

Dependencias de Comparación de Conjuntos de Valores: Un Enfoque Metodológico para la Representación de Reglas del Negocio

Viviana E. Ferraggine¹
Laura C. Rivero^{1,2}

¹ INTIA, Facultad de Ciencias Exactas, U.N.C.P.B.A.
Paraje Arroyo Seco, Campus Universitario
Tandil, Buenos Aires, Argentina
² LINTI, Facultad de Informática, U.N.L.P.
La Plata, Buenos Aires, Argentina
{vferra, lrivero} at exa.unicen.edu.ar

ABSTRACT

Se presenta un marco metodológico para identificar y formalizar patrones de reglas del negocio que puedan ser formuladas en base a dependencias de comparación de valores y que puedan plasmarse en un MER convencional utilizando las construcciones básicas universalmente aceptadas. Para aquellos casos en los que esto no es posible, se sugieren formas alternativas de especificación, sin pérdida de carga semántica. Las dependencias de comparación de conjuntos de valores es una clase que agrupa las dependencias de inclusión y de exclusión (primitivas) y las dependencias de igualdad y de superposición parcial (derivadas). Este trabajo se centra en el caso particular de las dependencias de igualdad en su condición de estructuras semánticas que permiten representar una nutrida variedad de reglas del negocio.

Palabras clave: Diseño conceptual, dependencias de comparación de conjuntos de valores, dependencias de inclusión, dependencias de igualdad, dependencias de exclusión, dependencias de superposición parcial, reglas del negocio.

1. INTRODUCCIÓN

Las reglas del negocio son componentes importantes de cualquier sistema de información. Dada su importancia práctica, es necesario su tratamiento consistente, a la vez que resulta sumamente útil contar con un criterio de clasificación.

La captura de *reglas del negocio* mediante el modelado de los datos como parte de la definición de las entidades, las relaciones, sus cardinalidades y los atributos no es una tarea sencilla, ya que éstas son generalmente complejas, tienen una estructura arbitraria, cambian frecuentemente y son difíciles de forzar.

El esquema o modelo conceptual de una base de datos relacional generalmente está descrito por una de las extensiones del muy difundido Modelo de Entidades y Relaciones (*MER*) [4], [5]. El *MER* está limitado en su poder expresivo, limitación que paradójicamente ha contribuido a su amplia aplicabilidad, a su formalización y simplicidad de uso [3]. De todas maneras es lo suficientemente flexible como para permitir la especificación de una amplia variedad de *reglas del negocio*.

Las dependencias entre datos son restricciones en las que valores en tuplas de relaciones en una base de datos dependen de igualdades o desigualdades respecto de otros valores [13]. Estas han sido tratadas, formalizadas y estudiadas, con diferente grado de profundidad y desde diferentes enfoques.

En el modelo relacional las dependencias de datos

primitivas son las dependencias funcionales y las *dependencias de inclusión (di)*. Las *di* tienen una versión opuesta en las *dependencias de exclusión (de)* [2], [10], [21]. Puesto que siempre que existan conjuntos de atributos compatibles y semánticamente vinculados es pertinente efectuar comparaciones entre ellos, puede definirse una clase que los englobe: *dependencias de comparación de conjuntos de valores (dccc)*. Esta clase incluye también las *dependencias de igualdad (dig)*, derivadas de las *dis* y las *dependencias de superposición parcial (dsp)*. Esta última dependencia no ha sido tratada como tal en la literatura especializada, aunque se ha analizado el problema de la superposición parcial en [20] bajo el nombre de correlación de datos, en el marco de un proceso de reingeniería de bases de datos.

El objeto de este trabajo es analizar las *digs* en función de las *dis*, enfatizando sus aspectos formales y semánticos. Estas dependencias permiten representar un variado conjunto de *reglas del negocio* elicitadas del Universo del Discurso (UdeD).

Con base en el estudio de diversos casos reales y en la aplicación de una heurística de transformación de esquemas no estructurados, definidos en [15] y [16], se pretende construir un catálogo de patrones, proponiendo formas alternativas de representación. El análisis de las *des* y *dsps* no se incluye en este trabajo.

Este artículo está organizado de la siguiente manera: la Sección 2 presenta la formalización de los conceptos involucrados; la Sección 3 resume el marco metodológico propuesto e ilustra un caso de motivación; la Sección 4 presenta el análisis de algunos casos particulares que ponen de manifiesto los problemas típicos, mostrando las transformaciones que permiten simplificar el esquema y obtener patrones de especificación; finalmente la Sección 5 contiene conclusiones, resultados esperados y trabajos futuros.

2. FORMALIZACIÓN DE LAS DCCV

Un esquema relacional queda definido por $\mathbf{R} = \langle \mathbf{R}, \mathbf{D} \rangle$, donde $\mathbf{R} = \{R_1, R_2, \dots, R_n\}$ representa el conjunto de esquemas de relaciones y \mathbf{D} el conjunto de restricciones. \mathbf{D} incluye las dependencias funcionales, *dis* [6], [7] *des* [1], *digs* y *dsps*. Subconjuntos de atributos de R_i se denotarán con letras del final del alfabeto, las tuplas se indicarán con t, t' y las subtuplas de t correspondientes a los atributos W , mediante $t.W$; r_i indica el estado o extensión de R_i ; y $r_i.W$ indica la proyección de r_i sobre las columnas W .

A continuación se presentan las definiciones de cada restricción.

Dependencias de Inclusión (*di*)

Una *dependencia de inclusión (di)* se define como la

existencia de atributos (simples o compuestos) en una relación, cuyos valores deben ser un subconjunto de los valores de atributos compatibles (simples o compuestos, respectivamente) en otra relación. Algunos autores solo exponen esta definición [10] [11], otros indican la posibilidad de que ambas relaciones puedan ser la misma [7] [6]. Formalmente una *di* es una expresión $R_i.W \overset{I}{=} R_d.Y$ donde R_i y R_d son nombres de relaciones (posiblemente el mismo); W e Y son conjuntos de atributos compatibles. La Figura 1 ilustra la representación gráfica de *dis* utilizada en las siguientes secciones. Cuando Y es la clave primaria (K_d) de R_d , la *di* es basada-en-clave y se denomina Restricción de Integridad Referencial (*rir*) o simplemente referencia. En este caso W constituye una clave extranjera (FK_i) para R_i . La *rir* se establece como $R_i.FK_i \ll R_d.K_d$. Las *rir*s pueden ser especificadas declarativamente mediante la cláusula *foreign key* en SQL y ser forzadas por la mayoría de los sistemas de bases de datos relacionales y posrelacionales actuales, mientras que las *dis* deben especificarse usualmente a través de cláusulas *check* con la semántica de aserciones. En adelante se usará *di* sólo para las dependencias de inclusión puras y *rir* para las basadas en claves o referencias.

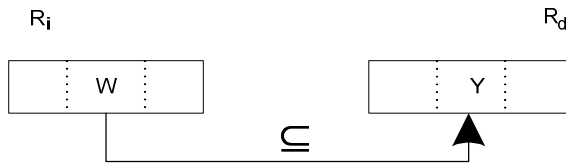


Fig. 1: Notación gráfica para *dis*

Dependencias de Igualdad (*dig*)

Una *dependencia de igualdad (dig)* resulta de establecer sendas dependencias de inclusión inversas. Esto implica la existencia de conjuntos de atributos en una relación, cuyos valores deben ser los mismos que los correspondientes al conjunto de atributos compatibles, en otra relación (o la misma). Una *dig* denotada como $R_k.W = R_l.Y$ es equivalente a: $R_k.W \subseteq R_l.Y \wedge R_l.Y \subseteq R_k.W$, siendo W e Y conjuntos de atributos compatibles. No corresponde definir términos izquierdo y derecho de la dependencia, dado que la igualdad se da en ambos sentidos.

Dependencias de Exclusión (*de*)

En simetría con la definición de una *di*, una *dependencia de exclusión (de)* se define como la existencia de atributos (simples o compuestos) en una relación, cuyos valores son mutuamente excluyentes respecto del conjunto de valores de atributos compatibles (simples o compuestos respectivamente) en otra relación. Ambas relaciones pueden ser la misma [2].

Formalmente una *de* es una expresión $R_k.W \parallel R_l.Y$ donde R_k y R_l son nombres de relaciones (posiblemente el mismo); W y Y son conjuntos de atributos compatibles. También las *des* se especifican usualmente a través de aserciones mediante cláusulas *check* de SQL.

Dependencias de Superposición Parcial (*dsp*)

Las *dependencias de superposición parcial (dsp)* tienen un significado equivalente a: $R_k.W \not\subseteq R_l.Y \wedge R_l.Y \not\subseteq R_k.W \wedge R_k.W \not\parallel R_l.Y$, siendo W e Y conjuntos de atributos compatibles. Esto significa que $r_k.W \cap r_l.Y \neq \emptyset \wedge r_k.W - r_l.Y \neq \emptyset \wedge r_l.Y - r_k.W \neq \emptyset$. Se denotarán $R_k.W \S R_l.Y$.

Estructura sintáctica de las *dcv*

En los últimos años la investigación de las *dis* ha categorizado muchos de sus aspectos. Un marco de referencia para su estudio puede encontrarse en [6] y [13], como así también el estudio de su caso particular, las *rir*s [15], [16] y [17]. Respecto de las *dis*, un análisis de su posible origen desde una perspectiva semántica se puede encontrar en [14] y un estudio preliminar de las *dis* puede encontrarse en [9].

Para el análisis y la comprensión semántica de las *dcvs* se ha seguido lo expuesto en [15] y [16], examinando la estructura sintáctica de los conjuntos de atributos involucrados independientemente del dominio de aplicación, y tomando en cuenta las diferentes ubicaciones de ambos términos de la dependencia, en relación con la disposición de la clave primaria.

Si W es un conjunto de atributos cualesquiera de R , K la clave primaria de R y Z un subconjunto de atributos secundarios, los cinco tipos o correlaciones posibles se ilustran en la Figura 2: I) $W \equiv K$; II) $W \equiv Z$; III) $W \equiv K^1$, siendo K^1 un subconjunto estricto de K , $K^1 \neq \emptyset$; IV) $W \equiv K \cup Z$; y finalmente V) $W \equiv K^1 \cup Z$, $K^1 \neq \emptyset$. En los casos II, IV y V, $Z \neq \emptyset$.

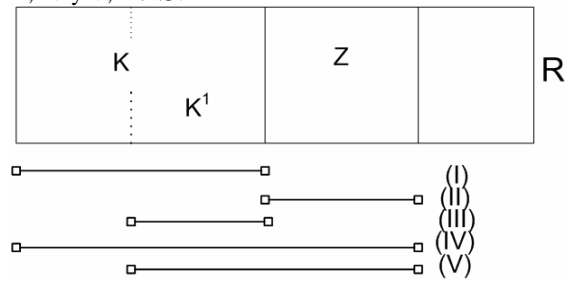


Fig. 2: Correlación de claves primarias y extranjeras

Teniendo en cuenta estas disposiciones, pueden obtenerse 25 conformaciones posibles para cada *dcv*, aunque en el caso de las *digs* y *des* la simetría reduce este número.

3. MARCO METODOLÓGICO: ANÁLISIS, TRANSFORMACIÓN Y ESPECIFICACIÓN

En esta sección se propone un análisis metodológico de las *dcvs*, basado en el estudio de la estructura sintáctica de los términos de las dependencias y su influencia en la posibilidad de aplicar transformaciones y detectar redundancias. Este análisis apunta a descubrir construcciones que sean representables en un *MER*, favoreciendo el planteo correcto y semánticamente enriquecido del esquema conceptual [15] [16]. En otros casos se busca identificar las *reglas del negocio* que motivan determinada especificación con el objeto de clasificarla y luego utilizar un patrón de especificación más adecuado, obteniendo de esta forma una documentación más precisa y homogénea.

Luego de analizar sucintamente un ejemplo a modo de motivación, en la Sección 4 se desarrolla un análisis de algunos casos de *digs*. Es importante destacar que las gráficas sólo ilustran los términos o atributos involucrados en la definición de una regla, no constituyendo una propuesta de extensión del *MER*. Esto es natural, puesto que estas dependencias, salvo que sean *rir*s, no son representables con los instrumentos que provee el modelo.

Ejemplo de motivación

Sean las tablas del esquema y sus correspondientes

restricciones de integridad referencial:

R_i : PROFESOR (id_P , id_D , NombreP)

R_k : CURSO (id_C , id_P , id_D Descripción)

DEPARTAMENTO (id_D , NombreD)

CURSO. id_P << PROFESOR. id_P

CURSO. id_D << DEPARTAMENTO. id_D

PROFESOR. id_D << DEPARTAMENTO. id_D

Adicionalmente las reglas del negocio en lenguaje natural:
Regla 1: Un profesor puede dictar un curso sólo si pertenece al mismo departamento que ofrece el curso.

Regla 2: el profesor pertenece a un departamento sólo si está dictando al menos un curso para ese departamento.

Su especificación algebraica a través de un par de *dis*, establece que los valores de los atributos id_P , id_D en la relación CURSO deben ser los mismos que los correspondientes al conjunto de atributos compatibles id_P , id_D , en la relación PROFESOR:

Curso (id_P , id_D) = Profesor (id_P , id_D)

Del análisis de la *dig* surge que:

Dependencias de inclusión Tipo IV: Curso (id_P , id_D) \subseteq Profesor (id_P , id_D) (di_1)

Dependencias de inclusión Tipo II: Profesor (id_P , id_D) \subseteq Curso (id_P , id_D) (di_2)

La Figura 3 esquematiza la situación presentando las dos tablas involucradas y los atributos que conforman los términos de las dependencias.

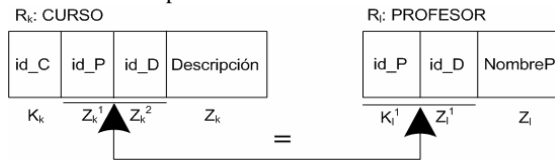


Fig. 3: Esquema de tablas y atributos involucrados

El primer paso consiste en determinar si existe alguna transformación de las *dis* que permita obtener un patrón simplificado representable mediante los mecanismos de abstracción estándar del MER, y si existe algún vínculo semántico redundante.

Si se resuelve la di_1 por la regla de pullback [13] se tiene que PROFESOR. id_P \rightarrow PROFESOR. id_D , entonces CURSO. id_C \rightarrow CURSO. id_D es una dependencia transitiva. Así en CURSO sólo quedaría CURSO. id_C \rightarrow CURSO. id_P y di_1 se transforma en la *rir*:

CURSO(id_P) << PROFESOR(id_P)

Esta transformación reduce di_2 a PROFESOR (id_P) \subseteq CURSO (id_P) (especificación), que continúa siendo de Tipo II pero con distinto término derecho que la original.

El análisis deberá concluir indicando si tiene sentido aplicar la heurística de transformación [15] [16]. La Figura 4 muestra las transformaciones pertinentes.

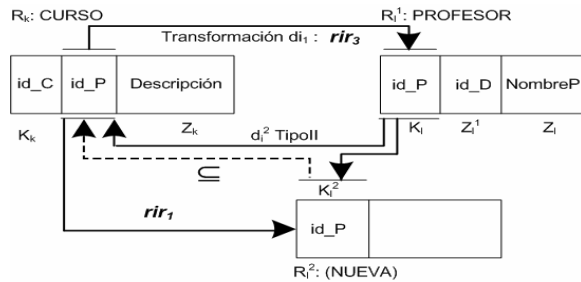


Fig. 4: Dependencias generadas en el esquema transformado (con redundancia)

La aplicación de una transformación con el objetivo de simplificar el esquema no parece producir el efecto

deseado ya que persiste una *di* de tipo II. Aparentemente se ha enriquecido el esquema al poner en evidencia NUEVA, pero en realidad es una proyección de las columnas claves de la tabla PROFESOR.

4. Digs: ANÁLISIS DE CASOS

Dado que una *dig* se interpreta mediante un par de dependencias de inclusión en sentido inverso, la estructura de una de las *dis*, condiciona el tipo o estructura de la restricción reversa.

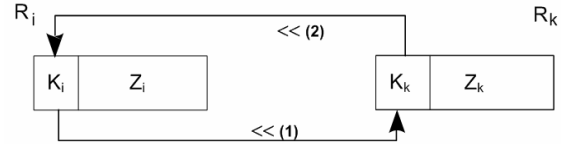


Fig 5: Dependencia de Igualdad en función de *dis*

En función de las 5 correlaciones posibles ilustradas en la Fig. 2 pueden establecerse las estructuras sintácticas de *digs* indicadas en la Tabla 1, en la que se han eliminado los casos simétricos.

W_k	K_k	Z_k	K_k'	$K_k \tilde{E} Z_k$	$K_k' \tilde{E} Z_k$
K_i	1 (1) I (2) I	2 (1) II (2) I	3 (1) III (2) I	4 (1) IV (2) I	5 (1) V (2) I
Z_i		6 (1) II (2) II	7 (1) III (2) II	8 (1) IV (2) II	9 (1) V (2) II
K_i'			10 (1) III (2) III	11 (1) IV (2) III	12 (1) V (2) III
$K_i \tilde{E} Z_i$				13 (1) IV (2) IV	14 (1) V (2) IV
$K_i' \tilde{E} Z_i$					15 (1) V (2) V

Tabla 1. Estructura de los términos de las *digs* en función de *dis* inversas. Referencias: K_* (Clave); Z_* (atributos no-clave); K_*' (subconjunto estricto de K_*). * = i, k

A continuación se desarrollan algunos casos particulares de *digs*, a partir de los cuales se puede inferir el análisis correspondiente al resto de las conformaciones.

Caso 1: Este caso es trivial (Fig. 6), ambas restricciones son *rirs*. Una partición vertical mediante proyecciones específicas de atributos y claves genera este tipo de construcciones. La necesidad de mantener la igualdad entre los dominios activos de la clave, presente en ambas proyecciones, genera esta *dig*. No hay transformación que pueda mejorar esta situación.

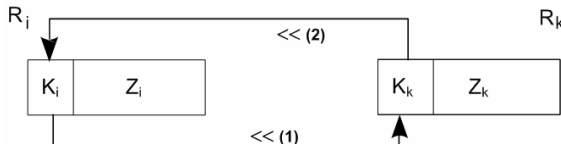


Fig. 6: Dig I-I

Caso 3: La Fig. 7 ilustra la estructura de un esquema que responde a este caso.

Para facilitar el análisis de este caso, es conveniente

interpretar la estructura de R_k . Se estudiarán los siguientes estados: materialización de una relación N:N, de una entidad débil o de una relación ternaria.

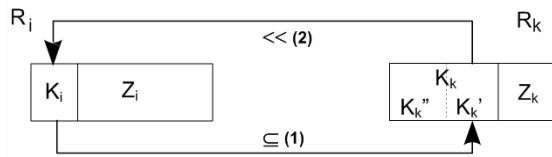


Fig. 7: Dig III-I

(a) Si representa una relación N:N.

Dado que la *di* (2) es en realidad una *rir*, la exigencia adicional que impone la *di* (1) indica la participación obligatoria de R_i en R_k . La Figura 8 muestra esta situación:

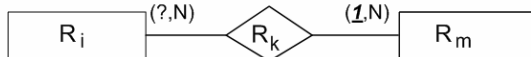


Fig. 8: Interpretación de una dig III-I como participación obligatoria en una relación N:N

La participación obligatoria se representa mediante el '1' que aparece como cardinalidad mínima en la relación de R_i con R_m .

(b) Si representa una entidad débil.

Si R_k es una entidad débil, la *di* (2) sería la *rir* que surge a partir de la derivación $MER \rightarrow$ esquema relacional. La *di* (1) indica que K_i es un subconjunto de la fracción fuerte de K_k (sea K_k'). Esto no tiene sentido, pero ambas *di* consideradas en conjunto representan la cardinalidad mínima '1' en la multiplicidad de R_i , respecto de R_k . Considerar K_i como subconjunto de la fracción débil de K_k , se considera sin sentido.

(c) Si representa una relación n-aria.

En este caso se pueden sacar conclusiones similares a las obtenidas en el caso de la relación N:N. Por ejemplo, si R_k es 1:1:1, relación ternaria entre R_i , R_j y R_m , el esquema relacional es $R_k(K_i, K_j, K_m, \dots)$ con claves alternativas K_i, K_j, K_i, K_m y K_j, K_m . En los dos primeros casos, la *di* (2) surge de la derivación y la *di* (1), que aisladamente no tiene sentido, indica la participación obligatoria de R_i en R_k . En estos casos no es sencillo visualizar la participación obligatoria, ya que la forma de lectura 'look across' y la semántica de estas relaciones, no permite representarla directamente.

Caso 8.- En este caso la *di* (1) puede resolverse por la regla de pullback, lo que resulta en la eliminación de Z_i y el replanteo de *di* (1) como $R_i[Z_i'] \ll R_k[Z_k]$. Obsérvese que la *di* (2) no es realmente de Tipo II, sino de Tipo IV, ya que Z_i en realidad es Z_i', Z_i'' y existe una dependencia funcional que los vincula. Nuevamente se puede aplicar la regla de pullback para esta dependencia y resulta $R_k[K_k] \ll R_i[Z_i', Z_i'']$ ó Z_k deben conservarse. Estas transformaciones convierten el Caso 8 en un Caso 2.

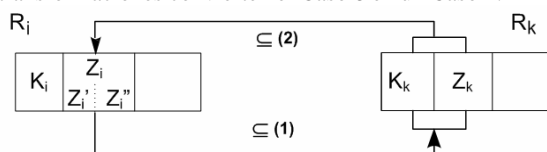


Fig. 9: Dig II-IV

Una transformación alternativa surge aplicando la transformación a R_i y luego la regla de pullback. No se

mejora la calidad del esquema respecto del resultado obtenido conservando R_k en su estado original.

Ventajas y desventajas de las transformaciones

Los casos de igualdad de conjuntos presentan la dificultad de generar ciclos de integridad referencial. Esto se suma a la imposibilidad de declararlos vía la cláusula *foreign key*. Esto implica que se requiera utilizar cláusulas *check* (aserciones, con acción de rechazo). Una actualización en una de las tablas produce el disparo y rechazo por falta de esta tupla en la otra tabla. Lo mismo ocurrirá si se da de baja. De implementarlo con *triggers*, la situación puede ser manejable de manera más artesanal. Un alta en cualquiera de las tablas puede despertar un alta en la otra, pero es necesario controlar que no se produzca el rebote de activaciones.

Una vez analizados todos los casos, las consecuencias de aplicar las transformaciones analizadas en esta sección pueden resumirse en:

- i) No hay casos en los que una manipulación de datos no realizable en el esquema original, se pueda llevar a cabo en el esquema transformado. Tampoco ocurre lo contrario: no hay operación posible de ser llevada a cabo en el esquema original que no pueda efectuarse en su correspondiente esquema transformado. Esto significa que la posibilidad de operar sobre los datos no se ve afectada por la transformación.
- ii) Con la excepción de los casos en los que existe una transformación analítica (mediante regla de pullback) o de aquellos casos con uno o ambos términos de Tipo III, la transformación estructural convierte el par de *dis* en dos *dis* y dos *rirs*, una de las cuales es redundante, en caso de generar ciclos de integridad referencial. Si se transforman ambas *dis*, se obtienen dos *dis* y dos *rirs* sin redundancia.
- iii) Hay casos en los que al menos una de las manipulaciones básicas es imposible de llevar a cabo. La utilización de claves surrogantes elimina este impedimento. Este no es un concepto privativo de estas transformaciones, sino un elemento de uso universal en el contexto de bases de datos.
- iv) Las *digs* que tienen una *di* de tipo I, se transforman cambiando el término izquierdo de la otra *di*, pero conservan su tipo. La excepción la constituyen las *digs* que tienen la otra *di* de tipo IV. El nivel semántico no mejora pues la entidad NUEVA ya está representada en otra de las existentes.
- v) Algunas estructuras que contienen uno o ambos términos de Tipo III permiten plasmar la igualdad simplemente a través de la participación obligatoria de una de las entidades, en la relación que la involucra. Otras pueden ser transformadas estructuralmente utilizando el concepto de inaplicabilidad a atributos o relaciones.
- vi) Todas las *digs* que tienen una *di* de Tipo IV pueden transformarse analíticamente por regla de pullback en *rirs*. Ésto convierte la *di* restante en otra de Tipo IV que también puede ser tratada analíticamente para simplificar el esquema. Los casos 4, 8, 11, 13 y 14 se convierten en casos 1, 2, 3, 1 y 3 respectivamente.
- vii) La 6ta. Forma normal [8] permite llevar una relación con nulos, a un conjunto de proyecciones obtenibles mediante particiones verticales y horizontales sin nulos. Las estrategias que se aplican para obtener estas proyecciones son similares a las aplicadas en los casos desarrollados.
- viii) Todas las transformaciones encontradas tienen como

base o permiten encontrar un esquema *MER* convencional o extendido con agregaciones, solamente. No ha sido necesario introducir nueva notación.

ix) Los casos restantes no pueden ser transformados de manera tal que la situación original sea mejorada en algún aspecto (operativo, nivel semántico, simplicidad).

5. CONCLUSIONES Y TRABAJO FUTURO

Este trabajo propone un análisis metodológico de las dependencias de comparación de conjuntos de valores. La propuesta está basada en el estudio de la estructura sintáctica de los términos de las dependencias y su influencia en la posibilidad de aplicar heurísticas conocidas para efectuar transformaciones y detectar redundancias. La aplicación de las transformaciones en los casos en los que sea pertinente, unida a un análisis de vínculos semánticos redundantes y a mecanismos de evaluación de corrección de esquemas constituyen un marco metodológico apropiado para lograr la reingeniería de esquemas de pobre calidad, desactualizados o erróneos, convirtiéndolos en otros apropiadamente estructurados y con su nivel semántico enriquecido. Las transformaciones propuestas contribuyen a la representación de aspectos inicialmente ocultos en modelo original del problema. Para aquellos casos en los que no se puede hallar una representación en un *MER* convencional con las herramientas conceptuales ya existentes, se buscan formas alternativas de especificación.

Si bien actualmente sólo se encuentra concluido el estudio de las *dis* y las *digs*, estos estudios preliminares indican que es prematuro y por lo tanto no aconsejable crear extensiones del *MER*, tal como sugieren algunos autores. Esta decisión debería postergarse al menos hasta haber estudiado previamente en forma exhaustiva las posibilidades de representación de las restantes dependencias con los mecanismos de abstracción ya estandarizados de facto.

La conveniencia o no de aplicar transformaciones se ha evaluado según el nivel semántico del esquema y la implementación de las operaciones de actualización básicas.

Adicionalmente, es necesario estudiar el efecto que las opcionalidades introducen en las relaciones en presencia de las *reglas del negocio*, dado que se presentan diferentes situaciones cuando la participación en las relaciones no es total. Ésto da lugar a relaciones redundantes en algunos casos y en otros pone en evidencia sobrecarga semántica de las entidades vinculadas. Este estudio está en sus comienzos.

Ciertamente, con el advenimiento de los modelos OO y OR, el estudio de las dependencias parece haber perdido cierto interés. Sin embargo, las transformaciones obtenidas y los conceptos subyacentes son aplicables en el contexto de otros paradigmas y a otras problemáticas, por ejemplo en migración y reingeniería de sistemas.

6. REFERENCIAS

[1] Albrecht, M.; "Informal and Intelligent Acquisition of Semantic Constraints in Database Design and Reverse Engineering", Rostocker Informatik-Berichte. 1994.
[2] Albrecht, M.; Buchholz, E.; Düsterhöft, A.; Thalheim, B.; "An Informal and Efficient Approach for Obtaining Semantic Constraints using Sample Data and Natural Language Processing", Lecture Notes In Computer Science, Vol. 1358, Selected Papers from a 1st. Workshop on Semantics in Databases, Prague, Czech Republic.

Thalheim, Bernhard and Libkin, Leonid (Eds.). 1998.
[3] Badía, A.; "Entity-Relationship Modeling Revisited", SIGMOD Record, Vol. 33, Nro. 1. 2004.
[4] Chen, P.; "The Entity-relationship Model: Toward a Unified View of Data" ACM Transactions on Database Systems. 1(1), pp. 9-36.1976.
[5] Chen, P.; "The Entity/Relationship Model: A bases for the enterprise view of data" AFIPS Conference Proceedings, Vol. 46. 1977.
[6] Codd, E.; "The Relational Model for Database Management" Version 2. Addison Wesley Publ.Co, Reading, MA. 1990.
[7] Connolly, T.; Begg, C.; Strachan, A.; "Database systems: A practical approach to design, implementation and management" 2ºEd. Addison Wesley. 1999.
[8] Date, C.; Darwen, H.; "How to Handle Missing Information without Using Nulls". <http://www.hughdarwen.freeola.com/TheThirdManifesto.web/Missing-info-without-nulls.pdf> (PDF). 2003.
[9] Doorn, J. & Rivero, L.; "Normalization of Non-BCNF Relations Integrity Constraints" Proceedings XII International Conference of Systems Engineering, ICSE-97, Coventry, UK, 1997, pp. 217-222.
[10] De Miguel, A.; Piattini, M.; Marcos, E. "Diseño de Bases de Datos Relacionales". Alfaomega -RaMa. 2000.
[11] Elmasri, R. & Navathe, S.; "Fundamentals of Database Systems". Addison-Wesley, 3º Ed. 2000.
[13] Levene, M. & Loizou, G.; "A Guided Tour of Relational Databases and Beyond" Springer-Verlag, London. 1999.
[14] Rivero, L.; Doorn, J.; Ferraggine, V.; "Inclusion Dependencies". En Developing Quality Complex Database Systems: Practices, Techniques and Technologies, Idea-Group Publishing, Hershey, PA. Editor Dr. Shirley A. Becker. Florida Institute of Technology, pp. 261-278. 2001.
[15] Rivero, L.; Ferraggine, V.; Doorn, J.; "Reingeniería de Bases de Datos: Un enfoque Basado en el Análisis de Dependencias de Inclusión" Revista Colombiana de Computación. Vol. 4. 2003.
[16] Rivero L.; Doorn J.; Ferraggine V.; "Enhancing Relational Schemas through the Analysis of Inclusion Dependencies". International Journal of Computer Research. Nova Science Publishers Company USA, Tomo 12, No 4. 2004.
[17] Rivero, L.; "An overview of Inclusion Dependencies". En Encyclopedia of Information Science and Technology. Khosrow-Pour M. (Editor). Idea Group. Inc. Hershey (USA). 2005.
[20] Tari, Z.; Buhkres, O.; Stokes, J; Hammoudi, S.; "The Reengineering of Relational Databases based on Key and Data Correlations" In Searching for Semantics: Datamining, Reverse Engineering, etc. Proceedings 7º IFIP 2.6 Working Conference on Database Semantics (DS-7). S. Scappapietra y F. Maryanski Ed. Chapman & Hall, Lausanne, 1998, pp. 183-214.
[21] Teorey, T. J.; "Database Modeling and Design: The Entity-Relationship Approach". Morgan Kaufmann Publishers, San Mateo, California. 1990.