Élodie-Jane Sims

6 août 2003

How to prevent interaction of funtionnal and inclusion dependencies

de Mark Levene[1] et Georges Loizou[2]

Je vous présente ici l'article "How to prevent interaction of functionnal and inclusion dependences" présenté par Mark Levene et George Loizou à la revue *Information Processing Letter* une première fois en octobre 1998 et accepté dans sa forme modifiée et définitive en juin 1999. Cet article est donc très récent, et les nombreux mois nécessaires pour sa révision peuvent montrer la difficulté à résoudre le problème traité. En fait nous verrons que ce papier propose une amélioration au problème mais ne le résoud pas complètement.

Résumé

Les dépendances fonctionnelles (FDs) et les inclusions de dépendances (INDs) sont les principales contraintes d'intégrité que l'on rencontre dans la pratique en base de données relationnelle. Un ensemble donné de FDs n'interagit pas avec un ensemble donné de INDs si toute implication logique d'une FD peut être uniquement déterminée par l'ensemble donné des FDs et si toute implication logique d'une IND peut être uniquement déterminée par l'ensemble donné des INDs. Les auteurs de l'article se sont appliqués à donner une condition nécessaire et deux nouvelles conditions suffisantes pour qu'un ensemble de FDs et un ensemble de "proper circular" INDs n'interagissent pas; ces deux conditions suffisantes sont "orthogonales" à des résultats connus dans la littératures des bases de données. L'article évoque aussi de la difficulté d'obtenir une condition nécessaire et suffisante pour éviter les interactions entre FDs et INDs.

1 Introduction

Le *problème d'implication* pour les FDs et INDs consiste en décider pour un ensemble Σ de FDs et INDs si Σ implique σ , où σ est une FD ou IND. Le problème d'implication est central dans la théorie des dépendances de données dans la reconnaissance de sous-classes de dépendances de données utilisables. Les solutions au problème d'implication sont utilisées pour la création de bases de données, elles sont utilisées pour tester si deux ensembles de dépendances sont équivalents ou si une dépendance dans un ensemble est redondante.

Le problème d'implication pour les FDs et INDs est connu pour être indécidable dans le cas général. Donc pour un ensemble Σ de FDs et INDs et une FD ou IND σ , si F les FDs et I les INDs n'interagissent pas le problème si Σ implique σ peut être décidé, par F seulement si σ est une

1. *Department of Computer Science, University College London, Gower Street, London WCIE 6BT, UK*
2. *Department of Computer Science, Birkbeck College, Malet Street, London WCIE 7HX, UK*

FD et par I seulement si σ est une IND.[3] Les auteurs de l'article ont donc cherché des conditions pour lesquelles, il n'y a pas de problème d'interactions.

L'article considère les interaction entre les FDs et les "proper circular" INDs qui sont une extension des INDs non circulaires comprenant des formes restreintes de cycles, appelées "proper cycle". Le principal résultat de l'article donne une condition nécessaire et deux nouvelles conditions suffisantes pour qu'il n'y ait pas d'interaction entre FDs et "proper circular" INDs. Les deux nouvelles conditions suffisantes sont "orthogonales" aux résultats déjà parus sur les interactions entre FDs et INDs non circulaires. Une condition nécessaire et suffisante de ne pas avoir d'interactions entre FDs et "proper circular" INDs n'existe pas actuellement et l'article montre deux façons d'éviter l'apparition de telles interactions. Ces deux façons utilisent le fait que la procédure *chase* peut être utilisée pour résoudre le problème d'implication pour les FDs et les "proper circular" INDs. La première approche, que l'on retrouve dans la littérature, consiste en éviter toute application de "règle FD" lors de l'appel de chase. La deuxième approche consiste en trouver une condition qui empêche toute ancienne valeur de la base de données originale d'être égalée par une autre valeur pendant l'exécution de chase (les nouvelles valeurs ajoutées pendant l'exécution de chase peuvent être égalée par d'autres valeurs). Le principal résultat de l'article empêche les interactions par la seconde approche.

Dans un premier temps l'article donne des définitions de FDs, INDs et leurs sous-classes et définit la procédure chase "comme un ensemble de tests de satisfaction des FDs et INDs". Dans un deuxième temps l'article donne une condition nécessaire pour qu'il n'y ait pas d'interaction entre FDs et "proper circular" INDs ainsi que plusieurs nouvelles conditions suffisantes pour éviter ces interactions. Dans la conclusion de l'article, les auteurs donnent une condition nécessaire et suffisante pour éviter les apparition d'interactions.

2 Définitions et procédure chase

Pour ce qui est des définitions de :
schéma de relation, schéma de base de données, relation, base de données, projection, dépendance fonctionnelle (FD), *FD triviales, p-standard, standard*,satisfaction de FD, inclusion de dépendances (IND), *INDs triviale, p-ary, unaires, circulaire, non circulaire*, satisfactions d'IND, graphe représentant des INDs, clés, surclés, Forme Normale de Boyce-Codd (BCNF), satisfaction d'un ensemble de FDs et INDs dans une base de données, clôture de Σ
l'article utilise les définitions et notations usuelles.
Notations :

- $\mathbf{R}$: schéma de base de données
- d : une base de données sur $\mathbf{R}$

3. L'article rappel que :
- on peut calculer F^+ grâce au système d'axiome de Armstrong
- on peut calculer I^+ grâce au système d'axiome de Casanova et Amaral
- si I est non circulaire, on peut calculer Σ^+ grâce au système d'axiome de Mitchell
- si I est un ensemble de INDs unaires, on peut calculer Σ^+ grâce au système d'axiome de Cosmadakis et al.

- $r \in d$: une relation sur le schéma de relation R
- F : un ensemble de FDs (dépendances fonctionnelles) sur **R**
- I : un ensemble de INDs (inclusions de dépendances) sur **R**
- $\Sigma = F \cup I$
- BCNF : forme normale de Boyce-Codd
- ...

Les principales notions introduites par les auteurs sont : INDs "proper circular" et les propriétés d'intersections et de "split-freeness".

Definition 1 *Un ensemble I d'**INDs** sur **R** est **"proper circular"***

- *soit si il est non circulaire*
- *soit si il existe* m *différents schémas de relations,* R_1, R_2,R_3, ..., $R_{\mathrm{m}} \in \boldsymbol{R}$, *avec* $\mathrm{m} > 1$, *tels que* I *contient les INDs :* $R_1[X_1] \subseteq R_2[Y_2]$, $R_2[X_2] \subseteq R_3[Y_3]$, $R_3[X_3] \subseteq R_4[Y_4]$,..., $R_{\mathrm{m}-1}[X_{\mathrm{m}-1}] \subseteq R_{\mathrm{m}}[Y_{\mathrm{m}}]$, $R_{\mathrm{m}}[X_{\mathrm{m}}] \subseteq R_1[Y_1]$, *et pour tout* $i \in \{1,2,\ ...,\ \mathrm{m}\}$ *on a* $X_i = Y_i$

une telle séquence d'IND est appelée "proper cycle".

Definition 2 ***propriété d'intersection*** *:* F *satisfait la propriété d'intersection si* $\forall F_j \in F$, $\forall A \in R_j$, *si il existe des FDs incomparables* $R_j : X \rightarrow A$, $R_j : Y \rightarrow A \in F_j^+$ *alors* $R_j : X \cap Y \rightarrow A \in F_j^+$

Definition 3 ***propriété "split-freeness"*** *:* F *satisfait la propriété de "split-freeness" si* $\forall F_j \in F$ *si il existe des FDs cycliques,*$R_j : XB \rightarrow A$, $YA \rightarrow B \in F_j^+$ *alors soit* $R_j : Y \rightarrow B \in F_j^+$ *soit* $R_J : (X \cap Y)A \rightarrow B \in F_j^+$

Les auteurs ensuite introduisent une proposition qui dit que la réduction de Σ peut être calculée efficacement par l'utilisation de la solution du problème d'implication pour les FDs.

L'article s'appuie sur l'utilisation de la procédure *chase* qui donne un algorithme forçant une base de données à satisfaire un ensemble de FDs et INDs.

Definition 4 Chase *(d, Σ) est le résultat de l'application des règles ci-dessous sur d aussi longtemps que possible :*

- ***Règle FD*** *: Si* $R_j : X \rightarrow Y \in F_j$ *et* $\exists t_1, t_2 \in r_j$ *tels que* $t_1[X] = t_2[X]$ *et* $t_1[Y] \neq t_2[Y]$ *alors* $\forall A \in Y$ *on change les occurrences dans d de max* $\{\ t_1[A], t_2[A]\ \}$ *par min* $\{\ t_1[A], t_2[A]\ \}$
- ***Règle IND*** *: Si* $R_i[X] \subseteq R_j[Y] \in I$ *et* $\exists t \in r_i$ *tel que* $t[X] \notin \pi_Y(r_j)$ *alors on ajoute le tuple* u *sur* R_j *dans* r_j *avec* u *tel que u[Y] = t[X] et* $\forall A \in R_j - Y$, *u[A] a une valeur plus grande que toutes les valeurs qui apparaissent déjà dans l'état courant de d.*

L'article énonce ces règles mais sans donner une implémentation de l'algorithme et n'explicite pas comment l'appliquer. D'ailleurs les auteurs remarquent qu'en général si on a I circulaire alors la procédure chase ne se termine pas toujours.

Les auteur ensuite énoncent un théorème conséquence de résultats extérieurs qui dit que :
si I est un INDs "proper circular" alors la procédure chase termine et satisfait Σ et la procédure peut être séparée en deux étapes, d'abord on applique les règles IND puis les règles FD .
(théorème 3.2.4)

3 Interactions entre FDs et INDs

Les auteurs redonnent une définition de non interactions *(motivée par l'exemple 3.1)*:

Definition 5 *F et I (sur $\boldsymbol{R}$)n'interagissent pas si :*

- *pour toute FDs α sur $\boldsymbol{R}$ et pour tout $G \subseteq F$*
  $G \cup I \models \alpha$ ssi $G \models \alpha$
- *pour toute INDs β sur $\boldsymbol{R}$ et pour tout $J \subseteq I$*
  $F \cup J \models \beta$ ssi $J \models \beta$

Dans cette partie l'article montre qu'il est nécessaire que Σ (avec I "proper circular") soit réduit pour éviter des interactions entre F et I mais que ce n'est pas suffisant.*(lemme 3.4)* La principale nouvelle condition suffisante de non interaction proposée est que Σ soit réduit et que F satisfasse la propriété d'intersection et la propriété de "split-freeness". *(théorème 3.3)*

Pour démontrer ces conditions les auteurs commencent par énoncer des lemmes :

Tout d'abord, ils donnent une définition de traçage : $t \in r_j$ est **tracé** par $u \in r_i$ si r_i est l'état d'origine de R_i et que t a été introduit par applications successives de règles INDs en commençant par u. *(déf 3.2)* t[A] est tracé par r_i si $t \in r_j$ est tracé par $u \in r_i$ et t[A]=u[B]

Premier lemme : si $d \models F$, R_i est un schéma de relation source respectant I non circulaire.Si pendant l'exécution de chase (d, Σ) lors de l'application de la règle FD on applique $R_j : X \rightarrow A(A \notin X)$ à l'état courant r_j et, soit $t_1[A]$ soit $t_2[A]$ est tracé par r_i , alors Σ est réduit, ou F viole et la propriété d'intersection et la propriété de "split-freeness". *(lemme 3.1)*

Le contre-exemple 3.2 montre le cas où F viole les propriétés d'intersection et de "split-freeness".

Le contre-exemple 3.3 montre que même si Σ^+ est réduite il peut y avoir des interactions entre F et I.

Le lemme 3.2 étend le lemme 2.1 au cas ou R_i est un schéma de relation source cyclique respectant un INDs "proper circular".

Après énoncé de ces lemmes, les auteurs peuvent énoncer leur principal théorème sur la non interaction entre F et I :

Theoreme 1 ***Si I est "proper circular", si F satisfait soit la propriété d'intersection soit la propriété de "split-freeness" et Σ est réduit alors F et I n'interagissent pas.***

(théorème 3.3)

Par la suite les auteurs énoncent et démontrent un lemme : Si F et I n'interagissent pas alors Σ est réduite. *(lemme 3.4)*

Puis ils énoncent une extension au cas "proper circular" de théorèmes déjà connus :
Cor 3.5 : Si F est p-standard et I "proper circular" et "p-ary" alors F et I n'interagissent pas.

Cor 3.6 : Si $\mathbf{R}$ est en BCNF avec I "proper circular" et Σ est réduit alors F et I n'interagissent pas.

Par la suite par le théorème suivant ils montrent qu'avec quelques restrictions "syntactiques" plus fortes sur Σ il suffit alors que Σ soit réduite pour éviter les interactions entre F et I.

Theoreme 2 *Avec I est "proper circular", si $\forall R_i[V] \subseteq R_j[U] \in I et \forall Y \to Z \in F_j$ on a $Y \subsetneq U$ ou si $\forall R_i[V] \subseteq R_j[U] \in I$ et $\forall Y \to Z \in F_j$, avec $B \in Z - Y$, on a $B \notin U$, alors il n'y a pas d'interactions entres F et I.*

(théorème 3.7)

Enfin les auteurs donnent des exemples montrants la limite des théorèmes et corollaires énoncés, qui ne permettent pas de fournir de condition nécessaire et suffisante à la non interaction.

4 Remarques de conclusion

Une condition suffisante pour qu'il n'y ait pas d'interaction, quand I est "proper circular", est encore un problème ouvert. L'article a montré que la réduction de Σ est une condition nécessaire pour éviter les interactions mais non suffisante. Les auteurs signalent qu'ils ont fait des gros progrès vers une solution au problème en identifiant deux nouvelles conditions suffisantes qui sont indépendantes aux résultats déjà parus, i.e Σ est réduit et F satisfait les propriétés d'intersection et de "split-freeness".

Leur principal résultat peut être vu a travers les lemmes 3.1 et 3.2 qui impliquent que, si I est "proper circular", les interactions peuvent être évitées par deux moyens en rajoutant des conditions extérieur à la réduction de Σ.
La première façon est d'ajouter des conditions pour éviter toute application d'une règle FD pendant l'exécution de chase pour une base de données satisfaisant F.
La seconde façon est d'ajouter des conditions qui empêchent l'application de règles FD uniquement dans le cas où une ancienne valeur est égalée par une nouvelle valeur.

Le problème de la caractérisation de non interaction est de trouver une condition nécessaire et suffisante qui couvrent ces deux situations. Les auteurs sont conscients que leurs propositions sont une avancée vers une solution mais qu'il faudra encore d'autres restrictions "syntactiques" sur les INDs afin d'obtenir des conditions nécessaires et suffisantes pour éviter toute interactions.