

# Construcción de una ontología para gramáticas formales y máquinas abstractas utilizando Protégé para la elicitación de requerimientos

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

Dr. Manuel Pérez Cota / Facultad / Universidad de Vigo

Mg. Leandro Antonelli / Universidad Nacional de La Plata

Dra. Roxana Giandini / Universidad Nacional de La Plata

Marina Cardenas/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional

## CONTEXTO

El presente proyecto se encuentra consolidado dentro de la línea de investigación encabezada por el proyecto “Modelos de especificación de requerimientos para la obtención de esquemas conceptuales en un dominio restringido: comparación de metodologías” desarrollado en el Dpto. de Sistemas de Información de la UTN FRC, 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 sobre un dominio de aplicación determinado: Máquinas Abstractas y las Gramáticas formales.

## RESUMEN

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.*

## 1. INTRODUCCION

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.

Paso 2. Considerar la reutilización de ontologías existentes.

Paso 3. Enumerar términos importantes para la ontología

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

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

Paso 6. Definir las facetas de los slots.

Paso 7. Crear instancias.

## 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.

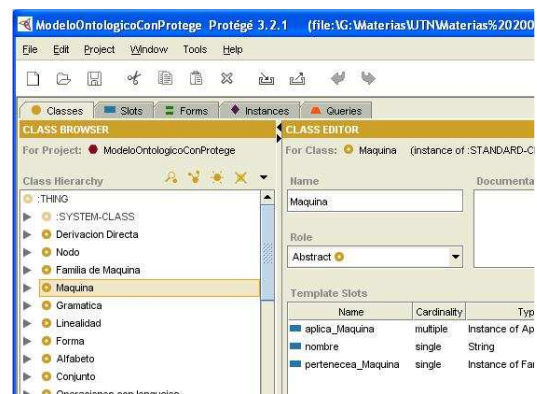


Figura 1. Modelo de clases en Protégé 2000

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).

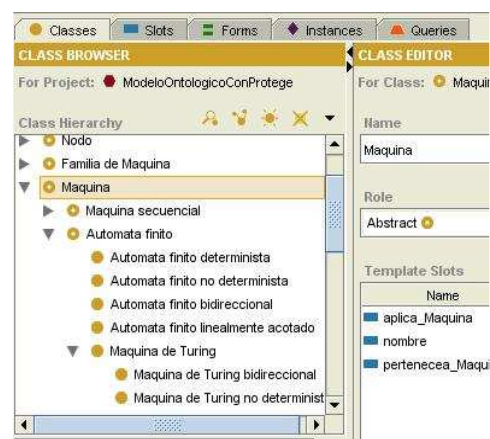


Figura 2. Taxonomía de clases en Protégé

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.

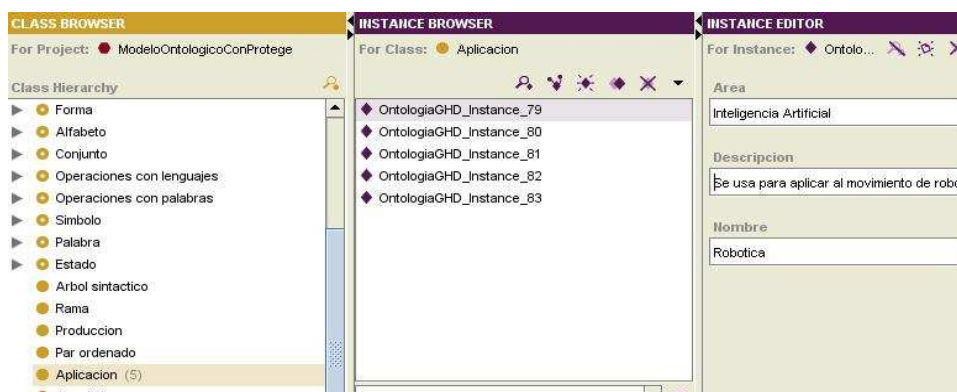


Figura 3. Creación de Instancias con Protégé 2000

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.

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

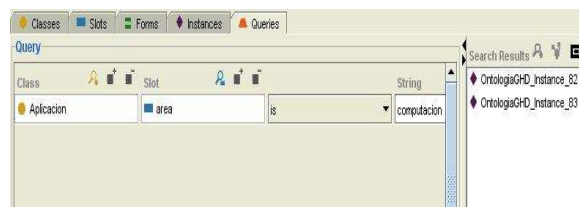


Figura 4. Creación de Queries con Protégé 2000

## 2. LINEAS DE INVESTIGACION y DESARROLLO

Como se mencionó inicialmente, el presente proyecto forma parte de la línea de investigación encabezada por el proyecto “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.

Dichas metodologías y herramientas derivarán en proyectos centrados en la obtención de modelos conceptuales comparables según una serie de parámetros que se definirán en base a una matriz comparativa predefinida oportunamente al terminar el desarrollo de dichos modelos. Dentro de la línea de investigación, cada proyecto se basará en las siguientes metodologías / herramientas respectivamente:

- ✦ 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.
- ✦ Modelo de Objetos y Diagrama de Clases obtenido a partir de una definición Ad-hoc.

### **3. RESULTADOS OBTENIDOS / ESPERADOS**

En base al modelo ontológico obtenido, se prevé la ampliación y profundización de los conceptos representados, continuando con el proceso iterativo que caracteriza a la metodología utilizada para su implementación. A partir de ello, se incorporará y refinará el dominio del problema que delimita el alcance de la ontología y se construirá un conjunto de queries de prueba del modelo ontológico para comprobar su performance. Una vez obtenido el modelo ontológico final, se procederá a compararlo con los obtenidos de otras metodologías y herramientas tales como las siguientes: LEL, Escenarios y Tarjetas CRC, utilizando como herramienta de descripción al BMW, Casos de Uso obtenidos a partir de la metodología RUP/UML con la utilización de Rational Rose; y un modelo de objetos y diagramas de clase obtenido a partir de una definición Ad-hoc.

### **4. FORMACION DE RECURSOS HUMANOS**

Esta línea de investigación forma parte del proyecto mencionado en la sección 2 el cual también se encuentra enmarcado dentro del ámbito de la temática adoptada por uno de sus integrantes como parte de su tesis de maestría en ingeniería de software y como parte de un informe técnico de investigación de su carrera de doctorado.

Además, se ha incorporado como parte integrante del proyecto, una beca BINIT para fomentar el acercamiento de jóvenes graduados a las actividades de investigación.

Por otra parte, también participan del proyecto, alumnos del último nivel la carrera de Ingeniería en Sistemas de Información de la UTN FRC, próximos a recibirse y con perspectivas de iniciarse en una carrera de posgrado o doctorado, con lo cual, uno de los objetivos del proyecto es el

contribuir a la formación de dichos alumnos.

### **5. BIBLIOGRAFIA**

[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.