

Lenguajes para Intercambio de Reglas en la Web Semántica

Pablo R. Fillottrani

Departamento de Ciencias e Ingeniería de la Computación

Universidad Nacional del Sur

Av. Alem 1253, Bahía Blanca 8000, Argentina

1. INTRODUCTION

El objetivo general de la investigación es la representación y el uso del conocimiento en sistemas computacionales, para lo cual es necesario desarrollar métodos eficientes para la representación, y mecanismos de razonamiento confiables que permitan su procesamiento. Para ser eficiente, la representación del conocimiento debe ser a la vez compacta y comprensible. Para ser confiable, el mecanismo de razonamiento debe no sólo ser exacto sino también intuitivo y computacionalmente tratable a la vez. El estudio de los formalismos que tienen estos objetivos aparentemente contradictorios involucra aspectos tanto teóricos como tecnológicos. Actualmente surge un renovado interés en estos temas debido a la necesidad emergente de compartir, actualizar y combinar el conocimiento de sistemas computacionales pre existentes.

El objetivo particular para el año próximo es la especificación de un lenguaje de intercambio de reglas para la Web Semántica. Las reglas representan una importante fuente de información en varias aplicaciones, por ejemplo proveen los fundamentos para representar la lógica de negocios en diversos sistemas de información modernos, y no existe todavía un estándar que permita compartir esta información entre diversas aplicaciones. La especificación de este lenguaje permitirá que las reglas escritas para una determinada aplicación, puedan ser publicadas, compartidas y re-usadas en otras aplicaciones posiblemente bajo otros motores de inferencia. Debido a la gran variedad existente en lenguajes de reglas y en las tecnologías usadas en sus motores de inferencia, este lenguaje deberá comprender un núcleo básico simple, y un conjunto de posibles extensiones para comunidades de usuarios específicas. Además, dado que el intercambio se realiza en el ámbito de la Web Semántica, será necesario considerar su interacción como extensión o complemento de otros lenguajes estándares ya existentes, como OWL, RDF(S) y SPARQL.

2. ANTECEDENTES

La Web ha sido diseñada como un espacio de información, con el objetivo de que sea útil no sólo para la comunicación hombre-hombre, sino también para que las computadoras puedan participar y ayudar [1, 2]. Uno de los mayores obstáculos para esto ha sido el hecho de que la mayoría de la información en la Web está diseñada para consumo humano, y aun si fue obtenida a partir de una base de datos con significado bien definido, la estructura de los datos no es evidente para un robot/agente que explora la Web. Dejando de lado el problema de la Inteligencia Artificial de entrenar a máquinas para que se comporten como personas, el objetivo de la Web Semántica en vez es el de desarrollar lenguajes para expresar la información en forma que sea procesable por las máquinas. La necesidad de automatizar la interoperación de los servicios accesibles mediante la Web, especialmente en aplicaciones B2B y de comercio electrónico, ha revalorizado este objetivo. Actualmente, esta interoperación se realiza a través de programas específicos que localizan y extraen información de las páginas e invocan servicios accesibles vía Web. Esta alternativa resulta poco útil ya que depende fuertemente del formato en el cual la información está representada (generalmente código HTML que contiene datos y presentación mezclados) y de las interfaces de los servicios (CGI, RMI, Corba, etc.). Con el propósito de dar solución a estos problemas, se ha creado el World Wide Web Consortium (W3C) [3] que reúne empresas (tanto proveedoras como consumidoras de servicios informáticos), instituciones académicas y de investigación, agencias de gobiernos e organismos sin fines de lucro con el objetivo de llevar la Web a su máximo potencial mediante el desarrollo de protocolos y recomendaciones para llevar estructura al contenido de las páginas Web, en forma no ambigua e interpretable por las computadoras. De esta forma se crea un contexto en donde herramientas automatizadas puedan rápidamente llevar a cabo tareas en representación de los usuarios. Desde 1994 los miembros de W3C han traba-

jado para producir estándares que proveen el fundamento para la Web de hoy, incluyendo HTML (HyperText Markup Language) [4], XML (Extensible Markup Language) [5], RDF (Resource Description Framework) [6, 7, 8], OWL (Web Ontology Language) [9] y otras tecnologías. Mediante la producción de estos estándares para lenguajes Web abiertos no-propietarios se busca evitar la fragmentación del mercado, y como consecuencia la fragmentación de la Web. Recientemente, W3C ha activado un grupo de trabajo [10] con el objetivo de especificar un lenguaje de reglas que permita su intercambio a través de la Web Semántica, el tema propuesto para la investigación. El propósito del grupo es balancear las necesidades de diversas comunidades de uso, tales como los usuarios de reglas de negocios, la comunidad de programación en lógica o los usuarios de otros estándares de la Web Semántica (RDF, OWL, etc), especificando extensiones para las cuales se pueda articular un consenso y que sean suficientemente motivadas por casos de uso. Como antecedente se tiene el gran interés generado por recientes trabajos de investigación sobre la combinación de lógicas descriptivas, lenguajes de ontologías y lenguajes de reglas [11, 12, 13, 14, 15, 16] y las conclusiones del Workshop on Rule Languages for Interoperability realizado en abril 2005 en Washington D.C., USA. Este grupo se encuentra ya trabajando en la caracterización de los casos de uso que definan las características esenciales para dicho lenguaje y sus posibles extensiones. Existen actualmente algunas propuestas relacionadas, tales como RuleML [17] y SWRL [18], que si bien no son totalmente satisfactorias de acuerdo a los objetivos mencionados, deben ser tenidas en cuenta para que los nuevos estándares no interfieran con inversiones en curso en dichas tecnologías. El tema es continuación directa de las actividades de investigación que se desarrollaron [19] dentro del grupo KRDB en la Free University of Bozen-Bolzano.

3. ACTIVIDADES Y METODOLOGÍA

En el contexto introducido, se propone investigar cómo integrar un lenguajes de intercambio de reglas con los lenguajes estándares ya existentes para la Web Semántica. La integración se realizará mediante la especificación de un lenguaje núcleo, y de un conjunto de extensiones específicas para cada aplicación. El detalle de las actividades y metodología para el primer año será el siguiente:

1. Estudio de alternativas para el lenguaje núcleo. Si bien es claro que el lenguaje núcleo debe estar basado en algún fragmento de la

lógica de Horn, no es tan claro cuál debe ser este fragmento. En principio, es sabido que si se agrega a OWL un lenguaje trivial de reglas el resultado es indecidible. Por lo tanto, el lenguaje núcleo debe ser suficientemente expresivo como para poder expresar la funcionalidad de la mayoría de las aplicaciones existentes, pero también debe ser suficientemente simple como para conservar las propiedades computacionales al incorporarse a un fragmento significativo de OWL.. Es necesario considerar en este punto los casos de uso que ya deberán estar caracterizados por el grupo de trabajo. Otro problema a resolver es la compatibilización entre el axioma de Closed World Assumption implícito en la gran mayoría de los lenguajes de regla, y el Open World Assumption presente en todas las lógicas descriptivas, sobre las que se base OWL. El soporte para tipos de datos (datatypes) de RDF(S) también debe ser considerado.

2. Análisis de tecnologías relacionadas: existen varias propuestas de estándares o lenguajes de reglas recientemente desarrollados que comparten parte de los objetivos mencionados: RuleML, SWSL, WSML, SWRL, Common Logic, Triple. Son principalmente esfuerzos académicos, con diversos enfoques, cuyas características, fortalezas y debilidades deben analizarse mediante su aplicación a los escenarios propuestos.
3. Desarrollo de un mecanismo de extensibilidad. Una vez finalizados los puntos anteriores, se deberá caracterizar un mecanismo mediante el cual se puedan definir extensiones al lenguaje núcleo, teniendo presente las funcionalidades dejadas afuera en primer lugar, y las necesidades de los distintas áreas de uso. Esta extensibilidad no debe afectar la usabilidad de la propuesta. Tiempo estimado: 3 meses.
4. Estudio y desarrollo de posibles extensiones: Varias extensiones al lenguaje núcleo serán necesarias. Algunas ya han sido identificadas, tales como Negación Clásica, Razonamiento por Default, Mundos Cerrados Locales, Lenguajes de Acción y Bases de Datos de Conocimientos. Otras como extensiones a mayores fragmentos de OWL o de la lógica de primer orden, razonamiento por restricciones, razonamiento rebatible, lógica de segundo orden o razonamiento con incertidumbre (fuzzy logic) son necesarias también. Se debe analizar la representación de cada una de estas posibilidades dentro del mecanismo de extensibilidad del punto anterior, cuánta funcionalidad es necesaria para los casos de uso, la complejidad resultante y la posible interacción entre

las diversas extensiones. No todas las extensiones podrán ser combinadas en un lenguaje unificado único.

4. CONTEXTO

El trabajo propuesto será llevado a cabo en el LIDIA, Laboratorio de Investigación y Desarrollo en Inteligencia Artificial del Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur. Dicho establecimiento cuenta con el hardware y el software apropiados para cumplir los objetivos del plan de trabajo propuesto. También cuenta con acceso a internet y acceso a bibliografía relativa a los temas involucrados en el plan de trabajo. El plan de trabajo se podrá desarrollar en el marco de las actividades de los proyectos PICT 15043 "Herramientas Inteligentes para la Web Semántica" y PAV 076 "Sistemas Inteligentes para Apoyo a los Procesos Productivos".

REFERENCES

- [1] Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. *Scientific American* (2001)
- [2] Decker, S., Melnik, S., van Harmelen, F., Fensel, D., Klein, M.C.A., Broekstra, J., Erdmann, M., Horrocks, I.: The semantic web: The roles of xml and rdf. *IEEE Internet Computing* (2000) 63–74
- [3] Jacobs, I.: World wide web consortium mission. (<http://www.w3c.org/Consortium/mission/>)
- [4] Raggett, D., Hors, A.L., Jacobs, I.: Html 4.01 specification. W3C Recommendation, <http://www.w3c.org/TR/REC-html40/> (1999)
- [5] Yergeau, F., Cowan, J., Bray, T., Paoli, J., Sperberg-McQueen, C.M., Maler, E.: XML 1.1 W3C recommendation. W3C Recommendation, <http://www.w3.org/TR/2004/REC-xml11-20040204/> (2004)
- [6] Manola, F., Miller, E.: RDF primer. W3C Recommendation, <http://www.w3.org/TR/rdf-primer/> (2004)
- [7] Klyne, G., Carroll, J.: RDF concepts and abstract syntax. W3C Recommendation, <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> (2004)
- [8] Hayes, P.: RDF semantics. W3C Recommendation, <http://www.w3.org/TR/rdf-mt/> (2004)
- [9] Patel-Schneider, P.F., Hayes, P., Horrocks, I.: Owl web ontology language semantics and abstract syntax. W3C Recommendation, <http://www.w3.org/TR/2004/REC-owl-semantic-20040210/> (2004)
- [10] Hawkes, S.: W3C rule interchange format working group charter. <http://www.w3.org/2005/rules/wg/charter> (2005)
- [11] Horrocks, I., Angele, J., Decker, S., Kifer, M., Grosz, B., Wagner, G.: Where are the rules? *IEEE Intelligent Systems* **18** (2003) 76–83
- [12] Grosz, B., Horrocks, I., Volz, R., Decker, S.: Description logic programs: Combining logic programs with description logic. In: *Proceedings WWW 2003, Budapest, Hungary* (2003)
- [13] Eiter, T., Lukasiewicz, T., Schindlauer, R., Tompits, H.: Combining answer set programming with description logics for the semantic web. In: *Proceedings KR 2004, Whistler, Canada* (2004) 141–151
- [14] Rosati, R.: On the decidability and complexity of integrating ontologies and rules. *Journal of Web Semantics* **3** (2005) 61–73
- [15] Franconi, E., Tessaris, S.: Rules and queries with ontologies: a unified logical framework. In: *Proceedings PPSWR 2004*. (2004) 50–60
- [16] Motik, B., Sattler, U., Studer, R.: Query answering for OWL-DL with rules. *Journal of Web Semantics* **3** (2005) 41–60
- [17] Boley, H., Tabet, S., Wagner, G.: Design rationale of RuleML: a markup language for semantic web rules. In: *Proceedings of the International Semantic Web Working Symposium, Stanford University* (2001)
- [18] Horrocks, I., Patel-Schneider, P.F.: A proposal for an OWL rule language. In: *Proceedings WWW 2004*. (2004) 723–731
- [19] Fillottrani, P., Franconi, E., Tessaris, S.: On the semantics of rules with ontologies. (UNIBZ-CS Technical Report)