

MODELO UNIFICADO DE REPRESENTACIÓN DE CONOCIMIENTO EN MAPAS CONCEPTUALES Y ONTOLOGÍAS

Alfredo Simón¹, Luigi Ceccaroni², Steve Willmott², Alejandro Rosete¹, Vivian Estrada³, Vladimir Lara⁴,
¹ Instituto Superior Politécnico “José Antonio Echeverría”, La Habana, Cuba, ¹ Universidad Politécnica de Cataluña, Barcelona, España, ³ Ministerio de Educación Superior, La Habana, Cuba, ⁴ Universidad de Costa Rica, San José, Costa Rica
Email: {asimon, rosete}@ceis.cujae.edu.cu, {luigi, steve}@lsi.upc.edu, vestrada@reduniv.edu.cu, vlara@ecci.ucr.ac.cr

Resumen. Se presenta una propuesta de integración entre un modelo conceptual gráfico, como es el caso de los mapas conceptuales, y ontologías codificadas en el lenguaje OWL. Los mapas conceptuales son una forma flexible de representación del conocimiento, muy útil en entornos de colaboración de humanos; OWL es un lenguaje de representación del conocimiento orientado a análisis semántico por parte de máquinas. La integración consiste en un conjunto de métodos de transformación entre mapas conceptuales y el correspondiente código OWL, en los dos sentidos. Estos métodos incluyen algunos procedimientos de eliminación de ambigüedades existentes y otros nuevos.

1 Introducción

En la representación del conocimiento orientada al análisis semántico por parte de máquinas (un contexto en el que se necesita cierto grado de formalidad) se está viendo cada vez más en ascenso el desarrollo y utilización de las ontologías. La formalización de ontologías por parte de diferentes agentes genera normalmente el surgimiento de conflictos en las conceptualizaciones que se pueden solucionar mediante la colaboración del hombre o por medio de métodos automatizados en casos simples (Hayes et al., 2003). Los procesos de diseño y creación de ontologías, así como las herramientas disponibles para su mantenimiento, como por ejemplo Protégé (The Protégé Project, 2000), son aún complejos para personas no expertas en el tema. Si a esto se le añade la complejidad inherente a los lenguajes de especificación, se puede afirmar que hoy existe cierta dificultad para la creación de un ambiente colaborativo para el desarrollo de ontologías por parte de personas no expertas.

Lo anterior sugiere la utilización de otra forma de representación, que de forma integrada con las ontologías pueda ser utilizada con naturalidad por el hombre y a la vez mantenga las características de poder ser interpretada computacionalmente. Se propone aquí la utilización de los **mapas conceptuales (MC)** como forma de representación a integrar con las ontologías. Los MC son una herramienta definida para el aprendizaje del hombre, de fácil construcción, flexible e intuitiva para el hombre. Teniendo en cuenta todo esto y a pesar de su bajo nivel de formalización, se ha logrado una integración de los MC con las ontologías, específicamente con aquellas codificadas en OWL (acrónimo del inglés *Web Ontology Language*, un lenguaje de marcado para publicar y compartir datos usando ontologías en Internet), facilitando así la gestión del conocimiento por parte de humanos. En este artículo se resumen los primeros resultados de dicha integración, que denominaremos *modelo unificado (MU)* y que está constituido de procedimientos de transformación de MC a OWL, y viceversa. Estos procedimientos incluyen la especificación de la formalización de los MC y algoritmos de desambiguación.

2 Las ontologías y sus lenguajes

En el campo de la inteligencia artificial, las ontologías fueron definidas para compartir y reutilizar conocimiento. Aportan un lenguaje de comunicación necesario en entornos distribuidos que involucran agentes software, como la Web semántica, y una descripción semánticamente formal para el procesamiento del conocimiento. Según la definición de Ceccaroni (2001), derivada de Gruber (1993), Borst (1997) y Studer et al. (1998), una ontología (o mapeo semántico) es *una especificación formal y explícita de una conceptualización compartida, orientada a análisis automático por parte de máquinas*. En sentido general, una ontología es la base del procesamiento semántico; es una red de **conceptos, relaciones y axiomas** para representar, organizar y entender un dominio de conocimiento; proporciona el marco de referencia común para todas las aplicaciones en cierto entorno. Los orígenes de las formalizaciones ontológicas se encuentran en las redes semánticas (RS) y la teoría de marcos de Minsky (1975), y en lenguajes como Ontolingua (Farquhar et al. 1995) y DAML+OIL (DARPA, 2001), entre otros, la mayoría de los cuales están basados en el más sencillo RDF. La última evolución de los lenguajes para codificar ontologías consiste en la definición de OWL (Smith et al, 2004), para el cual existen tres especificaciones, con

diferentes niveles de expresividad: OWL Lite, OWL DL y OWL Full. El modelo que se propone en la presente investigación tiene un alcance aplicable tanto a OWL Lite como a OWL DL.

3 La integración entre mapas conceptuales y ontologías: el modelo unificado

El MU propuesto se puede entender como la integración de un lenguaje gráfico poco formal (MC) y un sistema de ontologías codificado en el lenguaje formal OWL, un lenguaje de gran complejidad de interpretación por personas no expertas. Los MC fueron definidos dentro de la pedagogía y responden a una notación gráfica que ayuda a que el conocimiento pueda ser fácilmente compartido, manipulado e interpretado por los seres humanos. Desde el punto de vista estructural, existe una gran similitud entre MC, RS y ontologías basadas en RDF, siendo estas últimas más formalizadas: el lenguaje RDF está formalizado mediante *triples* (*sujeto, predicado, objeto*) mientras los MC usan la estructura de *proposiciones* (*nodo, palabra-enlace, nodo*).

Como base inicial de la integración entre MC y ontologías, se realizó un estudio comparativo entre sistema de representación estructurados de manera similar: las RS, los sistemas de marcos y del lenguaje RDF. Se encontró que para una integración entre MC y ontologías es necesaria una mayor formalización de las *palabras-enlace* (p-e), ya que son el elemento sintáctico que define el tipo de relación existente entre los conceptos, elemento imprescindible para la obtención del código OWL. Los resultados alcanzados en el análisis de las diferentes formas de representación del conocimiento permitieron definir que hay una correspondencia directa entre las *clases* y *propiedades* en las ontologías, y los *conceptos* y *relaciones* en los MC. Conjuntamente, se han agrupado las relaciones de los MC por categorías de p-e, que deben satisfacer las necesidades semánticas requeridas para la obtención del código OWL. Específicamente, se proponen las siguientes categorías de p-e:

Categorías	Ejemplos de palabras-enlace
Subclasificación (CSC)	es_un ¹ , tiene_por_subclase, tiene_parte_a, tiene_dependencia, incluye, agrupa, se_compone_de, comprende_a, puede_ser
Instanciación (CI)	tiene_por_instancia, instancia_como, tiene_ejemplo, instancia_de
Propiedad (CP)	tiene, posee, tiene_propiedad, toma_valor, tiene_valor, se_compone_de
Propiedad-Valor Directa (CPVD)	Sustantivos, por ejemplo: tipo, pared, rueda, etc.
Propiedad-Valor Indirecta (CPVI)	Formas Verbales, por ejemplo derivadas de: contener (contenido), ejercer (ejerce), representar (representa), etc.

Tabla 1. Categorías de palabras de enlace

En esta propuesta de p-e se ha tenido en cuenta la posibilidad de relacionar dos conceptos en las dos direcciones, incluyendo un conjunto de p-e simétricas ha algunas de las incluidas en esa categoría pero a ser utilizadas en el sentido contrario, tal es el caso de: es_un e instancia_de. Lo anterior es de gran utilidad ya que amplía las posibilidades de relación entre conceptos, por ejemplo: (Auto, tiene_ejemplo, Fiat 126p) y (Fiat 126p, instancia_de, Auto); (Auto, agrupa, Fiat) y (Fiat, es_un, Auto).

La cuestión principal aquí está en *¿cómo identificar la semántica que transmite la p-e que se utiliza para construir el código OWL?* Una solución que se propone es la de analizar al MC como un texto estructurado, ya que, a pesar de que responde a una forma gráfica de representación de conocimiento, cada una de las proposiciones en el incluidas es una frase en lenguaje natural. Es posible así complementar el MU con la aplicación de técnicas de *procesamiento de lenguaje natural* (PLN), como es el caso de algoritmos de desambiguación léxica para, a partir del sentido correcto de cada concepto del MC, determinar la naturaleza de las relaciones entre ellos, centrando la atención en las relaciones del tipo: *hiperonimia*¹-*hiponimia*², *meronimia*³-*holonimia*⁴ (Fellbaum, 1998). La integración entre el MC (lenguaje visual) y OWL (lenguaje formal) se puede formular como un proceso de

¹ Relación semántica de inclusión entre unidades léxicas, que va de lo más general a lo más específico.

² Relación semántica de inclusión de unidades léxicas que va de lo más específico a lo más general, lo opuesto de *hiperonimia*. Ej: “perro” se encuentra en relación de hiponimia con “animal”.

³ Relación semántica asociada a: parte de, es sustancia de, es miembro de, etc. Ej: X es merónimo de Y si X forma parte de Y; si X es una sustancia de Y; si X es un miembro de Y; por ejemplo: “azul” es merónimo de “color”; “doctor” es merónimo de “oficio”; “dedo es un merónimo de mano”.

⁴ Relación semántica opuesta a la meronimia, Ej: “bicicleta” es un holónimo mientras que “sillín” y “pedal” son merónimos de bicicleta.

transformación bidireccional; mecanismos complementarios de PLN para el análisis del MC contribuyen a que este no pierda una de sus principales bondades para el hombre que es su flexibilidad.

4 Algoritmos de desambiguación

El procedimiento que propone la presente investigación para la desambiguación de los conceptos en MC es una adaptación del procedimiento de Pons-Porrada et al. (2003), agregando algunos elementos de la propuesta de Cañas et al. (2003), como la medida de distancia entre conceptos para definir el contexto. El MU se basa en estos y otros trabajos previos y en la ontología léxica orientada a análisis por parte de máquinas WordNet (Fellbaum, 1998). WordNet tiene una estructura basada en *synsets* (conjuntos de uno o más sinónimos) relacionados como en una RS, los cuales están relacionados por medio de diferentes tipo de relación (*sinonimia*, *antonimia*⁵, *hiperonimia*, *meronimia*, *implicación*, etc.). Cada *synset* tiene asociado uno o varios dominios muy útil en el proceso de desambiguación (Magnini et al., 2002, Pons-Porrada et al., 2003). Se pueden mencionar dos aspectos que, a opinión de los autores, son de gran importancia en un proceso de desambiguación conceptual en un MC: (1) el trabajo con los dominios a los cuales pertenecen los conceptos, en correspondencia con los *Dominios de Magnini* (Magnini et al., 2002) y (2) el uso de la mayor cantidad de tipos de relación de las presentes en WordNet. Ambos aspectos permiten aumentar la precisión en la identificación del sentido de un concepto dentro de un contexto y contrarrestan en alguna medida las limitantes presentes en las ontologías léxicas, particularmente en lo referente a la cantidad de términos almacenados. En el trabajo se utiliza un WordNet en español incluido en EuroWordNet (Vossen, 1997), para extender su campo de aplicación a este idioma.

En este sentido la propuesta descrita en (Cañas et al. 2003) muestra algunas limitantes, por ejemplo: no hace uso de los dominios y se utilizan solo relaciones de *hiperonimia*, además no especifica como se lleva a cabo la desambiguación del *concepto-principal*. Por el contrario, en el procedimiento de desambiguación descrito en (Pons-Porrada et al. 2003) sí se tiene en cuenta los dominios asociados a cada *synset* y una mayor cantidad de tipos de relación, sin embargo su aplicación directa a un MC no aprovecharía la simplificación de la complejidad de ese proceso que se pudiera lograr a partir del hecho de que un MC muestra una estructura resumida de un texto. El procedimiento propuesto en la presente investigación se describe a continuación.

Sea la notación $C_c(c)$ el contexto de un concepto c en un MC que tiene por significado: el conjunto de conceptos que están incluidos en las proposiciones del MC en las se encuentra c y aquellos que se encuentran a una distancia $D = 2$ a partir de c . Siendo D la distancia del camino (proposición) a recorrer de un concepto c_1 a un concepto c_2 cualquiera. Sean $G = \langle S, R \rangle$, el grafo no orientado asociado al grafo de WordNet y $s \in S$, un *synset*. Llamaremos *d-vecindad* de s al conjunto de todos los *synset* s' tal que existe una cadena de s a s' , siendo d la distancia a recorrer de s a s' en G . En acuerdo con Pons cada relación semántica de WordNet tiene un impacto diferente en la desambiguación de un término, por tal razón se ha tomado la función de peso $w(R)$ sobre dichas relaciones que se propone en (Pons-Porrada et al, 2003): se asigna un peso de 10 a la relación de *sinonimia*, 8 a las relaciones *near_synonym*, *xpos_near_synonym*, 5 a las de *hiperonimia* o *meronimia*, 1 a la de *antonimia*⁵ y 2 a las restantes. La votación que recibe un *synset* s del concepto c con respecto al contexto $C_c(c)$ y la *d-vecindad* de s se define como:

$$V_s(C_c, d) = \sum_{s' \in \Omega} v(s')/d$$

$$\text{Donde } \Omega = C_c(c) \cap d\text{-vecindad}(s) \text{ y } v(s') = \prod_{j=1}^d w(R_j)$$

siendo R_j las relaciones que etiquetan las aristas que forman el camino de longitud j entre s y s' .

Relacionar los synsets a los conceptos del MC.

1. Asignar a cada concepto c una lista de todos los *synset* $S(c)$ en los que aparece en WordNet e incorporarlos a una lista de “*Conceptos no desambiguados*” (CND).
2. Se pasa a una lista de “*Conceptos ya desambiguados*” (CD) todos aquellos conceptos que tengan un solo *synset* en $S(c)$ conjuntamente con el *synset* correspondiente.

Determinar el dominio del concepto-principal.

3. Si *concepto-principal* \in CD entonces:

⁵ Relación semántica de oposición entre los significados de dos palabras: las palabras “útil” e “inútil” tienen una relación de antonimia.

- a. ese dominio sería el correspondiente a ese *synset* y será identificado como DCP (Dominio Concepto Principal)
- si no:
 - b. Para cada $s \in S(c)$ calcular su votación ($V_s(C, d)$). El *synset* correspondiente al *concepto-principal* sería el que mayor votación tendría. En este caso DCP sería el dominio correspondiente al *synset* resultante que coincida con el que más se repita dentro de los s' incluidos en $d\text{-vecindad}(s)$. Es poco probable que varios *synsets* tengan la misma votación, en ese caso se tomaría como DCP el dominio incluido en cada uno de ellos que más se repita dentro del conjunto de s' incluidos en esa vecindad.

Desambiguar por el dominio.

4. Para cada concepto $c \in \text{CND}$, si alguno de sus $s \in S(c)$ incluye por dominio a DCP, pasar este concepto conjuntamente con ese *synset* a la lista de CD.

Desambiguar los conceptos pendientes (en caso de que CND no esté vacía).

5. Para cada $c \in \text{CND}$, realizar el mismo procedimiento relacionado con la votación de los *synset* $s \in S(c)$ correspondiente a cada concepto c que se describe en el paso 3b.

5 El proceso de transformación en el Modelo Unificado

En la explicación del proceso de transformación que se proponen en el MU serán utilizados los dos ejemplos de MC que se muestran en las siguientes figuras. El primero (Figura 1) es un MC de un dominio general dentro del contexto de la anatomía del cuerpo humano y el segundo (Figura 2) es un MC obtenido de un fragmento de la ontología *Activities.owl* utilizada en el contexto del proyecto @LIS TechNET⁶ (Ceccaroni et al., 2005).

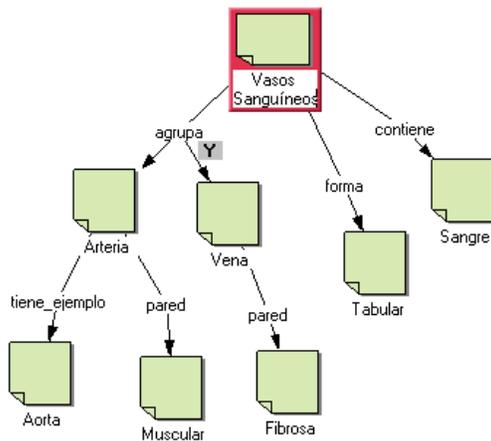


Figura 1. Mapa conceptual de un dominio general (Vasos Sanguíneos)

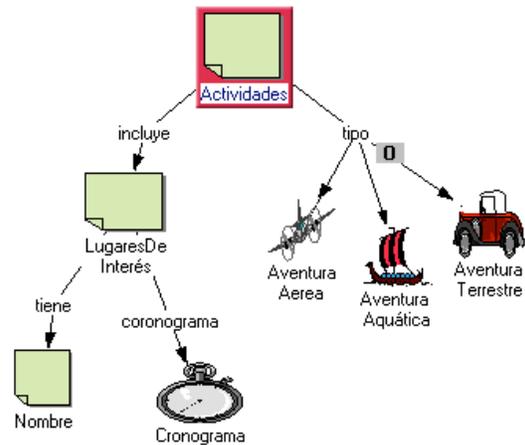


Figura 2. Mapa conceptual obtenido de un fragmento de la ontología *Activities* utilizada en el proyecto @Lis TechNET.

5.1 La obtención del OWL a partir del MC

Fase 1. Desambiguación de conceptos. En esta etapa inicial se realiza el proceso de identificación de los sentidos correctos de todos los conceptos del MC utilizando el método descrito en la sección anterior. Al culminar el proceso, se tendrá el dominio al que pertenece el MC que generalmente se corresponde con el del *concepto-principal* y que a su vez debe coincidir con el que más se repite entre los conceptos del MC, aunque esto depende mucho del nivel de especialización del conocimiento. Además del dominio del MC se conocen los sentidos (*synset*) de todos o la mayoría de los conceptos del MC, también en dependencia del nivel de especialización del MC. Con esta información se pasa la próxima fase con una lista **LP** que incluye todas las proposiciones que conforman el MC y cada concepto es acompañado de su *synset*.

⁶ Proyecto @LIS TechNET del programa EuropeAid [http://europa.eu.int/comm/europeaid/projects/alis/index_es.htm]

Fase 2. Conceptualización y lista de términos OWL. Todos los *nodos-conceptos* son expresados inicialmente como *clases* en OWL (*owl:Class*). Tomando como ejemplos a conceptos que se muestran en la Figura 2 la codificación en OWL sería:

```
<owl:Class rdf:ID = "Actividades">
  <rdfs:subClassOf rdf:resource = "Thing">... />
<owl:Class rdf:ID = "Aventura Aerea">
  ...
```

El *concepto-principal*, en este ejemplo “Actividades” se expresa como *subclase* de la clase “Thing” (clase más abstracta dentro de de cualquier ontología).

Fase 3. Identificación de las Relaciones de Subclasificación. Para cada proposición $p \in LP$ con sintaxis (c1, p-e, c2):

1. Si $p-e \in CSC$, **c1** es expresado en OWL como clase y **c2** como subclase de **c1**. Codificación en OWL utilizando como ejemplo el MC de la Figura 1:

```
<owl:Class rdf:ID = "Vena" >
  <rdfs:subClassOf rdf:resource = "Vasos Sanguíneos" />
  ...
```

2. Si $p-e \notin CSC$ es necesario recurrir al uso de WordNet.

Sea **s1** y **s2** los sentidos de **c1** y **c2** respectivamente:

- a. Generar todas las cadenas de *synset* de relaciones de *hiponimia* a partir de **s1** y almacenarlas en la lista **LHipo(c1)** y todas las cadenas de *synsets* de relaciones de *hiperonimia* a partir de **s2** y almacenarlas en la lista **LHiper(c2)**
- b. Si $s2 \in LHipo(c1)$ o $s1 \in LHiper(c2)$, significa que existe algún camino de *hiponimia* o de *hiperonimia* entre **s1** y **s2** respectivamente, lo cual permite inferir que **c2** es una *subclase* de **c1**.

Esto se puede comprobar en las proposiciones (Vaso Sanguíneo, agrupa, Arteria) y (Vaso Sanguíneo, agrupa, Vena) en MC de la Figura 1, donde Arteria y Vena tienen una relación de *hiponimia* con Vaso Sanguíneo y se codificaría en OWL de la misma forma que se muestra en el punto anterior.

Fase 4. Identificación de las Relaciones de Instancia. Se hace cuando en las proposiciones con sintaxis (c1, p-e, c2) $p-e \in CI$ y **c2** es un concepto hoja (no se encuentra en ninguna proposición como concepto origen) y se codificaría en OWL a **c2** como una *instancia* de **c1**. Tomando como ejemplo la proposición (Arteria, tiene_ejemplo_a, Aorta) (Ver Figura 1.) el código OWL que se genera sería:

```
<owl:Class rdf:ID = "Arteria">
  ...
<Arteria rdf:ID = "Aorta" >
  ....
</Arteria>
```

Fase 5. Identificación de las Relaciones de Propiedad. Dada la proposición $p \in LP$, con sintaxis (c1, p-e, c2):

1. Si $p-e \in CP$, se interpretaría con una sintaxis (clase, p-e, propiedad) y se puede inferir que **c2** es una es un tipo de propiedad de **c1**. Utilizando la proposición (LugaresDeInterés, tiene, Nombre) incluida en el MC de la Figura 2 la codificación sería:

```
<owl:Class rdf:ID = "LugaresDeInterés" >
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:DataTypeProperty rdf:ID = "Nombre"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
```

2. Si $p-e \in CPVD$, se interpretaría con una sintaxis (clase, propiedad, valor) y se puede inferir que la p-e es una propiedad de **c1** y c2 es el valor que toma dicha propiedad. Un ejemplo de esta situación sería la proposición (LugaresDeInterés, conograma, Cronograma) y la codificación en OWL sería:

```

<owl:ObjectProperty rdf:about = "#cronograma">
  <rdf:type rdf:resource="&owl; FunctionalProperty"/>
  <rdfs:domain rdf:resource="#LugaresDeInterés"/>
  <rdfs:range rdf:resource="#Cronograma"/>
</owl: ObjectProperty>

```

Si no existe otra relación de **c1** con otro concepto por medio de la misma **p-e**, se puede inferir que la propiedad podrá tomar un solo valor por lo que se ha considerado incorporar a la codificación de la propiedad la característica: `<rdf:type rdf:resource="&owl; FunctionalProperty"/>`

3. Si **p-e** ∈ **CPVI**, se interpretaría con una sintaxis (clase, propiedad indirecta, valor) y se puede inferir que **c2** es el valor de la propiedad derivada de la **p-e** y perteneciente a **c1** y su codificación en OWL sería, tomando como ejemplo la proposición (Vasos Sanguíneos, contiene, Sangre) de la Figura 1:

```

<owl:Class rdf:ID = "Sangre"> . . . </owl:Class>
<owl:ObjectProperty rdf:ID = "contenido">
  <rdfs:domain rdf:resource = "#Vasos Sanguíneos"/>
  <rdfs:range rdf:resource = "#Sangre"/>
</owl:ObjectProperty>

```

En el MC pueden existir relaciones no binarias, como sucede en el caso de la Figura 1. Esta situación en particular es interpretada como que una propiedad puede tener asociado al menos uno de los que relaciona la **p-e**, y sería codificado en OWL de la siguiente forma:

```

<owl:Class rdf:ID = "Actividades">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource = "#tipo">
        <owl:someValueFrom rdf:resource = "#Aventura Aerea"/>
        <owl:someValueFrom rdf:resource = "#Aventura Terrestre"/>
      ....
    </owl:Class>

```

En Gómez et al. (2004) se tiene en cuenta este mismo criterio pero para todos los tipos de propiedades, no siendo así en nuestro caso.

4. Si **p-e** ∉ {**CP**, **CPVD**, **CPVI**} es necesario recurrir al uso de WordNet.
 - Sea **s1** y **s2** los sentidos de **c1** y **s2** respectivamente:
 - a. Generar todas las cadenas de *synset* de relaciones de *meronimia* a partir de **s1** y almacenarlas en la lista **LMero(c1)** y todas las cadenas de *synset* de relaciones de *holonimia* a partir de **s2** y almacenarlas en la lista **LHolo(c2)**
 - b. Si **s2** ∈ **LMero(c1)** o **s1** ∈ **LHolo(c2)**, significa que existe algún camino de *meronimia* o de *holonimia* entre **s1** y **s2** respectivamente, lo cual permite inferir que entre **c1** y **c2** existe una relación de propiedad. La propiedad pertenece a **c1**, es identificada como la **p-e** y toma como valor a **c2**. La codificación en OWL coincide con lo expresado en el punto 2 dentro de esta misma fase.

5.2 La obtención del MC a partir de OWL

La obtención de un MC a partir de código OWL se realiza a través de las cinco fases siguientes:

- 1. Representación de la arquitectura básica.** Según las convenciones vistas, las clases (*owl:Class*) y las instancias (*X rdf:ID="Y"*) se representan como conceptos en el MC.
- 2. Identificación del concepto-principal.** El *concepto-principal* corresponderá a aquella *clase* que solo sea *subclase* de "Thing" y no sea una *propiedad* de ninguna *clase* o *subclase*. En caso de que exista más de una clase que cumpla con estos requerimientos, donde es mayormente frecuente que suceda en contextos muy generales, se seleccionaría la clase que más subclases tendría.
- 3. Representación de relaciones entre clases y subclases.** Se crean enlaces dirigidos, del nodo correspondiente a la *clase* al nodo correspondiente a la *subclase* y el enlace sería etiquetado con cualquiera de las *palabras-enlaces* incluidas en **CSC**.

4. Representación de las relaciones de instancia. Se crean enlaces dirigidos, del *nodo-concepto* asociado a la instancia en OWL al concepto asociado con la clase a la que corresponde dicha instancia. El enlace es etiquetado con cualquiera de las *palabras-enlaces* incluidas en **CI**.

5. Identificación de las relaciones según las propiedades (*owl:ObjectProperty*). Una propiedad puede tener uno o varios valores. En el primer caso, pueden ser expresada mediante en una proposición en el MC: los términos especificados en *rdfs:range* y en *rdfs:domain* serían los conceptos, siendo el origen el concepto especificado en *rdfs:domain*, y la *palabra-enlace* sería el término especificado en *owl:ObjectProperty*. En el segundo caso, se crearía un *nodo-concepto* por cada uno de los elementos especificados con *owl:someValueFrom*, también se representaría con un *nodo-concepto* la clase a la cual corresponde dicha propiedad y se crearía una relación dirigida a cada uno de los nodos-conceptos creados por la especificación de *owl:someValueFrom*, con origen en el *nodo-concepto* asociado a la clase y con una *palabra-enlace* compartida especificada según el término usado en *owl:onProperty*.

6 Implementación y trabajo futuro

Los procedimientos descritos han sido implementados en lenguaje C++. El proceso de transformación implementado parte de tener un MC del que sea posible generar una representación en XML o una ontología codificada en OWL en dependencia de cual sea la dirección de la transformación. En el caso de MC, este es elaborado con la herramienta MACOSOFT (Simón, 2006) y por medio de ella misma es generado el código XML correspondiente con un formato específico; con respecto a la ontología en OWL, puede ser obtenida de alguna de las herramientas para la construcción de ontologías existentes. Las fases concebidas en cada uno de los procedimientos de transformación son ejecutadas internamente de forma iterativa e incremental. Se cuenta con una versión de WordNet en español la cual es consultada durante el proceso de transformación por el componente. Los resultados preliminares obtenidos de las pruebas realizadas se ven reflejados en los ejemplos utilizados aquí. Se pueden dar casos en los cuales el proceso de transformación fracasa. Uno de estos casos es cuando se consulta WordNet y no se encuentra ningún *synset* que se corresponda con algún concepto, muy común en contextos de conocimiento muy específico y problema presente en todas las herramientas que consultan WordNet. En respuesta a esto se trabaja en dos variantes, una primera es hacer un mayor uso de las *palabras-enlaces* de dominios específicos y la segunda variante es hacer uso de la Base de Conocimiento que posee el Servidor de Mapas Conceptuales (SERVIMAP) en la cual se almacenan una gran cantidad de MC (Simón, 2006). Esos MC pueden ser utilizados para buscar proposiciones similares a la que se analiza y con esta nueva información hacer nuevas consultas en WordNet.

7 Trabajos Relacionados

En Gómez et al. (2004) se propone un mecanismo de transformación de un MC al lenguaje OWL. Se utiliza XTM como un lenguaje intermedio entre el MC y el OWL, relazándose de la siguiente forma: el MC se lleva a XTM, luego son aplicados un conjunto de reglas al XTM (Ej: (1) los conceptos expresados mediante *topic* en el XTM son expresados como *owl:Class rdfs:...* en el OWL y (2) los tipos de relaciones expresadas en XTM como *association* son expresadas como *owl:ObjectProperty rdfs:...*) y se obtiene el código OWL. Aquí se identificaron dos inconvenientes principales: (1) no se tiene en cuenta la semántica de la relación, todas las *associations* en XTM no siempre deben ser expresadas como propiedades en OWL; (2) el XTM es un lenguaje desprovisto de semántica en los elementos que expresan los tipos de relaciones del MC, por lo que mantienen en mismo problema presente en el MC de formalización del tipo de relación. En el caso de la propuesta de Hayes et al. (2004) se muestra un entorno de desarrollo colaborativo de ontologías basado en MC. Se parte de un RDF que es llevado a un MC, aprovechando la semejanza entre el *triples* de RDF (sujeto-predicado-objeto) y las *proposiciones* de los MC (concepto-enlace-concepto). La obtención del MC se realiza utilizando restricciones OWL para la depuración del RDF, *templates*, *estereotipos* para los *nodos* y un vocabulario restringido de p-e. Algunas diferencias identificadas: (1) se trabaja con documentos RDF y no OWL, (2) la representación del conocimiento en el MC mediante un vocabulario más restringido y menos natural para el hombre, (3) no se manifiesta una clara integración entre el lenguaje de ontologías y el MC.

8 Agradecimientos

Este trabajo ha sido elaborado como consecuencia de convenios de colaboración interuniversitarios con la ayuda económica de la Unión Europea, en el marco del Proyecto @LIS TechNET (ALA2002/049-055/2209) del programa EuropeAid [http://europa.eu.int/comm/europeaid/projects/alis/index_es.htm] y del Proyecto CORDIAL-2 (II0021FA) del Programa Alfa para el desarrollo de profesionales de América Latina y el Caribe. Los contenidos del presente documento, sin embargo, son sólo responsabilidad de los autores y pueden no reflejar la posición de la Unión Europea. También se agradece a Jordi Atserias del Departamento de Lenguajes y Sistemas Informáticos (LSI) de la Universidad Politécnica de Cataluña (UPC) por sus contribuciones con relación a EuroWordNet.

9 Conclusiones

Se obtuvieron las siguientes conclusiones en la investigación: (1) se demostró que existe una estrecha relación entre mapas conceptuales y ontologías; (2) la integración de mapas conceptuales (lenguaje visual) y OWL (lenguaje formal) crea las bases para el desarrollo de entornos de colaboración hombre-máquina en el trabajo con ontologías; (3) utilizar mecanismos complementarios de procesamiento de lenguaje natural para el análisis de mapas conceptuales permite lograr un nivel de interpretación semántica de sus relaciones muy útil para la formalización de estos en código OWL manteniendo la flexibilidad en su representación; (4) los procedimientos definidos garantizan una estrecha integración de mapas conceptuales con el código OWL en aquellos contextos en los que no se requiera de restricciones cuantificables, ni elementos de negación.

10 Referencias

- Borst, P. (1997). *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. Ph. D. thesis, University of Twente, The Netherlands.
- Cañas, A., Valerio, A., Lalinde, J., Carvalho, M. y Arguedas, M. (2003). *Using WordNet for Word Sense Disambiguation to Support Concept Map Construction*. Proc. of 10th International Symposium on String Processing and Information Retrieval. Ed. Springer-Verland. Manaus, Brasil.
- Ceccaroni L. (2001). *ONTOWEDSS - An Ontology-Based Environmental Decision-Support Systems for the Management of Wastewater Treatment Plants*. Ph. D. thesis, Universidad Politécnica de Cataluña, Barcelona, España. pp. 14.
- Ceccaroni, L., Willmott, S., Cortés García, U. y Barbera-Medina, W. (2005). @LIS TechNET: Hacia la enseñanza práctica de las tecnologías de Internet de la próxima generación. *Actas de la 5ª Conferencia Internacional de la Educación y la Formación basada en las Tecnologías (Online Educa Madrid 2005)*. Madrid, España. pp. 139-142 URL: <http://docs.alis-tech.net/alis/pubs/technical-papers/alis-tech.net-proxima-generacion/>
- DARPA (2001). DAML+OIL ontology markup language. *Defense Advanced Research Projects Agency*. URL: <http://www.daml.org/2001/03/daml+oil.daml>
- Farquhar, A., Fikes, R., Rice, J. (1995). The Ontolingua Server: a Tool for Collaborative Ontology Construction. URL: <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/farquhar/farquhar.htm>
- Fellbaum, Ch. (1998). *WordNet: An Electronic Lexical Database*, The MIT Press, University of Cambridge.
- Gómez, H., Díaz B. y González, A. (2004). Two-layered approach to knowledge representation using conceptual maps description logic, *Proc. of the First International Conference on Concept Mapping*. España.
- Gruber, T. (1993) *Toward Principles for the Design of Ontologies used for Knowledge Sharing*. Formal Ontology in Conceptual Analysis and Knowledge representation. Guarino y Poli Ed., Kluwer Academic Publishers.
- Hayes, P., Eskrindge, T., Reichherzer, T., Saavedra, R. (2004). A Framework for Constructing Web Ontologies using Concept Maps. *Proc. DALM Meeting*.
- Magnini, B., Strapparava, C. y Glifio, A. (2002). *The Role of Domain Information in Word Sense Disambiguation*. Natural Language Engineering. Cambridge University Press. pp. 359-373.
- Minsky, M. (1975). *A Framework for Representing Knowledge*. The Psychology of Computer Vision, McGraw-Hill, pp. 211-277.

- Novak J. D. y Gowin, D. B. (1984). *Learning how to learn*, Cambridge Press. New York.
- Pons-Porrada, A., Berlanga, R., Ruiz-Shulcloper, J. (2003). Un nuevo método de desambiguación del sentido de las palabras usando WordNet. *Proc. X Conferencia de la Asociación Española para la Inteligencia Artificial*. País Vasco.
- The Protégé Project, *University of Stanford*, 2000, URL: <http://protege.stanford.edu/>
- Smith, M., Welty, Ch., McGuinness, D. (2004). OWL Web Ontology Language Guide. W3C. URL: <http://www.w3c.org/TR/owl-guide/>
- Simón, A., Willmott, S., Rosete, A., Ceccaroni, L. (2005). Modelo Unificado para la integración entre Ontologías y Mapas Conceptuales. *Reporte de Investigación*. Universidad Politécnica de Cataluña. LSI-05-64-R.
- Simón, A. (2006). Plataforma para la Gestión del Conocimiento con Mapas Conceptuales. *Tesis de Maestría*. (Tutores: Rosete, A, Estrada, V.) Centro de Estudios de Ingeniería y Sistemas (CEIS). Instituto Superior Politécnico “José Antonio Echeverría” (Cujae), C. Habana, Cuba.
- Studer, R., Benjamins, V. R., Fensel D. (1998). *Knowledge engineering: principles and methods*. Data and Knowledge Engineering, 25(1-2), 161–197.
- Vossen P. (1997). EuroWordNet: a multilingual database for information retrieval. *Proc. DELOS Workshop on Cross-Language Information Retrieval*. Switzerland.