

Examen de *Systèmes d'information*

Fabien Coelho et Claire Medrala – Mines Paris – PSL

Lundi 10 janvier 2022 – 10h45

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 des correcteurs. **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 :

Un coureur à pied, éventuellement membre d'un club, peut s'inscrire à des trails (courses dans la nature) de différents niveaux. Un trail est composé d'étapes ordonnées. Une étape est un point géographique caractérisé par ses coordonnées et un délai maximal de passage pour ne pas être disqualifié de la course. La première et la dernière étape du trail constituent son départ et son arrivée. Lors d'une course, le système permet d'enregistrer l'heure de passage des coureurs à chaque étape. Un coureur peut abandonner sa course.

Vous pouvez également discuter brièvement les motivations des choix de modélisation faits et les limites du modèle.



2 Traduction relationnelle

/3

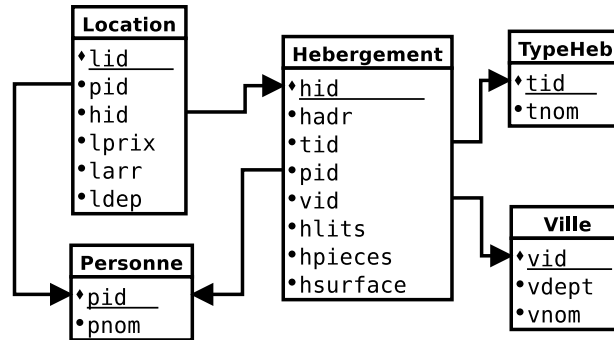
À partir du modèle E/A précédent, construisez un modèle relationnel. Vous prendrez soin de bien préciser les champs utiles et les contraintes pertinentes sur vos relations. Vous commenterez les contraintes que vous ne pourriez exprimer directement dans le modèle.

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

3 Requêtes

/8

On considère le modèle relationnel suivant, qui enregistre les locations d'hébergements de différents types dans différentes villes effectuées par des personnes. Les attributs soulignés sont les clefs primaires, les points noirs désignent des attributs NOT NULL, les flèches sont des clefs étrangères.



Les personnes sont identifiées par leur nom supposé unique (**pnom**). Un type d'hébergement (**tnom**) peut être une *maison*, un *appartement*, une *cabane*... Une ville est dans un département et a un nom (**vdept**, **vnom**), la paire étant unique. Un hébergement a une adresse unique (**hadr**) dans la ville, un type (**tid**), un propriétaire (**pid**), une ville (**vid**), un nombre de lits (**hlits**), un nombre de pièces (**hpieces**), une surface (**hsurface**). Enfin, la location d'un hébergement (**hid**) par une personne (**pid**) a une date d'arrivée (**larr**), une date de départ (**ldep**) et le prix total payé (**lprix**).

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

1. Quelles sont les villes et adresses ordonnées des hébergements de type *maison* d'au moins 100 m^2 loués par *Calvin* (locataire) dans le département *971* ?
Pour cette requête uniquement, suggérez au moins 3 index (hors clefs primaires ou uniques signalées) potentiellement utiles pour en améliorer les performances.
2. Combien d'hébergements différents ont été loués en juillet 2020 (date d'arrivée) à un prix moyen par jour inférieur à 120 € ?
3. Quelles sont les noms ordonnés des propriétaires qui n'ont jamais été locataires ?
4. Quels sont les logements ordonnés (département, ville, adresse) avec leur durée moyenne de location et leur nombre total de locations, pour lesquels la durée moyenne de location est supérieure à la moyenne globale des durées de toutes les locations ?
5. Quels sont les noms ordonnés des propriétaire qui ont loué leurs biens plusieurs fois à eux-mêmes ?
6. Pour tous les hébergements ordonnés (département, ville, adresse), quel est le nombre (y compris 0) de locations commencées dans l'année 2020 et celles commencées dans l'année 2021 ?
7. Quels hébergements ordonnés (département, ville, adresse) ont été loués *en continu* (locations successives sans interruptions) pendant au moins 300 jours ?

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, REST, Python DB API, Injection SQL, Décisionnel, Formes normales* et expliquez en moins de 100 mots ce que vous en avez retenu.

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