

# Ing1 - Examen de Bases de Données

## EISTI 2012-2013 - CORRIGE

Lundi 14 janvier 2013 - Durée : 2h

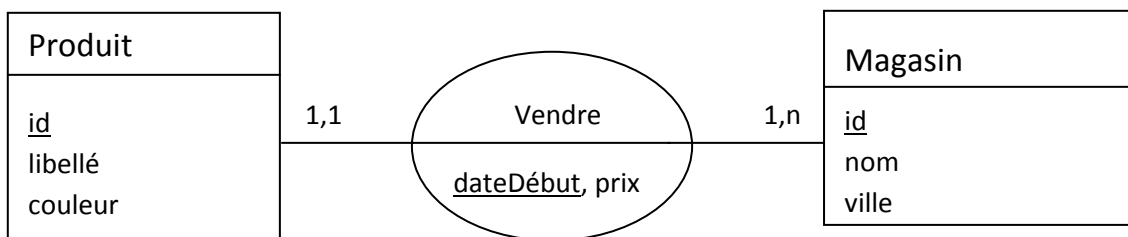
1 feuille A4 recto-verso autorisée. Ordinateur et calculatrice interdits.

### 1<sup>ère</sup> Partie – Questions de cours – 7 points

Toutes les questions sont relatives à un énoncé et indépendantes entre deux énoncés différents.

#### Enoncé 1 :

On considère le MCD suivant :



#### Question 1.

Donner la signification des 4 cardinalités de ce MCD.

#### Réponse (0,5p)

Un produit est vendu par un et un seul magasin.

Un magasin peut vendre plusieurs produits et au moins un.

#### Question 2.

- Est-il possible d'avoir les 2 instances d'associations suivantes :
  - produit d'id 1 vendu au magasin d'id 4 à une dateDébut du 9 janvier 2013 au prix de 15€ ;
  - produit d'id 1 vendu au magasin d'id 4 à une dateDébut du 13 février 2013 au prix de 25€ ?
- Est-il possible d'avoir les 2 instances d'associations suivantes :
  - produit d'id 2 vendu au magasin d'id 5 à une dateDébut du 9 janvier 2013 au prix de 15€ ;
  - produit d'id 2 vendu au magasin d'id 6 à une dateDébut du 9 janvier 2013 au prix de 25€ ?

#### Réponse (0,5p)

- Oui
- Non

#### Question 3.

- Quelle forme normale s'attaque à la problématique suivante : « un attribut dépend de la clé, de toute la clé » ?
- Quelle forme normale s'attaque à la problématique suivante : « un attribut dépend de la clé, rien que la clé » ?

**Réponse (0,5p)**

- a) 2FN
- b) 3FN

**Enoncé 2 :**

On considère R de schéma suivant : R(A, B, C, D, E). On a les dépendances fonctionnelles suivantes :

1.  $A \rightarrow B$
2.  $A \rightarrow C$
3.  $A \rightarrow E$
4.  $F \rightarrow E$
5.  $B, C, D \rightarrow E$
6.  $B, C \rightarrow F$

Toute autre dépendance non triviale ou non obtenue à partir des précédentes est considérée comme fausse.

**Question 4.** Donner le numéro des dépendances élémentaires ?

**Réponse (1p)**

1, 2, 3, 4, 6

**Question 5.** Donner le numéro des dépendances directes ?

**Réponse (1p)**

1, 2, 4, 6

**Enoncé 3 :**

Soit une base de données relationnelle simplifiée de booking.com qui permet de gérer les hôtels et les commentaires de clients sur leurs séjours :

- Hotel(id, nom, nombreEtoiles, adresse, ville, codePostal, pays)
- Client(id, nom, prénom, catégorie, ville, pays)
- Evaluation(*idHotel*, *idClient*, note, commentaire)
- Chambre(id, type)
- Tarif(*idHotel*, *idChambre*, tdate, prix)
- idHotel référence Hotel.id
- idClient référence Client.id
- idChambre référence Chambre.id

Les clés primaires sont soulignées ; les clés étrangères en italique.

**Question 6.** On considère que cette base contient :

- 100 millions d'hôtels dont 100 pour la ville de Biarritz en France.
- 10 milliards d'évaluation répartis de manière homogène par hôtel.

On souhaite estimer le coût des requêtes suivantes en nombre d'enregistrements produits :

1.  $\sigma_{ville='Biarritz'}(Hotel \bowtie_{Hotel.id=idHotel} Evaluation)$
2.  $\sigma_{ville='Biarritz'}(Hotel) \bowtie_{Hotel.id=idHotel} Evaluation$

Quel est le nombre d'enregistrements du résultat pour ces 2 requêtes ?

Quel est le nombre maximum d'enregistrements produits au maximum (en considérant les relations intermédiaires) pour ces 2 requêtes ?

**Réponse (1,5p = 0,5 + 0,5 + 0,5)**

Il y a 100 évaluations par hôtel (10 milliards / 100 millions), soit 100 x 100 à Biarritz. Donc le nombre d'enregistrements du résultat de 2 requêtes est 10000.

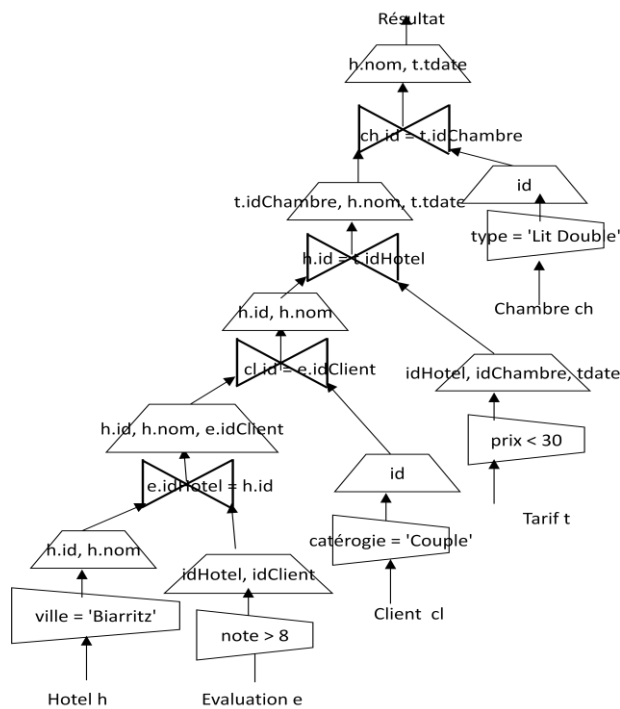
Nombre maximum d'enregistrements produits par la 1<sup>ère</sup> requête : 10 milliards (après la jointure)

Nombre maximum d'enregistrements produits par la 2<sup>ème</sup> requête : 100 (après la restriction)

**Question 7.** Donner un arbre algébrique optimisé en utilisant l'heuristique du cours pour la requête SQL suivante qui donne la liste des hôtels de *Biarritz* ayant été notés au moins 8 par des clients de catégorie *Couple* ainsi que les dates où ces hôtels proposent des chambres avec *Lit Double* pour un tarif inférieur à 30 euros.

```
SELECT h.nom, t.date
FROM Hotel h, Client cl, Evaluation e, Chambre ch, Tarif t
WHERE h.id=e.idHotel
      AND e.note>=8
      AND cl.id=e.idClient
      AND h.id=t.idHotel
      AND t.prix<30
      AND ch.type = 'Lit Double'
      AND h.ville = 'Biarritz'
      AND ch.id=t.idChambre
      AND cl.catégorie = 'Couple';
```

**Réponse (2p)**



## 2<sup>ème</sup> Partie – Requêtes – 13 points

On reprend la base de données relationnelle simplifiée de booking.com (énoncé 3).

**Ecrire chaque requête :**

1. en Algèbre relationnelle (AR) ou en Arbre algébrique (AA) selon demande ;
2. en SQL.

**Attention, vous avez au total 15 réponses.**

1. Trier les hôtels parisiens selon leur nombre d'étoiles en ordre croissant. (AR + SQL)

**Réponse (1p = 0,5 + 0,5)**

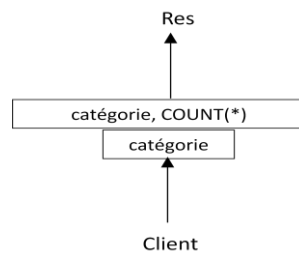
```
SELECT *  
FROM Hotel  
WHERE ville = 'Paris'  
ORDER BY nombreEtoiles ;
```

Res = Tri( $\sigma_{ville = 'Paris'}$ (Hotel), nombreEtoiles $\uparrow$ )

2. Compter le nombre de clients de chaque catégorie. (AA + SQL)

**Réponse (1p = 0,5 + 0,5)**

```
SELECT catégorie, COUNT(*)  
FROM Client  
GROUP BY catégorie ;
```



3. Quel est le prix d'une chambre double de New York Palace le soir du réveillon 31/12/2012 ? (AR + SQL)

**Réponse (2p = 1 + 1)**

```
SELECT prix  
FROM Hotel H, Chambre C, Tarif  
WHERE H.id = idHotel  
AND idChambre = C.id  
AND nom = 'New York Palace'  
AND tdate = to_date('31/12/2012', 'DD/MM/YYYY')  
AND type = 'double' ;
```

$R1 = \sigma_{\text{tdate} = \text{to\_date}('31/12/2012', 'DD/MM/YYYY')}(\text{Tarif})$

$R2 = \sigma_{\text{nom} = \text{'New York Palace'}}(\text{Hotel})$

$R3 = \sigma_{\text{type} = \text{'double'}}(\text{Chambre})$

$R4 = R1 \bowtie_{\text{id} = \text{idHotel}} R2$

$R5 = R4 \bowtie_{\text{idChambre} = R3.\text{id}} R3$

$\text{Res} = \Pi_{\text{prix}}(R5)$

4. Calculer la note moyenne de chaque hôtel. On veut afficher dans le résultat le nom, la ville et le pays de chaque hôtel ainsi que la note moyenne. **(AR + SQL)**

**Réponse (2p = 1 + 1)**

```
SELECT nom, ville, pays, AVG(note)
FROM Hotel, Evaluation
WHERE id = idHotel
GROUP BY idHotel, nom, ville, pays ;
```

$R1 = \pi_{\text{idHotel}} \rho_{\text{idHotel}, \text{AVG}(\text{note})}(\text{Evaluation})$

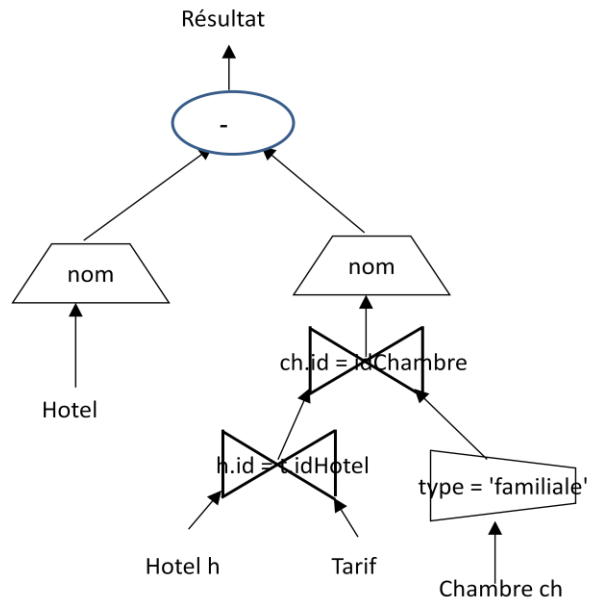
$R2 = R1 \bowtie_{\text{idHotel} = \text{id}} \text{Hotel}$

$\text{Res} = \Pi_{\text{nom, ville, pays, AVG}(\text{note})}(R2)$

5. Donner la liste des hôtels qui n'ont pas de chambre familiale. **(AA + SQL)**

**Réponse (2p = 1 + 1)**

```
SELECT nom
FROM Hotel
MINUS
SELECT nom
FROM Hotel H, Chambre C, Tarif
WHERE H.id = idHotel
AND idChambre = C.id
AND type = 'familiale' ;
```



6. Quel est le client qui a évalué tous les hôtels de la ville Durbuy ? Donner son nom et son prénom.  
**(AR + 1 SQL avec COUNT + 1 SQL avec EXISTS)**

**Réponse (3p = 1 + 1 + 1)**

$R1 = \rho_{idHotel} (\Pi_{id} (\sigma_{ville = 'Durbuy'} (Hotel)))$

$R2 = Evaluation \div R1$

$Res = \Pi_{nom, prenom} (R2 \bowtie_{idClient = id} Client)$

```

SELECT C.nom, prénom
FROM Client C, Hotel H, Evaluation
WHERE C.id = idClient
AND H.id = idHotel
AND ville = 'Durbuy'
GROUP BY idClient, C.nom, prénom
HAVING COUNT(idHotel) = (SELECT COUNT(id)
                          FROM Hotel
                          WHERE ville = 'Durbuy');

```

```

SELECT nom, prénom
FROM Client C
WHERE NOT EXISTS
  (SELECT *
   FROM Hotel H
   WHERE NOT EXISTS
     (SELECT * FROM Evaluation
      WHERE C.id = idClient
      AND H.id = idHotel
      AND H.ville = 'Durbuy'));

```

7. Classer en ordre décroissant les hôtels selon le nombre de commentaires laissés par les clients, y compris les hôtels qui n'ont pas eu d'évaluation. On affichera le nom, la ville et le pays des hôtels ainsi que le nombre d'évaluations. **(AR + SQL)**

**Réponse (2p = 1 + 1)**

```
SELECT nom, ville, pays, COUNT(commentaire)
FROM Hotel LEFT JOIN Evaluation ON id = idHotel
GROUP BY id, nom, ville, pays
ORDER BY COUNT(commentaire) DESC ;
```

R1 = Hotel  Evaluation

R2 = id, nom, ville, pays  $F_{\text{nom, ville, pays, COUNT(commentaire)}}$  (R1)

Res = Tri(R2, COUNT(commentaire) DESC)