

Informatique pour tous - PC*

Rappels d'algèbre relationnelle

Florent Pompigne
pompigne@crans.org

Lycée Buffon

année 2017/2018

Vocabulaire

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.
- Le **domaine** d'un attribut est l'ensemble dans lequel cet attribut peut prendre ses valeurs.

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.
- Le **domaine** d'un attribut est l'ensemble dans lequel cet attribut peut prendre ses valeurs.
- Un **enregistrement** (ou ligne) est un n-uplet de valeurs des différents attributs.

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.
- Le **domaine** d'un attribut est l'ensemble dans lequel cet attribut peut prendre ses valeurs.
- Un **enregistrement** (ou ligne) est un n-uplet de valeurs des différents attributs.
- Une **relation** (ou table) est une collection d'enregistrements correspondant aux mêmes attributs.

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.
- Le **domaine** d'un attribut est l'ensemble dans lequel cet attribut peut prendre ses valeurs.
- Un **enregistrement** (ou ligne) est un n-uplet de valeurs des différents attributs.
- Une **relation** (ou table) est une collection d'enregistrements correspondant aux mêmes attributs.
- Le **schéma** d'une relation est la liste des couples (attribut, domaine) utilisés par la relation.

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.
- Le **domaine** d'un attribut est l'ensemble dans lequel cet attribut peut prendre ses valeurs.
- Un **enregistrement** (ou ligne) est un n-uplet de valeurs des différents attributs.
- Une **relation** (ou table) est une collection d'enregistrements correspondant aux mêmes attributs.
- Le **schéma** d'une relation est la liste des couples (attribut, domaine) utilisés par la relation.
- Le **cardinal** d'une relation est le nombre d'enregistrements qui la composent.

Vocabulaire

- Un **attribut** (ou colonne) est une caractéristique des entités que l'on veut représenter dans la base de données.
- Le **domaine** d'un attribut est l'ensemble dans lequel cet attribut peut prendre ses valeurs.
- Un **enregistrement** (ou ligne) est un n-uplet de valeurs des différents attributs.
- Une **relation** (ou table) est une collection d'enregistrements correspondant aux mêmes attributs.
- Le **schéma** d'une relation est la liste des couples (attribut, domaine) utilisés par la relation.
- Le **cardinal** d'une relation est le nombre d'enregistrements qui la composent.
- Une **clé primaire** d'une relation est un attribut distinct pour chaque enregistrement.

Opérateurs de l'algèbre relationnelle

Opérateurs de l'algèbre relationnelle

- Une **projection** permet de sélectionner un sous-ensemble des attributs. On écrit par exemple $\pi_{\text{nom}, \text{prenom}}(A)$. La clause correspondante en SQL est le `SELECT`.

Opérateurs de l'algèbre relationnelle

- Une **projection** permet de sélectionner un sous-ensemble des attributs. On écrit par exemple $\pi_{\text{nom}, \text{prenom}}(A)$. La clause correspondante en SQL est le `SELECT`.
- Une **sélection** permet de sélectionner un sous-ensemble des enregistrements à l'aide d'un prédicat. On écrit par exemple $\sigma_{\text{classe}='PC*'}(A)$. La clause correspondante en SQL est le `WHERE`.

Opérateurs de l'algèbre relationnelle

- Une **projection** permet de sélectionner un sous-ensemble des attributs. On écrit par exemple $\pi_{\text{nom}, \text{prenom}}(A)$. La clause correspondante en SQL est le `SELECT`.
- Une **sélection** permet de sélectionner un sous-ensemble des enregistrements à l'aide d'un prédicat. On écrit par exemple $\sigma_{\text{classe}=\text{'PC*'}}(A)$. La clause correspondante en SQL est le `WHERE`.
- Un **renommage** permet de changer le nom d'un attribut. On écrit par exemple $\rho_{\text{nom} \leftarrow \text{nom de famille}}(A)$. La clause correspondante en SQL est le `AS`.

Opérateurs de l'algèbre relationnelle

- Une **projection** permet de sélectionner un sous-ensemble des attributs. On écrit par exemple $\pi_{\text{nom}, \text{prenom}}(A)$. La clause correspondante en SQL est le `SELECT`.
- Une **sélection** permet de sélectionner un sous-ensemble des enregistrements à l'aide d'un prédicat. On écrit par exemple $\sigma_{\text{classe}='PC'}(A)$. La clause correspondante en SQL est le `WHERE`.
- Un **renommage** permet de changer le nom d'un attribut. On écrit par exemple $\rho_{\text{nom} \leftarrow \text{nom de famille}}(A)$. La clause correspondante en SQL est le `AS`.
- Le **produit cartésien** de deux relations A et B est la relation $A \times B = \{(a_1, \dots, a_n, b_1, \dots, b_m) \mid (a_1, \dots, a_n) \in A \text{ et } (b_1, \dots, b_m) \in B\}$
La clause correspondante en SQL est le `JOIN`. On le suit en général d'un `ON` pour effectuer une sélection optimisée sur des égalités d'attributs.

Agrégation

En SQL, la clause `GROUP BY` *attribut* permet de rassembler les lignes ayant la même valeur pour cet attribut dans un **agrégat**. Dans la table renvoyée, chaque ligne correspond alors à un agrégat.

Agrégation

En SQL, la clause `GROUP BY` *attribut* permet de rassembler les lignes ayant la même valeur pour cet attribut dans un **agrégat**. Dans la table renvoyée, chaque ligne correspond alors à un agrégat.

On peut utiliser avec ces agrégats des **fonctions d'agrégation** telles que `COUNT`, `SUM`, `MAX`, `MIN` et `AVG`.

Agrégation

En SQL, la clause `GROUP BY` *attribut* permet de rassembler les lignes ayant la même valeur pour cet attribut dans un **agrégat**. Dans la table renvoyée, chaque ligne correspond alors à un agrégat.

On peut utiliser avec ces agrégats des **fonctions d'agrégation** telles que `COUNT`, `SUM`, `MAX`, `MIN` et `AVG`.

La clause `WHERE` sélectionnant des lignes, elle doit être écrite avant un `GROUP BY`, et est effectuée d'abord (sélection **en amont**). La sélection sur les agrégats (**en aval**) se fait avec la clause `HAVING`, et utilise en général des fonctions d'agrégation.