

# Bases de données: Devoir final

Pierre Senellart ([pierre.senellart@ens.fr](mailto:pierre.senellart@ens.fr))

27 mai 2020

Vous avez trois heures pour réaliser ce devoir final, depuis chez vous. Tous les documents sont autorisés, mais la communication entre étudiants est interdite et toute preuve de collusion entre étudiants sera sévèrement pénalisée. Avant l'issue des trois heures, vous devez soumettre votre devoir sur le site Moodle du cours, à l'endroit indiqué, sous la forme d'un unique PDF (qui peut contenir un document produit électroniquement ou des scans ou photos de documents papier, tant que le résultat est lisible et prend moins de 8 MB). Tout retard dans le rendu sera pénalisé de points en moins. Le devoir est constitué d'un unique problème de 10 questions, noté sur 20 points.

On souhaite modéliser une partie du système d'informations de gestion des patients d'un hôpital. Un patient a normalement un nom, un prénom, un identifiant unique (son numéro INSEE), une date de naissance, un genre. Cependant, toute ou partie de cette information peut être vide (par exemple, patient étranger ou d'identité inconnue). Un personnel de l'hôpital dispose des mêmes informations (cette fois-ci obligatoirement présentes) plus d'un rôle (aide soignant, infirmier, médecin, administratif) et un service (un département de l'hôpital). Chaque visite d'un patient de l'hôpital est associée à une date d'admission et de départ (pour les patients qui sont déjà partis), ainsi qu'à un service. Pour chaque départ, on indique si le patient est considéré guéri, a été transféré dans un autre service ou hôpital, ou si le patient est décédé. (On considère ainsi qu'un patient passant par plusieurs services a une visite par service successif.) À chaque visite d'un patient est associée une liste non vide de symptômes, ainsi qu'une liste de diagnostics (une liste de maladies), chaque diagnostic étant établi à une date précise par un médecin de l'hôpital.

1. (2 points) Proposer un schéma entité-association pour ce système d'informations. Si vous êtes amené à faire des hypothèses pour interpréter les besoins décrits plus haut, les indiquer explicitement.
2. (2 points) En déduire un schéma relationnel et l'exprimer avec des ordres SQL `CREATE TABLE` ; y faire figurer les clefs primaires (`PRIMARY KEY`), clefs candidates (`UNIQUE`), clefs étrangères `REFERENCES`, contraintes `CHECK`, et indiquer quels attributs sont nécessairement `NOT NULL`.
3. (1 point) Quelles sont les contraintes logiques du système d'informations qui ne sont pas impliquées par les contraintes SQL ? Les exprimer sous forme de formule de la logique du premier ordre, si possible sous forme de TGD ou d'EGD.
4. (3 points) Écrire une requête calculant les noms et prénoms des patients ayant été diagnostiqués de la maladie X, avec leur état (en cours de traitement, guéri, transféré, décédé) :
  - a) en algèbre relationnelle ;
  - b) en calcul relationnel ;
  - c) en SQL.
5. (2 points) Proposer trois plans d'exécution pour cette requête, et discuter du plan le plus efficace, en fonction des tailles des différentes relations.

6. (1 point) Donner une borne supérieure sur la complexité asymptotique, en les données, de l'évaluation de cette requête.
7. (2 points) Écrire en SQL une requête calculant, pour une maladie X donnée, la liste des symptômes associés à cette maladie et le nombre de malades ayant quitté l'hôpital ayant exhibé ce symptôme, triée par nombre de malades décroissants.
8. (2 points) Proposer, dans ce schéma d'informations, une requête utile pour laquelle il est nécessaire d'employer un langage de requêtes récursif. Décrire cette requête en français et l'écrire dans le langage de requête récursif de votre choix.
9. (2 points) Proposer une manière de gérer les droits d'accès aux données : comment faire en sorte qu'un personnel administratif n'ait pas accès aux symptômes et maladies des patients, mais puisse avoir accès à l'ensemble des patients, de leurs visites, et, indépendamment, à des statistiques sur les maladies et symptômes fréquents ? Sans écrire de code, détailler précisément comment ce contrôle de droits serait mis en place au sein de la base de données et dans une application exploitant ces données.
10. (3 points) On souhaite utiliser la base de données ainsi construite pour donner des statistiques en temps réel sur le site Web de l'hôpital du nombre de patients actuellement en cours de traitement dans chacun des services de l'hôpital. Indiquer comment procéder (y compris en indiquant la ou les requêtes SQL utiles) et quelles précautions employer pour éviter tout problème de sécurité potentiel, en donnant des exemples précis.