

LIF10 – FONDEMENTS DES BASES DE DONNÉES

TD8 – conception & normalisation

Licence informatique – Automne 2013–2014

L'algorithme de synthèse pour les DFs prend en entrée une relation R la relation (universelle) à décomposer et F un ensemble minimal de dépendances sur R . Son principe est le suivant :

1. Construire une couverture canonique, c'est-à-dire une couverture minimum de F , et réduire ses parties gauches et droites.
2. Générer une relations XY pour chaque DF $X \rightarrow Y$;
3. Générer une relations XY' pour chaque DMV $X \twoheadrightarrow Y$ avec $F \models Y' \rightarrow Y$;
4. On supprime les schémas de relation qui ne sont pas maximaux par inclusion.
5. S'il y a perte de jointure, alors on rajoute une relation composée d'une clé de F .

Exercice 1 : modélisation de services chirurgicaux

1 *Un service chirurgical est identifié par son nom, et possède un numéro de téléphone. Il a un chef*
2 *de service unique, lui même identifié par son nom.*

3 *Une opération chirurgicale concerne un patient. Elle sera identifiée dans la suite par sa date, heure*
4 *et la salle d'opération ; elle est pratiquée par un anesthésiste et un ensemble de chirurgiens qui en*
5 *retirent un salaire, et possède une durée.*

6 *Un code d'admission unique est attribué à chaque patient, dont on connaît également le sexe,*
7 *l'âge et le médecin traitant. Notez que, à une date donnée, le patient ne subit qu'une seule*
8 *intervention.*

9 *Tout personnel des services (dont les chirurgiens et anesthésistes) a un nom unique et un numéro*
10 *de téléphone où il peut être joint à tout instant. Certains de ces personnels sont autorisés à*
11 *consulter les données, et possèdent ainsi un mot de passe.*

12 *Enfin, notons qu'un chirurgien appartient à un seul service ; le chef de service est d'ailleurs pris*
13 *parmi ses chirurgiens.*

1. Dresser la liste des dépendances fonctionnelles et multivaluées de l'énoncé précédent.
2. Donner un schéma de base de données normalisé en appliquant l'algorithme de synthèse. Préciser la forme normale du schéma obtenu.
3. Donner un schéma entité-association qui modélise cet énoncé.

Exercice 2 : modélisation d'emplois du temps

Les données sont les suivantes :

- 14 – Un élève possède un numéro d'étudiant unique et un nom.
- 15 – Un groupe possède un code unique et un nom, correspond à un ensemble d'élèves, et est inséré
- 16 dans une formation.
- 17 – Une formation est libellée par un code unique, et se compose d'un ensemble de groupes.
- 18 – Une UE possède un code unique et une description.
- 19 – Un professeur, identifié par son nom, est responsable d'une liste d'UE depuis un certain nombre
- 20 d'années.
- 21 – Une salle possède un code unique, et un nombre de places.
- 22 – un créneau horaire a une heure de départ, une durée, et une date.
- 23 – Enfin, l'association d'une UE, d'un groupes, d'une salle et d'un créneau horaire donne une
- 24 séance de cours, à laquelle on attribue un commentaire.

1. Dresser la liste des dépendances fonctionnelles et mutualuées de l'énoncé précédent.
2. Donner un schéma de base de données normalisé en appliquant l'algorithme de synthèse. Préciser la forme normale du schéma obtenu.
3. Donner un schéma entité-association qui modélise cet énoncé.

Corrections

Solution de l'exercice 1

1. On peut en déduire l'ensemble de DF suivant :

1. $NomService \rightarrow NumTelService, NomChef$, ligne 1–2,
2. $DateOp, HeureOp, SalleOp \rightarrow Anesthésiste, CodePatient, Duree, SalaireAnesthésiste$, ligne 3–5,
3. $DateOp, HeureOp, SalleOp, NomChirurgien \rightarrow SalaireChirurgien$, ligne 4–5,
4. $CodePatient \rightarrow Sexe, Age, MedecinTraitant$, ligne 6–7,
5. $CodePatient, DateOp \rightarrow HeureOp, SalleOp$, ligne 7,
6. $NomPersonnel \rightarrow NumTel$, ligne 9–10,
7. $PersonnelAutorise \rightarrow MotdePasse$, ligne 10–11,
8. $NomChirurgien \rightarrow NomService$, ligne 12,
9. $NomChef \rightarrow NomService$, ligne 12–13,

Pour les dépendances d'inclusions on trouve :

1. $R[NomChirurgien] \subseteq R[NomPersonnel]$, ligne 9,
2. $R[Anesthésiste] \subseteq R[NomPersonnel]$, ligne 9,
3. $R[PersonnelAutorise] \subseteq R[NomPersonnel]$, ligne 10–11,
4. $R[NomService, NomChef] \subseteq R[NomService, NomChirurgien]$, ligne 12–13,

Modélisation avec les DI Il y a une réelle difficulté de modélisation des lignes 12–13 du texte avec les DFs (a) et (i) et la DI (d). On peut remarquer que d'une part préciser simplement $NomService \rightarrow NomChef$, $NomChirurgien \rightarrow NomService$, $NomChef \rightarrow NomService$ et $R[NomChef] \subseteq R[NomChirurgien]$ n'est pas suffisant, car rien n'indique qu'on ne peut être chef que de son service d'affectation, comme le montre l'instance suivante :

	<u>NomService</u>	<u>NomChef</u>	<u>NomChirurgien</u>
t_1	s_1	a	b
t_2	s_2	b	a

D'autre part, en présence de $R[NomChef, NomService] \subseteq R[NomChirurgien, NomService]$ et de la DF $NomChirurgien \rightarrow NomService$ la dépendance $NomChef \rightarrow NomService$ peut être déduite, c'est un cas d'application de la propriété $\{R[XY] \subseteq S[TU], S : T \rightarrow U\} \models R : X \rightarrow Y$ (exercice 1.3 du TD4). C'est donc cette dernière réponse qui paraît être la bonne. . .

2. Je pense qu'il faut leur faire admettre (et vérifier chez eux) que l'ensemble précédent est réduit et qu'on peut donc synthétiser un schéma en forme normale. En construisant une relation par DF, on obtient un ensemble de schémas ; on supprime ceux inclus dans d'autres, on nomme les schémas restant. On souligne les clés. On obtiendra :

- $Service(\underline{NomService}, NumTelService, NomChef)$,
 $NomChef$ comme clé alternative de $Service$,
- $Intervention(\underline{DateOp}, \underline{HeureOp}, \underline{SalleOp}, Anesthésiste, CodePatient, Duree, SalaireAnesthésiste)$,
 $CodePatient, DateOp$ comme clé alternative de $Intervention$.
- $Intervenant(\underline{DateOp}, \underline{HeureOp}, \underline{SalleOp}, \underline{NomChirurgien}, SalaireChirurgien)$
- $Patient(\underline{CodePatient}, Sexe, Age, MedecinTraitant)$
- $Personnel(\underline{NomPersonnel}, NumTel)$
- $PersonnelAutorise(\underline{PersonnelAutorise}, MotdePasse)$
- $Affectation(\underline{NomChirurgien}, \underline{NomService})$

Questions de modélisations

Création d'identifiants on pose la question s'il faut créer des identifiants $IDOperation$, $IDService$, $IDPersonnel$ en bijection avec les clés non-numériques existantes. On peut argumenter les deux points de vue :

non Tout d'abord ce n'est pas demandé dans la modélisation, le concepteur prend ainsi une certaine liberté qui ne modifie pas la sémantique mais alourdi la base. De plus, la multiplication des identifiants ne facilite pas toujours l'utilisabilité du système. C'est une réponse du *point de vue fonctionnel*.

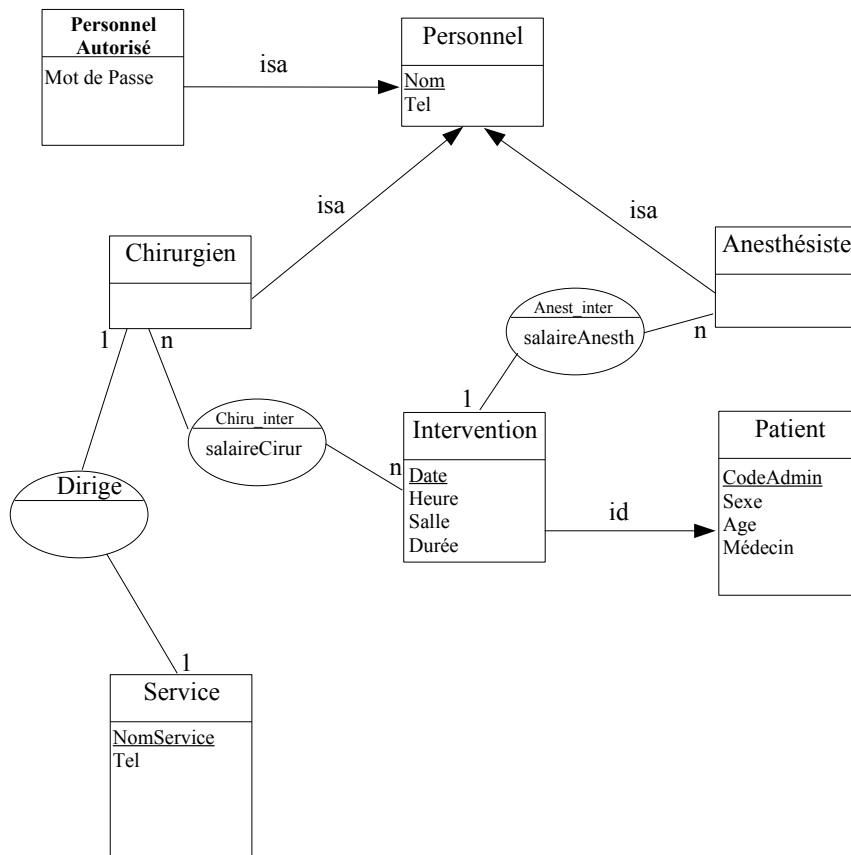
oui L'indexation sur des clés composées de grandes tailles ne sera pas efficace voire très coûteux (exemple, recherche avec LIKE). De plus l'identification sur des chaînes soulève des problèmes de normalisation (exemples, majuscules, fautes de frappes, accents, gestion des réponses multiples d'un LIKE). C'est une réponse du *point de vue technique*.

C'est plutôt le premier point de vue qu'on retient dans cet exercice car on est encore à la conception de la base de données et les éléments en faveur du *oui* ne sont pas encore pertinents à cette étape.

Valeurs nulles Il est assez naturel de fusionner les relations *Personnel* et *PersonnelAutorise* et d'enregistrer un mot de passe si le personnel est autorisé et NULL sinon. Techniquement c'est faisable et sans doute à privilégier. Toutefois les NULL ne sont pas traités par la théorie utilisée, pour les intégrer il faudrait tout enrichir depuis les définitions de bases, mais ce n'est pas nécessairement simple et même faisable. Par exemple, qu'en est-il de la définition de la satisfaction d'une DF en présence de valeurs nulles ?

Clés alternatives La relation *Intervention* comporte deux clés minimales *DateOp, HeureOp, SalleOp* et *CodePatient, DateOp*. Par synthèse on va ainsi obtenir deux schémas dont un sera inclus dans l'autre et donc redondant. En pratique, on utilisera une clé primaire sur *DateOp, HeureOp, SalleOp* et une contrainte d'unicité sur *CodePatient, DateOp*.

3. Voir la figure



Solution de l'exercice 2

Nota bene : ceci est un exemple de réponse qui pourrait être plus amplement discuté comme l'a été l'exercice précédent.

1. 1. DFs
 - NumEtud → NomEtud
 - CodeGroupe → NomGroupe
 - CodeUE → Description
 - NomProf, CodeUE → NbAnnee
 - CodeSalle → NbPlace
 - CodeUE, CodeGroupe, CodeSalle, HeureDep, Duree, Date → Commentaire
2. MVDs

- $CodeGroupe \rightarrow NumEtud, NomEtud$
- $CodeForm \rightarrow CodeGroupe, NomGroupe$

L'ensemble de DF est minimal et réduit. Dire aux étudiants de le vérifier par eux mêmes chez eux.

- En appliquant l'algo, on obtient une relation par dépendance. Enfin, il faut leur faire nommer les relations, et reporter les clés à partir des DF. Attention, pour les DMV, on obtient des relations avec 2 attributs, puisque qu'on ne garde que ce qui est nécessaire de la partie gauche.
- Voir la figure

