Un Lenguaje Lógico para especificar Restricciones Explícitas en el Modelo Entidad- Interrelación y sus Extensiones

El lenguaje de la lógica de primer puede ser tomado como base para la definición de un lenguaje particular adecuado para expresar, sin ambigüedades, las diferentes restricciones explícitas que permiten describir aquellos aspectos de un mini-mundo particular que no pueden representarse a través de las estructuras y restricciones implícitas del modelo entidad-interrelación y de las extensiones estudiadas en el curso.

El basamento del lenguaje lógico propuesto busca extender la lógica de primer orden con:

- Operadores que permitan la manipulación directa de conjuntos y sus elementos, ya que el modelo entidad-interrelación se basa en la teoría de conjuntos.
- Cuantificadores que permitan expresar, de una forma concisa, la existencia de un número específico de elementos en el universo específico de elementos que se están estudiando.

Para un esquema E desarrollado utilizando el modelo entidad-interrelación y sus extensiones se tendrá un lenguaje específico que utilizará los siguientes elementos de lógica de primer orden:

- Cualquier símbolo alfabético (letra) con o sin subíndices constituirá una **variable**. En este lenguaje, las variables serán utilizadas para indicar los elementos de un conjunto que, para el caso del modelo entidad-interrelación estudiado, podrá indicar una entidad, una interrelación, o un valor de un atributo multivaluado (cuya representación formal en el modelo será un conjunto de valores). En ningún momento, podrá utilizarse como variable un símbolo alfabético que corresponda al nombre de un conjunto de entidades, un conjunto de interrelaciones, el nombre de un atributo, o el nombre de un rol.
- Las constantes permitidas en el lenguaje de especificación de restricciones explícitas de un esquema entidad-interrelación E serán las siguientes:
    - Las constantes numéricas y cadenas de caracteres, cuya finalidad será representar valores específicos de un atributo simple.
    - Los conjuntos de constantes numéricas y cadenas de caracteres, cuya finalidad será representar valores específicos de un atributo multivaluado.
    - Las tuplas de cualquier aridad, con elementos en el dominio de constantes numéricas (naturales, enteros, racionales e irracionales y reales) y secuencias de caracteres, para representar valores de un atributo compuesto.
    - Constantes particulares de interés que representan los nombres de los conjuntos de entidades y conjuntos de interrelaciones definidos en el esquema.
    - Constantes particulares de interés que representan los nombres asignados a los roles que califican la participación de un conjunto de entidades en un conjunto de interrelaciones.
    - Constantes de interés que representan los nombres de los atributos definidos para un conjunto de entidad o un conjunto de interrelaciones.
    - La constante **IS_A** será utilizada para representar explícitamente las diferentes interrelaciones que pueden establecerse entre las subclases y superclases en un esquema desarrollado utilizando las extensiones del modelo entidad-interrelación estudiadas. Las estructuras en las que podrá utilizarse la constante **IS_A** serán las interrelaciones subclase/superclase, las generalizaciones, las especializaciones y las categorizaciones.
    - En lo que respecta a las constantes permitidas, es importante hacer la siguiente salvedad: Aún cuando la definición matemática del modelo entidad-interrelación establece que los valores de un

atributo serán considerados como elementos en el conjunto de partes del dominio de valores, se relajará la notación de valores de la siguiente forma:

- Si **{v }** es el valor de interés para un atributo simple en una restricción explícita, se utilizará sólo **v** para indicar dicho valor en lugar de utilizar la verdadera notación matemática del valor de un atributo simple en el modelo.
- Para representar el valor nulo se utilizará la constante especial **null** en lugar de {}, el valor designado para representar el valor nulo en la definición formal del modelo.

- El único predicado permitido para asociar o trabajar con conjuntos será el predicado pertenece y su negación. De igual manera se utilizará el símbolo matemático $\in$ (o alternativamente $\notin$) para representar estos predicados. Adicionalmente se deberá utilizar la notación infija para indicar el uso de estos predicados.

Ejemplo: Para indicar que la variable *a* será utilizada para describir una entidad del conjunto de entidades ESTUDIANTE, se utilizará la fórmula

$$a \in \text{ESTUDIANTE}$$

- Para la definición de fórmulas podrán utilizarse los comparadores habituales (los operadores =, ¹ , ³ , £, >, <),
- Se podrán *"modularizar"* fórmulas para no complicar la estructura de una fórmula lógica. En tal sentido se podrán definir predicados de la siguiente forma:
  - Si *P* es un predicado de aridad *n* que representa la fórmula lógica *t*, entonces *P* puede ser definido de la siguiente forma:

$$P(t_1, t_2, \ldots, t_n) \; º \; t\,(t_1, t_2, \ldots, t_n)$$

Para una explicación mas detallada del significado de las constantes, variables, predicados y funciones, puede consultar el apéndice que se encuentra el final de este documento.

Notación "punto"

Utilizaremos el símbolo "**.**" para referenciar el(los) valor(es) de un atributo.

Observaciones:

- Para hacer referencia a un atributo compuesto se hace uso de la referenciación múltiple, navegando a través de todos los atributos que le anteceden en la composición. En general la forma de acceder a un atributo se puede representar a través de la expresión regular

<variable> [.<nombre_atributo>] +

donde $[x]^+$ indica la repetición de x una o mas veces.
Supongamos, por ejemplo, que la variable f representa una entidad del conjunto de entidades A, cuyo único atributo es a_1 y está compuesto por los atributos a_2 y a_3, donde a_3 también es un atributo

compuesto por a_4, a_5, y a_6. Para indicar el valor de a_5 en la entidad f se utilizará el siguiente término:

$$f.a_1.a_3.a_5$$

Notación "corchetes"

Se utilizarán los símbolos "*[*" y "*]*" para referenciar las entidades participantes en una interrelación de bajo cierto rol. Suponiendo que se tiene la interrelación c_1 del conjunto de interrelaciones C, el cual enlaza a los conjuntos de entidades A y B (ver Figura 1), para describir la entidad del conjunto de entidades A que está enlazada por la interrelación c_1 se utilizará el término c_1[A].

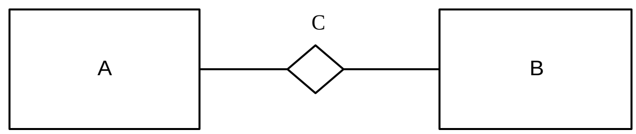

Figura 1. Diagrama ERE para explicar la notación de corchetes. En este caso si $c_1=(a_1,b_1)$ Î C, luego $c_1[A] = a_1$.

Para el caso de interrelaciones donde se califique la participación de un conjunto de entidades utilizando roles, la notación de corchetes será extendida para mantener dos parámetros: el primero representará el conjunto de entidades enlazado, mientras que el segundo indicará el rol de participación.

Este tipo de notación es particularmente útil en interrelaciones recursivas, donde las entidades participantes son en verdad una misma entidad como se muestra en la Figura 2, y donde se distingue la participación de la entidad por el uso de un rol. En tal sentido, y utilizando el ejemplo de la Figura 2, para indicar la entidad del conjunto de entidades *A* que se enlaza bajo el rol r_1 a una interrelación *c* del conjunto de interrelaciones R se utilizará el término *c[A,r1].*

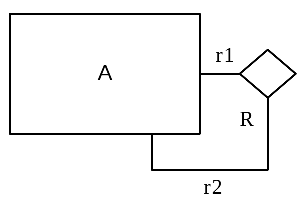

Figura 2. Diagrama ERE para explicar la notación de corchetes en el caso de relaciones recursivas. En este caso si $c_1=(a_1,a_2)$ Î C, luego $c_1[A,r_1] = a_1$.

Notación de "tuplas"

Utilizaremos los símbolos "*(*",*"*)*" para denotar los objetos que conforman una instancia de un conjunto de interrelaciones. Cada objeto en la tupla se separará por comas, y no se asumirá ningún orden particular entre ellos.

Operadores Matemáticos de Agregación

Adicionalmente, se podrán utilizar expresiones aritméticas para definir términos que represente el proceso de resumir de ciertos datos. En particular, se permitirá el uso de las siguientes funciones:

- $\sum$ En caso de querer representar sumatorias.
- $\prod$ En caso de querer representar productorias.
- AVG En caso de querer representar el promedio.
- MIN En caso de querer representar el mínimo valor de un conjunto.
- MAX En caso de querer representar el máximo valor de un conjunto.

Extensión del Cuantificador Existencial

El conjunto de cuantificadores se extiende con los cuantificadores numéricos:

- "Al menos n" ($\$_{n}$): Denota la existencia de *n* objetos diferentes del universo modelado que cumplen una fórmula lógica
- "A lo más n" ($\$^{n}$): Denota *n* como una cota superior del número de objetos diferentes del universo modelado que cumplen una fórmula lógica
- "Exactamente n" ($\$!_{n}$): Denota la existencia de exactamente *n* objetos diferentes del universo modelado que cumplen una fórmula lógica
- "Entre n y m" ($\$_{n}^{m}$): Denota la existencia de un rango entre *n* y *m*, de objetos del universo modelado que cumplen con la fórmula lógica. En este caso *n* representa el límite inferior de ese rango (que incluye a tal número *n*), mientras que *m* representa el límite superior de ese rango (que incluye a tal número *m*).

Ejemplos

En lo sucesivo, todos los ejemplos versarán sobre el diagrama ERE que se muestra en la Figura 3. En este diagrama se representa la interrelación *dicta* que existe entre un profesor, una asignatura y un período, así como la interrrelación *es_requisito* que existe entre dos asignaturas.

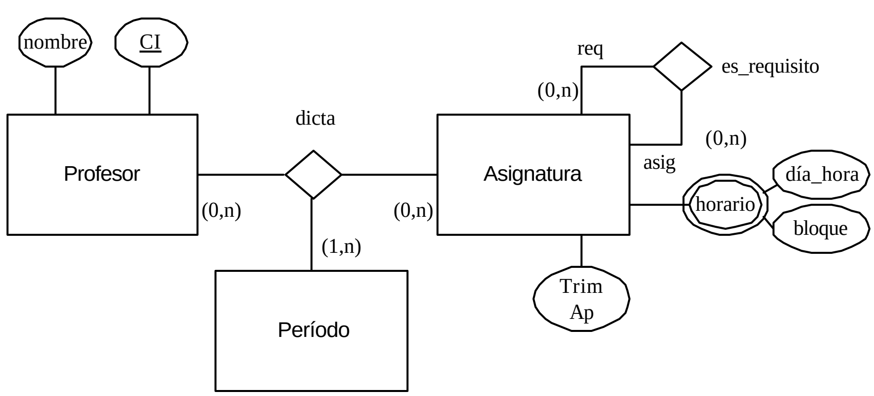

Figura 3. Diagrama ERE sobre el que se basan los ejemplos de uso de lenguaje de restricciones explícitas

En este ejemplo, todo profesor tiene como atributos su cédula y nombre, mientras que toda asignatura posee como atributos TrimAp (atributo simple monovaluado que expresa trimestre de apertura, el cual es un entero del 1 al 15 que indica en cuál de los 15 trimestres de la carrera se recomienda cursar esta asignatura). De igual forma, el atributo horario (atributo compuesto y multivaluado) describe los datos asociados a los bloques de horario de una asignatura, i.e. los días y horas en los que se dicta la misma.

Ejemplo 1:

Motivación: El siguiente ejemplo muestra el uso de la notación punto para un atributo simple.

La restricción **"No existe asignatura que aparezca en el pensum en un trimestre mayor al 15"** permite representar el dominio de valores válidos para el atributo *TrimAp* del conjunto de entidades *Asignatura*.

Utilizando el lenguaje de restricciones explícitas visto, la restricción antes enunciada puede representarse como

$$\neg\exists a\left(a \in Asignatura \wedge \left(\left(aTrimAp > 15\right) \vee \left(aTrimAp < 1\right)\right)\right)$$

o alternativamente como

$$\forall a\left(a \in Asignatura \rightarrow \left(\left(a.TrimAp \geq 1\right) \wedge \left(aTimAp \leq 15\right)\right)\right)$$

Ejemplo 2:

Motivación: El siguiente es un ejemplo donde se emplean algunos de los elementos de la notación más básicos, como los cuantificadores de la lógica de primer orden, la notación de tuplas y la notación de conjuntos.

La restricción ***"No existe una asignatura, tal que sea requisito de si misma"*** permite caracterizar la interrelación es_requisito como una interrelación antirreflexiva.

Utilizando el lenguaje de restricciones explícitas visto, la restricción antes enunciada puede representarse como

$$\neg\exists a(a \in Asignatura \wedge (a,a) \in es\_requisito)$$

o también puede ser representado de la siguiente forma:

$$\forall a(a \in Asignatura \rightarrow (a,a) \notin es\_requisito)$$

Ejemplo 3:

Motivación: En este ejemplo se usa la notación de corchetes para indicar roles.

"No existe asignatura que tenga algún requisito tal que el mismo se oferte en un trimestre posterior a ella".

Utilizando el lenguaje de lógica de primer orden esto se puede expresar como:

$$\neg\exists e(e \in es\_requisito \wedge e[Asignatura, req].TrimAp \geq e[Asignatura, asig].TrimAp)$$

o alternativamente como

$$\forall e\left(e \in es\_requisito \rightarrow \left(e[Asignatura, req].TrimAp < e[Asignatura, asig].TrimAp\right)\right)$$

Ejemplo 4:

Motivación: En este ejemplo se usa uno de los cuantificadores númericos definidos como extension de los cuantificadores básicos.

"Un profesor no puede dictar mas de dos asignaturas en un mismo período".

Esta restricción se puede expresar de la siguiente forma:

$$\forall p\, \forall f\, \exists^2 a(p \in Período \wedge f \in \Pr ofesor \wedge a \in Asignatura \wedge dicta(p, a, f))$$

Ejemplo 5:

Motivación: En este ejemplo se ilustra el uso de la notación punto para un atributo multivaluado. Este ejemplo ilustra cómo los atributos multivaluados son tratados como conjuntos.

"Para una misma asignatura, no se puede tener dos horarios con diferente letra de bloque y exactamente el mismo horario en días y horas".

Esta restricción se puede expresar como

$$\forall a\, \forall h_1 \neg\exists h_2(a \in Asignatura \wedge h_1 \in a.horario \wedge h_2 \in a.horario \wedge h_1.bloque = h_2.bloque \wedge h_1 \neq h_2)$$

o alternativamente como

$$\forall a\, \forall h_1 \forall h_2\left(\left(a \in Asignatura \wedge h_1 \in a.horario \wedge h_2 \in a.horario \wedge h_1.bloque = h_2.bloque\right) \rightarrow h_1 = h_2\right)$$

Apéndice: Repaso de Lógica de Primer Orden

Un Lenguaje de Primer Orden ***LPO*** está constituido de los siguientes conjuntos de símbolos:

- ***CON***: (Constantes) símbolos que representan objetos particulares del universo modelado
- ***VAR***: (Variables) símbolos que representan objetos genéricos del universo modelado
- ***PRE***: (Predicados) símbolos que representan relaciones o acciones o propiedades de los objetos del universo modelado. Al número de objetos que asocia la relación se le llama "la aridad del predicado".
- ***FUN***: (Funcionales) símbolos que representan funciones que se aplican a los objetos del universo modelado (cuyo resultado es otro objeto del universo), permiten denotar valores de propiedades de los objetos. Al número de objetos que toma como argumento la función se le llama "la aridad de la función"

Los conectores de la Lógica de Primer Orden son:

- ***UNA***: conectores unarios: ***UNA***={Ø}
  - Negación (Ø): Permite expresar la negación de una sentencia.
- ***BIN***: conectores binarios: ***BIN***={Ù,Ú,®, « }
  - Conjunción (Ù): Permite expresar el "y" lógico de dos fórmulas lógicas.
  - Disyunción (Ú): Permite expresar el "o" lógico (inclusivo) de dos fórmulas lógicas.
  - Condicional (®): Permite expresar la "implicación" lógica de dos fórmulas lógicas.
  - Bicondicional («): Permite expresar la "equivalencia" lógica de dos fórmulas lógicas.

Símbolos comparadores

Adicionalmente al conjunto de símbolos antes mencionados, se asume la existencia de predicados especiales, los cuales son los comparadores usuales y que están definidos en el conjunto ***CMP.***

- ***CMP:*** comparadores: ***CMP***={=,¹,³,<,>,£}

El conjunto de cuantificadores de la Lógica de Primer Orden ***CTF*** ={$,"}:

- Existencial ($): Denota la existencia de algún objeto del universo modelado que cumple una fórmula lógica.
- Universal ("): Denota que todos de objetos del universo modelado cumplen una fórmula lógica.

El conjunto de los Términos ***TER*** se define según las siguientes reglas:

- Si ***c***∈***CON***, entonces ***c***∈ ***TER***
- Si ***x***∈***VAR***, entonces ***x***∈***TER***
- Si ***F***∈***FUN*** y la aridad de ***F*** es n y $t_1, t_2, .. t_n$∈***TER*** entonces ***F***($t_1, t_2, .., t_n$) ∈***TER***

El conjunto de las Fórmulas Bien Formadas ***FBF*** se define según las reglas

- Si t_1∈***TER*** y t_2∈ ***TER*** y Q∈***CMP***, entonces (t_1 Q t_2)∈***FBF***
- Si ***P***∈***PRE*** y la aridad de ***P*** es n y $t_1, t_2, .. t_n$∈***TER***, entonces ***P***($t_1, t_2, .., t_n$) ∈***FBF***
- Si ***A***∈***FBF*** y Q ∈***UNA***, entonces Q***A*** ∈***FBF***
- Si ***A***∈***FBF*** y ***B***∈***FBF*** y Q ∈***BIN***, entonces (***A***Q***B***) ∈***FBF***
- Si ***A*** ∈***FBF*** y ***x***∈***VAR*** y Q∈***CTF***, entonces Q_x***A*** ∈***FBF***

Formas Aristotélicas

Las sentencias cuantificadas del lenguaje natural son combinaciones de las siguientes cuatro formas básicas:

- "Todos los P son Q": $\forall x(P(x) \rightarrow Q(x))$
- "Algunos P son Q": $\exists x(P(x) \wedge Q(x))$
- "Algunos P no son Q": $\neg \forall x(P(x) \rightarrow Q(x))$
- "Ningún P es Q": $\neg \exists x(P(x) \wedge Q(x))$