

## THE EXPRESSIVE *SHIQ* LIKE ONTOLOGY LANGUAGE FOR THE SEMANTIC WEB

### LA EXPRESIVA *SHIQ* COMO LENGUAJE ONTOLÓGICO PARA LA WEB SEMANTICA

Eduard G. Puerto Cuadros, Luz Marina Santos Jaimes, Wilson A. Pérez Miranda

#### Universidad de Pamplona

Grupo de Investigación Ciencias Computacionales  
{eduardpuerto, lsantos, wilp\_s}@unipamplona.edu.co

**Abstract:** This article describes Description Logic *SHIQ* like ontology language that it can support the Semantic Web. The vision of the Semantic Web is to make that the information existing in the current Web can to be interpreted by the computers without the human intervention. So that this happens, it is necessary that the information of the pages Web is codified by means of ontology. The description logics are excellent candidates like ontology language to represent the knowledge of the resources Web, they formally defining the concepts and their relations, with capacity to make deductions with this knowledge.

**Resumen:** Este artículo describe la Lógica Descriptiva *SHIQ* como un lenguaje ontológico que puede soportar la Web Semántica. La visión de la Web Semántica es hacer que la información existente en la actual web pueda ser interpretada por los ordenadores sin intervención humana. Para que esto suceda, es necesario que la información de las páginas web se codifique mediante una ontología. Las lógicas descriptivas son excelentes candidatas como lenguaje ontológico para representar el conocimiento de los recursos web, definiendo formalmente los conceptos y sus relaciones, con capacidad para realizar deducciones con este conocimiento.

**Keywords:** Ontología, Lógicas Descriptivas, Web Semántica, *SHIQ*, DAML+OIL.

## INTRODUCCIÓN

La actual Web adolece de una manera de encontrar información de forma precisa y de poder realizar deducciones con la información existente. Una de las razones es que los actuales agentes de búsqueda no están diseñados para “comprender” los recursos de información que reside en la Web.

Tim Berners-lee y colegas [Berners, 1999; Berness, et al 2001] proponen abordar esta y otras deficiencias transformado la red desde un espacio de información a un espacio de conocimiento.

La idea es que los recursos web puedan ser utilizados y “comprendidos” por herramientas automatizadas, tales como motores de búsqueda, y por usuarios humanos. Este compartir información entre diferentes agentes requiere de un mark-up semántico del contenido de información del recurso. Estas anotaciones son hechas en un expresivo lenguaje estándar y ciertos términos. Para garantizar que diferentes agentes tengan una comprensión común de los términos, se requiere de una ontología para describirlos.

Las ontologías permiten establecer una terminología compartida entre agentes. Una definición formal es la propuesta por Gruber [Gruber, 1993] y extendida por Studer y colegas [Studer, 1998] que se describe como “una especificación explícita y formal sobre una conceptualización compartida”.

El uso de ontologías en este contexto requiere de un buen diseño, bien definido, y un lenguaje ontológico intuitivo a los usuarios humanos, compatible con estándares Web (e.g., XML, RDFS, OWL), expresivo y razonable (el razonamiento asegura la calidad de la ontología). Este artículo presenta en la sección 2 el por qué las lógicas descriptivas son buenas candidatas como lenguajes ontológicos, en la sección 3 se muestra la relación *SHIQ-OWL* junto con la parte formal de *SHIQ*. La sección 4 presenta un paralelo entre *SHIQ*- DAML y por último las conclusiones.

## 2. LAS DLS COMO LENGUAJE ONTOLÓGICO

Las Lógicas Descriptivas (DLs: Description Logics) son formalismos para la representación y gestión de conocimiento conceptual de un dominio de aplicación [Lizcano, 2001], cubren los esquemas de representación de conocimiento de las redes semánticas de Quillian y los sistemas de Marcos de Minski y poseen una semántica bien definida al estilo de Tarski.

Las DLs usualmente están provistas de un formalismo terminológico y aseverativo. En su forma más elemental, un *axioma terminológico* (TBox) puede ser usado para introducir los nombres (abreviaciones) para descripciones complejas. E.g., HombreFeliz Y Humano \*¬Mujer \*  $\exists$ casado.Medica \* ( $\geq 5$  tieneHijos) \*  $\forall$ hijo.Profesor. El *formalismo aseverativo* (ABox) puede ser usado para indicar características de los individuos. E.g., HombreFeliz(EDGAR).

Los sistemas de DLs proporcionan a los usuarios varias capacidades de inferencia a través de algoritmos de *subsumición*, de *instanciación*, de *consistencia* [Baader, 2003]. En orden asegurar un comportamiento razonable y fiable de un sistemas DL, estos problemas de inferencia deben por lo menos ser decidibles para la DL empleada por el sistema y preferiblemente de baja complejidad. El razonamiento en una ontología usualmente lo que busca es computar la jerarquía de conceptos.

## 3. RELACIÓN DE SHIQ-OWL

Las ontologías son fundamentales para la Web Semántica y su construcción, integración, y evolución dependen en gran parte de una semántica clara y eficientes herramientas de razonamiento. Las DLs proveen ambas cosas, por tanto éstas son ideales como lenguaje ontológico.

SHIQ es una DLs que sirve como lenguaje ontológico y es igual a ALC [Schmidt, 1991]+  $R_+$  + jerarquía de rol + roles inverso + QNR y es la base para OWL.

OWL DL  $\approx$  SHIQ extendida con individuos (SHOIQ).

OWL Lite  $\approx$  SHIQ con solo restricciones funcional (SHIF).

Las siguientes características dan soporte a lo anterior:

1. *SHIQ* provee restricciones de número y es más expresiva que sus predecesoras. En *SHIQ*, podemos expresar: “persona que tiene a lo mas dos hijos” (sin mencionar las características de estos hijos),  
( $\beta 2$  tieneHijo),
2. *SHIQ* permite la formulación de complejos axiomas terminológicos tales como: “los humanos tiene padres humanos”,  
Humano  $\mu \exists$ tienePadre.Humano.
3. *SHIQ* permite roles inversos, roles transitivos, y subroles. E.g., de acuerdo con tieneHijo uno puede usar asimismo una forma inversa como tienePadre,

Se ha discutido en la DL y la comunidad ontológica que estas características juegan un papel central al describir propiedades de agregación de objetos y construcción de ontologías [Fensel, 2000b; Stevens, 2001; Sattler, 2000]. Las actuales DLs y en particular *SHIQ* proporcionan estas características como formalismos lógicos de los lenguajes ontológicos para la Web como OIL, DAML+OIL, OWL [Stevens, 2001].

### 3.1 Aspectos Formales de SHIQ

*Sintaxis SHIQ-roles.* Sea  $R$  un conjunto de nombre de roles dividido en un conjunto  $R_+$  de roles transitivos y un conjunto  $R_p$  de roles normales. El conjunto de todos los SHIQ-roles es  $R \cap \{ r^- \mid r \in R \}$ , donde  $r^-$  es el rol inverso de  $r$ . Un axioma de

inclusión de rol es de la forma  $r \mu s$ , donde  $r, s$  son *SHIQ*-roles. Una arquitectura de rol es un conjunto finito de *axiomas de inclusión de rol*.

**Semántica SHIQ-roles.** Una interpretación  $I = (\Delta^I, \cdot^I)$  consiste de un conjunto  $\Delta^I$ , llamado el dominio de  $I$ , y una función  $\cdot^I$  que relaciona cada rol a un subconjunto de  $\Delta^I \times \Delta^I$  tal que, para toda  $p \in R$  y  $r \in R_+$ ,  $\langle x, y \rangle \in p^I$  sii  $\langle y, x \rangle \in (p^-)^I$ , si  $\langle x, y \rangle \in r^I$  y  $\langle y, z \rangle \in r^I$  entonces  $\langle x, z \rangle \in r^I$ . Una interpretación  $I$  satisface una jerarquía de rol  $R$  sii  $r^I \subseteq s^I$  para cada  $r \mu s \in R$ ; cada interpretación es un modelo de  $R$ .

**SHIQ-conceptos.** Sea  $N_c$  un conjunto de nombres de conceptos. El conjunto de *SHIQ*-conceptos es el mínimo conjunto tal que:

1. Cada nombre de concepto  $A \in N_c$  es un *SHIQ*-concepto.
2. si  $C$  y  $D$  son *SHIQ*-conceptos y  $r$  es un *SHIQ*-rol, entonces  $C * D, C + D, \neg C, \forall r.C, \exists r.C$  son *SHIQ*-conceptos.
3. si  $C$  es un *SHIQ*-concepto,  $r$  es un simple *SHIQ*-rol, y  $n \in \subseteq$ , entonces  $(\beta n r.C)$  y  $(\rho n r.C)$  son *SHIQ*-conceptos.

Los conceptos pueden ser usados para describir las nociones relevantes de un dominio de aplicación.

**GCI.** Una inclusión General de Concepto o GCI es de la forma  $C \mu D$ , donde  $C, D$  son *SHIQ*-conceptos. Un conjunto finito de GCIs es llamado un TBox. Una interpretación  $I$  es un modelo de un TBox  $T$  si satisface todos los GCIs en  $T$ , i.e.,  $C^I \subseteq D^I$  considerado para cada  $C \mu D \in T$ .

Un concepto definido es de la forma  $AYC$ , donde  $A$  es un nombre de concepto. Esto puede verse como una abreviación para los dos GCIs  $A \mu C$  y  $C \mu A$ . Los problemas de inferencia son definidos con respecto a TBox y una jerarquía de roles.

**Definición.** El concepto  $C$  es satisfactible con respecto a una jerarquía de roles  $R$  y el TBox  $T$  si existe un modelo  $I$  de  $R$  y  $T$  con  $C^I \neq \emptyset$ . Cada interpretación es un modelo de  $C$  con respecto a  $R$  y  $T$ . El concepto  $D$  subsume al concepto  $C$  con respecto a  $\langle R, T \rangle$  (escrito  $C \mu_{\langle R, T \rangle} D$ ) sii  $C^I \subseteq D^I$  considerado para todo modelo  $I$  de  $R$  y  $T$ . Dos conceptos  $C, D$  son equivalentes con respecto a  $R$  (escrito  $C \equiv_{\langle R, T \rangle} D$ ) si cada uno de ellos se subsume.

Por definición, la equivalencia puede ser reducida a la subsumición. Además, la subsumición puede ser reducida a la satisfactibilidad ya que  $C \mu_{\langle R, T \rangle} D$  sii  $C * \neg D$  es insatisfactible con respecto a  $R$  y  $T$ .

### 3.2 Cómo SHIQ puede ser usado para definir ontologías?

En general, una ontología puede ser formalizada en un TBox como sigue: En primer lugar, delimitamos los mundos posibles introduciendo restricciones en las interpretaciones permitidas. e.g., para expresar que en nuestro mundo, queremos considerar “humanos, los cuales son magos o hechiceros”, podemos usar las GCIs,

$$\text{Humano} \mu \text{Mago} + \text{Hechicero y} \\ \text{Mago} \mu \neg \text{Hechicero}$$

Ahora, para expresar “humanos que tienen exactamente dos padres y todos los padres e hijos de humanos son humanos”, lo escribimos como:

$$\text{Humano} \mu \forall \text{tienePadre.Humano} * \\ (\beta 2 \text{tienePadre} \blacklozenge) * (\rho 2 \text{tienePadre} \blacklozenge) * \\ \forall \text{tienePadre}^- . \text{Humano},$$

$\blacklozenge$  es una abreviación para el concepto más general  $A + \neg A$ .

Una vez delimitado el mundo, podemos definir las nociones relevantes de nuestro dominio de aplicación usando definición de conceptos. Un nombre de concepto es *definido* si aparece en la parte izquierda de la definición y *primitivo* en otro caso.

El conjunto de conceptos definidos junto con los GCIs debe satisfacer las siguientes tres condiciones:

1. No existan definiciones múltiples; cada nombre de concepto definido debe aparecer a lo mas una vez como parte izquierda en una definición de concepto.
2. No deben existir definiciones cíclicas; no deben haber dependencias cíclicas entre los nombre definidos en el conjunto de definiciones de conceptos.
3. Los nombres definidos no deben aparecer en cualquiera de las GCI adicionadas.

Como un simple ejemplo de un conjunto de definiciones de conceptos que satisfacen las restricciones anteriores, definimos el concepto de Abuelo y Padre,

$$\text{Padre} \equiv \text{Humano} * \exists \text{tienePadre}^{-} \cdot \diamond,$$

$$\text{Abuelo} \equiv \exists \text{tienePadre}^{-} \cdot \text{padre},$$

La TBox consistente de las anteriores definiciones de conceptos y GCIs, junto con el hecho que Abuelo es un súper rol transitivo de tienePadre, implica la siguiente relación de subsumición.

$$\text{Abuelo} * \text{Hechicero} \mu \exists \text{tienePadre}^{-} \cdot \exists \text{tienePadre}^{-} \cdot \text{Hechicero},$$

En síntesis *SHIQ*-TBox puede, como primera medida, axiomatizar las nociones básicas de un dominio de aplicación (los conceptos primitivos) por GCIs, transitividad de aseveraciones, e inclusión de roles, en la sentido que estas aseveraciones restringen la posible interpretación de las nociones básicas. Por otra parte, nociones más complejas (la definición de conceptos) puede ser introducida por definiciones de conceptos.

La taxonomía de un TBox es dada por la subsumición jerárquica de la definición de conceptos. Esto puede ser computado usando un algoritmo de subsumición para *SHIQ*. La ingeniería de conocimiento puede comprobar si el TBox captura su conocimiento (intuición) por medio del chequeo de la satisfactibilidad de la definición de conceptos (ya que no tiene sentido dar una definición compleja para el concepto vacío), y para chequear si su lugar en la taxonomía corresponde con su posición intuitiva. El poder expresivo de *SHIQ* junto con el hecho que se puede “verificar” la TBox en el sentido mencionado anteriormente es la principal razón por la cual *SHIQ* esta bien situada como un lenguaje ontológico [Sattler, 2000; Fensel, 2000 Stevens, 2001].

#### 4. *SHIQ* Y PRIMEROS LENGUAJES ONTOLÓGICOS

DAML+OIL es un lenguaje ontológico para la web semántica cuya semántica puede ser definida vía una traducción en una expresiva DL. Esta relación permite que DAML+OIL utilice los resultados formales obtenidos en las investigaciones de DL (decidibilidad y complejidad) y usar los razonadores implementados en DL (FaCT y Racer) en orden a probar los servicios de razonamiento para aplicaciones DAML+OIL. Una Ontología DAML+OIL se corresponde con la TBox de una DL junto con una jerarquía de roles, describiendo el dominio en términos de *clases* (que corresponden a conceptos) y *propiedades* (correspondientes a los roles).

Constructor	DL Syntax	Example
IntersectionOf	$C_1 * \dots * C_n$	Humano * Hombre
UnionOf	$C_1 + \dots + C_n$	Doctor + Abogado
ComplementOf	$\neg C$	$\neg$ Hombre
OneOf	$\{x_1 \dots x_n\}$	{Juan, María}
ToClass	$\forall P.C$	$\forall$ tieneHijo.Doctor
hasClass	$\exists r.C$	$\exists$ tieneHijo.Abogado
hasValue	$\exists r.\{x\}$	$\exists$ ciudadanoDe.{Col}
minCardinalityQ	$(\rho n r.C)$	$\rho$ 2 tieneHijo.Abogado
MaxCardinalityQ	$(\beta n r.C)$	$\beta$ 1 tieneHijo.Hombre
InverseOf	$r^{-}$	tieneHijo $^{-}$

Tabla 1. DAML+OIL Constructores

Una ontología consiste de un conjunto de *axiomas* que expresan relaciones de subsumición entre clases o propiedades.

Las clases DAML+OIL pueden ser nombres o expresiones compuestas de simples clases y propiedades usando una variedad de constructores. El conjunto de constructores soportados por DAML+OIL, junto con la equivalencia de sintaxis en DL se muestra en el cuadro 1.

Sentencias de lógicas descriptivas pueden ser escritas usando una combinación de XML y RDF.

E.g., Humano \* Hombre en DAML

```
<daml: Class>
<daml: intersectionOf rdf: parseType="daml:
collection">
<daml: Class rdf: about="#Humano"/>
<daml: Class rdf: about="#Hombre"/>
</daml: intersectionOf >
</daml:Class>
```

y ( $\rho$ 2tieneHijo.Abogado)

```
<daml: Restriction daml:minCardinalityQ="2">
<daml:onProperty rdf: resource="#tieneHijo"/>
<daml: hasClassQ rdf: resource="#Abogado"/>
</daml:Restriction>
```

Una importante característica de DAML+OIL es que, además de las clases “abstractas” definidas por la ontología, se pueden usar tipos de datos XML esquema en hasClass, hasValue, y cardinalidad.

Por ejemplo, la clase Adulto puede ser expresada por la equivalente Persona \*  $\exists$ edad.Mayor17, donde Mayor17 es un tipo de dato esquema XML basado en decimal, con la restricción que los

valores deben ser por lo menos 18. Usando una combinación de esquema XML y RDF esto puede escribirse como:

```
<xsd:simpleType name="Mayor17">
  <xsd:restriction base="xsd:positiveInteger">
    <xsd:minInclusive value="18"/>
  </xsd:restriction>
</xsd:simpleType>
<daml:Class rdf:ID="Adulto">
  <daml:intersectionOf rdf:parseType="daml:
    collection">
    <daml:Class rdf:about="#Persona"/>
    <daml:Restriction>
      <daml:onProperty rdf:resource="#age"/>
      <daml:hasClass rdf:resource="#sobre17"/>
    </daml:Restriction>
  </daml:intersectionOf>
</daml:Class>
```

Una ontología DAML+OIL consiste de un conjunto de axiomas. El cuadro 2 resume los axiomas soportados por DAML+OIL.

Tabla 2. DAML+OIL axiomas

Axioma	DL Sintaxis	Ejemplo
subClassOf	$C_1 \mu C_2$	Humano $\mu$ Animal * Bipedo
sameClassAs	$C_1 \equiv C_2$	Hombre $\equiv$ Humano * Masculino
subPropertyOf	$P_1 \mu P_2$	tieneHija $\mu$ tieneHijo
samePropertyAs	$P_1 \equiv P_2$	costo $\equiv$ precio
disjointWith	$C_1 \mu \neg C_2$	Masculino $\mu \neg$ Femenin
sameIndividualAs	$\{x_1\} \equiv \{x_2\}$	{presidente_Fidel} $\equiv$ {F_Castro}
differentIndividualFrom	$\{x_1\} \mu \neg \{x_2\}$	{Jesus} $\mu \neg$ {Pedro}
transitiveProperty	$P \in R_+$	tieneAncestro <sup>+</sup> $\in R_+$
uniqueProperty	$\diamond \mu (\beta 1 P \cdot \diamond)$	$\diamond \mu (\beta 1 \text{ tieneMadre} \cdot \diamond)$
unambiguousProperty	$\diamond \mu (\beta 1 P^- \cdot \diamond)$	$\diamond \mu (\beta 1 \text{ esMadreDe}^- \cdot \diamond)$

Estos axiomas permiten determinar la subsumición o equivalencia entre clases o propiedades, la disyunción de clases, la equivalencia o no equivalencia de individuos, y varias propiedades de propiedades (roles DLs). En particular, es posible establecer que una propiedad es única (funcional), sin ambigüedades (su inverso es funcional) o transitiva. Esto muestra que, excepto para individuos y tipos de datos, los constructores y axiomas de DAML+OIL pueden ser trasladados en SHIQ. DAML+OIL es equivalente a la extensión de SHIQ más individuos.

## 5. CONCLUSIONES

El énfasis en la investigación de las DLs sobre una semántica formal, basada en la lógica y una investigación cuidadosa sobre los problemas básicos de razonamiento, junto con la disponibilidad de sistemas altamente optimizados para muy expresivas DLs, hacen esta familia de formalismos de representación de conocimientos un punto de partida ideal para la definición de lenguajes ontológicos para la web semántica.

Los servicios de razonamiento requeridos para soportar la construcción, integración, y evolución de ontologías de alta calidad son proporcionados por el estado del arte de los sistemas de DL para lenguajes muy expresivos. Al ser estos lenguajes utilizados en la práctica, requieren herramientas basadas en DL que soporten la adquisición de conocimiento (construcción de ontologías), mantenimiento (evolución de la ontología), e integración de ontologías.

Los primeros pasos en esta dirección ya se han dado. Por ejemplo OilEd es una herramienta que soporta el desarrollo de ontologías OIL y DAML+OIL, y ICOM es una herramienta que soporta el diseño e integración de relación-entidad y diagramas UML.

Como trabajo futuro para este estudio esta el determinar formalmente el tipo de razonamiento en SHIQ, presentar un análisis de sus extensiones y variaciones.

## REFERENCIAS

- Baader, I. Horrocks, and U. Sattler. Description logics as ontology languages for the semantic web. In D. Hutter and W. Stephan, editors, Festschrift in honor of Jorg Siekmann. LNAI. Springer-Verlag, 2003.
- Berners-Lee, T. (1999). Weaving the Web. Harper, San Francisco.
- Berners-Lee, J. Hendler, and O. Lassila. The semanticWeb. Scientific American,284(5): 34-43, 2001.
- Fensel, F. van Harmelen, M. Klein, H. Akkermans, J. Broekstra, C. Fluit, J. van der Meer, H.-P. Schnurr, R. Studer, J. Hughes, U. Krohn, J. Davies, R. Engels, B. Bremdal, F. Ygge, T. Lau, B. Novotny, U. Reimer, and I. Horrocks. On-To-Knowledge: Ontology-based tools for knowledge management. In Proceedings of the

- eBusiness and eWork 2000 (eBeW'00) Conference, 2000
- Gruber T., "A Translation Approach to Portable Ontology Specification", Knowledge Acquisition, 5 (2), 199-220, CA, 1993.
- Lizcano, L., Ojeda, R.: Fundamento formal de las Lógicas Descriptivas. Revista de Ingeniería de UNAL (2001)
- Schmidt-Schaub, M., Smolka, G.: Attributive concept descriptions with complements. Artificial Intelligence. 48 (1991) 1-26.
- Sattler, U. (2000). Description logics for the representation of aggregated objects. In Horn, W. (Ed.), Proceedings of the 14th European Conference on Artificial Intelligence (ECAI2000), pp. 239-243 Berlin, Germany. IOS Press Amsterdam.
- Studer S, Benjamins R., and Fensel D., "Knowledge Engineering: Principles and Methods," Data and Knowledge Engineering, 25, 161-197, 1998.
- Stevens, I. Horrocks, C. Goble, and S. Bechhofer. Building a reason-able bioinformatics ontology using OIL. In Proceedings of the IJCAI-2001 Workshop on Ontologies and Information Sharing, pages 81-90, 2001.
- Sattler. Description logics for the representation of aggregated objects. In W.Horn, editor, Proceedings of the 14th European Conference on Artificial Intelligence. IOS Press, Amsterdam, 2000.