

• 14 00191 300
copia 1

**Comprobació de restriccions d'integritat
en bases de dades deductives bitemporals**

Carme Martín
Jaume Sistac

Report LSI-94-9-T

 **UPC**
Facultat d'Informàtica
de Barcelona - Biblioteca
16 JUN. 1995

COMPROVACIÓ DE RESTRICCIONS
D'INTEGRITAT EN BASES DE DADES
DEDUCTIVES BITEMPORALS

Carme Martín, Jaume Sistac.
LSI, Març 1994.

INDEX

	Pàg.
1.- Introducció.	2
2.- Bases de dades deductives bitemporals.	3
3.- Mètodes de comprovació de restriccions d'integritat en bases de dades deductives bitemporals.	5
4.- Problemàtica de la incorporació del temps en bases de dades deductives per a la comprovació de la integritat.	11
5.- Conclusions.	12
6.- Bibliografia.	13

1.- Introducció.

Actualment, es pot considerar que les bases de dades relacionals han estat aplicades amb èxit a la indústria, però cada vegada més pateixen un fort increment en complexitat i tamany. Per facilitar el desenvolupament d'aplicacions i proporcionar capacitats de raonament a les bases de dades relacionals caldria integrar la definició dels coneixements compartits pels diferents usuaris. Un SGBD deductiu és un sistema que permet integrar aquests coneixements en les bases de dades, derivant noves informacions a partir de les introduïdes explícitament en la base de dades.

D'altra banda, el temps és una part essencial de la informació sobre la constant evolució del món real. Les dades emmagatzemades necessiten ser interpretades dins del context de temps i per poder obtenir això, haurem de treballar amb bases de dades bitemporals.

Una base de dades deductiva bitemporal ens ofereix els avantatges de la capacitat de raonament d'una base de dades deductiva i l'evolució en el temps d'una base de dades bitemporal.

En bases de dades deductives bitemporals, així com en bases de dades deductives, s'han de satisfer en tot moment les restriccions d'integritat definides per l'usuari.

Una restricció d'integritat d'una base de dades deductiva bitemporal es violada si esdevé certa després de fer una actualització sobre predicats bàsics o predicats derivats, en un cert temps, que succeeix durant el temps definit per la restricció. Per tant, la comprovació de la integritat de la base de dades deductiva bitemporal consistirà en veure si qualsevol actualització cau dins l'àmbit temporal d'una restricció d'integritat, i si això succeeix, si es possible que produeixi una violació d'alguna restricció que porti a la base de dades a un estat inconsistent.

Aquest treball s'inscriu en el context de les bases de dades deductives bitemporals, que descriurem breument en l'apartat següent, podent aprofitar, alhora, aportacions desenvolupades tant en l'àmbit de bases de dades deductives, com en el de bases de dades bitemporals. A continuació, a l'apartat 3, es descriu l'únic mètode per a comprovació de restriccions d'integritat en bases de dades deductives bitemporals, aparegut fins ara.

2.- Bases de dades deductives bitemporals.

Per definir una base de dades deductiva bitemporal cal que veiem, prèviament, què és una base de dades deductiva i què és una base de dades bitemporal.

Bases de dades deductives:

Una base de dades deductiva D , definida, entre d'altres autors, per: [Oli91], [Ten92] i [Urp93], està formada per tres conjunts finits: un conjunt F de fets, un conjunt R de regles de deducció i un conjunt I de restriccions d'integritat.

Regles de deducció:

Les regles de deducció ens permeten derivar noves informacions, anomenades fets derivats, a partir de fets (predicats bàsics) i/o altres fets derivats.

Una regla de deducció és una fórmula del tipus: $A \leftarrow L_1 \wedge \dots \wedge L_n$ amb $n \geq 1$, on A és un àtom i els L_1, \dots, L_n són literals (àtoms o àtoms negats). A, L_1, \dots, L_n estan universalment quantificats davant la fórmula. 'A' rep el nom de conclusió i els L_1, \dots, L_n s'anomenen condicions. Un predicat derivat pot estar definit per una o més d'una regla de deducció.

Restriccions d'integritat:

Una restricció d'integritat és una fórmula tancada de primer ordre que qualsevol estat de la base de dades ha de satisfer. Veiem aquesta fórmula: $\leftarrow L_1 \wedge \dots \wedge L_n$ amb $n \geq 1$, on L_i són literals i totes les variables es suposen universalment quantificades davant la fórmula.

Bases de dades bitemporals:

Una base de dades bitemporal, definida, entre d'altres autors, per [JSS93], [SA86] i [TCG+93] és una base de dades que suporta un temps vàlid i un temps de transacció, on el temps vàlid és el temps en que la informació va ser, és o serà vàlida, i el temps de transacció és el temps en que la informació es donada a conèixer a la base de dades. Una relació bitemporal es pot veure, doncs, com una seqüència d'estats històrics on cadascun d'ells és una relació històrica completa (gràficament representat a la figura 2.1).

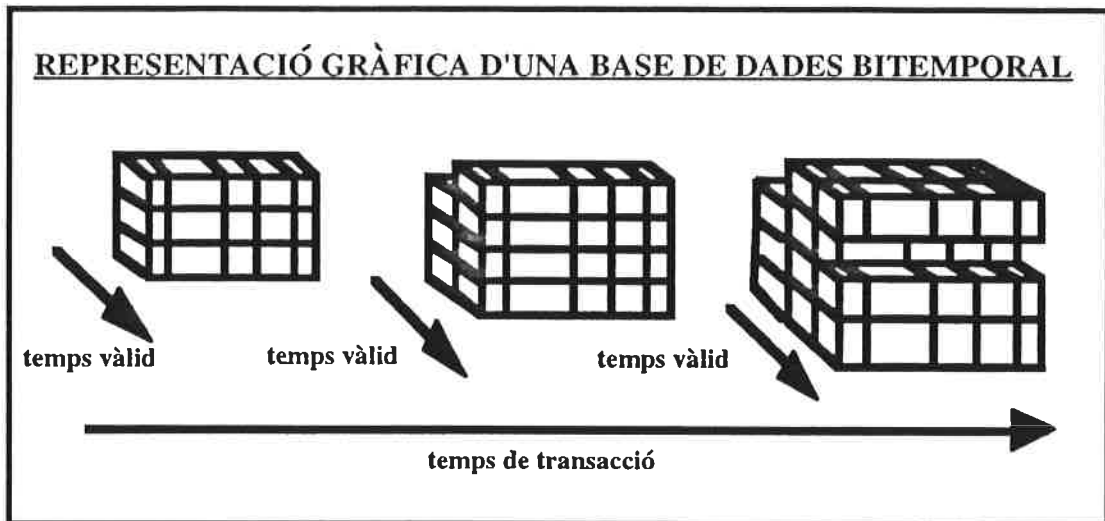


Figura 2.1

Bases de dades deductives bitemporals:

Una base de dades deductiva bitemporal, definida per [Sri88], és una base de dades deductiva que suporta el temps vàlid i el temps de transacció definits prèviament.

Per poder treballar amb el concepte de temps, en bases de dades deductives, necessitem escollir una representació bitemporal [JSS93] i una sèrie d'operadors que ens permetin treballar amb aquesta representació [All83].

3.- Mètodes de comprovació de restriccions d'integritat en bases de dades deductives bitemporals.

Els treballs realitzats en comprovació d'integritat en bases de dades deductives bitemporals incorporen temps vàlid i temps de transacció a una base de dades deductiva obtenint una base de dades deductiva bitemporal i desenvolupen un mètode de comprovació de restriccions d'integritat aplicables a aquest tipus de bases de dades.

Sobre el tema de la comprovació de la integritat de bases de dades temporals o deductives podem trobar diversos mètodes, però en el camp de les bases de dades deductives bitemporals no han aparegut gaires mètodes que tractin el tema de la comprovació de la integritat. L'únic mètode que, explícitament, tracta aquest tema és: [Ple93], però hi ha d'altres treballs, que no veurem en aquest apartat, que faciliten la comprovació de la integritat en bases de dades deductives bitemporals, com el marc lògic que ens proveeix [Sri88] basat en l'Event Calculus de [KS86] o el mètode de [CI88] que només incorpora temps vàlid o el mètode de [HS90] que només incorpora temps de transacció.

Mètode d'en Plexousakis [Ple93]:

Aquest mètode utilitza com a llenguatge de representació TELOS, definit a [MBJ+90].

El mètode, es basa en trobar una *estructura simplificada parametritzada* per a cada literal que formi part d'una restricció d'integritat o d'una regla de deducció. L'estructura simplificada parametritzada té les components que es mostren a la figura 3.1. i que s'expliquen a continuació:

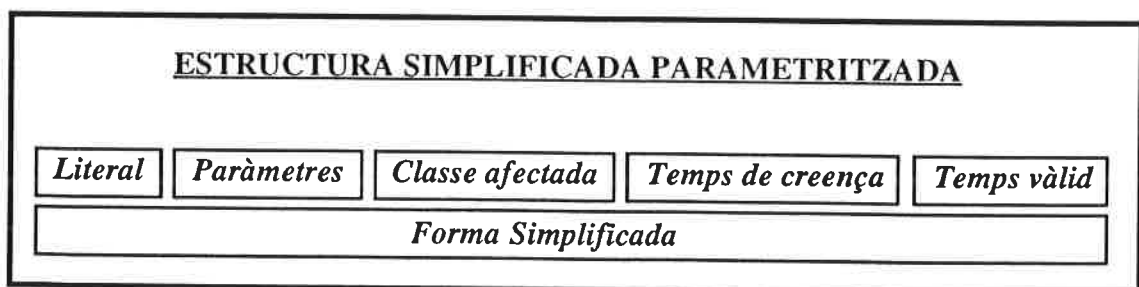


Figura 3.1

Literal és el literal de la restricció pel que es genera l'estructura simplificada parametritzada.

Paràmetres és la llista de variables que es poden instanciar.

Temps històric i temps de creença són el temps vàlid i el temps de transacció associats a la restricció d'integritat.

Classe afectada per un literal L és la classe C tal que, inserint o esborrant una instància de C podem afectar la certesa de L.

Per trobar la *classe afectada* d'un literal podem fer servir les següents metaregles:

- Literals d'instanciació:

Per cada literal de la forma *instanceOf* ($x, y, t1, t2$), si y està instanciada llavors y és la classe afectada, en altre cas la classe afectada és la classe *instanceOf*.

- Literals de generalització:

Per cada literal de la forma *isA* ($x, y, t1, t2$), on x i y representen classes, la classe afectada és la classe *isA*.

- Literals d'atributs:

Per cada literal de la forma *att* ($x, y, t1, t2$) on att és un atribut de la classe x, si x i y no estan instanciades llavors la classe afectada és la classe atribut Q amb components: *from* (Q) = X, *to* (Q) = Y, *label* (Q) = att i *when*(Q) = T, on x és una instància de X en t1, y és una instància de Y en t1 i ambdues es coneixen en t2.

- Literals de proposició:

Per cada literal de la forma *prop* (p, x, y, z, t), si les components x i z són iguals llavors la classe afectada és la classe *Individual*, si no la classe afectada és la classe *Attribute*. Si x i z no estan instanciades llavors la classe afectada és la classe *Proposition*.

Per a obtenir la *forma simplificada* d'un literal cal seguir els següents passos:

- Pas1: Eliminar els quantificadors lligats a les variables instanciades.
Les variables instanciades esdevenen paràmetres.
- Pas2: Les variables temporals es restringeixen amb respecte al temps històric i de creença de la restricció i el resultat obtingut s'afegeix a la restricció, mitjançant el predicat *during*.
- Pas3: Els literals que inserten (o esborren) una tupla es substitueixen per la constant booleana cert (o fals), ja que després de l'actualització es coneix que el fet expressat pel literal és cert (o fals).
- Pas4: S'apliquen les regles d'absorció de la figura 3.2, si es poden aplicar.
- Pas5: S'apliquen les regles de simplificació temporal, si es poden aplicar. Les regles de simplificació temporal s'apliquen, mitjançant una taula temporal definida a [Ple93], a conjuncions del tipus:

$$\text{during}(t, i_1) \wedge r_1(t, i_2) \wedge r_2(i_1, i_2)$$

Podem veure un exemple d'aplicació d'aquest pas a la figura 3.3

REGLES D'ABSORCIÓ:			
$\emptyset \wedge T = \emptyset$	$\emptyset \wedge F = F$	$\emptyset \Rightarrow T = T$	$\emptyset \Rightarrow F = \neg \emptyset$
$\emptyset \vee T = T$	$\emptyset \vee F = \emptyset$	$T \Rightarrow \emptyset = \emptyset$	$F \Rightarrow \emptyset = T$
$\neg T = F$	$\neg F = T$	$\emptyset \Leftrightarrow T = \emptyset$	$\emptyset \Leftrightarrow F = \neg \emptyset$

Figura 3.2

Un cop trobada l'estructura simplificada parametritzada haurem de construir un graf, on s'expressin les dependències existents entre les restriccions d'integritat i les regles de deducció, per a fer eficient la comprovació d'integritat.

El graf de dependències representa com els fets derivats implícits de les regles de deducció poden afectar la integritat de la base de dades.

Els nodes del graf són les estructures simplifiades parametritzades que hem calculat abans.

Les *fletxes* del graf denoten la dependència de restriccions a regles. Si una restricció C depèn directament d'una regla R hi posem una fletxa de R a C.

EXEMPLE D'APLICACIÓ DE SIMPLIFICACIONS TEMPORALS:

Aplicarem simplificació temporal a la conjunció:

during (t, 1/88..9/88) \wedge before (t, 5/88..12/88) \wedge
overlaps (1/88..9/88, 5/88..12/88)

Recordem que apliquem la simplificació a conjuncions del tipus:

during (t,i1) \wedge r1 (t,i2) \wedge r2 (i1,i2)

Si unifiquem obtenim:

r1 = before

r2 = overlaps

i1 = 1/88..9/88

i2 = 5/88..12/88

Cerquem a la taula, la intersecció de before i overlaps i obtenim: during (t, i1-i1*i2)

i1*i2 = 5/88..9/88

i1-i1*i2 = 1/88..5/88

Per tant el resultat és:

during (t, 1/88..5/88)

Gràficament:

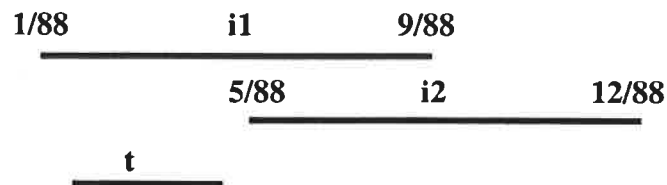


Figura 3.3

Exemple:

Veiem un exemple amb dues restriccions d'integritat IC1, IC2 i dues regles de deducció DR1, DR2:

Regles de deducció:

DR1: Una persona vinculada a la universitat treballa en el departament que té a mateixa adreça que ell.

$$\forall d / \text{Departament } \forall u / \text{University } \forall s / \text{String } \forall t1, t2 / \text{TimeInterval} \\ (\text{univ} (d, u, t1, t2) \wedge \text{location} (u, d, t1, t2) \Rightarrow \text{d_addr} (d, s, t1, t2))$$

DR2: Un departament d'una universitat té la mateixa adreça que consta a la localització de la universitat.

$$\forall u / \text{UnivAffiliate } \forall d / \text{Department } \forall s, s' / \text{String } \forall t1, t2 / \text{TimeInterval} \\ (\text{address} (u, s, t1, t2) \wedge \text{d_addr} (d, s', t1, t2) \wedge (s=s' [\text{at } t1, \text{believed at } t2]) \\ \Rightarrow \text{works_in} (u, d, t1, t2))$$

Restriccions d'integritat:

IC1: Un autor d'un article no pot ser revisor d'aquest mateix article.

$$\forall c / \text{ConfPaper } \forall r / \text{Referee } \forall a / \text{Author } \forall t1, t2 / \text{TimeInterval} \\ (\text{ref} (c, r, t1, t2) \wedge \text{author} (c, a, t1, t2) \Rightarrow (r \neq a [\text{at } t1, \text{believed at } t2]))$$

IC2: Un autor no pot presentar un article en una conferència organitzada pel departament a on hi treballa.

$$\forall c / \text{Conference } \forall p / \text{ConfPaper } \forall a / \text{Author } \forall d / \text{Department} \\ \forall t1, t2 / \text{TimeInterval} \\ (\text{submitted_to} (p, c, t1, t2) \wedge \text{organized_by} (c, d, t1, t2) \wedge \text{author} (p, a, t1, t2) \\ \wedge \text{works_in} (a, d, t1, t2) \Rightarrow \text{False})$$

A la figura 3.4, hi ha un exemple de construcció de l'estructura simplificada parametritzada per IC1. A la figura 3.5 trobem el graf de dependències, on cada node contindrà les estructures simplifiades parametritzades d'IC1, IC2, DR1 i DR2. A partir, d'aquest graf, podem veure si qualsevol actualització afecta a la integritat de la base de dades deductiva bitemporal.

EXEMPLE D'APLICACIÓ DEL MÈTODE DE CONSTRUCCIÓ DE L'ESTRUCTURA SIMPLIFICADA PARAMETRITZADA.

Agafem IC1:

$\forall c / \text{ConfPaper} \quad \forall r / \text{Referee} \quad \forall a / \text{Author} \quad \forall t1, t2 / \text{TimeInterval}$
 $(\text{ref}(c, r, t1, t2) \wedge \text{author}(c, a, t1, t2) \Rightarrow (r \neq a [\text{at } t1, \text{believed at } t2]))$

amb $T = (1/1/88..*)$ i $T' = (2/1/88..*)$

Suposem que cerquem l'estructura simplificada parametritzada de la restricció IC1 pel literal de la restricció *author* :

Literal: author

Parametres: c, a, t1, t2

Classe afectada: Es la classe atribut definida per una proposició *P* amb components:

from(*P*) = Paper , *to*(*P*) = Author ,

label(*P*) = author , *when*(*P*) = t

Gràficament:



Temps històric: T

Temps de creença: T'

Figura 3.4

GRAF DE DEPENDÈNCIES DE L'EXEMPLE

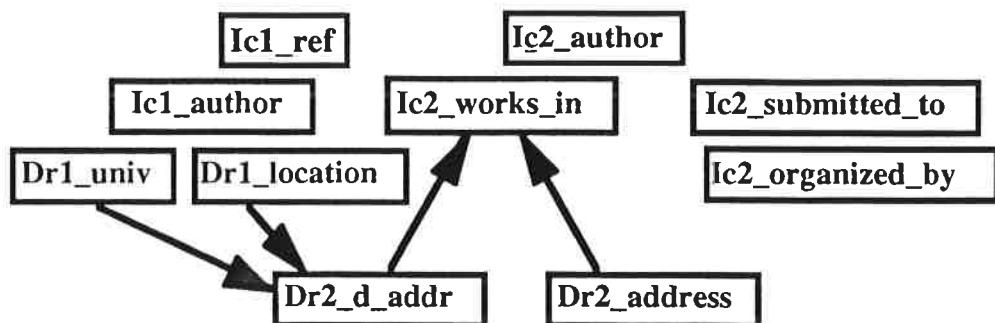


Figura 3.5

4.- Problemàtica de la incorporació de temps en bases de dades deductives per a la comprovació de la integritat.

Tenint en compte que només hem vist un mètode que tracti la comprovació de la integritat en bases de dades deductives bitemporals, ràpidament, ens ve a la ment, la següent pregunta: Quins problemes hi ha en fer comprovació d'integritat en bases de dades deductives bitemporals? En aquest apartat intentarem respondre aquesta pregunta, agrupant els problemes en tres categories:

❶ *Explosió de les dades a tractar.*

En afegir temps a una base de dades deductiva, obtenim una gran quantitat d'informació que hem d'emmagatzemar, per sempre. Això, vol dir que els mètodes que tractin la comprovació d'integritat en bases de dades deductives bitemporals han d'estar preparats per tractar una gran quantitat de dades (moltes més que altres tipus de bases de dades), en el menor temps possible.

❷ *Complexitat de les operacions a realitzar.*

Quan incorporem temps a una base de dades deductiva estem afegint operacions addicionals que ens permetin treballar amb aquest temps i que per tant, compliquen força la comprovació de la integritat en bases de dades deductives, sobretot tenint en compte que hem d'estar disposats a treballar amb temps que, potser no coneixem, tenen diferents granularitats de temps, etc.

❸ *Trobar prou potència de representació.*

Per poder treballar amb temps en una base de dades deductiva necessitem utilitzar un bon llenguatge de representació. Molts dels bons mètodes de comprovació d'integritat en bases de dades deductives no s'han adaptat a bases de dades bitemporals degut a la dificultat de trobar un llenguatge de representació adequat.

5.- Conclusions.

Les bases de dades temporals són necessàries per modelitzar la realitat de la forma més fidel possible, i dins d'aquestes les bases de dades deductives bitemporals són les que ens permeten interpretar millor les dades dins del context del temps amb capacitats de deducció.

En bases de dades deductives bitemporals hi ha un gran camp obert d'investigació que rebrà un gran impuls en els propers anys. Concretament, en el tema de comprovació de les restriccions d'integritat en bases de dades deductives bitemporals, tenint en compte que en aquests moments existeix només un mètode que intenta tractar aquest problema i que es poden aprofitar moltes investigacions realitzades en l'àmbit de bases de dades deductives i bases de dades temporals, estem segurs que, properament, surgiran nous mètodes que aportin més llum per a resoldre aquesta problemàtica.

6.- Bibliografia.

- [All83] Allen, J.F. "Maintaining knowledge about temporal intervals". Communications of the ACM. 1983. Volum 26. Num 11. Pàgs 832-843.
- [CI88] Chomicki, J.; Imielinski, T. "Temporal deductive databases and infinite objects". ACM SIGACT-SIGMOD-SIGART. Symposium on principles of database systems. Març 1988. Pàgs 61-73.
- [HS90] Hülsmann, K.; Saake, G. "Representation of the historical information necessary for temporal integrity monitoring". Proc. of the 2nd. Int. Conf. on EDBT'90. Pàgs 378-392.
- [JSS93] Jensen, C.S.; Soo, M.D.; Snodgrass, R.T. "Unification of temporal data models". Proc. of the IEEE Int. Conf. on Data Engineering'93. Pàgs 262-271.
- [Kou93] Koubarakis, M. "Representation and querying in temporal databases: the power of temporal constraints". IEEE. 1993. Pàgs 328-334.
- [KS86] Kowalski, R. A.; Sergot, M. J. "A logic-based calculus of events". New generation computing. Número 4. 1986. Pàgs 67-95.
- [MBJ+90] Mylopoulos, J.; Borgida, A.; Jarke, M.; Koubarakis, M. "Telos: Representing knowledge about information systems". ACM Transactions on information systems. Volum 8. Número 4. 1990. Pàgs 324-362.
- [Oli91] Olivé, A. "Integrity constraints checking in deductive databases". Proc. of the 17th. Int. Conf. on VLDB'91. Pàgs 513-523.
- [Ple93] Plexousakis, D. "Integrity constraint and rule maintenance in temporal deductive knowledge bases". Proc. of the 19th. Int. Conf. on VLDB'93. Pàgs 146-157.

- [SA86] Snodgrass, R.; Ahn, I. "Temporal databases". IEEE Computer. Setembre, 1986. Pàgs 35-42.
- [Sri88] Sripada, S.M. "A logical framework for temporal deductive databases". Proc. of the 14th. Int. Conf. on VLDB'88. Pàgs 171-182.
- [Ten92] Teniente, E. "El mètode dels esdeveniments per a l'actualització de vistes en bases de dades deductives". Tesi doctoral. U.P.C. 1992.
- [TCG+93] Tansel, A.U.; Clifford, J.; Gadia, S.; Jajodia, S.; Segev, A.; Snodgrass, R. "Temporal databases: theory, design and implementation". Benjamin/Cummings. 1993.
- [Urp93] Urpí, T. "El mètode dels esdeveniments interns per al càlcul de canvis en bases de dades deductives". Tesi doctoral. U.P.C. 1993.