

# Examen de *Systèmes d'information*

Fabien Coelho et Claire Medrala – MINES Paris

Lundi 11 janvier 2021 – 13h45

## Conditions de l'examen

La durée de l'examen est de 1 heure 30 minutes tout compris. **Aucun document ni appareil électronique** n'est autorisé. **Aucune communication** entre élèves n'est permise. Un dictionnaire peut être consulté par les élèves non francophones. Le barème est indicatif, la notation sur 20. Des points de bonus sont susceptibles d'être attribués au loisir du correcteur. **Attention**, l'exercice 2 dépend de l'exercice 1.

### 1 Modélisation entité-association /5

Un modèle entité-association (E/A) se compose d'**entités** (rectangles) reliées entre elles par des **associations** (losanges). Chaque lien a une **cardinalité** (? 1 \* +). Entités et associations ont des **attributs** (ovales) qui représentent des données.

Construire un tel modèle pour représenter la situation suivante : une plateforme télévisuelle propose des épisodes de séries de différents types. Chaque série est aussi caractérisée par un ensemble de mots clés. Les abonnés sont caractérisés par leur genre, leur année de naissance, leur région, leur niveau de revenu annuel. Le système garde trace de toutes leurs visualisations, en particulier la durée de visionnage afin de savoir si ils ont regardé un épisode en entier, pour aider à mieux cibler les nouvelles propositions du moteur de recommandations.

Vous pouvez aussi ajouter quelques commentaires si vous les estimez nécessaires, ou discuter brièvement les motivations des choix de modélisation faits et les limites du modèle.

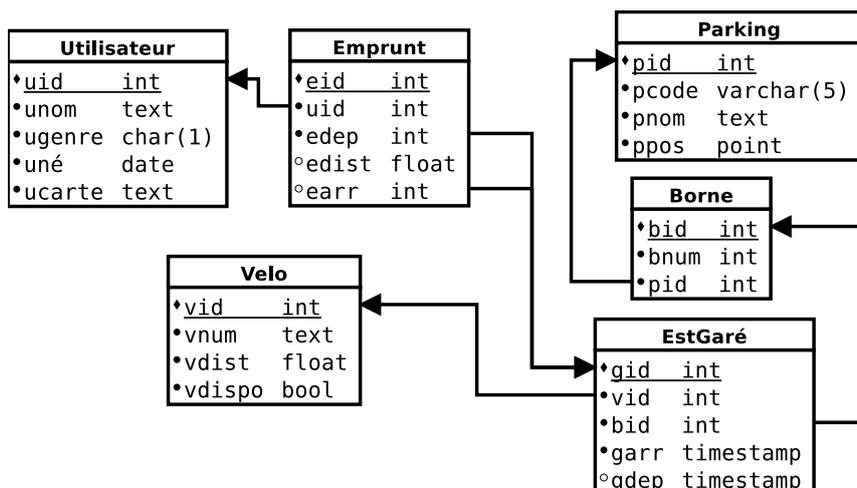
### 2 Traduction relationnelle /3

À partir du modèle E/A précédent, construisez un modèle relationnel. Vous prendrez soin de bien préciser les champs utiles et les contraintes pertinentes sur vos relations. Vous commenterez les contraintes que vous ne pourriez exprimer directement dans le modèle.

Pour gagner du temps de rédaction, vous utiliserez les abréviations suivantes : **CT** pour CREATE TABLE, **I** pour INTEGER, **S** pour SERIAL, **T** pour TEXT, **C** pour CHAR, **VC** pour VARCHAR, **D** pour DATE, **Ts** pour TIMESTAMP, **In** pour INTERVAL, **F** pour FLOAT, **B** pour BOOLEAN **PK** pour PRIMARY KEY, **U** pour UNIQUE, **NN** pour NOT NULL, **R** pour REFERENCES. . .

### 3 Requêtes /8

On considère le modèle relationnel suivant, qui représente les données d'un système de location de vélos de type *Vlib*. Les utilisateurs du système ont un nom, un genre, une date de naissance (triplet supposé unique), une carte d'abonnement (unique également). Les vélos ont un numéro (unique), une distance parcourue, peuvent être disponibles ou non (par exemple ils sont en réparation, la batterie est en cours de rechargement. . .). Les bornes d'accueil (numérotées) des vélos sont rattachées à des parkings. Les parkings ont un code (unique), un nom (unique), une position GPS (unique). Un vélo est garé à une borne avec un instant d'arrivée et éventuellement un instant de départ. Enfin, les utilisateurs empruntent des vélos entre deux bornes où ils sont garés (départ et arrivée), et parcourent lors de cet emprunt une certaine distance connue à l'arrivée.



Les attributs soulignés sont les clés primaires, les points noirs désignent des attributs NOT NULL, les flèches sont des clés étrangères.

Proposez **une** requête (en cas d'ambiguïté choisir une interprétation) ou à défaut sa structure, pour chacune des questions suivantes :

- Quels vélos, identifiés par leurs numéros, ont été empruntés par l'utilisateur *Calvin* le 11 décembre 2020 au parking *Pont Neuf* ?  
Pour cette requête uniquement, suggérez des index (hors clés primaires ou unique signalés) potentiellement utiles pour en améliorer les performances.
- Quels vélos, identifiés par leurs numéros, ont été garés dans au moins quatre parkings différents, par ordre de leurs numéros ?
- Compter le nombre de vélos actuellement garés et le nombre de vélos actuellement empruntés, sous forme de deux tuples, y compris si une réponse est zéro.
- Pour *tous* les utilisateurs, la distance parcourue, et leur rang dans l'ordre alphabétique de leurs noms, par ordre des distances décroissantes, puis des noms.
- Pour *tous* les vélos, donner le temps total et le nombre d'emprunts, par ordre décroissant des temps, puis par ordre de leurs numéros.
- Redondance...
  - Le modèle présenté comporte au moins une redondance. Laquelle ?
  - Si c'est volontaire, quelle raison a pu pousser le concepteur à cela ?
  - Afficher les objets pour lesquels l'écart entre les deux informations redondantes est supérieur à 5%. On prendra garde à traiter correctement les cas particuliers.

## 4 Questions de cours

/4

Choisissez un thème parmi les deux tirés aléatoirement en début d'examen dans la liste *Postgres, Relationnel, Optimisation, Droits, Transaction, MVCC, PL/pgSQL, Formes normales* et expliquez en moins de 100 mots ce que vous en avez retenu.

Citez les noms de trois scientifiques ayant obtenu le prix Turing pour leurs travaux de recherche sur les bases de données.