

Construcción basada en ontologías del Léxico Extendido del Lenguaje

Héctor J. Ruidías^{1,2}, María Laura Caliusco², and María R. Galli³

¹ Departamento de Ingeniería y Ciencias de la Producción, Universidad Gastón Dachary, Posadas, Argentina,

² CIDISI, Universidad Tecnológica Nacional, Facultad Regional Santa Fe, Argentina

³ INGAR, Consejo Nacional de Investigaciones Científicas y Técnicas
{hruidias,mcaliusc}@frsf.utn.edu.ar, mrgalli@santafe-conicet.gov.ar
<http://www.frsf.utn.edu.ar/cidisi/>

Resumen El éxito de un proceso de desarrollo de software depende en gran medida de la calidad de los documentos de especificación de sus requerimientos. Si bien, la especificación de los requerimientos en lenguaje natural facilita la comunicación con los “involucrados” (stakeholders), se dificulta la evaluación de la consistencia de los mismos. Por lo tanto, se requiere de modelos formales que faciliten esta evaluación. Las ontologías han surgido como herramientas para crear dichos modelos formales, sin embargo, la mayoría de las ontologías de documentos de requerimientos existentes son sólo simples taxonomías. En esta propuesta se presenta un marco de trabajo basado en ontologías para la construcción del Léxico Extendido del Lenguaje (LEL). Dicho marco de trabajo se compone de un metamodelo del LEL basado en una ontología “pesada”, es decir compuesta por axiomas de integridad y de derivación; y de una heurística para la construcción de un LeL a partir de dicha ontología. Además, se plantea una discusión sobre las ventajas y dificultades de dicho marco de trabajo mediante un caso de estudio.

Keywords: metamodelos en requerimientos, ontologías, LEL

1. Introducción

La primera medida de éxito para un sistema de información es el grado en el que éste logra el propósito para el que fue desarrollado. La Ingeniería de Requerimientos (IR) proporciona el mecanismo apropiado para descubrir dicho propósito entendiendo los deseos del cliente, sus necesidades, evaluando la factibilidad de la satisfacción de éstas, y finalmente negociando una solución razonable, especificando la solución sin ambigüedades, validando la especificación y gestionando los requerimientos conforme éstos se transforman en un sistema operacional [13].

La IR propone adquirir conocimiento sobre el Universo del Discurso (UdeD) capturando las demandas, necesidades y problemas presentes en él y, luego, determinar que los requisitos especificados para el sistema de software sean los correctos. La IR tiene una interacción muy fuerte con aquellos que demandan

un producto de software. Por lo tanto, el proceso de la IR implica un esfuerzo cooperativo que involucra a usuarios, clientes, ingenieros de software, consultores y expertos, los que en la literatura se denominan “stakeholders” o “involucrados” en español. Por lo tanto, los ingenieros de requisitos deben comprender, modelar y analizar el UdeD donde el software operará y los “involucrados” deben confirmar que la visión de los ingenieros es correcta [6].

Para llevar adelante sus actividades el ingeniero de requisitos cuenta con una serie de métodos y técnicas que involucran la construcción de documentos de especificación de requerimientos, tales como el LEL (Léxico Extendido del Lenguaje), escenarios y casos de uso. Dichos documentos se escriben principalmente en lenguaje natural dado que es una forma de ayudar a la comunicación entre el ingeniero de requisitos y los “involucrados” [9]. Como los requerimientos se especifican mayormente en forma textual, se presenta uno de los primeros problemas en relación a la verificación de los documentos de especificación de requerimientos. Dichos documentos especificados en forma textual suponen una enorme dificultad al momento de ser interpretados apropiadamente por una máquina, lo que conduce a que los mismos no puedan ser sometidos a un proceso automático o semiautomático que verifique su correctitud, claridad y completitud.

Para solucionar los problemas anteriormente planteados se requiere de modelos formales. Con el advenimiento de la Web Semántica, surge una nueva forma de modelos que son las ontologías[5]. Desde una perspectiva pragmática, las mismas pueden ser definidas como un artefacto de representación de la semántica de un dominio dado. En el área de la IR se han definido diferentes trabajos que hacen uso de las ontologías[2]. En cuanto al uso de las ontologías como modelos de los documentos de especificación de requerimientos se pueden encontrar trabajos que proponen simples taxonomías y que por lo general requieren de un conocimiento especializado en el formalismo subyacente, y que además no presentan reglas de derivación que permitan verificar la correctitud, claridad y completitud de los mismos.

El objetivo del presente trabajo es presentar un marco de trabajo para la construcción del LEL, basado en una ontología “pesada”, es decir, compuesta por axiomas de integridad y derivación, y de una heurística de construcción del mismo. Para ello, el paper se organiza de la siguiente forma: en la Sección 2 se presenta la problemática del LEL. En la sección 3 se introduce la ontología construida a partir del metamodelo del LEL, así como los aspectos formales considerados. En la Sección 4 se presenta la propuesta de utilizar ontologías en la etapa de IR mediante una heurística que permite aprovechar las ventajas de la formalidad subyacente. En la Sección 5 se plantea un caso de estudio para ilustrar la aplicación de la propuesta y se analizan las ventajas y dificultades que se han presentado. Finalmente, en la Sección 6 se presenta un último apartado con las conclusiones del trabajo.

2. Vinculación entre ontologías y LEL

2.1. Ontologías y usos en la ingeniería de requerimientos

La definición más popular de una ontología es aquella que la define como “una especificación formal y explícita de una conceptualización compartida” [4]. Se destacan en esta definición dos ideas claves, la de conceptualización compartida y la de especificación formal. La primera refiere a un modelo abstracto, con sus elementos y la vinculación entre éstos explícitamente definida, que además cumple con el requisito de ser consensuado y aceptado por una comunidad. La segunda, en tanto, refiere a la cuestión del lenguaje con que se representa, de tal manera que sea interpretable de igual manera tanto por personas como por computadoras. Por lo tanto, las ontologías aportan un esquema formal de especificaciones explícitas sobre las conceptualizaciones del dominio. Éste es el motivo principal por el que las ontologías se consideran tecnologías apropiadas para la interoperabilidad de sistemas heterogéneos, el desarrollo y uso de sistemas inteligentes, la gestión del conocimiento y la comprensión de la información que se difunde, como una forma de reducir la brecha existente entre los usuarios y los expertos de un dominio particular.

Desde un punto de vista pragmático, las ontologías son herramientas que permiten modelar el conocimiento a través de la clasificación de conceptos, reglas y axiomas propios del dominio de aplicación de manera formal, especificando el significado, el contexto y las relaciones de los términos de un vocabulario asociado a un determinado dominio, brindando un modelo abstracto de cómo se piensan determinadas cosas en el mundo real. Para explicitar dicha representación, una ontología se basa en los siguientes elementos [5]:

- Clases: Son las ideas a formalizar y representan los conceptos en el sentido más amplio. Las clases en una ontología se suelen organizar en taxonomías a las que se les pueden aplicar los mecanismos de herencia.
- Atributos: Representan propiedades de las clases.
- Relaciones: Representan las asociaciones, generalmente binarias, entre las clases. El primer argumento de la relación se conoce como el dominio y el segundo como el rango.
- Instancias: Se usan para representar elementos o individuos de una clase en una ontología.
- Axiomas: Los axiomas formales sirven para modelar sentencias que son siempre ciertas. Normalmente, se usan para representar conocimiento que no puede ser formalmente definido por los componentes descritos anteriormente. Además, también se usan para verificar la consistencia de la propia ontología.

Al modelar un dominio con ontologías, cuando se introduce un nuevo concepto es importante verificar si el mismo es consistente o contradictorio con lo definido previamente. Esta propiedad se conoce como el *concepto satisficible* (o respectivamente *insatisficible*). También, puede ser necesario saber si un concepto es más general que otro (subsunción), si son equivalentes o si son disjuntos.

Una vez definida una ontología, al definir las instancias, las propiedades más importantes que habrá que verificar son las de la consistencia de las instancias respecto a las clases definidas, y la derivación de una instanciación a partir de las instancias definidas previamente .

Desde un punto de vista de modelado, las ontologías pueden ser utilizadas como modelos o metamodelos[7]. La diferencia reside en que una ontología vista como un metamodelo necesita ser instanciada con las principales entidades del dominio en tanto que en una ontología vista como un modelo las clases representan las entidades principales de un dominio dado y son instanciadas con los objetos del dominio.

En la IR se maneja información proveniente de distintas fuentes, con una impronta de discrepancias en la convergencia de las visiones de los “involucrados” que a menudo difieren ya no únicamente en su perspectiva sobre un mismo aspecto del problema, sino también en el vocabulario empleado para describirlo. Las características de esta problemática ha atraído el interés sobre el uso de ontologías en esta disciplina [1] y se ha explorado su aplicación para 1) describir documentos de especificación de requerimientos, 2) la representación formal del conocimiento del dominio de la aplicación y 3) la representación formal de los requerimientos. Por lo tanto, la ontología presentada en este trabajo pertenece al segundo caso, con un metamodelo del LEL cuyas instancias son las principales entidades del UdeD que se está analizando. En este sentido, la ventaja reside en que luego dichas entidades pueden utilizarse para generar una ontología que represente un modelo del dominio analizado.

2.2. Metamodelo del LEL

Puede entenderse al Léxico Extendido del Lenguaje como un glosario cuya finalidad es la comprensión del lenguaje del UdeD. Cuenta con un metamodelo que define cómo se organizan los términos de dicho vocabulario, así como un proceso de construcción y una heurística que guían dicho proceso, estableciendo así un conjunto de criterios de calidad mínimos. En su confección se persiguen dos principios, el de *circularidad*, por lo que se busca maximizar las referencias entre términos, y el de *vocabulario mínimo* que alude a la necesidad de mantener al mínimo posible la referencia a términos que no pertenecen al dominio. Un uso frecuente que se les da a los LELs es en la construcción de *escenarios*[10], por lo que resulta de mucha importancia que se cuente con una representación sólida sin discrepancias, errores u omisiones (DEO) [6][3].

La sintaxis del LEL está especificada a partir del metamodelo del listado 1.1, donde se expresa que el LEL está compuesto de uno o más elementos SIMBOLO. Por su parte un SIMBOLO puede tener una o más ocurrencias de NOMBRE (lo que nos remite a la presencia de sinónimos), una NOCION y un IMPACTO. Por su parte, el NOMBRE puede ser una PALABRA, FRASE o ACRÓNIMO. La NOCION es la denotación del símbolo y, por lo general, consiste de una oración o único párrafo que la describe. Si se requieren más oraciones puede ser un indicador de que hay un solapamiento de definiciones. El IMPACTO explica la connotación del símbolo en el UdeD, enfatizando en las acciones que realiza

o recibe (dependiendo si se trata de un objeto o sujeto), o bien, de tratarse de un SIMBOLO de tipo verbo, con cuáles otros está relacionado, por lo general, en cuestión de secuenciación de actividades.

Listado 1.1. Metamodelo del Léxico Extendido del Lenguaje en sintaxis pseudo BNF

```

<LEL> :=          {<SIMBOLO>} [1:n]
<SIMBOLO>:=      {<NOMBRE>} [1:n] + <NOCION>+<IMPACTO>
<NOMBRE> :=      <PALABRA> |<FRASE>| <ACRONIMO>
<NOCION>        :=  NOCION: + {<ORACION>} [1:n]
<IMPACTO>       :=  IMPACTO: + {<ORACION>} [1:n]

```

2.3. Actividades del proceso de construcción del LEL

En la Fig 1 se aprecia el proceso general de construcción de un LEL que incluye las actividades de planear, recolectar, describir, verificar y validar[6]. Dichas actividades siguen un flujo principal, aunque con frecuencia se retroalimentan. Por ejemplo, como resultado de la revisión de la lista de DEO puede pasarse de la actividad de VALIDAR a la de RECOLECTAR.

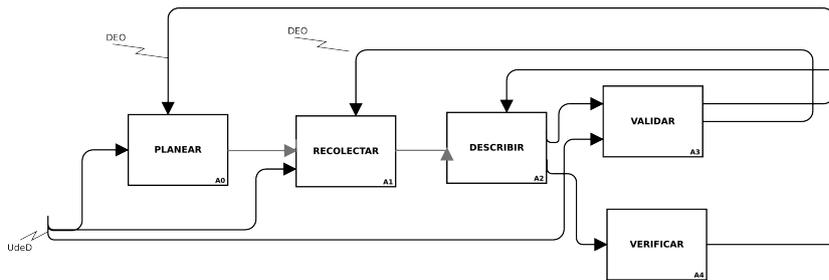


Figura 1. SADT simplificado con las actividades del proceso de elaboración del LEL

A continuación se ofrece una breve descripción de dichas actividades y las principales heurísticas empleadas en cada etapa.

A0. Planear: consiste en identificar fuentes de información, priorizarlas según factibilidad y relevancia para luego seleccionar una estrategia de elicitación de los símbolos del dominio de la aplicación.

A1. Recolectar: los símbolos se seleccionan a partir de lo recabado en los documentos de las entrevistas o de los documentos formales de la organización, y siguen los siguientes lineamientos:

- Se deben seleccionar exclusivamente frases y términos empleados en el dominio de la aplicación.
- Se debe discriminar entre aquellos términos y frases según su frecuencia de uso en entrevistas y documentos de clientes y usuarios.

- Es conveniente excluir aquellas expresiones o términos que formen parte del dominio público y no sean específicos del dominio de la aplicación.
- Es recomendable incluir los términos dudosos para luego ahondar en ellos.
- Incluir símbolos que sean expresiones, no importa su extensión, como así también aquellos que sean acrónimos.

Con los símbolos organizados en una lista priorizada por la relevancia de éstos en el UdeD, se procede a establecer las pautas de clasificación que obedece a los siguientes criterios preestablecidos que se mantendrán salvo que se requiera un cambio según las necesidades particulares del dominio:

Tipo sujeto: se trata de una entidad activa, ya sea una persona, un sistema, maquinarias, etc. Un símbolo de este tipo se caracteriza por actuar y provocar cambios en el dominio de la aplicación.

Tipo objeto: entidades de tipo pasivo a las que se le aplican las acciones de los símbolos de tipo sujeto.

Tipo Verbo: constituye toda acción o actividad llevada a cabo en el dominio.

Tipo estado: un estado supone una configuración particular del dominio de la aplicación a la que se puede arribar por aplicación de una acción y en la que se pueden encontrar tantos sujetos como objetos.

A3. Describir: esta actividad completa la descripción del símbolo al especificar su *noción* y su *impacto*. La *noción* es la definición del símbolo adoptada en el UdeD, describiendo en el caso del tipo sujeto quién y qué lugar ocupa en el UdeD; para el tipo objeto qué representa, sus características y relación con otros objetos. Para el tipo verbo, se establece quién ejecuta dicha actividad (relación con tipo sujeto), en qué tiempos y los aspectos que la definen como actividad en el UdeD, así como su relación con otras actividades.

A4. Verificar: se busca encontrar errores que tienen que ver con la sintaxis del metamodelo del LEL, y del cumplimiento de las heurísticas definidas para la descripción de los símbolos, el cumplimiento del principio de circularidad y vocabulario mínimo, y si fueron adecuadamente clasificados los símbolos.

A5. Validar: esta actividad se realiza en conjunto con los “involucrados”. Se emplean entrevistas cortas donde se busca corroborar las nociones e impactos de cada símbolo con los usuarios o actores claves del negocio. Esta actividad repercutirá en las actividades A1, A2 y eventualmente en la A0.

Una de las pautas globales que guían el proceso, principalmente en las dos primeras actividades, consiste en tomar en consideración los objetivos y alcances del sistema a desarrollar, puesto que ello supondrá cotas sobre el propio UdeD, y tendrá un especial peso en la relevancia de los símbolos que deberán ser incluidos necesariamente y aquellos que en última instancia serán dejados fuera.

2.4. Heurísticas de verificación y validación del LEL

La heurística de construcción del LEL consiste en una serie de verificaciones y validaciones que se realizan sobre el LEL para detectar discrepancias, errores u omisiones. Éste constituye un compromiso mínimo de calidad que se busca alcanzar. Las heurísticas que se consideraron para este trabajo son las definidas

por [6]. A continuación, se describirán sólo algunas por cuestiones de espacio, pero enfatizando en las que son de interés para este trabajo:

- Cada símbolo debe tener un nombre y al menos una oración en el impacto y una en la noción.
- Los nombres de símbolos deben ser usados pese a su longitud.
- Las oraciones de la noción y el impacto deben ser simples y directas.
- Las oraciones deben expresar una sola idea.
- Si dos símbolos comparten oración en la noción o en el impacto, deben repetirse en ambas entradas.
- Toda noción o impacto debe tener una referencia a otro símbolo.
- Todo símbolo debe tener una referencia a otro símbolo.

2.5. El problema de la incompletitud

La completitud respecto al LEL hace alusión a si se ha indentificado todos los símbolos posibles en el UdeD y si se ha realizado una descripción apropiada de éstos con el suficiente nivel de detalle, lo que se conoce como completitud interna. Esta es una característica difícil de lograr y más aún de medir, aunque hay algunos esfuerzos al respecto[11], sobre la base de los cuales se hace una pequeña contribución en este trabajo.

3. Metamodelo del LEL basado en Ontologías

Dado que el proceso de elaboración del LEL es inherentemente iterativo e incremental, se ha tenido en cuenta en la construcción de la ontología el hecho de contar con información incompleta, y es por ello que resulta muy apropiado el paradigma del mundo abierto (Open Word Assumption o OWA) de la *lógica descriptiva*, formalismo que da sustento al estándar OWL (Ontology Web Language) sobre el que se ha construido esta propuesta. Sin embargo, debido a algunas limitantes relacionadas con la funcionalidad esperada, se ha recurrido a una solución mixta que integra dicho formalismo con el del lenguaje de reglas de lógica de primer orden, lo que se trata en la sección 3.3.

La ontología propuesta tiene como objetivo asistir al analista en la construcción de los LEL sin tener conocimientos sobre formalizaciones con ontologías sino empleando sus propias competencias en el análisis del UdeD y su documentación a partir de tales plantillas. Si bien ya se ha desarrollado una herramienta para favorecer la amigabilidad del uso de dicho metamodelo en la identificación y especificación de los símbolos [12], este trabajo se enfoca en los aspectos relativos a las inferencias y detección de inconsistencias en el modelo generado por el analista.

3.1. Ontología del metamodelo del LEL

Para describir el metamodelo del LEL se empleará el formato funcional[8] para ontologías, el cual es más legible que el basado en el del estándar OWL/XML.

El listado 1.2 establece que para formar parte de la clase 'LEL'(:LEL) un individuo tiene que contar con una relación 'tiene símbolo' (:has_symbol). En el listado 1.3 se establece que un individuo será clasificado como 'Símbolo' (:Symbol) cuando tenga al menos una relación 'tiene noción' (:hasNotion) y al menos una 'tiene impacto' (:hasBehavioralResponse). Los listados 1.5 y 1.6 se corresponden con la declaración de noción e impacto respectivamente, sólo que no se representan a partir de clases, sino de relaciones (ObjectProperty) con individuos de la clase 'Oración'(:Sentence).

La definición de la clase más interesante es sin duda la de 'Oración', mediante la cual, como se podrá ver más adelante, se producen la mayoría de las clasificaciones. En el listado 1.4 se puede observar cómo está declarada, con varias restricciones sobre relaciones con individuos de la clase 'Objeto'(:Object), 'Sujeto'(:Subject) y 'Verbo' (:Verb). La relación más crítica y tratada como relación de equivalencia que condiciona la clasificación de un individuo como 'Oración' es la que indica que una oración debe tener al menos un verbo (ObjectExactCardinality(1 :has_verb :Verb)). No obstante, conviene mencionar que se realizó una salvedad en el caso de que se trate de una oración de la noción, la que podrá ser exceptuada de la condición previa al cumplir que tiene la relación 'describe a un símbolo' (:describe_to :Symbol). Se ha definido también restricciones para la vinculación de la 'Oración' con un símbolo de tipo 'Objeto' (:has_object :Object) y con la referencia a un 'Sujeto' (:performed_by :Subject).

Un aspecto que es muy importante señalar respecto a la construcción de ontologías es la declaración de clases disyuntas, es decir, aquellas que no pueden compartir individuos. En el listado 1.7 se indica que un individuo sólo puede pertenecer a una de las clases listadas.

Listado 1.2. Definición de la clase LEL

```
Declaration ( Class (:LEL))
EquivalentClasses (:LEL
  ObjectSomeValuesFrom (: has_symbol :Symbol))
```

Listado 1.3. Definición de la clase Símbolo(Symbol)

```
Declaration ( Class (:Symbol))
EquivalentClasses (:Symbol
  ObjectIntersectionOf (
    ObjectSomeValuesFrom (
      :hasNotion :Sentence) ObjectSomeValuesFrom (
      :hasBehavioralResponse :Sentence)))
```

Listado 1.4. Definición de la clase ORACION (Sentence)

```
Declaration ( Class (: Sentence))
EquivalentClasses (: Sentence
  ObjectIntersectionOf (
    ObjectExactCardinality (1 :has_verb :Verb)
    ObjectAllValuesFrom (: has_verb :Verb)))
```

```

EquivalentClasses (: Sentence ObjectSomeValuesFrom (
  : describe_to : Symbol))
EquivalentClasses (: Sentence
ObjectExactCardinality (1 : describe_to : Symbol))
SubClassOf (: Sentence
  ObjectIntersectionOf (ObjectAllValuesFrom (: has_object : Object)
    ObjectSomeValuesFrom (: has_object : Object)))
SubClassOf (: Sentence
  ObjectIntersectionOf (ObjectAllValuesFrom (: performed_by : Subject)
    ObjectSomeValuesFrom (: performed_by : Subject)))

```

Listado 1.5. Definición de la relación 'TIENE NOCIÓN' (hasNotion)

```
Declaration (ObjectProperty (: hasNotion))
```

Listado 1.6. Definición de la relación 'TIENE IMPACTO' (hasBehavioralResponse)

```
Declaration (ObjectProperty (: hasBehavioralResponse))
ObjectPropertyRange (: hasBehavioralResponse : Sentence)
```

Listado 1.7. Se declara las subcategorías de símbolos como clases disyuntas

```
DisjointClasses (ontolel: Object ontolel: State
  ontolel: Subject ontolel: Verb)
```

3.2. Axiomatización a partir de la heurística de construcción

La propuesta aquí presentada consiste en auxiliar al ingeniero de requisitos en la tarea de verificación y validación del LEL, lo cual se logra por medio de la detección de inconsistencias en la ontología, de manera tal que se puedan realizar acciones correctivas. No obstante, conviene exponer primeramente el razonamiento subyacente en los mecanismos de detección de las inconsistencias y cómo deben ser interpretadas.

La principal característica, como se ha señalado, es la adopción de un modelo de Mundo abierto (OWA), cuya principal característica es que si se desconoce un valor, éste no asume el valor lógico de *falso* sino *desconocido*. Ésto tiene interesantes implicancias, en algunos dominios no se desea que el razonador responda si una sentencia es verdadera o falsa, sino simplemente qué individuos cumplen ciertas restricciones y en función de estas cómo serán clasificados. En el dominio de la IR, que es altamente dinámico y puede distribuirse incluso a través de distintos grupos de análisis, es deseable un enfoque más abierto, más orientado a admitir información incompleta.

3.3. Más allá de la lógica descriptiva

Debido a que se ha presentado la necesidad de inferir algunas relaciones dentro de la ontología del LEL, se ha recurrido al lenguaje de reglas SWRL (Semantic

Web Rule Language). Si bien éste emplea lógica de primer orden, lo que supone la adopción de un modelo de mundo cerrado (Close Word Assumption), su uso está contemplado en el estándar en conjunto con OWL-DL (cuyo formalismo se encuadra dentro de OWA) y no reviste un problema dentro del alcance de esta propuesta. En el caso de la ontología del LEL presentada se ha optado por inferir relaciones a partir de otras definidas, que no pueden ser resueltas por inferencias sobre lógica descriptiva (por ejemplo con relaciones inversas). Éste es el caso de establecer la relación entre dos símbolos a partir de las menciones a éstos en las oraciones empleadas en la noción o impacto. Ésta relación es fundamental, puesto que es la que define si un símbolo forma parte o no de un LEL. En este punto es importante destacar que en la ontología se ha previsto una relación de suscripción de un individuo a otro que potencialmente es un individuo de tipo LEL, éste es el caso de la relación 'corresponde a' (:correspond_to).

En el listado 1.8 se presenta una de las reglas utilizadas para establecer la relación entre dos símbolos del LEL (is_related_to), definida a partir del sujeto presente en una oración (indicado por el segundo argumento de performed_by). A su vez, se incluye la relación donde se indica que dicha oración describe a un símbolo a través del impacto de éste (describe_to). Con la ejecución de la regla antes señalada se encuentra en condiciones de ser disparada la regla del listado 1.9. La presencia de ésta última relación inferida (part_of) es fundamental en esta propuesta, pues es la que otorga relevancia al símbolo como parte efectiva del LEL.

Listado 1.8. Una relación 'Esta relacionada con' (:is_related_to) se obtiene mediante la ejecución de una regla

```
Sentence(?stc), Symbol(?s1), Symbol(?s2),
  correspond_to(?s1, ?lel1), correspond_to(?s2, ?lel2),
  describe_to(?stc, ?s1), performed_by(?stc, ?s2),
  SameAs(?lel1, ?lel2), DifferentFrom(?s1, ?s2)
-> is_related_to(?s1, ?s2)
```

Listado 1.9. Una relación 'forma parte de' (:part_of)

```
Symbol(?s), Symbol(?x), correspond_to(?s, ?l),
is_related_to(?s, ?x), DifferentFrom(?s, ?x)
-> part_of(?s, ?l)
```

3.4. Clasificación e inferencias

Las inconsistencias pueden aparecer por problemas de clasificación errónea, lo cual, en el contexto de esta propuesta significa que un símbolo fue asignado en forma inconsistente a lo largo de varias oraciones. En la Fig 2 puede verse el resultado del razonador el cual indica que la ontología es inconsistente. La explicación que ofrece el razonador en este caso es que el individuo ha tratado de clasificarse en dos clases disyuntas (ver Listado 1.7), que en el ejemplo son

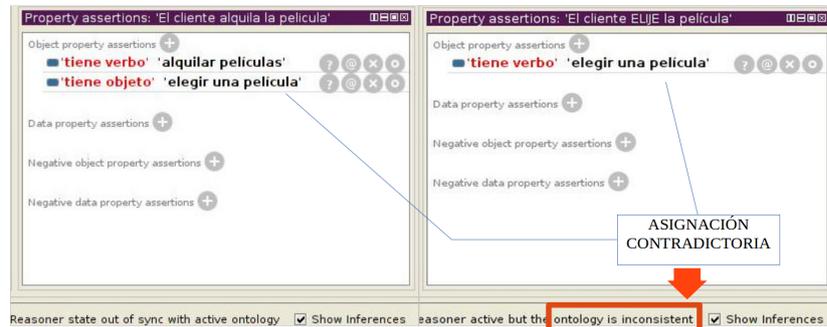


Figura 2. Ontología Inconsistente: Se ha empleado la relación “tiene objeto” y “tiene verbo” para designar al mismo individuo, por lo que se trató de clasificarlo en clases disyuntas

la clase Verbo y Objeto, debido al uso de una asignación errónea por parte del analista.

La detección de sinónimos puede ocurrir en las primeras revisiones del analista sobre el empleo de los términos del UdeD. La forma en que se efectúa la detección de sinónimos es mediante el razonamiento con lógica descriptiva. Dicho mecanismo es activado a partir del uso de la relación *describe_to*, que vincula una oración a un individuo de tipo símbolo al cual describe, de manera tal que una oración que forma parte de la noción sólo puede clasificar a un símbolo según la declaración del listado 1.4. Siguiendo el modelo del mundo abierto, si ocurre que se asignara esta relación a dos individuos el razonador inferirá que se trata del mismo individuo y establecerá una relación *sameAs* entre estos.

3.5. Tratamiento de la incompletitud con ontologías

Si bien es un tema difícil de determinar, pues se establece sobre la validación con los usuarios, se ha optado por una métrica basada en [11], que explica, en función de la ontología del LEL y las relaciones establecidas mediante mecanismo de inferencia, cuál es la completitud a través de un mecanismo de verificación.

La ecuación 1 muestra cómo se determina este valor mediante el cociente entre las cardinalidades de las relaciones 'parte de' (*part_of*) y 'corresponde a' (*correspond_to*). Es por ello que una máxima completitud se interpreta, no desde el punto de vista de los términos o del corpus elicitado efectivamente desde el dominio, sino justamente del LEL en relación con dicho corpus, el cual se entiende que es resultado de una actividad llevada a cabo por un analista, empleando para ello su juicio, pericia técnica y una heurística acorde a las características de la ontología presentada aquí.

$$CL = \frac{|part_of|}{|correspond_to|} \quad (1)$$

4. Heurística para la construcción del LEL basado en ontologías

La propuesta que se realiza en este trabajo consiste en la definición de una heurística de construcción de un LEL que se basa en la propuesta original de Hadad [6], pero que se orienta a otorgar una mayor flexibilidad aprovechando la ventaja que supone el uso de ontologías para definir individuos y clasificarlos en forma automática en las clases que correspondan. Es así que se ha utilizado como criterio orientador el proceso referido en la sección 2.3, considerando algunos puntos de la heurística de la sección 2.4 que se realizarán en forma semiautomática.

Actividad 1. Identificación de las oraciones del LEL: En esta actividad las oraciones son extraídas directamente del corpus de texto de los documentos de elicitación y de las entrevistas. Si bien, pueden surgir modificaciones (como por ejemplo el paso de la voz activa a la pasiva), se debe mantener en la medida de lo posible la terminología específica empleada en el dominio de la aplicación, lo que permitirá luego su posterior validación con el usuario.

Actividad 2. Identificación de los símbolos: Los símbolos son extractos del texto, que incluye tanto palabras aisladas, como construcciones que incluyen núcleo verbal, modificadores, objetos, entre otras que son relevantes para el UdeD, pero para lo que se requiere el examen del analista, por ejemplo: Película de acción.

Actividad 3. Asignación de las oraciones a los símbolos: La asignación de oraciones a símbolos es el acto de atribuirle al símbolo su denotación (noción) y connotación (impacto). Para ello, se presenta la posibilidad de hacerlo mediante la noción, mediante una relación 'tiene noción', asignación que establece la descripción de símbolo mediante la oración, y la asignación del impacto, para lo que se seguirán como criterios orientativos los fijados por la heurística de descripción de los símbolos de la actividad A3 del proceso y las heurísticas de la sección 2.4 y 2.5.

Actividad 4. Ejecución del razonamiento automático: Esta actividad es crucial, pues es donde la calidad de construcción del LEL se pone a prueba. Se contemplan dos tipos de inferencia, el primer tipo se basa en el formalismo de lógica descriptivas, y supone la clasificación de individuos en las clases si cumplen los restricciones (axiomas) de la ontología, a la vez que se detectan posibles inconsistencias según las definiciones establecidas. El otro tipo de inferencia corresponde al espectro "duro" de la ontología, que es mediante el uso de reglas en SWRL presentadas en la sección 3.4, donde se efectúan afirmaciones (asserts) de nuevas relaciones entre los individuos. Esta actividad es factible efectuarla en simultáneo con las 3 actividades antes mencionadas, lo que supondría la detección temprana de inconsistencias.

Actividad 5. Revisión de la ontología inferida: El analista en esta etapa emplea la heurística de descripción de los LEL para solucionar posibles inconvenientes que causan símbolos que no son identificados como tal, es decir, individuos que no son clasificados como SIMBOLO, individuos clasificados en categorías incorrectas, posiblemente a partir de la aplicación de una relación equivocada en

un individuo de tipo ORACION o por no incluir una relación que denote una relación con un símbolo VERBO (como se ha visto ésto derivaría en la clasificación automática del individuo dentro de dicha categoría).

5. Caso de Estudio

El ejemplo que se presenta en este caso de estudio está planteado sobre un caso de una tienda online de contenidos digitales ficticia, “Megastore Digital Contents”. El siguiente es un extracto del texto sobre el que se basa el análisis:

“Megastore Digital Contents” es una empresa Multinacional con sede central en EEUU. Se dedica a la venta de contenidos digitales (libros, películas, series) a sus suscriptores en todo el mundo. El usuario del servicio accede al portal web de la empresa introduciendo un nombre de usuario y contraseña. Si no está registrado, puede iniciar el proceso de registro al servicio. Para el alquiler de una película, el usuario debe estar logueado. El proceso de compra consiste en elegir de la lista de items organizados por categorías el producto deseado, luego el usuario inicia el proceso de compra, para lo que introduce los datos de su tarjeta y adquiere el derecho a acceder al contenido.

Siguiendo la heurística planteada en la sección 4, se procede a identificar las oraciones en el texto, las que se muestran en la captura de Protégé de la Fig 3, las que son clasificadas directamente dentro de la clase “Oraciones”.

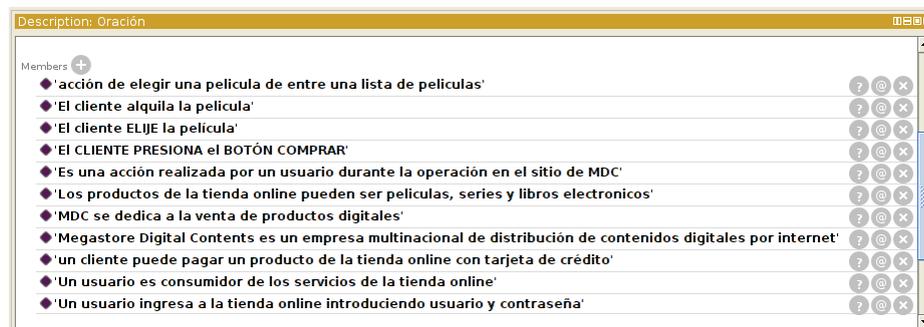


Figura 3. Las oraciones identificadas dentro del caso de estudio

El símbolo “alquilar películas” es presentado en la Fig 4, donde, como puede apreciarse, resulta clasificado por el razonador como Verbo y además se infieren otras relaciones. Una de las más importantes es la relación “está relacionado con el símbolo”, donde se puede notar que se cumple la condición de circularidad. Por otra parte, se confirma que el símbolo ha sido incorporado como una entrada válida del LEL de MDC, mediante la relación inferida “ símbolo forma parte del

LEL de”, a partir de lo cual puede ser empleado en otras representaciones, como es el caso de escenarios.

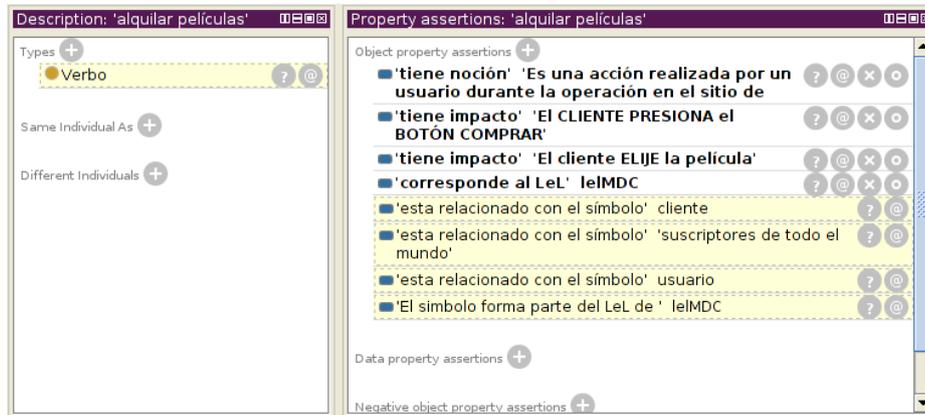


Figura 4. Relaciones y clasificaciones inferidas a partir de los axiomas y las reglas

6. Discusión y trabajos futuros

En este trabajo se presentó una ontología como metamodelo de un documento de especificación de requerimientos como es el caso del LEL, así como una heurística para su construcción. Este enfoque facilitaría la detección temprana de errores mediante mecanismos de inferencia aplicados sobre la ontología del LEL, de manera que tales errores sean rectificadas o ratificados directamente por los analistas según corresponda. Asimismo, la delimitación establecida en este trabajo sobre los tipos de razonamientos aplicables para cada problemática identificada en el caso de estudio ofrece una propuesta interesante desde el punto de vista de la integración entre OWA y CWA.

Lo anterior se pone de manifiesto con el caso de estudio presentado, donde el razonamiento con ontologías para la verificación del LEL se emplea en la detección de inconsistencias y en la obtención de nuevas afirmaciones sobre el dominio, lo que pone en evidencia potenciales DEO. Puede ocurrir también que eventualmente influya en la validación debido a la presencia de relaciones que por las características del mundo abierto, deriven en afirmaciones (la inferencia de relaciones *sameAs* que se expuso en la sección 3.4) que a priori puedan parecer erróneas o al menos sospechosas, pero que va a significar para el analista una necesidad de revisión de las descripciones efectuadas sobre los símbolos del LEL.

Es así que como modelos formales de los Documentos de Especificación de Requerimientos, las ontologías facilitan la construcción de los mismos ya que definen y restringen el uso de los elementos que los componen. Pero además, al ser

un modelo formal se podrían emplear para generar a partir de ellas una ontología que represente formalmente el conocimiento del dominio de la aplicación. Al contar con modelos formales en la etapa de especificación de requerimientos se podrían plantear a su vez heurísticas que transformen los mismos en modelos de diseño reduciendo así la brecha existente entre los requerimientos del usuario y el proceso de desarrollo de software.

Referencias

1. Castañeda, V., Ballejos, L., Caliusco, M.L., Galli, M.R.: The use of ontologies in requirements engineering. *Global journal of researches in engineering* 10(6) (2010)
2. Castañeda, V., Ballejos, L.C., Caliusco, M.L.: Improving the quality of software requirements specifications with semantic web technologies. In: WER (2012)
3. Graciela D. S. Hadad, J.H.D.: Explicitar requisitos del software usando escenarios. (2009)
4. Gruber, T.R.: A translation approach to portable ontology specifications. *Knowl. Acquis.* 5, 199–220 (Jun 1993), <http://portal.acm.org/citation.cfm?id=173743.173747>
5. Gómez-Pérez, A., Fernández-López, M., Corcho, O.: *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web*. Springer, Berlin Heidelberg New York (2004)
6. Hadad, G.D.S.: Uso de escenarios en la derivación de software (2010), <http://hdl.handle.net/10915/19641>, eje: Concurso de tesis
7. Henderson-Sellers, B.: Bridging metamodels and ontologies in software engineering. *Journal of Systems and Software* 84(2), 301–313 (Feb 2011), <http://www.sciencedirect.com/science/article/pii/S0164121210002906>
8. Motik, B., Patel-Schneider, P.F., Parsia, B., Bock, C., Fokoue, A., Haase, P., Hoekstra, R., Horrocks, I., Ruttenberg, A., Sattler, U., Smith, M.: OWL 2 web ontology language: Structural specification and functional-style syntax. Last call working draft, W3C (2008), <http://www.w3.org/2007/OWL/draft/owl2-syntax/>, (to be published, may be superseded)
9. Pohl, K.: *Requirements Engineering: Fundamentals, Principles, and Techniques*. Springer (Jul 2010)
10. do Prado Leite, J.C.S., Hadad, G.D.S., Doorn, J.H., Kaplan, G.N.: A scenario construction process. *Requirements Engineering* 5(1), 38–61 (2000), <http://www.springerlink.com/index/8TBRCYJCPAJQ0V31.pdf>
11. Ridaio, M., Doorn, J.H.: Estimación de completitud en modelos de requisitos basados en lenguaje natural. In: Alencar, F.M.R., Sánchez, J., Werneck, V. (eds.) *Anais do WER06 - Workshop em Engenharia de Requisitos*, Rio de Janeiro, RJ, Brasil, Julho 13-14, 2006. pp. 146–152 (2006)
12. Ruidías, H.J., Caliusco, M.L., Galli, M.R.: Towards the integration of ontologies in the context of MDA at CIM level (Oct 2012), <http://hdl.handle.net/10915/23735>, eje: Workshop Ingeniería de software (WIS)
13. Sommerville, I.: *Software Engineering* (7th Edition). Pearson Addison Wesley (2004)