

<input type="checkbox"/>	0	<input type="checkbox"/>	0	<input type="checkbox"/>	0
<input type="checkbox"/>	1	<input type="checkbox"/>	1	<input type="checkbox"/>	1
<input type="checkbox"/>	2	<input type="checkbox"/>	2	<input type="checkbox"/>	2
<input type="checkbox"/>	3	<input type="checkbox"/>	3	<input type="checkbox"/>	3
<input type="checkbox"/>	4	<input type="checkbox"/>	4	<input type="checkbox"/>	4
<input type="checkbox"/>	5	<input type="checkbox"/>	5	<input type="checkbox"/>	5
<input type="checkbox"/>	6	<input type="checkbox"/>	6	<input type="checkbox"/>	6
<input type="checkbox"/>	7	<input type="checkbox"/>	7	<input type="checkbox"/>	7
<input type="checkbox"/>	8	<input type="checkbox"/>	8	<input type="checkbox"/>	8
<input type="checkbox"/>	9	<input type="checkbox"/>	9	<input type="checkbox"/>	9

← codez ci-contre le numéro donné par l'enseignant, et inscrivez le aussi ci-dessous.

NOM Prénom :

Durée : 60 minutes. Aucun document n'est autorisé.

Les téléphones portables sont interdits.

Les questions faisant apparaître le symbole ♣ peuvent présenter zéro, une ou plusieurs bonnes réponses. Les autres questions ont une unique bonne réponse.

En général, une bonne case cochée compte pour 1 point et une mauvaise case cochée compte pour -1 point.

Vous devez noircir complètement les cases.

Si vous utilisez du blanco correcteur pour décocher une case ne re-dessinez pas la case.

---

## DF, FN et Algorithmes de décomposition

---

Nous souhaitons concevoir une base de données pour gérer des films. L'ensemble des dépendances fonctionnelles est le suivant (ensemble non exhaustif) :

1.  $\text{idArtiste}, \text{nom} \rightarrow \text{annéeNaissance}, \text{prenom}$
2.  $\text{idArtiste}, \text{annéeNaissance} \rightarrow \text{nom}, \text{prenom}$
3.  $\text{idFilm} \rightarrow \text{titre}, \text{idRéalisateur}, \text{genre}, \text{résumé}$
4.  $\text{idFilm}, \text{idArtiste} \rightarrow \text{nomRôle}$
5.  $\text{idFilm}, \text{idArtiste}, \text{genre} \rightarrow \text{nomRôle}$
6.  $\text{idFilm}, \text{titre} \rightarrow \text{année}, \text{idRéalisateur}, \text{titre}$
7.  $\text{idArtiste} \rightarrow \text{nom}, \text{prenom}$

On considère la relation R munie de l'ensemble  $\text{DF} = \{(1), (2), (3), (4), (5), (6), (7)\}$

$R(\text{idArtiste}, \text{nom}, \text{prenom}, \text{annéeNaissance}, \text{idFilm}, \text{titre}, \text{année}, \text{idRéalisateur}, \text{genre}, \text{résumé}, \text{nomRôle})$

**Question 1** Quelles sont les clés possibles pour R ?

- ☐  $\{\text{idArtiste}, \text{annéeNaissance}, \text{titre}, \text{idFilm}\}$ 
☐  $\{\text{idArtiste}, \text{idFilm}, \text{genre}\}$   
☒  $\{\text{idArtiste}, \text{idFilm}\}$ 
☐ Aucune de ces réponses n'est correcte.  
☐  $\{\text{idArtiste}\}, \{\text{idArtiste}, \text{idFilm}\}$

**Question 2** Selon DF, est-ce que la dépendance fonctionnelle 5 est élémentaire ?

- ☒ Non
 ☐ Oui
 ☐ On ne peut pas savoir

**Question 3 ♣** Parmi les dépendances fonctionnelles de l'ensemble DF, lesquelles apportent de la redondance (même partiellement) ?

- ☒ 1    ☒ 2    ☐ 3    ☐ 4    ☒ 5    ☒ 6    ☐ 7  
☐ Aucune de ces réponses n'est correcte.

**Question 4** Quelle est la couverture minimale CV(DF) ?

- ☐ idArtiste→nom; idArtiste→prénom; idArtiste→annéeNaissance; idArtiste, annéeNaissance→nom; idArtiste, annéeNaissance→prénom; idFilm→titre; idFilm,titre→année; idFilm→idRéalisateur; idFilm→genre; idFilm→résumé; idFilm, idArtiste→nomRole  
☐ idArtiste→nom, prénom, annéeNaissance; idFilm,titre→année; idFilm→titre, idRéalisateur, genre, résumé; idFilm, idArtiste→nomRôle  
☐ Aucune de ces réponses n'est correcte.  
☒ idArtiste→nom; idFilm→titre; idFilm→genre; idFilm→resume; idFilm,idArtiste→nomRole; idFilm→annee; idFilm→idRealisateur; idArtiste→anneeNaissance; idArtiste→prenom

**Question 5 ♣** Les formes normales fournissent un cadre formel permettant de limiter les anomalies de redondance et celles rencontrées à l'ajout, à la suppression ou encore à la mise à jour des tuples. Les dépendances fonctionnelles (DF) entre les données sont utilisées pour guider la solution à ces anomalies. La forme normale FNBC permet d'assurer l'élimination de certaines anomalies, quel type de DF est possible dans une relation en FNBC ?<sup>a</sup>

- ☒ C → NC    ☒ C → PC    ☐ PC → PC    ☐ NC → PC  
☐ NC → NC    ☐ PC → NC    ☐ Aucune de ces réponses n'est correcte.

<sup>a</sup>Pour rappel : NC (attribut ou ensemble d'attributs ne faisant pas parti d'une clé), C (attribut ou ensemble d'attributs étant une clé), PC (attribut ou ensemble d'attributs faisant parti d'une clé).

**Question 6** Les algorithmes de normalisation permettent d'obtenir un ensemble de relations vérifiant une forme normale souhaitée. L'algorithme de Bernstein, permet d'obtenir de relations en 3FN. Réaliser la décomposition de R munie de son ensemble DF avec l'algorithme de Bernstein. Combien de relations sont obtenues ?

- ☒ 3    ☐ 4    ☐ 5    ☐ Aucune de ces réponses n'est correcte.

---

## Evaluation de requêtes

---

Rappel de la sélectivité  $S$  de la sélection :

$$S_{\sigma(A=valeur)} = 1/|\pi_A(R)| ; S_{\sigma(A > valeur)} = (max(A) - valeur) / (max(A) - min(A) + 1) ;$$

$$S_{\sigma(p1 \wedge p2 \wedge \dots \wedge pn)} = \prod_{i=1}^{pn} S_{\sigma(i)} ; S_{\sigma(p1 \vee p2 \vee \dots \vee pn)} = \sum_{i=1}^{pn} S_{\sigma(i)} - \prod_{i=1}^{pn} S_{\sigma(i)}.$$

Soit une base de données simplifiée composée des relations :

**Film**(idFilm number, titre varchar2(20), année number, genre varchar2(20)).

**Roles**(idFilm number, idArtiste number, nomRole varchar2(20)).

Remarquez que la relation Roles a une clé étrangère vers la relation Film.

On suppose que :

- il existe 2 implémentations pour la sélection (n étant le nombre de pages) : parcours séquentiel ( $O(n)$ ) et utilisation d'arbre B+ ( $O(\log n)$ ), on les appellera respectivement  $\sigma^{O(n)}$  et  $\sigma^{O(\log n)}$
- il existe 2 implémentations de la projection : sans élimination des doublons ni tri ( $O(n)$ ) et avec élimination des doublons et tri ( $O(n \cdot \log n)$ ), on les appellera respectivement  $\pi^{O(n)}$  et  $\pi^{O(n \cdot \log n)}$  ;
- il existe 3 implémentations pour la jointure : boucle imbriquée ( $O(n \cdot m)$ ), boucle imbriquée avec index ( $O(n \cdot \log m)$ ), et tri-fusion avec une des relations déjà triée si jointure par clé primaire et clé étrangère ( $O(n + m \log m)$ ); on les appellera respectivement  $\bowtie^{O(n \cdot m)}$ ,  $\bowtie^{O(n \cdot \log m)}$  et  $\bowtie^{O(n + m \cdot \log m)}$ ;

- les attributs clés primaires sont indexés avec des index B+ ;
- les types date, number et varchar2(m), occupent respectivement 12, 6 et m octets ;
- le nombre de tuples de la relation Film est de 5 000 ;
- le nombre de tuples de la relation Roles est de 100 000 ;
- la taille d'un bloc est de 8192 octets ;
- les valeurs des attributs sont indépendantes et la distribution de ces valeurs dans leur domaine est uniforme.

**Question 7** Soit la requête suivante. A priori quelle serait l'expression algébrique correspondant à un plan d'exécution correct ET le plus susceptible d'être choisi par le SGBD (le plus optimisé) ?  
SELECT titre, nomRole

FROM Film NATURAL JOIN Roles WHERE année=2016;

- ☐ Aucune de ces réponses n'est correcte.
- ☐  $\pi_{titre, nomRole}^{O(n)}(\sigma_{annee=2016}^{O(\log n)}(Film) \bowtie^{O(n \cdot \log m)} Roles)$
- ☒  $\pi_{titre, nomRole}^{O(n)}(\sigma_{annee=2016}^{O(n)}(Film) \bowtie^{O(n \cdot \log m)} Roles)$
- ☐  $\pi_{titre, nomRole}^{O(n \cdot \log n)}(\sigma_{annee=2016}^{O(n)}(Film) \bowtie^{O(n \cdot m)} Roles)$
- ☐  $\pi_{titre, nomRole}^{O(n \cdot \log n)}(\sigma_{annee=2016}^{O(\log n)}(Film \bowtie (\pi_{idFilm, nomRole}^{O(n \cdot \log n)}(Roles))))$
- ☐  $\pi_{titre, nomRole}^{O(n)}(\sigma_{annee=2016}^{O(n)}(Film \bowtie^{O(n \cdot m)} Roles))$
- ☐  $\pi_{titre, nomRole}^{O(n \cdot \log n)}(\sigma_{annee=2016}^{O(n)}(Film \bowtie^{O(n \cdot \log m)} Roles))$

**Question 8** Quelle est l'estimation du nombre de tuples résultat de la requête :  
SELECT \*

FROM Film NATURAL JOIN Roles;

- ☐ |Film|\*|Roles|      ☐ |Film|      ☐ |Film|\*|Roles|\*selectivité de la jointure
- ☒ |Roles|

**Question 9** Supposez que le nombre de valeurs distinctes de l'attribut genre est de 20 (e.g., drame, action, historique, etc). Quelle est l'estimation du nombre d'octets du résultat de la requête suivante ?

SELECT \*

FROM Film

WHERE genre='science-fiction';

- ☒ 13 000      ☐ 250      ☐ 14 000      ☐ Aucune de ces réponses n'est correcte.

**Question 10** Quel est le nombre de plans d'exécution possibles de la requête suivante :  
SELECT \*

FROM Film NATURAL JOIN Roles

WHERE année>2010;

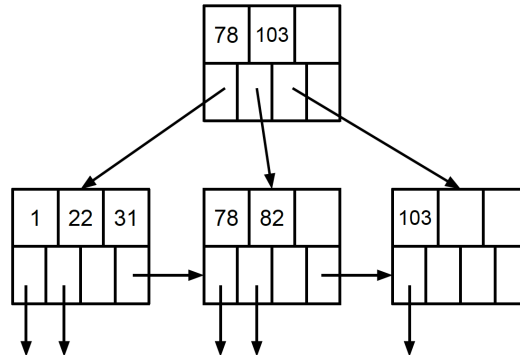
- ☒ 24      ☐ Aucune de ces réponses n'est correcte.      ☐ 12      ☐ 6

**Question 11** Combien de blocs seraient nécessaires pour stocker la relation Roles, si le calcul d'un enregistrement doit être un multiple de 4 octets et l'entête prend 8 octets (la taille d'un enregistrement doit comprendre l'entête) ?

- ☐ Aucune de ces réponses n'est correcte.      ☐ 390      ☐ 781      ☒ 537

## Indexation

**Question 12 ♣** Soit l'arbre B+ suivant.



Choisir la séquence dont l'insertion des clés dans l'ordre donné entraîne l'ajout d'un niveau supplémentaire à l'arbre initial.

- ☒ (26,32,76)   
 ☐ (52,70,85)   
 ☐ (200,70,100,150)  
☐ Aucune de ces réponses n'est correcte.

**Question 13** En supposant qu'un *index clairsemé* (ou non dense) contienne 100 pointeurs et que chaque bloc puisse contenir au plus 20 enregistrements. Combien d'enregistrements peut contenir au plus la relation ?

- ☒ 2000   
 ☐ 100   
 ☐ 200

**Question 14** Soit une taille de bloc de 4096 octets. Les clés sont des entiers de 4 octets et les pointeurs sont des entiers de 8 octets. Combien de clés peut stocker un bloc ?

- ☒ 340   
 ☐ 511   
 ☐ 360   
 ☐ Aucune de ces réponses n'est correcte.

## Transactions

Notation :  $l_{Ti}(x)$  = lecture sur  $x$  de la transaction  $T_i$ ,  $e_{Ti}(x)$  = écriture sur  $x$  de la transaction  $T_i$ .

**Question 15 ♣** Parmi les ordonnacements suivants, cocher ceux qui sont sérialisables ?

- ☐  $l_{T1}(A); e_{T2}(A); e_{T1}(A); l_{T2}(A)$   
☐  $l_{T1}(A); e_{T2}(A); l_{T2}(A); e_{T1}(A)$   
☒  $l_{T1}(A); e_{T1}(A); l_{T2}(A); e_{T2}(A)$   
☒  $l_{T1}(A); e_{T2}(B); l_{T2}(B); e_{T1}(A)$   
☒  $l_{T2}(A); l_{T1}(A); e_{T2}(A); e_{T1}(B)$   
☒  $l_{T1}(A); e_{T2}(A); l_{T2}(A); e_{T1}(B)$   
☐ Aucune de ces réponses n'est correcte.

**Question 16** Soit le protocole de verrouillage à 2 phases (Two-Phase Lock). Soit l'exécution de deux transactions  $T1$  et  $T2$  :  $l_{T1}(A); e_{T1}(A); l_{T2}(B); e_{T2}(B); l_{T1}(B); e_{T1}(B); l_{T2}(A); e_{T2}(A)$ . A quelle opération se produit un deadlock ?

- ☒  $l_{T2}(A)$    
 ☐  $e_{T1}(B)$    
 ☐  $e_{T1}(A)$    
 ☐  $l_{T2}(B)$   
☐ Un deadlock ne se produit pas.

# CORRECTION

**Question 17** Soit le protocole de verrouillage à 2 phases (Two-Phase Lock). Soit l'exécution de deux transactions T1 et T2 :  $l_{T1}(A); e_{T1}(A); l_{T2}(B); e_{T2}(B); l_{T1}(B); e_{T1}(B); l_{T2}(A); e_{T2}(A)$ . A quelle opération s'interrompt T1 ?

- ☒  $l_{T1}(B)$ 
☐  $e_{T1}(B)$ 
☐  $e_{T1}(A)$ 
☐  $l_{T2}(B)$   
☐ T1 ne s'interrompt pas.

**Question 18** Cochez l'ordonnancement qui produit une anomalie de perte de mise à jour.

- ☐  $l_{T1}(A); e_{T1}(A); l_{T2}(A); e_{T2}(A)$ 
☐  $l_{T1}(A); l_{T2}(A); e_{T2}(A); l_{T1}(A)$   
☒  $l_{T1}(A); l_{T2}(A); e_{T2}(A); l_{T1}(B); e_{T1}(A)$ 
☐  $l_{T1}(A); l_{T1}(B); e_{T1}(A); l_{T2}(B); e_{T2}(B)$

**Question 19** Si un ensemble d'opérations de deux transactions T1 et T2 est sérialisable alors :

- ☐ On est sûr de pouvoir exécuter T1 puis T2.  
☐ On est sûr de pouvoir exécuter T2 puis T1.  
☒ On est sûr de pouvoir exécuter (T1 puis T2) OU (T2 puis T1).  
☐ On n'est sûr de rien.

**Question 20 ♣** Quelle propriété parmi les propriétés ACID assurent les mécanismes de tolérance aux pannes ?

- ☐ Cohérence  
☐ Isolation  
☒ Atomicité  
☒ Durabilité  
☐ Aucune de ces réponses n'est correcte.