

Hacia un Lenguaje Gráfico para SPARQL-DL

Christian Gimenez¹

Germán Braun^{1,2,3}

Laura Cecchi¹

Pablo Fillottrani^{2,4}

email: {christian.gimenez, german.braun, lcecchi}@fi.uncoma.edu.ar,
prf@cs.uns.edu.ar

¹*Grupo de Investigación en Lenguajes e Inteligencia Artificial*
Departamento de Teoría de la Computación - Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

²*Laboratorio de I&D en Ingeniería de Software y Sistemas de Información*
Departamento de Ciencias e Ingeniería de la Computación
UNIVERSIDAD NACIONAL DEL SUR

³*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

⁴*Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC)*

Resumen

Esta línea de investigación se desarrolla en forma colaborativa entre docentes-investigadores de la Universidad Nacional del Comahue y de la Universidad Nacional del Sur, en el marco de proyectos de investigación financiados por las universidades antes mencionadas.

El objetivo general de este trabajo es el desarrollo de metodologías que integren consultas textuales y gráficas con razonamiento automático, en ambientes de modelado conceptual y ontológico. Así como el uso de razonamiento para la identificación de errores en la ingeniería ontológica es vital y tiene consecuencias evidentes, es esencial que sea también considerado en la generación de consultas, junto con el soporte gráfico para la interacción de los usuarios con los modelos.

Palabras Clave: Consultas gráficas, Ingeniería de Software basada en Conocimiento, Lógicas Descriptivas, Ontologías.

Contexto

Este trabajo está parcialmente financiado por la Universidad Nacional del Comahue, en el

marco del proyecto de investigación Agentes Inteligentes y Web Semántica (04/F014) y de una Beca de Investigación de en la categoría Graduado de Iniciación; por la Universidad Nacional del Sur a través del proyecto de investigación Integración de Información y Servicios en la Web (24/N027) y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. Los proyectos de investigación tienen una duración de cuatro años y la beca doctoral una duración de 5 años, finalizando esta última en abril de 2019. La Beca de Investigación UNCo en la categoría Graduado de Iniciación tiene una duración de un año.

1. Introducción

En trabajos previos [1, 2, 3, 4] se ha presentado a *crowd*, una herramienta para el modelado conceptual y ontológico. El objetivo es el de permitir a los usuarios realizar modelos de forma visual y colaborativa, con asistencia de servicios de razonamiento para determinar inconsistencias. Esto es posible debido a la formalización en Lógica Descriptiva propuesta por [5],

la cual puede ser expresada de forma textual y consultada por consistencia utilizando razonadores como Racer [6] o Koncude [7].

Se ha mostrado que aplicar técnicas visuales para el acceso de datos es particularmente exitoso. Esto se debe al incremento de usuarios no expertos que no están familiarizados con lenguajes de consultas de bases de datos como SQL. Además, en comparación con el uso del lenguaje natural, las consultas visuales superan el problema de depender en el lenguaje del usuario y las limitaciones impuestas por el área de aplicación [8]. Es importante proveer a los expertos de dominios con herramientas para la formulación de consultas para expresar sus necesidades de información en términos de consultas sobre las ontologías [9].

Las ontologías, ya sean modeladas gráficamente o incorporadas textualmente utilizando alguna sintaxis formal (por ejemplo, OWL 2 [10]), pueden ser consultadas en cuanto a su nivel conceptual para poder estudiarlas o para hacer consultas que integren tanto el nivel terminológico (TBox), el de relaciones (RBox) como el de aserciones (ABox). Esto se puede llevar a cabo por medio de consultas SPARQL-DL [11, 12]. Este lenguaje de consulta se define para OWL-DL y puede combinar consultas del TBox, RBox y ABox, a diferencia de DIG [13] y nRQL [14] que están limitados a consultas atómicas. Además, se alinea con SPARQL [15] para mejorar la interoperabilidad de aplicaciones con la Web Semántica.

Sin embargo, basados en las ventajas de las consultas visuales expresadas en los párrafos anteriores, es posible expresarlas por medio de primitivas visuales.

La estructura del presente trabajo es la siguiente. En la sección 2 presentamos los objetivos de los proyectos de investigación en los que se enmarca este trabajo y describimos la línea de investigación actual. En la sección 3 indicamos algunos resultados obtenidos y trabajos futuros. Finalmente, comentamos aspectos referentes a la formación de recursos humanos en esta temática.

2. Línea de Investigación y Desarrollo

El proyecto de investigación Agentes Inteligentes y Web Semántica, UNCo, tiene varios objetivos generales. Uno de ellos es el de generar conocimiento especializado en el área de agentes inteligentes y en lo referente a la representación y el uso del conocimiento en la Web Semántica, para entre otros, dar soporte a comunidades de desarrollo de ontologías. En este sentido, y compartido con el proyecto de investigación Integración de Información y Servicios en la Web, de la UNS, se estudian técnicas de representación de conocimiento y razonamiento, metodologías de modelado conceptual y mecanismos para la interoperabilidad de aplicaciones, tanto a nivel de procesos como de datos.

Ambos proyectos confluyen en la línea de investigación de este trabajo, en la que se explora entre otros, sobre temas afines a la representación del conocimiento, las Lógicas Descriptivas [16], las Ontologías, la Ingeniería de Software basada en Conocimiento y la Ingeniería de Conocimiento.

En esta línea de investigación se propone desarrollar un lenguaje gráfico que represente las consultas SPARQL-DL a realizarse sobre una ontología. Inicialmente, se describirán primitivas visuales para consultar las estructuras del TBox. Las consultas visuales se traducirán a su versión textual, las cuales en conjunto con una ontología serán ejecutadas sobre un motor SPARQL-DL para obtener las respuestas. Este proceso se describe en el siguiente párrafo con un ejemplo.

Suponemos un modelo extraído del ejemplo de [5] de un diagrama UML mostrado en la Figura 1. Se realiza una consulta visual con la que se desea buscar las clases "Y" que sean subclases de "X". Ésta se transforma en una consulta SPARQL-DL la cual se utilizará para alimentar al motor junto con la formalización del diagrama UML a DL. Las respuestas obtenidas, son mostradas al usuario en forma textual, pudiendo ser interpretadas gráficamente reemplazando las variables por los nombres de clases encontrados.

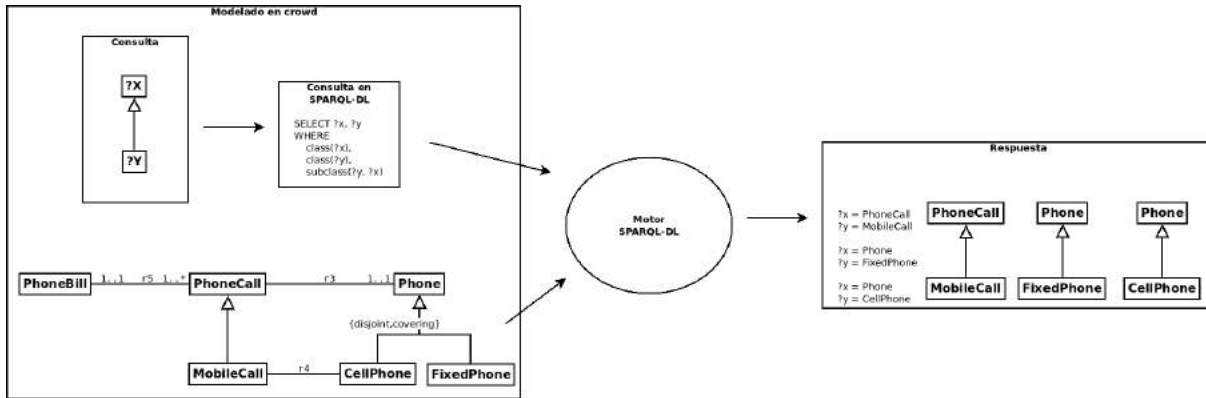


Figura 1: Ejemplo de consulta visual.

Para implementar este lenguaje, se continuará con el desarrollo de la herramienta Web denominada *crowd*, la cual contribuye al modelador en el diseño y la visualización de ontologías durante su evolución, actualmente por medio del lenguaje UML, asistidos por técnicas de razonamiento automáticas. Para ello, se implementará una extensión de la arquitectura que consistirá en un diseño de interfaz que permita el modelado de consultas utilizando dicho lenguaje gráfico. En el servidor, se llevará a cabo la traducción y su ejecución sobre una ontología ofrecida por el usuario, mostrando los resultados en el cliente.

3. Resultados Obtenidos y Trabajo Futuro

Actualmente, *crowd* permite crear modelos basados en un subconjunto de primitivas de UML. Puede determinar las inconsistencias del modelo debido a la formalización a Lógica Descriptiva del modelo basándonos en la propuesta en [5], y posteriormente utilizando un programa razonador. También, es posible obtener la representación textual del modelo formalizado en OWL 2 [17, 10] para ser utilizado con otras herramientas. Un prototipo está disponible en la URL crowd.fi.uncoma.edu.ar.

Además, está en desarrollo el soporte al metamodelo descrito en [18] para brindar interoperabilidad entre los lenguajes UML, EER y ORM considerando sus diferencias semánticas. Esto

implica la necesidad de incorporar las primitivas gráficas correspondientes a los lenguajes indicados y la implementación del metamodelo en el servidor.

Se propone definir un lenguaje visual para la realización de consultas sobre las ontologías modeladas. Dicho lenguaje se pretende que posea primitivas que sean similares a los lenguajes de modelado ampliamente usados en la ingeniería de Software para minimizar la curva de aprendizaje. De esta manera, se espera poder buscar patrones que se presenten en la ontología mezclando consultas de nivel conceptual (TBox) como del nivel de datos (ABox).

Finalmente, como trabajo futuro a este desarrollo se considera implementar una extensión de *crowd* que permita al usuario el modelado de las consultas y la ontología, como así también la importación de los datos necesarios para obtener las respuestas.

4. Formación de Recursos Humanos

Uno de los autores de este trabajo está inscripto en el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET).

En la Universidad Nacional del Comahue, Facultad de Informática, se otorgaron Becas de Investigación en la categorías de Graduado de Iniciación. Una de esas becas fue otorgada a uno de los autores de este trabajo. En el marco de es-

ta beca, se espera que el investigador comience un posgrado.

Referencias

- [1] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Fillottrani. Una Arquitectura Cliente-Servidor para Modelo Conceptual Asistido por Razonamiento Automático. In *XVIII Workshop de Investigadores en Ciencias de la Computación*, 2016.
- [2] Christian Gimenez, Germán Braun, Laura Cecchi, and Laura Fillottrani. crowd: A Tool for Conceptual Modelling assisted by Automated Reasoning - Preliminary Report. In *the 2nd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '16 JAIIO '16*, 2016.
- [3]
- [4] Germán Braun, Christian Gimenez, Pablo Fillottrani, and Laura Cecchi. Towards conceptual modelling interoperability in a web tool for ontology engineering. In *the 3rd Simposio Argentino de Ontologías y sus Aplicaciones SAOA '17 JAIIO '17*, 2017.
- [5] Daniela Berardi, Diego Calvanese, and Giuseppe De Giacomo. Reasoning on UML class diagrams. *Artif. Intell.*, 168(1-2):70–118, 2005.
- [6] V. Haarslev and R. Möller. Racer system description. In R. Goré, A. Leitsch, and T. Nipkow, editors, *International Joint Conference on Automated Reasoning, IJ-CAR'2001, June 18-23, Siena, Italy*, pages 701–705. Springer-Verlag, 2001.
- [7] Andreas Steigmiller, Thorsten Liebig, and Birte Glimm. Konclude: System description. *Journal of Web Semantics (JWS)*, 27:78–85, 2014.
- [8] Amineh Fadhil. Ontovql: Ontology visual query language. Master's thesis, Concordia University Montreal, Quebec, Canada, September 2008.
- [9] Ahmet Soylu, Evgeny Kharlamov, Dmitriy Zheleznyakov, Ernesto Jiménez-Ruiz, Martin Giese, and Ian Horrocks. OptiqueVQS: Ontology-Based Visual Querying. In *Proceedings of the International Workshop on Visualizations and User Interfaces for Ontologies and Linked Data co-located with 14th ISWC*, 2015.
- [10] World Wide Web Consortium (W3C). OWL 2 Web Ontology Language Document Overview (Second Edition), 2012. <http://www.w3.org/TR/owl2-overview/>, accedido en Junio de 2013.
- [11] Ilianna Kollia, Birte Glimm, and Ian Horrocks. Sparql query answering over owl ontologies. In Grigoris Antoniou, Marko Grobelnik, Elena Simperl, Bijan Parsia, Dimitris Plexousakis, Pieter De Leenheer, and Jeff Pan, editors, *The Semantic Web: Research and Applications*, pages 382–396, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
- [12] Evren Sirin and Bijan Parsia. Sparql-dl: Sparql query for owl-dl. In *In 3rd OWL Experiences and Directions Workshop (OWLED-2007)*, 2007.
- [13] S. Bechhofer, R. Moller, and P. Crowther. The DIG Description Logic Interface. In *In Proc. of International Workshop on Description Logics (DL2003)*, 2003.
- [14] Volker Haarslev, Ralf Möller, and Michael Wessel. Querying the semantic web with racer + nrql. In *In Proceedings of the KI-2004 International Workshop on Applications of Description Logics (ADL'04)*, 2004.
- [15] W3C OWL Working Group. *SPARQL 1.1 Query Language*. W3C Recommendation, 21 March 2013. Available at <http://www.w3.org/TR/sparql11-query/>.

- [16] Diego Calvanese, Maurizio Lenzerini, and Daniele Nardi. Description logics for conceptual data modeling. In *Logics for Databases and Information Systems*, pages 229–263. Kluwer, 1998.
- [17] Thorsten Liebig, Marko Luther, Mariano Rodriguez, Diego Calvanese, Michael Wessel, Ralf Möller, Matthew Horridge, Sean Bechhofer, Dmitry Tsarkov, and Evren Sirin. OWLlink: DIG for OWL 2. In *In Proceedings of the 5th OWL Experiences and Directions Workshop (OWLED-2008)*, page 48, 2008.
- [18] Pablo R. Fillottrani and C. Maria Keet. KF metamodel formalization. *CoRR*, abs/1412.6545, 2014.