

EXAMEN - BASES DE DONNEES
IFSIC DIIC2 LSI,
2006-2007
Durée : 2 heures
Documents autorisés

Exercice 1 (4 points)

On désire gérer les emplois du temps des différents personnels (*enseignants, enseignants-chercheurs et chercheurs*) de l'UFR, pour savoir à tout moment s'il est possible de les joindre, et où. Pour cela, on considère que, en dehors des périodes où ils peuvent être joints dans leur bureau, les personnels peuvent être en réunion, quel que soit leur statut. Une réunion est désignée par une date précise, une tranche horaire et une salle de réunion. On veut connaître les autres personnes participant à la réunion.

Chaque personne est désignée par son nom, son prénom, le bureau où on peut la joindre.

Les *enseignants* peuvent, de plus, être en cours. Un cours est identifié par la matière enseignée à laquelle est affectée toujours la même salle. Il est désigné par une période de début et de fin (ex. de février à mai), un jour de la semaine, une tranche horaire et une salle de cours. Plusieurs *enseignants* peuvent enseigner la même matière dans l'année, à des jours et créneaux horaires différents. Un enseignant peut enseigner plusieurs fois la même matière dans l'année, à des périodes différentes.

Les *chercheurs* peuvent être à certaines périodes de l'année en mission en dehors de l'UFR. Une mission est désignée par une date de début et de fin, un lieu de mission avec le numéro de téléphone correspondant. Les chercheurs appartiennent à un laboratoire dont on peut joindre le secrétariat en cas d'urgence.

Les *enseignants-chercheurs* sont à la fois enseignants et chercheurs, avec un pourcentage plus ou moins grand d'enseignement (par rapport à la recherche) à effectuer. Ils peuvent donc être soit en réunion, soit en mission, soit en cours.

- 1) Etablir le schéma Entité-Association de cette application.
- 2) En déduire un schéma de base de données relationnelle.

Ne pas oublier de préciser les cardinalités pour les associations et les identificateurs des entités.

Exercice 2 (6 points)

Soit la base de schéma :

JOUEUR (NUJOUEUR, NOM, PRENOM, ANNAISS, NATIONALITE)
RENCONTRE (NUGAGNANT, NUPERDANT, LIEUTOURNOI, ANNEE, SCORE)
GAIN (NUJOUEUR, LIEUTOURNOI, ANNEE, PRIME, SPONSOR)

Les attributs NUGAGNANT, NUPERDANT et NUJOUEUR y sont définis sur le même domaine.

Les clés des relations sont soulignées. On suppose que tous les joueurs, engagés dans un tournoi, touchent une prime.

Exprimer les requêtes suivantes : **1)** par une expression algébrique (quand elle est possible) et **2)** en SQL :

- a) Nom et nationalité des joueurs ayant participé à la fois au tournoi de Roland Garros et à celui de Wimbledon, en 1992.
- b) Nom et nationalité des joueurs ayant été sponsorisés par Peugeot et ayant gagné à Roland Garros au moins un match (avec un sponsor quelconque).
- c) Numéros des joueurs qui ont toujours perdu à Wimbledon, et toujours gagné à Roland Garros.
- d) Liste des vainqueurs français de tournoi, mentionnant le nom du joueur, son sponsor, le lieu et l'année des tournois qu'il a gagné.
- e) Nom des joueurs ayant participé à tous les tournois disputés en 1994, avec leur prime.

Exercice 3 (4 points)

Soit S le schéma de base de données relationnelle suivant :

$$S = \{ R(A,B,C,D) \}$$

sur lequel on a défini un ensemble F de dépendances fonctionnelles tel que :

$$F = \{ BC \rightarrow DA, D \rightarrow C, C \rightarrow A \}$$

- 1) Donnez la clé minimale de R et la couverture minimale.
- 2) S est-il en forme normale de Boyce-Codd ? S est-il en 3ème forme normale ? Expliquez votre raisonnement.
- 3) La décomposition de S en $S' = \{ R1(A,D), R2(B,C,D) \}$ est-elle valide et sans perte d'information ? Expliquez votre raisonnement.
- 4) La décomposition de S en S' préserve-t-elle les dépendances fonctionnelles ? Expliquez votre raisonnement.

Exercice 4 (6 points)

D'après le schéma l'exercice 2 :

JOUEUR (NUJOUEUR, NOM, PRENOM, ANNAISS, NATIONALITE)
RENCONTRE (NUGAGNANT, NUPERDANT, LIEUTOURNOI, ANNEE, SCORE)
GAIN (NUJOUEUR, LIEUTOURNOI, ANNEE, PRIME, SPONSOR)

On a obtenu les statistiques suivantes :

Nom de table	Nombre de tuples	Facteur de blocage
JOUEUR	NJ=500	FbJ=50
RENCONTRE	NR=3000	FbR = 300
GAIN	NG = 3000	FbG = 300

Pour la table **JOUEUR** :

- un index primaire **J1** existe sur l'attribut **NUJOUEUR** avec **nNiveaux_NUJOUEUR(J1) = 3**

JOUEUR	NUJOUEUR	NOM	PRENOM	ANNAISS	NATIONALITE
Nombre de valeurs distinctes	500	100	100	50	10
Valeur minimale	0	-	-	1920	-
Valeur maximale	499	-	-	1990	-

Pour la table **RENCONTRE** :

- un index primaire existe sur la clé avec un seul niveau.
- un index secondaire **R** existe sur **SCORE** avec **nNiveaux_SCORE(R) = 2**

RENCONTRE	NUGAGNANT	NUPERDANT	LIEUTOURNOI	ANNEE	SCORE
Nombre de valeurs distinctes	200	200	10	30	10
Valeur minimale	0	0	-	1920	-
Valeur maximale	499	499	-	2006	-

Pour la table **GAIN** :

- un index primaire **G1** existe sur l'attribut **NUJOUEUR** avec **nNiveaux_NUJOUEUR(G1) = 2**
- un index secondaire **G2** en arbre B+ existe sur **PRIME** avec **nNiveaux_SPONSOR(G2) = 3** et **nBlocsF_PRIME(G2) = 50** *prime*

GAIN	NUJOUEUR	LIEUTOURNOI	ANNEE	PRIME (en k€)	SPONSOR
Nombre de valeurs distinctes	500	10	30	20	30
Valeur minimale	0	-	1920	1	-
Valeur maximale	499	-	2006	100	-

- 1) Calculer le nombre de blocs occupés par chaque table.
- 2) Que signifie la requête suivante ?

```
SELECT NOM, PRENOM, SPONSOR
FROM JOUEUR J, RENCONTRE R, GAIN G
WHERE
    PRIME > 90 AND
    NATIONALITE='USA' AND
    R.SCORE='FORFAIT' AND
    G.NUJOUEUR = R.NUPERDANT AND
    J.NUJOUEUR = G.NUJOUEUR AND
    R.ANNEE=G.ANNEE AND
    R.LIEUTOURNOI= G.LIEUTOURNOI;
```

J X G
R X G

A
G

- 3) Calculer la cardinalité de sélection des attributs utilisés dans la requête précédente.
- 4) Représenter l'arbre algébrique de la requête obtenu par l'optimisation algébrique.
- 5) Donner une estimation du coût et la taille des résultats pour chaque opération de sélection et de jointure de la requête donnée en 2). Indiquer en particulier pour les jointures, quel algorithme devrait être appliqué pour optimiser le coût en entrée/sortie de la jointure.