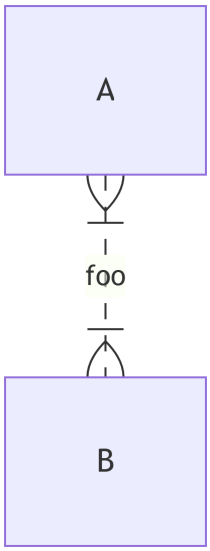
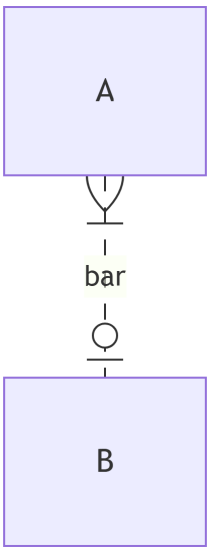
- **L3 MIASHS/Ingémath**
- **Université Paris Cité**
- Année 2023-2024
- Course Homepage

- Moodle

> **Rappel de notation**
>
> Dans le formalisme Entité-Association, nous utilisons les *pattes de corbeau* (*crowfoot*) pour décrire les contraintes de cardinalités.



L'entité `A` est reliée à l'entité `B` par l'association `foo`. Une instance de `A` peut apparaître de 1 à n fois parmi les instances de l'association `foo`, de même une instance de `B` peut apparaître de 1 à n fois parmi les instances de l'association `foo`.



L'entité `A` est reliée à l'entité `B` par l'association `bar`. Une instance de `A` peut apparaître de 0 à 1 fois parmi les instances de l'association `bar`, une instance de `B` peut apparaître de 1 à n fois parmi les instances de l'association `bar`.

> Notez que la contrainte portant sur le nombre de participations des instances d'une entité à une association apparaît *à l'opposé* de l'entité.

Un système d'information pour l'administration

Un pays (par exemple *France*) est organisé en régions (par exemple *Bretagne*, …). Les régions sont identifiées par leur nom. Ces régions sont-elles même divisées en départements (par exemple *Finistère*). Les départements sont identifiés par de numéros (par exemple *29* pour le *Finistère*). Les départments possèdent aussi un nom qui les distingue.

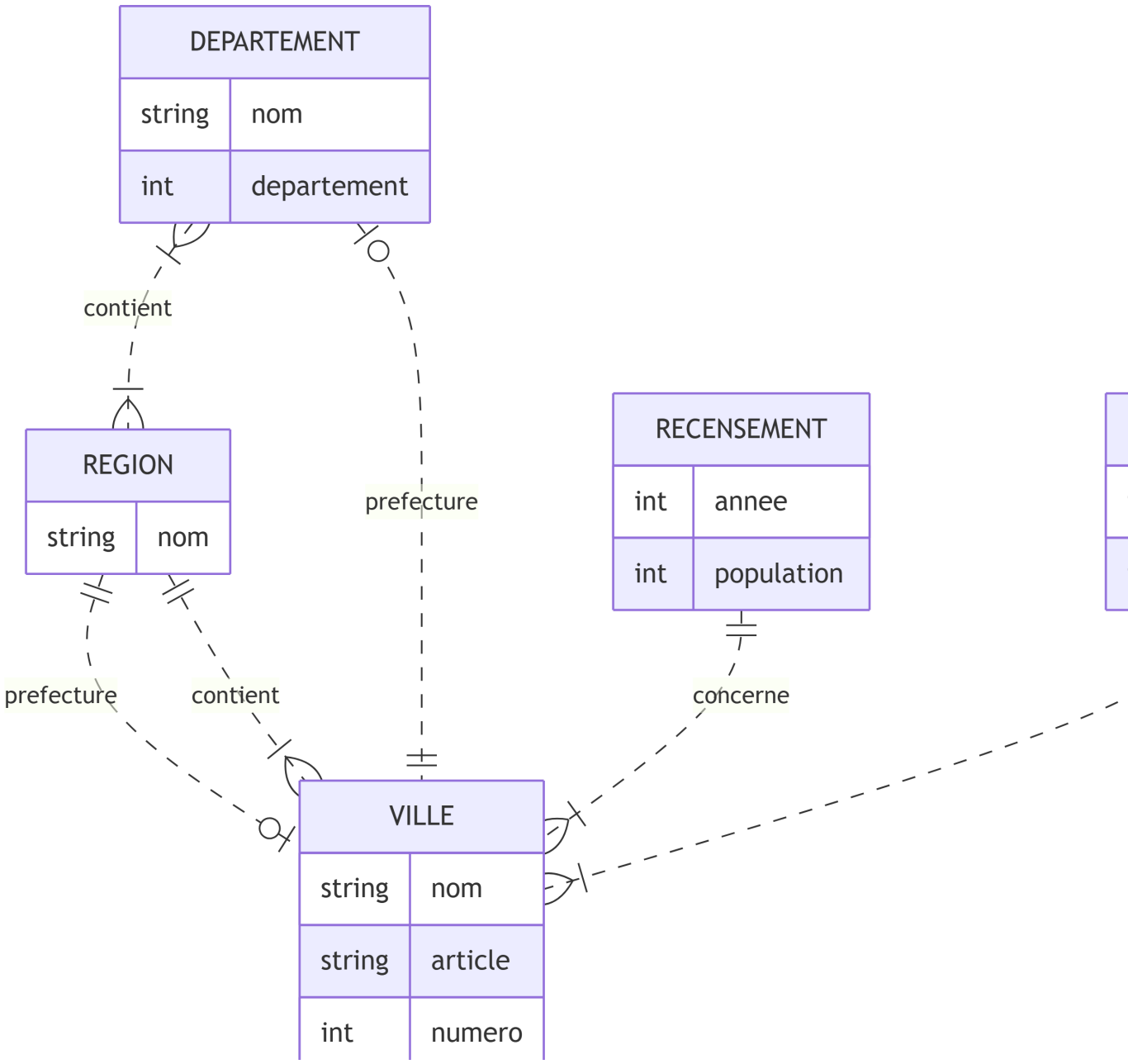
Dans chaque département, il y a des communes qui portent des noms, éventuellement précédés d'un article (par exemple *Le Mans*, *Le Bourget*, …). Pour chaque département, les communes sont numérotées à l'aide de leur rang dans l'ordre alphabétique construit sur les noms privés de l'éventuel article (dans la *Sarthe*, 72, *Le Mans* est au rang 181 entre *Mamers* et *Mansigné*).

Pour chaque commune et chaque année, on connaît le montant total perçu au titre de la taxe foncière.

Chaque année, on procède à un recensement dans chaque commune. On en connaît la population.

Chaque région a une préfecture de région, et chaque département a une préfecture de département.

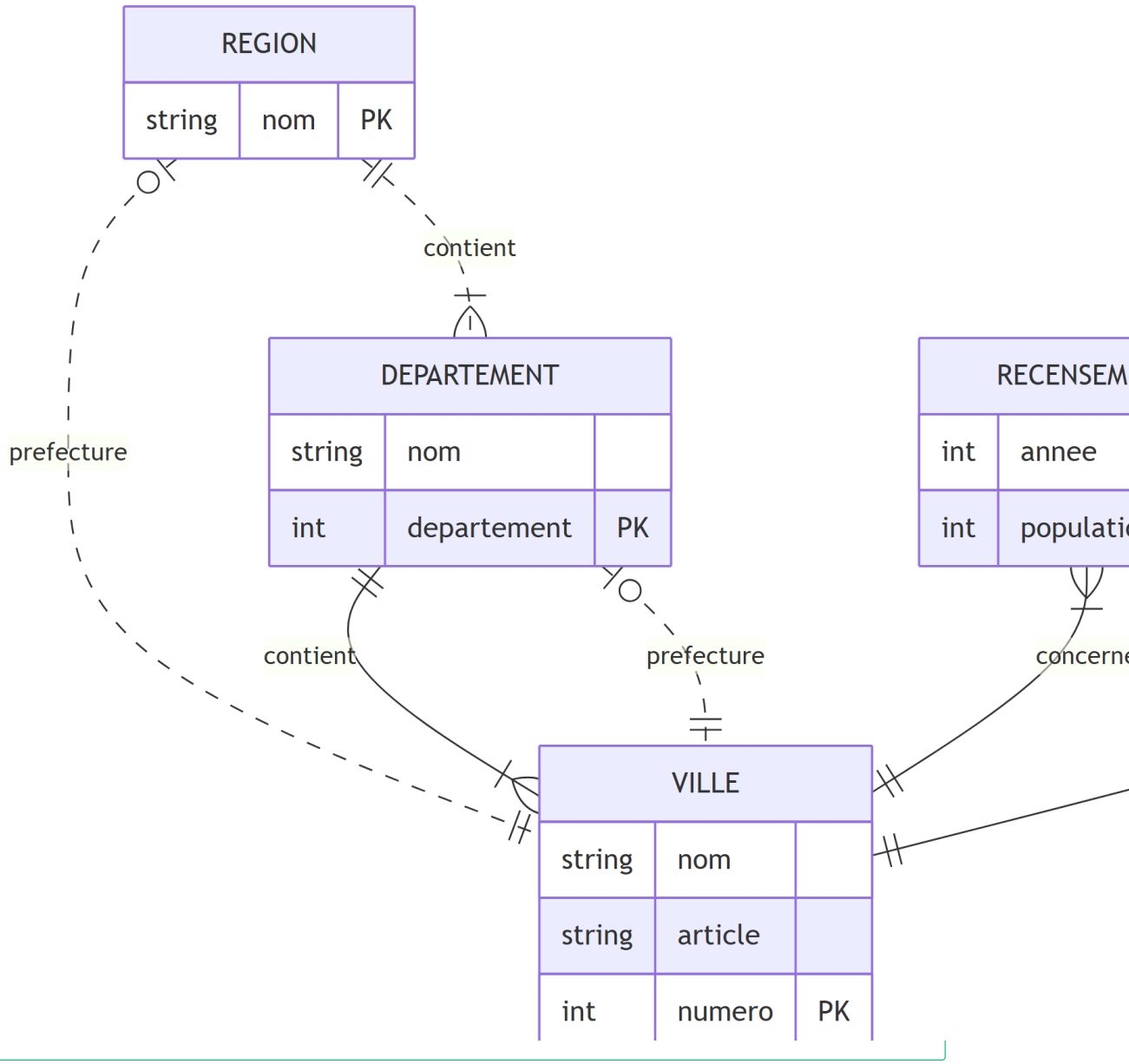
On a cherché à modéliser cette description en formalisme Entité-Association comme suit.



Question 1

1. Un certain nombre d'erreurs se sont glissées dans ce schéma Entité-Association. Corrigez-le.
2. Proposez un identifiant pour chaque entité.
3. Désignez les entités faibles.

Solution



Solution (suite)

Les numéros des communes sont des identifiants relatifs. Ils apparaissent sur nos numéros de sécurité sociale (chiffres en position 4, 5, 6).
`VILLE`, `RECENSEMENT` et `TAXE` sont des entités faibles.
L'identifiant de `VILLE` est le couple `departement, numero`
L'identifiant de `TAXE` est le triplet `departement, numero, annee`. Idem pour `RECENSEMENT`.
Les associations notées en trait plein sont dites *identifiantes*. L'identifiant d'une entité faible est formé en assemblant l'identifiant de l'entité de cardinalité $0:n$ et l'identifiant relatif de l'entité faible (cardinalité $1:1$).

Soit le schéma $\mathcal{A} = \{\mathtt{A}, \mathtt{B}, \mathtt{C}, \mathtt{D}, \mathtt{E}, \mathtt{F}\}$ et l'ensemble Σ de DF

```
A,B → C
  B → F
C,E → A
  F → E
```

Question 2

Calculer la clôture/fermeture $\{\texttt{A},\texttt{B}\} = [\{\texttt{A},\texttt{B}\}]^+_\Sigma$ de $\{\texttt{A},\texttt{B}\}$ en utilisant l'ensemble de DF Σ.

Solution 2

$\{\texttt{A},\texttt{B}\}^+$ contient forcément `A` et `B`. Par la DF `A,B → C`, elle contient aussi `C`. Comme `B → F`, elle contient `F`. Comme `F → E`, elle contient `E`. Donc elle contient les attributs : $\{\texttt{A},\texttt{B}\}^+ = \{\texttt{A},\texttt{B},\texttt{C},\texttt{E},\texttt{F}\}$ (tous sauf `D`).

Question 3

Quelles sont les clefs du schéma $\mathcal{A}$ en utilisant l'ensemble de DF Σ ?

Solution 3

Comme `D` n'apparaît dans aucune DF (en membre gauche, ou en membre droit), toute clé est formée de la réunion de $\{\texttt{D}\}$ et d'une clé pour $\mathcal{A} \setminus \{\texttt{D}\}$ muni de Σ.
On vient de voir que $\{\texttt{A},\texttt{B}\}$ est une super-clef de $\mathcal{A} \setminus \{\texttt{D}\}$ muni de Σ. C'est aussi une clef de $\mathcal{A} \setminus \{\texttt{D}\}$ puisque $[\{\texttt{A}\}]^+_\Sigma = \{\texttt{A}\}$ et $[\{\texttt{B}\}]^+_\Sigma = \{\texttt{B},\texttt{E},\texttt{F}\}$ pour ses deux sous-ensembles.
Donc $\{\texttt{A},\texttt{B},\texttt{D}\}$ est une clé.
Ce n'est pas la seule clef : $\{\texttt{B},\texttt{C}\}$ en est aussi une.
Toutes les super-clefs doivent contenir `B` (qui ne dépend d'aucun autre attribut), et comme $\texttt{B}^+ = \{\texttt{B},\texttt{E},\texttt{F}\}$, `E` et `F` n'apportent rien par rapport à `B` seul.
Les clefs sont donc $\{\texttt{A},\texttt{B},\texttt{D}\}$ et $\{\texttt{B},\texttt{C},\texttt{D}\}$.

On travaillera par la suite sur le schéma relationnel `villes` et les tables définies par le code `SQL` suivant.

```sql
CREATE SCHEMA villes;

CREATE TABLE villes.region (
  nom varchar NOT NULL PRIMARY KEY,
  prefecture_num int4 NULL,
  prefecture_dep int4 NULL
);

CREATE TABLE villes.departement (
  nom varchar NULL UNIQUE,
  departement int4 NOT NULL PRIMARY KEY,
  region varchar NULL REFERENCES villes.region(nom)
    ON DELETE SET NULL ON UPDATE CASCADE,
  numero int4 NULL
);

CREATE TABLE villes.ville (
  nom varchar NULL,
  article varchar NULL,
  departement int4 NOT NULL REFERENCES villes.departement
    ON DELETE RESTRICT ON UPDATE CASCADE,
  numero int4 NOT NULL,
  CONSTRAINT ville_pkey PRIMARY KEY (departement, numero)
);

ALTER TABLE villes.region ADD
  CONSTRAINT region_prefecture_fk FOREIGN KEY (prefecture_dep,prefecture_num)
  REFERENCES villes.ville(departement,numero)
```

```
  ON DELETE RESTRICT ON UPDATE CASCADE;

ALTER TABLE villes.departement ADD
  CONSTRAINT departement_prefecture_fk FOREIGN KEY (departement,numero)
  REFERENCES villes.ville(departement,numero)
  ON DELETE SET NULL ON UPDATE CASCADE;

CREATE TABLE villes.recensement (
  annee year NOT NULL,
  departement int4 NOT NULL,
  numero int4 NOT NULL,
  population int4 NULL,
  CONSTRAINT recensement_pk PRIMARY KEY (annee, departement, numero),
  CONSTRAINT recensement_fk FOREIGN KEY (departement,numero)
    REFERENCES villes.ville(departement,numero)
    ON DELETE RESTRICT ON UPDATE CASCADE
);

CREATE TABLE villes.taxe (
  annee year NOT NULL,
  departement int4 NOT NULL,
  numero int4 NOT NULL,
  valeur int4 NULL,
  CONSTRAINT taxe_pk PRIMARY KEY (annee, departement, numero),
  CONSTRAINT taxe_fk  FOREIGN KEY (departement,numero)
    REFERENCES villes.ville(departement,numero)
    ON DELETE RESTRICT ON UPDATE CASCADE
);
```

Toutes les questions qui suivent se rapportent au schéma `villes`.
Pour chaque question, proposer une requête écrite en algèbre relationnelle ou en `SQL`.

Question 4

Lister par région le nombre de communes.

Solution 4

```
SELECT departement.region, COUNT(*)
  FROM villes.departement JOIN
       villes.ville USING (departement)
 GROUP BY departement.region ;
```

`nom` n'est pas une clé pour `ville`. `COUNT(DISTINCT ville.nom)` ne donnera pas nécessairement le résultat attendu.

Question 5

Quel département contient le plus de communes ?

Solution 5

Version avec `>= ALL` :

```
WITH counts AS (
  SELECT departement, COUNT(*)
  FROM villes.ville
  GROUP BY departement
)

SELECT departement
FROM counts
WHERE count >= ALL (SELECT count FROM counts);
```

Version avec `RANK()` :

```
WITH counts AS (
  SELECT departement.nom, departement, COUNT(ville.numero)
    FROM villes.departement JOIN villes.ville USING (departement)
   GROUP BY departement),
rangs AS (
  SELECT nom, departement,
         RANK() OVER (PARTITION BY departement ORDER BY count DESC)
    FROM counts)
SELECT nom, departement
  FROM rangs
 WHERE rank = 1;
```

Question 6

Quel département a connu la plus forte croissance démographique relative entre 2010 et 2020 ?

Solution 6

```
WITH dpop10 AS (SELECT departement, SUM(population) as pop10
    FROM villes.ville NATURAL JOIN villes.recensement
   WHERE annee = 2010
   GROUP BY departement),
dpop20 AS (SELECT departement, SUM(population) as pop20
    FROM villes.ville NATURAL JOIN villes.recensement
   WHERE annee = 2020
   GROUP BY departement),
croissance AS (SELECT departement, (pop20 - pop10)/pop10 as delta
    FROM dpop10 NATURAL JOIN dpop20)
SELECT departement
  FROM croissance
 WHERE delta >= ALL (SELECT delta FROM croissance);
```

Question 7

Pour chaque région, quelle est la proportion de la population qui vit dans des préfectures départementales en 2020 ?

Solution 7

```sql
WITH totalpop AS (
  SELECT d.region, SUM(r.population) AS total
  FROM villes.departement d JOIN
       villes.recensement r ON (d.departement = r.departement)
  WHERE r.annee = '2020'
  GROUP BY d.region),

  prefpop AS (
  SELECT d.region, SUM(population) AS pref
  FROM villes.departement d JOIN
       villes.recensement r ON (d.departement=r.departement AND
                                d.numero=r.numero)
  WHERE annee = '2020'
  GROUP BY region.nom)

SELECT region, (pref + 0.0)/total
FROM totalpop NATURAL JOIN
     prefpop;
```