

Le modèle de données relationnel

Origine: E.F. Codd (1970).

Objets et classes pour la modélisation:

→ présentation synthétique du modèle relationnel

Intérêt (pour nous):

- SGBD relationnels = tendance actuelle. Par exemple Access, Oracle, SQL Server, My SQL ...
- SQL = langages d'interaction avec les bases de données relationnelles
- modèle accompagné d'une algèbre relationnelle → interrogation de la base de données
- structure de données la plus simple et la plus uniforme; bases théoriques solides (théorie des ensembles).

Concepts de bases: relation, attribut et domaine.

Domaine

Définition: "un domaine est un ensemble dénombrable de valeurs caractérisé par un nom".

Exemple:

- sortes-de-raisin = {chardonnay, sémillon, pinot, cabernet-sauvignon}
- couleurs = {rosé, blanc, rouge}
- noms = l'ensemble des noms de personnes
- string = toutes les chaînes de caractères possibles
- ...

Rappel produit cartésien:

Le produit cartésien d'une liste de domaines D_1, D_2, \dots, D_n noté $D_1 \times D_2 \times \dots \times D_n$ est l'ensemble des n-uplets ou tuples $\langle v_1, v_2, \dots, v_n \rangle$ où v_1 est une valeur de D_1 , v_2 une valeur de D_2, \dots et v_n une valeur de D_n .

Domaine et produit cartésien (suite)

Exemple: le produit cartésien sortes-de-raisin \times couleurs est composé des 12 tuples suivants

| | |
|--------------------|-------|
| chardonnay | rosé |
| chardonnay | blanc |
| chardonnay | rouge |
| sémillon | rosé |
| sémillon | blanc |
| sémillon | rouge |
| pinot | rosé |
| pinot | blanc |
| pinot | rouge |
| cabernet-sauvignon | rosé |
| cabernet-sauvignon | blanc |
| cabernet-sauvignon | rouge |

Relation

Définition: “une relation est un sous-ensemble du produit cartésien d’une liste de domaines caractérisé par un nom”.

Représentation (commode):

table à 2 dimension. Chaque ligne correspond à un tuple, chaque colonne correspond à un domaine.

Exemple:

La relation “sorte-de-vin” est un sous ensemble du produit cartésien sortes-de-raisin \times couleurs.

domaine

sorte-de-vin

| | |
|--------------------|-------|
| chardonnay | blanc |
| sémillon | blanc |
| pinot | rosé |
| pinot | blanc |
| pinot | rouge |
| cabernet-sauvignon | rouge |

← tuple

Remarque: afin de rendre l'ordre des colonnes sans importance, on associe un nom à chaque colonne
→ notion d'attribut.

Attribut (constituant)

Définition: "colonne d'une relation caractérisée par un nom. Rôle joué par un domaine dans la relation".

Le nom associé à un attribut est souvent porteur de sens et souvent différent du domaine sur lequel est défini l'attribut.

Exemple:

sorte-de-vin

| cépage | couleur-vin |
|--------------------|-------------|
| chardonnay | blanc |
| sémillon | blanc |
| pinot | rosé |
| pinot | blanc |
| pinot | rouge |
| cabernet-sauvignon | rouge |

Analogie avec le modèle des classes

- les colonnes (attributs du modèle relationnel) ne sont rien d'autres que les attributs du modèle des classes
- les lignes correspondent aux occurrences d'une classe ou d'une association du modèle des classes

Extension (instance), intention et schéma d'une relation

Extension, instance de relation ou relation : ensemble des tuples d'une relation qui existent dans la base de données à un instant donné.

Intention: structure d'une relation. La structure d'une relation est décrite par un schéma de relation.

Définition: "le schéma d'une relation est constitué par le nom de la relation suivi de la liste des attributs et de la définition de leurs domaines".

Exemples:

SORTE-DE-VIN

| | |
|-------------------------|-----------------------|
| cépage:sortes-de-raisin | couleur-vin: couleurs |
|-------------------------|-----------------------|

CRU

| | | | |
|-----------------|---------------|----------------|-----------------------|
| nom-cru: string | région:string | commune:string | couleur-cru: couleurs |
|-----------------|---------------|----------------|-----------------------|

VIN

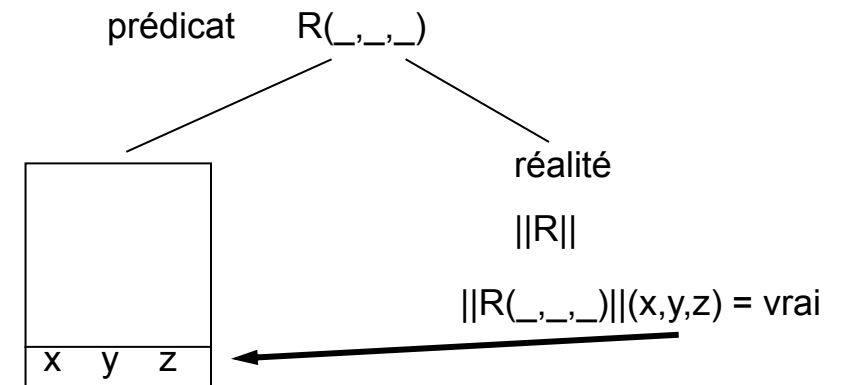
| | | | | |
|---------------|-----------------------|-----------------|-----------------------|-----------------------|
| région:string | couleur-vin: couleurs | millésime:année | appréciation:qualités | maturité:conservation |
|---------------|-----------------------|-----------------|-----------------------|-----------------------|

Notation du modèle relationnel

- Les lettres Q, R, S dénotent des noms de schéma de relation.
- Les lettres q, r, s dénotent des relations (ou instance de relation).
- Les lettres t, u, v dénotent des tuples.
- Un schéma de relation d'arité n est représenté par $R(A_1: D_1, A_2: D_2, \dots, A_n: D_n)$. Lorsque les domaines des attributs sont évidents, on notera $R(A_1, A_2, \dots, A_n)$.
- R^+ désigne l'ensemble des attributs de R.
- Etant donné un attribut A_i , $\text{dom}(A_i)$ indique son domaine, c'est-à-dire D_i .
- $r(R)$ dénote une relation (ou instance de relation) sur un schéma R. On peut aussi noter iR ou encore r .
- Un n-tuple de $r(R)$ est représenté par $t = \langle v_1, v_2, \dots, v_n \rangle$ où v_i est la valeur correspondant à l'attribut A_i . Dans ce cas, $t(A_i)$ se réfère à la valeur v_i du tuple t.
- $\|R(A_1, A_2, \dots, A_n)\|$ représente l'interprétation du prédicat associé à R.

L'interprétation d'une relation

Le prédicat d'une relation R, noté $R(A_1, A_2, \dots, A_n)$, définit la sémantique de la relation R.



Exemple, l'interprétation du prédicat de la relation SORTE-DE-VIN:

$\| \text{SORTE-DE-VIN}(\text{cépage}, \text{couleur-vin}) \|$: "Avec le cépage cépage on peut produire du vin de couleur couleur-vin".

Pour une relation, il n'existe qu'un seul prédicat -> une seule interprétation des données.

Pour chaque tuple t de $r(R)$, $\|R(A_1, A_2, \dots, A_n)\|(t)$ est vrai.

Relation (remarque)

Les relations peuvent représenter des classes ou des associations du monde réel.

Exemples

La relation “auteur(nom-auteur, âge, adresse)” représente le fait que chaque objet de la classe auteur a un nom, un âge et une adresse.

La relation “article(titre-article, résumé)” représente le fait que chaque objet de la classe article a un titre et un résumé.

La relation “écriture(nom-auteur,titre-article)” représente le fait qu’un auteur écrit un article. Un tuple de la relation écriture associe un tuple de la relation auteur à un tuple de la relation article.

→ le modèle relationnel représente de manière uniforme les classes et les associations sous forme de relations.

Règles d’intégrité

Définition: “les règles d’intégrité (R.I.) sont des assertions qui doivent être vérifiées à tout moment par les données contenues dans la base de données”.

Remarque: notion équivalente à celle de contrainte du modèle des classes

Trois type de R.I. complètent obligatoirement la description d’un schéma relationnel:

- (1) contrainte de clé
- (2) contrainte de référence
- (3) contrainte d’entité

(1) Contrainte de clé

Définition: “la clé d’une relation est un ensemble d’attributs minimum dont la connaissance permet d’identifier un tuple unique de la relation considérée”.

Définition: “La contrainte de clé impose que chaque relation possède une clé”.

Remarques:

- il peut exister plusieurs clés → on en choisit une arbitrairement

Règles d'intégrité obligatoires (suite)

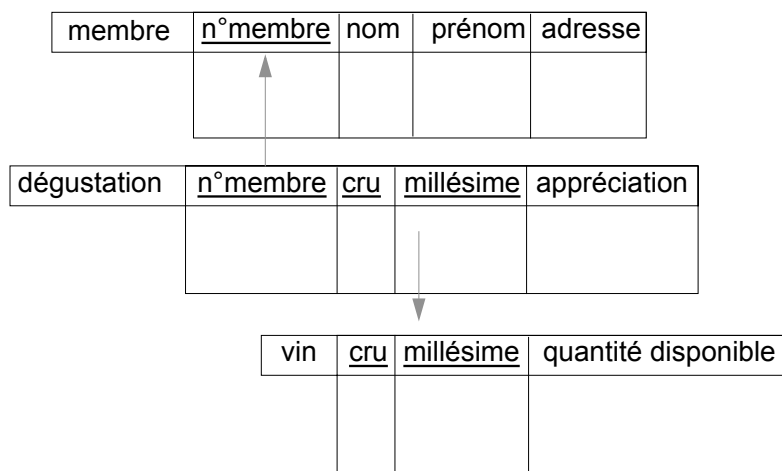
(2) Contrainte de référence (C.R.)

Définition: “une contrainte de référence porte sur une relation r et impose que la valeur d'un groupe d'attributs apparaisse comme valeur de clé dans une autre relation s”.

→ une C.R. impose que tout tuple d'une relation r qui se réfère à une relation s doit se référer à un tuple existant dans la relation s.

Une telle contrainte s'applique en général sur des relations qui décrivent des associations de la réalité perçue.

Exemple: la relation dégustation ci-dessous a deux C.R



Règles d'intégrité obligatoires (suite)

(3) Contrainte d'entité

Définition: “la contrainte d'entité impose que tout attribut faisant partie de la clé d'une relation soit non nul”.

La vérification de cette contrainte doit se faire chaque fois que l'on insère un tuple dans la base de données.

Schéma et instance d'une base de données relationnelle

Une base de données contient généralement plusieurs relations.

Définition: “le schéma d'une base de données relationnelle est constitué d'un ensemble de schémas de relation $SR=\{R_1, R_2, \dots, R_n\}$ et d'un ensemble de règles d'intégrité $RI=\{R.I._1, R.I._2, \dots, R.I._m\}$ ”.

Définition: “l'instance d'une base de données relationnelle est constituée d'un ensemble de relations (ou instances de relation) $BD=\{r(R_1), r(R_2), \dots, r(R_n)\}$ t.q. chaque relation $r(R_i)$ satisfait l'ensemble des règles d'intégrité RI ”.