ensuite sont :

- infinis dénombrables,
- 2 à 2 disjoints

constantes utilisées

soient les ensembles suivants :

- **att** un ensemble d'attributs totalement ordonnés par $\leq$
- **dom** ensemble de constantes du domaine
- **relname** ensemble de noms de relation
- **var** ensemble de variables à valeur sur **dom**

exemple

$\{$titre, realisateur, année, nom, nationalité$\} \subset$ **att**

$\{$lucas, lynch, dune, 1977, nikita, starwars, 1990, américaine, 1984, française, coréenne, besson, ki-duk, locataires, 2005 $\} \subset$ **dom**

$\{$films, réalisateurs$\} \subset$ **relname**

sorte

soit une fonction *sorte* : $\textbf{relname} \rightarrow \mathcal{P}(\textbf{att})$

sorte associe à une relation de nom $R \in \textbf{relname}$ un ensemble d'attributs

sorte

soit une fonction *sorte* : $\textbf{relname} \rightarrow \mathcal{P}(\textbf{att})$

sorte associe à une relation de nom $R \in \textbf{relname}$ un ensemble d'attributs

exemple :
$\mathit{sorte}(\text{films}) = \{\text{titre, réalisateur, année}\}$

schéma, arité

le *schéma d'une relation* de nom R est $R[U]$ où $U = sorte(R)$

l'*arité* d'une relation R est la cardinalité de $sorte(R)$ notée $|sorte(R)|$ ou arité(R)

schéma, arité

le *schéma d'une relation* de nom R est $R[U]$ où $U = sorte(R)$

l'*arité* d'une relation R est la cardinalité de $sorte(R)$ notée $|sorte(R)|$ ou arité(R)

films est une relation

- de schéma films[titre, réalisateur, année]
- et d'arité 3

schéma d'une base de données

D le *schéma d'une base de données* est un ensemble fini de noms de relation

schéma d'une base de données

D le *schéma d'une base de données* est un ensemble fini de noms de relation

donc $D \subset$ **relname**

schéma d'une base de données

D le *schéma d'une base de données* est un ensemble fini de noms de relation

donc $D \subset$ **relname**

exemple :
CINEMA $= \{$films, réalisateurs$\}$

les différentes approches

introduction

pourquoi différentes approches ?

introduction

pourquoi différentes approches ?

pour expliquer les notions qui vont suivre, une des approches pourra être plus facile à utiliser que l'autre

introduction

pourquoi différentes approches ?

pour expliquer les notions qui vont suivre, une des approches pourra être plus facile à utiliser que l'autre

la structure est la même, c'est juste une différence de point de vue !

est-ce que les noms d'attribut sont importants?

c'est à dire : seront utilisés pour manipuler les relations?

approche *nommée*

est-ce que les noms d'attribut sont importants ?

c'est à dire : seront utilisés pour manipuler les relations ?

approche *nommée*

peut-on se contenter d'utiliser leur position ?

approche *non nommée*

approche non nommée

un tuple t est un élément de $\mathbf{dom}^n$

approche non nommée

un tuple t est un élément de $\mathbf{dom}^n$

si t est un tuple, la i^{eme} coordonnée de t est notée $t(i)$

si R est une relation d'arité n

- un tuple de R est un tuple d'arité n
- $sorte(t) = sorte(R)$

exemple

$$t = (\text{starwars,lucas,1977})$$

$$t(3) = 1977$$

$$\mathit{sorte}(t) = \{\text{titre,réalisateur,année}\}$$

approche nommée

si U est un ensemble d'attributs, un tuple t est vu comme une fonction de U vers **dom**

alors $sorte(t) = U$

si t est un tuple sur U, et $V \subseteq U$, on note $t[V] = t|_V$ la *restriction* de t à V

exemple

soit la relation films

- de schéma films[titre,réalisateur,année]
- *sorte*(films) = {titre, réalisateur, année}

le tuple (starwars,lucas,1977) correspond à une fonction
t de *sorte*(films) vers **dom**

avec t(titre) = star wars, t(réalisateur) = lucas, t(année) = 1977

t sera noté (titre : starwars, réalisateur : lucas, année : 1977)

correspondance entre les approches

comme **att** est ordonné, il y a une correspondance naturelle entre ces approches

donc on adopte celle que l'on souhaite !

comment manipule-t-on les relations ?

comme des ensembles de tuples

approche *conventionnelle*

comment manipule-t-on les relations ?

comme des ensembles de tuples

approche *conventionnelle*

comme des ensembles de faits

approche *logique*

approche conventionnelle

une *instance de relation* de schéma $R[U]$ est un ensemble fini $\mathbf{I}$ de tuples de sorte U

une *instance de base de données* de schéma D est une fonction I de domaine D telle que :

$I(R)$ est une instance de relation pour chaque $R \in D$

exemple (approche non nommée)

l'ensemble $\mathbf{I}_1$ ={(starwars,lucas,1977), (dune,lynch,1984), (nikita,besson,1990), (locataires,ki-duk,2005)}

est une instance de films

l'ensemble $\mathbf{I}_2$ ={(lucas,américaine), (lynch,américaine), (ki-duk,coréenne), (besson,française)}

est une instance de réalisateurs

exemple (approche nommée)

l'ensemble $\mathbf{I}_1 = \{f_1, f_2, f_3, f_4\}$ est une instance de films

avec f_1(titre)=starwars, f_1(réalisateur)=lucas, f_1(année)=1977, f_2(titre)=(dune),f_2(réalisateur)=lynch,f_2(année)=1984, ...

exemple (suite)

soit la base de données CINEMA de schéma
CINEMA={fims,réalisateurs}

une instance de CINEMA est I avec

$I(films) = \mathbf{I}_1$

$I(réalisateurs) = \mathbf{I}_2$

approche logique

R est une relation d'arité n

un *fait* de R est une expression de la forme $R(a_1,\ldots,a_n)$ avec $a_i \in \textbf{dom}$ pour $i \in [1,n]$

on note $R(u)$ si $u = (a_1,\ldots,a_n)$

approche logique

R est une relation d'arité n

un *fait* de R est une expression de la forme $R(a_1,\ldots,a_n)$ avec $a_i \in \mathbf{dom}$ pour $i \in [1,n]$

on note $R(u)$ si $u = (a_1,\ldots,a_n)$

exemple : films(starwars,lucas,1977) est un fait de films

instances

une *instance de relation* $\mathbf{I}$ est un ensemble fini de faits de R

une *instance de base de données* de schéma D est un ensemble fini $\mathbf{D} = \bigcup_{R \in D} \mathbf{I}$

c'est à dire l'union des instances de chaque relation de D

exemple (approche non nommée)

$\mathbf{I}_1$ ={films(starwars,lucas,1977), films(dune,lynch,1984), films(nikita,besson,1990), films(locataires,ki-duk,2005)}

est une instance de films

$\mathbf{I}_2$ = {réalisateurs(lucas,américaine), réalisateurs(lynch,américaine), réalisateurs(ki-duk,coréenne), réalisateurs(besson,française)}

est une instance de réalisateurs

exemple (suite)

$\mathbf{D} = \mathbf{I}_1 \bigcup \mathbf{I}_2$ est une instance de la la base de données CINEMA de schéma CINEMA={fims,réalisateurs}

exemple (suite)

$\mathbf{D} = \mathbf{I}_1 \bigcup \mathbf{I}_2$ est une instance de la la base de données CINEMA de schéma CINEMA={fims,réalisateurs}

on pourra la représenter comme ça :

films	titre	réalisateur	année
	starwars	lucas	1977
	nikita	besson	1990
	locataires	ki-duk	2005
	dune	lynch	1984

realisateurs	nom	nationalité
	lucas	américaine
	lynch	américaine
	besson	française
	ki-duk	coréenne