



## IGE487 – Modélisation de bases de données

### Examen #1 (questionnaire) – Automne 2020

---

#### Enseignant

Luc Lavoie

#### Date

2025-10-03

#### Notes et consignes

- ◇ Cet examen d'une durée de 110 minutes est individuel.
- ◇ La correction est basée, entre autres, sur le fait que chacune des réponses soit claire (c'est-à-dire lisible et compréhensible), exacte (c'est-à-dire précise et sans erreurs), complète (c'est-à-dire comprenant tous les éléments requis et tous les éléments nécessaires) et concise (c'est-à-dire sans éléments superflus).
- ◇ La documentation personnelle est limitée à une seule feuille recto verso au format ISO A4 ou US Letter.
- ◇ Aucun équipement informatique, électronique ou de communication n'est autorisé.
- ◇ **Justifier les réponses et remettre le questionnaire et le cahier avant de quitter la salle d'examen.**

#### Pondération

Question	Q1	Q2	Q3	Q4	Q5	Total
Poids	/20	/20	/20	/20	/20	/100

#### Questions

- (1) Le modèle (utilisé par le langage) SQL diffère du modèle relationnel en plusieurs points. Présenter cinq de ces différences en les illustrant chacune d'un exemple permettant d'en mesurer la portée effective. Compléter votre exemple d'un jeu de données et d'expressions relationnelles dont les résultats montrent la différence entre l'approche SQL et l'approche relationnelle. Donner les avantages et les inconvénients qui en découlent, dans un modèle et dans l'autre.
- (2) Un résultat de test en laboratoire (RTL) est caractérisé par un numéro unique, la date de prise de l'échantillon et la date de l'analyse. Il existe plusieurs types de RTL : DFG, HbA1c, etc. Un DFG comporte deux mesures : la créatinine (en g/L) et le DFG lui-même (en (mL/min)/1,73 m<sup>2</sup>). Un HbA1c ne comporte que la mesure HbA1c elle-même (un taux exprimé en pourcentage). Une mesure est caractérisée par sa valeur (numérique), son unité de mesure et le nom du protocole de mesurage (pMes). Concevoir un schéma relationnel en forme normale de Boyce-Codd (FNBC) permettant de représenter différents types de RTL, de stocker des RTL et de valider les RTL en regard de leur type. Identifier chacune des dépendances fonctionnelles (en ajoutant au besoin les dépendances implicites communes). Identifier chacune des clés (candidates et référentielles). Donner le prédicat correspondant à chacune des relations.
- (3) Soit les relations R (x, a) et S (y, b) et T (x, y, c, d, e) où {x} est l'unique clé candidate de R, {y} est l'unique clé candidate de S et {x, y} est l'unique clé candidate de T. En outre, {x} est une clé référentielle de T vers R et {y} est une clé référentielle de T vers S. Supposer que l'attribut d est inapplicable pour certaines valeurs de c et que l'attribut e peut être absent (annulable). Concevoir un schéma relationnel en FNBC représentant R, S et T sans faire appel à des attributs annulables. Le schéma est-il en cinquième forme normale ? Pourquoi ? Montrer comment obtenir les relations d'origine R, S et T à partir des relations de votre schéma.



- (4) Les deux structures de données les plus fréquemment utilisées pour représenter les index sont les arborescences B+ et l'adressage dispersé (*hash coding*). Donner leurs avantages et inconvénients respectifs en regard des index primaires et des index secondaires, lorsque mis en oeuvre en *mémoire primaire (RAM)*.
- (5) Proposer les meilleures stratégies pour calculer la jointure de deux relations dans le contexte suivant :
- les relations sont représentées horizontalement (tuple par tuple) ;
  - les index peuvent être entièrement stockés en mémoire interne ;
  - les tuples sont stockés en mémoire externe ;
  - le temps d'accès à la mémoire externe est dix-mille fois plus grand que celui de la mémoire interne ;
  - toutes les clés candidates sont indexées, mais elles seulement.

Considérer les trois cas suivants relativement à l'ensemble des attributs de jointure :

- il forme une clé candidate dans chacune des relations ;
- il s'agit d'une clé référentielle ;
- il n'est indexé par aucune des relations.

Considérer les deux cas suivants relativement aux cardinalités des relations :

- elles sont approximativement égales ;
- l'une est dix fois plus grande que l'autre.

Pour chacun des six cas, proposer la meilleure stratégie : donner les algorithmes utilisés et l'ordre de complexité en temps.