

Construcción de una ontología utilizando Protégé para la elicitación de requerimientos

Ing. Marcelo Marciszack/ Facultad Reg. Córdoba/ Universidad Tecnológica Nacional

Dr. Manuel Pérez Cota / Facultad / Universidad de Vigo

Mg. Leandro Antonelli / Universidad Nacional de La Plata

Marina Cardenas/ Facultad Reg. Córdoba/ Universidad Tecnológica Nacional

Abstract

El presente trabajo describe el desarrollo de un modelo de representación conceptual del dominio de la teoría de gramáticas formales y máquinas abstractas, a través del modelado ontológico, con el objetivo de evaluar su utilización como herramienta de soporte a la elicitación de requerimientos de software basándose en una metodología de modelado ontológico.

Este desarrollo es parte de un proyecto que nos permitirá comparar entre sí las diferentes metodologías y herramientas que suelen utilizarse en la etapa de relevamiento de requerimientos del proceso de desarrollo de software, y por otro lado, permitirá establecer el grado de correspondencia entre la conceptualización de las Máquinas y Gramáticas, ya que al existir un isomorfismo entre ambos dominios, éste debería continuar en los esquemas conceptuales resultantes de la aplicación de cada una de las metodologías comparadas.

Palabras clave: Ontología, modelo conceptual, elicitación, maquinas abstractas, gramáticas formales, captura de requerimientos, Protégé 2000.

Introducción

Una ontología es un sistema de representación del conocimiento acerca de un dominio o ámbito específico, con el fin de obtener una representación formal de los conceptos que contiene y de las relaciones que existen entre dichos conceptos.

Además, una ontología se construye en relación a un contexto de utilización especificando una conceptualización, por lo que cada ontología incorpora un punto de vista. Todas las conceptualizaciones (definiciones, categorizaciones, jerarquías, propiedades, herencia, etc.) de una ontología pueden ser procesables e

interpretadas por una computadora o un ser humano.

Un concepto también asociado a este sistema de representación es la Web Semántica, la cual tiene como objetivo principal la creación de un medio universal para el intercambio de información basado en representaciones del significado de los recursos que se encuentran en la Web, de una manera inteligible para las máquinas. Para que esto pueda llevarse a cabo, se necesita que el conocimiento de la web esté representado de forma que sea legible por las computadoras, esté consensuado, y sea reutilizable y es por ello que las ontologías proporcionan la vía para representarlo.

Para el diseño del modelo ontológico se establece como dominios a modelar las Gramáticas Formales y Máquinas Abstractas. La elección de estos dominios tiene un doble propósito: por un lado nos permitirá comparar entre sí las diferentes metodologías y herramientas de la etapa de relevamiento de requerimientos del proceso de desarrollo de software y por otro lado, permitirá establecer el grado de correspondencia entre la conceptualización de las Máquinas y Gramáticas, ya que al existir un isomorfismo entre ambos dominios, éste debería continuar en los esquemas conceptuales resultantes de la aplicación de cada una de las metodologías comparadas.

Elementos de las ontologías

Las ontologías proporcionan un vocabulario común de un área y definen, a diferentes

niveles de formalismo, el significado de los términos y relaciones entre ellos.

El conocimiento en ontologías se formaliza principalmente usando cinco tipos de componentes: conceptos, relaciones, funciones, axiomas e instancias.

Los conceptos, entidades o clases en la ontología se suelen organizar en taxonomías. Se suele usar tanto el término clases como conceptos. Un concepto puede ser algo sobre lo que se dice algo y, por lo tanto, también podría ser la descripción de una tarea, función, acción, estrategia, proceso de razonamiento, etc.

Las relaciones representan un tipo de interacción entre los conceptos del dominio. Como ejemplos clásicos de relaciones binarias podemos mencionar: “subclase de” y “conectado a”.

Las funciones son un tipo especial de relaciones en las que el n-ésimo elemento de la relación es único para los “n-1” precedentes.

Los axiomas son expresiones que son siempre ciertas. Pueden ser incluidas en una ontología con muchos propósitos, tales como definir el significado de los componentes ontológicos, definir restricciones complejas sobre los valores de los atributos, argumentos de relaciones, etc. verificando la corrección de la información especificada en la ontología o deduciendo nueva información.

Las instancias se usan para representar elementos específicos de la ontología.

Herramienta de modelado de ontologías: Protégé

Las ontologías requieren de un lenguaje lógico y formal para ser expresadas.

En la inteligencia artificial se han desarrollado numerosos lenguajes para este fin, algunos basados en la lógica de predicados y otros basados en frames (taxonomías de clases y atributos), que tienen un mayor poder expresivo, pero menor poder de inferencia; e incluso existen lenguajes orientados al razonamiento. Todos estos lenguajes han servido para desarrollar otros lenguajes

aplicables a la Web. Es por ello que para el desarrollo de este proyecto se ha optado por utilizar la herramienta Protégé [2] que implementa el lenguaje OWL (Ontology Web Language) [3][4] para el modelado de ontologías basadas en Frames. La misma ha sido desarrollada por la Universidad de Stanford y se utiliza para el desarrollo de Ontologías y Sistemas basados en el conocimiento por medio de una interfaz de usuario que facilita la creación de estructuras de frames con clases, slots e instancias de una forma integrada.

Protégé permite:

- Modelar una ontología de las clases que describen un tema particular.
- Creación de una herramienta de adquisición de conocimiento para recoger conocimiento.
- Entrar en casos específicos de datos y de la creación de una base de conocimiento.
- La ejecución de usos.

Metodología

Debido a la existencia de diversas metodologías de desarrollo de ontologías, se ha optado por la descrita en [1] ya que se adapta mejor a los requerimientos de modelado, desde el punto de vista de la simplicidad y completitud de la misma.

Antes de proceder a la explicación de dicha metodología, es preciso definir una serie de reglas que ayudarán a tomar decisiones de diseño y modelado:

1. No existe una forma correcta y única de modelar un dominio, es por ello que la mejor solución casi siempre depende de la aplicación que se le dará a la ontología una vez modelada.
2. El desarrollo de ontologías es un proceso necesariamente iterativo. La ontología inicial evoluciona y se refina a través de las sucesivas iteraciones.
3. Los conceptos en la ontología deben ser cercanos a los objetos (ya sean físicos o lógicos) y relaciones en el dominio de interés.

Tomando como base lo dicho anteriormente, a continuación se explicará la metodología adoptada para realizar el modelado de la ontología:

Paso 1. *Determinar el dominio y alcance de la ontología.*

Este paso ayuda a determinar el alcance del modelo, definiendo el ámbito de aplicación de la ontología con respecto a su dominio específico. Para ello se requiere la formulación de preguntas de competencia que ayudarán a limitar el alcance. Estas preguntas servirán después como prueba de control de calidad: ¿La ontología contiene suficiente información para responder esos tipos de preguntas? ¿Las respuestas requieren un nivel particular de detalle o representación de un área particular? Las preguntas de competencia son solamente un bosquejo y no necesitan ser exhaustivas.

Paso 2. *Considerar la reutilización de ontologías existentes.*

Reusar ontologías existentes puede ser un requerimiento si nuestro sistema necesita interactuar con otras aplicaciones que ya se han dedicado a ontologías particulares o vocabularios controlados. Muchas ontologías ya están disponibles en forma electrónica y pueden ser importadas dentro un entorno de desarrollo de ontologías.

Paso 3. *Enumerar términos importantes para la ontología*

En este paso es necesario escribir una lista con todos los términos que potencialmente servirán para el modelo de la ontología, estos términos son extraídos de la descripción del dominio de aplicación.

Paso 4. *Definir las clases y la jerarquía de clases.*

Hay varios posibles enfoques para desarrollar una jerarquía de clases:

- *Un proceso de desarrollo top-down* comienza con la definición de los conceptos más generales en el dominio la

subsecuente especialización de los conceptos.

- *Un proceso de desarrollo bottom-up* comienza con la definición de las clases más específicas, las hojas de la jerarquía, con el subsecuente agrupamiento de esas clases en conceptos más generales.

- Un *proceso de desarrollo* combinado es el resultado de una combinación de los enfoques top-down y bottom-up: primero definimos los conceptos más sobresalientes y luego los generalizamos y especializamos apropiadamente.

Ninguno de esos tres métodos es inherentemente mejor que cualquiera de los otros. El enfoque a tomar depende fuertemente de la visión personal del dominio.

Paso 5. *Definir las propiedades de las clases: slots.*

Las clases aisladas no proveerán suficiente información para responder las preguntas de competencia del Paso 1. Una vez que hemos definido algunas de las clases, debemos describir la estructura interna de los conceptos.

Ya hemos seleccionado clases de la lista de términos creada en el Paso 3. La mayoría de los términos restantes son muy probablemente propiedades de esas clases.

En general, hay varios tipos de propiedades de objeto que pueden llegar a ser slots en una ontología:

- Propiedades “intrínsecas”.
- Propiedades “extrínsecas”.
- partes, si el objeto es estructurado; pueden ser “partes” físicas y abstractas.
- relaciones con otros individuos; éstas son las relaciones entre miembros individuales de una clase y otros ítems.

Paso 6. *Definir las facetas de los slots.*

Los slots pueden tener diferentes facetas que describen el tipo de valor, valores admitidos, el número de los valores (cardinalidad), y otras características de los valores que los slots pueden tomar.

Paso 7. Crear instancias.

El último paso consiste en crear instancias individuales de clases en la jerarquía. La definición de una instancia individual de una clase requiere (1) elegir una clase, (2) crear una instancia individual de la clase y (3) rellenar los valores del slot.

Aplicación

Debido a que las gramáticas proporcionan las reglas utilizadas en la generación de las cadenas de los lenguajes, es inmediata la relación entre estas gramáticas y las Maquinas Abstractas capaces de aceptarlos. Es así que los lenguajes son el puente que vincula gramáticas y maquinas. En el modelo ontológico obtenido a partir de la metodología anterior, se puede observar que se ha representado este tipo de relaciones con el objetivo de determinar el isomorfismo entre los conceptos mencionados anteriormente.

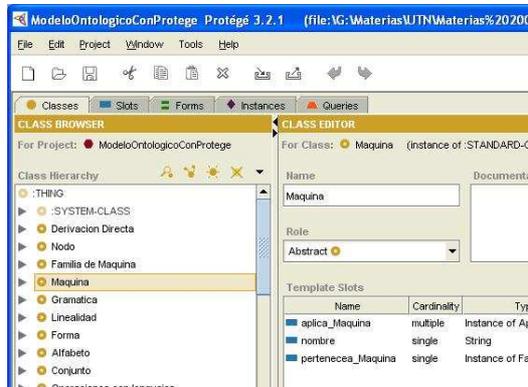


Figura 1. Modelo de clases en Protégé 2000

Utilizando Protégé 2000 para el modelado de la ontología diseñada en el presente trabajo, se procedió a la construcción de la jerarquía de clases conceptuales vinculadas al dominio.

Subsecuentemente se incorporaron los Slots a las clases conjuntamente con las facetas asociadas a cada uno. Las relaciones jerárquicas de herencia se pueden visualizar al hacer clic sobre una clase Padre (o también llamadas clases base).

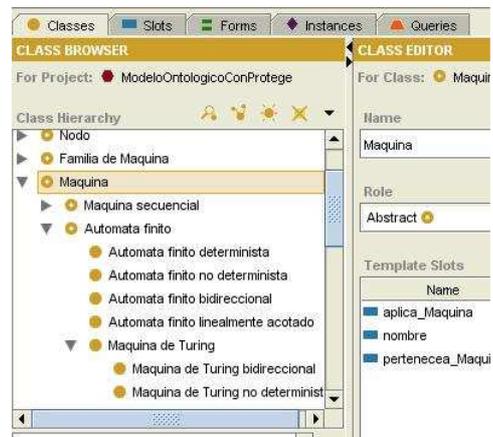


Figura 2. Taxonomía de clases en Protégé

Aquellas relaciones que no son de herencia entre los conceptos, son representadas a través de slots del tipo Instance, en el cual se define en su faceta, el tipo de clase con la cual se relaciona.

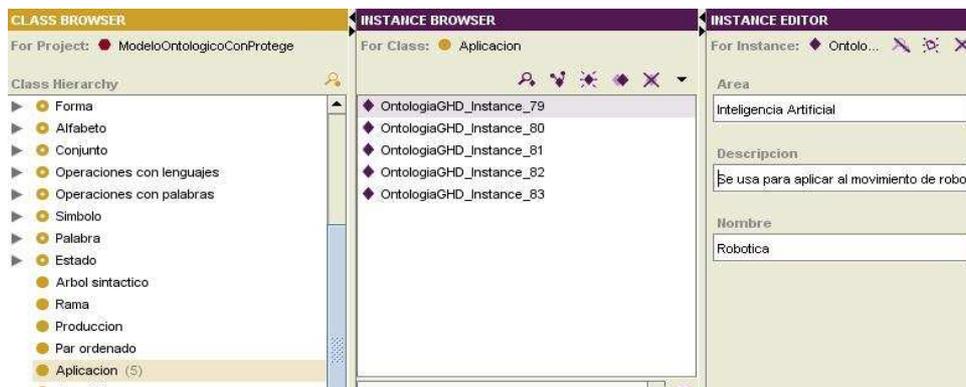


Figura 3. Creación de Instancias con Protégé 2000

Una vez que se realizó la carga del modelo completo, se crearon nuevas instancias para algunas clases con el objeto de efectuar pruebas, para lo cual se ingresó una consulta en la pestaña de la ventana principal llamada Queries.

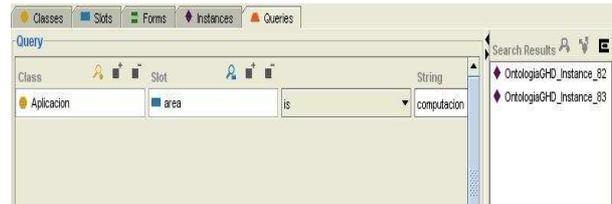


Figura 4. Creación de Queries con Protégé 2000

Se prevé en trabajos futuros, la utilización de programas razonadores de ontologías para poder inferir nuevo conocimiento a partir el modelo ontológico.

Resultados : Valoración de criterios.

A continuación se presentan un cierto conjunto de criterios seleccionados para la validación de la aplicación de la metodología y la herramienta.

Valoración Metodologías/Herramientas	
Metodología	<i>Ontology Development 101</i>
Herramienta	Protégé-2000 Ver. 1.7

Metodología		
Criterio evaluado	Valoración	Observaciones
Claridad Conceptual.	MB	Se puede determinar claramente la representación y vinculación de los conceptos del modelo.
Potencialidad para abstraer esencia del dominio.	MB	Sería necesario poder incorporar información referida a la etapa de diseño e implementación.
Identificación de la fuente	MB	La fuente es identificada pero no existen formalidades acerca de su documentación.
Reducción de ambigüedades sobre conceptos y manejo de sinónimos	B	No se pueden representar conceptos sinónimos de manera explícita.
Facilidad de aplicación y flexibilidad para adopción de criterios de diseño	B	La metodología de modelado es fácil de aplicar pero se presentan inconvenientes en el momento de definir criterios de diseño que influyen en las decisiones de modelado.
Facilidad de entendimiento con el usuario del Dominio.	MB	El modelo es entendible por el usuario del Dominio pues es una simplificación conceptual del mismo.
Mantenibilidad del Modelo	MB	El modelo es mantenible a través del tiempo sin demasiada complejidad.
Reutilización del Modelo	MB	El modelo es reutilizable pero es necesario adaptarlo al dominio de aplicación.
Documentación del modelo.	MB	Es posible contar con una documentación detallada del modelo en base a la metodología como así también en la herramienta.
Jerarquización de los requerimientos del modelo	E	Permite una taxonomía de conceptos que facilita el entendimiento y su modelado.

Validación del modelo resultante.	E	Es posible validar el modelo resultante con respecto a las Preguntas de Competencia a través de Queries que permitan determinar la relación entre ambos.
Versionado en proceso iterativo.	R	No provee las facilidades para realizar un versionado y gestión de la configuración del modelo.
Facilidad de trazabilidad de requerimientos.	MB	Se puede determinar la trazabilidad de los requerimientos a través de las preguntas de competencia.
Producto como insumo para la construcción del sistema modelado.	R	Los frames o clases obtenidas es frecuente que difieran del diseño de clases en una futura implementación Orientada a Objetos.

Herramienta		
Criterio evaluado	Valoración	Observaciones
Facilidad de instalación y configuración.	E	Es fácil de instalar y configurar, muy intuitivo.
Curva de Aprendizaje de la Herramienta	MB	De muy fácil aprendizaje
Capturar en forma fiel y precisa la abstracción del modelo.	B	Existen ciertas consideraciones de implementación que no se pueden plasmar en el modelo.
Facilidad de introducir cambios en el diseño.	MB	Es fácil la incorporación de nuevos conceptos, como así también modificar sus relaciones.
Soporte a proceso iterativo y manejo de versionado.	B	No brinda prestaciones compatibles con el manejo de versionado, pero soporta el proceso iterativo a partir de las sucesivas mejoras que es posible incorporarle al modelo.
Visualización a través de diferentes vistas del modelo.	B	No existen vistas en el modelo.
Uso de notaciones y simbología que faciliten el entendimiento del usuario.	E	La notación es fácil de entender por el usuario.
Uso de notaciones y simbología que faciliten el entendimiento del equipo de desarrollo en la construcción del modelo.	E	La notación es fácil de entender por el equipo de desarrollo en la construcción del modelo.
Facilidad de mapeo directo del modelo para la construcción del sistema	R	No existe mapeo directo del modelo en la construcción del sistema.
Portabilidad	MB	Es posible instalar la herramienta en diversas plataformas.

Resultados: Conclusión de la valoración de criterios de la Metodología / Herramienta .

A modo de conclusión preliminar sobre la valoración de la Metodología/herramienta desarrollada sobre este ejemplo, se puede

concluir que la Metodología “ Ontology Development 101” que se basa fuertemente en poder dar respuestas a preguntas de competencia de la ontología sobre el dominio a modelar y el uso de la herramienta “Protégé-2000 Ver. 1.7”,

pueden ser clasificados en aspectos positivos y negativos.

Aspectos Negativos:

- La herramienta no soporta la representación de sinónimos de conceptos de manera explícita.
- Ciertas consideraciones de implementación como autoincrementar el valor de un slot o valores calculables, no pueden ser representados más que con solo un comentario en el campo del slot.
- No permite el borrado en cadena, es decir, si se borra un slot de un concepto, el mismo seguirá vigente en el entorno del proyecto, por lo cual hay que borrarlo también del proyecto.
- No permite representar las instancias a través de alias que permitan su identificación unívoca.
- No proporciona utilidades adaptadas al dominio de elicitación de requerimientos, tales como aspectos a tener en cuenta para el análisis, diseño e implementación del software.
- No se interesa desde la perspectiva de punto de vista del cliente/usuario, aunque puede ser subsanado con las respuestas o preguntas de competenciadas que debe ser capaz de responder la ontología.
- No brinda prestaciones compatibles con el manejo de versionado y gestión de la configuración, pero soporta el proceso iterativo a partir de las sucesivas mejoras que es posible incorporarle al modelo.

Aspectos Positivos:

- El modo de inspección que ofrece la herramienta por medio de la funcionalidad de Queries permite evaluar el diseño de la ontología posibilitando la verificación de las Preguntas de Competencia formuladas a

través del desarrollo de la metodología, sin embargo es acotado a una serie de relaciones entre los conceptos según la funcionalidad de consultas provista por la herramienta. Cabe destacar que es posible utilizar otras herramientas más específicas para la generación de consultas sobre la ontología, pero en este caso, al evaluar Protégé, hemos optado por utilizar el módulo de consultas que el mismo ofrece para evaluar su funcionalidad.

- El entorno visual ofrece una fácil, rápida e intuitiva interacción con el usuario.
- Permite determinar claramente la representación y vinculación de los conceptos del modelo.
- En cuanto a la metodología de modelado, es fácil su aplicación pero se presentan inconvenientes en el momento de definir criterios de diseño que influyen en las decisiones de modelado.
- El modelo ontológico puede ser reutilizado por otras ontologías.
- Es posible contar con una documentación detallada del modelo en base a la metodología como así también en la herramienta.
- Permite una taxonomía de conceptos que facilita el entendimiento y su modelado.
- La representación visual de los conceptos del modelo permite su fácil entendimiento del usuario.

Discusión

Si bien se evidencia que, además del isomorfismo entre los modelos de Gramáticas Formales y Máquinas abstracta se mantiene en los Metamodelos obtenidos a partir de la aplicación de las Ontologías editadas con la herramienta Protégé, recién al realizar las actividades de modelado con las demás herramientas y metodologías que forman parte del proceso general de comparación recién estaremos en

condiciones de concluir la propuesta del proyecto de investigación “Modelos de especificación de requerimientos para la obtención de esquemas conceptuales en un dominio restringido: comparación de metodologías”, el cual tiene como objetivo la comparación de diferentes metodologías y herramientas para la especificación de requerimientos con el fin de determinar un esquema conceptual [5] sobre el dominio de aplicación de las Máquinas Abstractas y Gramáticas Formales.

El conjunto de Metodologías/Herramientas a comparar dentro del proyecto son:

- ✦ LEL (Léxico Extendido del Lenguaje) [6], Escenarios y Tarjetas CRC, utilizando como herramienta de descripción al BMW (Baseline Mentor Workbench).
- ✦ Casos de Uso obtenidos a partir de la metodología RUP/UML [7] (Rational Unified Process) con la utilización de Rational Rose.
- ✦ Ontologías utilizando Protégé-2000 como herramienta de modelado y edición de Ontologías.

Luego de la construcción de los modelos con todos los conjuntos de metodologías y herramientas, recién se podrá establecer una conclusión definitiva sobre todos los modelos

Conclusión sobre Ontologías/Protégé

En resumen, el conjunto de metodología/herramienta seleccionada como son las Ontologías y Protégé como editor de las mismas, es útil para la exploración del dominio en la etapa de elicitación, pero al no poder precisar una correspondencia de los frames con los objetos directos en una futura implementación hacen que la misma resulte deficiente, ya que al no haber un mapeo directo entre frames y clases en la

construcción del sistema y por lo tanto la ontología deberá ser adaptada en el proceso de desarrollo.

Bibliografía

[1] Ontology Development 101: A Guide to Creating Your First Ontology. Noy, Natalya F., McGuinness, Deborah L.(2005). Stanford University, Stanford.

[2] Ontologías 2. Ontologías en acción. Protégé – OWL. Alberto Barrón Cedeño (2005).

Homepage: <http://theory.lcs.mit.edu/~rajsbaum/cursos/web/ontologias2.pdf>

[3] W3C. OWL Web Ontology Language. Overview. <http://www.w3.org/TR/owl-features/>

[4] OWL. Homepage: <http://www.hipertexto.info/documentos/owl.htm>

[5] Modelado de Requisitos para la Obtención esquemas conceptuales. Emilio Insfrán, Isabel Díaz y Burbano Margarita. <http://www.dsic.upv.es/~einsfran/papers/39-ideas2002.pdf>

[6] Herramienta para implementar LEL y Escenarios (TILS). Gustavo Gil, Alejandro Oliveros, Gustavo Rossi. Tesis de Maestría Universidad Nacional de la Plata.

[7] The Unified Language User Guide. G. Booch, J. Rumbaugh, I. Jacobson. Addison Wesley 1998.

Datos de Contacto:

Marcelo Martín Marciszack
marciszack@gmail.com