View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Servicio de Difusión de la Creación Intelectual

Métodos Relacionales para la Especificación, Verificación, y Composición de Servicios Semánticos en la Web

Carlos José Gonzalía, Dpto. de Cs. e Ing. de la Computación, Universidad Nacional del Sur, Av. Alem 1253, 8000 Bahía Blanca, Argentina. Correo electrónico: cjg@cs.uns.edu.ar

Palabras clave: ingeniería del software, servicios semánticos en la web, métodos formales, verificación asistida por computadora, lógicas para la programación.

1. Introducción

El objetivo de nuestra investigación consiste en la aplicación de métodos formales de la ingeniería del software, específicamente aquellos basados en sistemas formales relacionales, a los problemas de desarrollo de servicios semánticos en la web. Los sistemas formales relacionales que se tienen en mente son descendientes de las álgebras relacionales de Tarski, y su uso será asistido por herramientas de software existentes o a crear durante la investigación propuesta.

Los servicios en web son una clase particular de programas, y como tales su desarrollo se beneficiaría de la aplicación de métodos formales de la ingeniería del software. Los problemas usuales de asegurar su corrección y otras propiedades deseables se suman a las propiedades específicas que deben brindar como servicios (en particular la composición de los mismos para obtener comportamientos más complejos y/o acordes a las necesidades específicas de la aplicación), más aún cuando se añade a ellos un nivel semántico de anotaciones que deben ser procesables por computadoras.

2. Objetivos generales de la investigación

El propósito de nuestra investigación es la aplicación de métodos relacionales a los problemas particulares de especificación y verificación de servicios semánticos en la web, tanto al nivel de servicios individuales como en sus interacciones por medio de composición. Se desean obtener formalizaciones asistidas por computadora de casos de estudio realistas, y también verificaciones de modelos de tales casos. Necesariamente, la especificación formal de tales casos es un requisito para proceder a pruebas de corrección o verificaciones de propiedades formales deseables en los servicios. Es de especial interés la formalización de pruebas de propiedades sobre los aspectos de algoritmos distribuidos que presenta la interacción y composición de servicios semánticos en la web.

Dada la complejidad computacional tanto de los problemas inherentes a los servicios semánticos en web, como de los sistemas formales a usar para investigarlos, el uso de herramientas de software es sensato. Tales herramientas se tomarán de entre las ya existentes y aplicadas con éxito en tareas de métodos formales, como así también nuevas herramientas a desarrollar (particularmente en lo que concierne a propiedades probabilísticas de los servicios semánticos). De especial interés es el posible uso de herramientas y marcos formales basados en la teoría de tipos constructiva, dada la importante vinculación entre la idea de demostración constructiva de una propiedad y la idea de algoritmo (según lo sugerido por correspondencias basadas en el isomorfismo de Curry-Howard).

3. Breve descripción de los fundamentos teóricos y técnicos del problema

Los servicios en web [10] son componentes de software que proveen una actividad de interés a usuarios que quieran acceder a tales servicios a través de internet. Esto por supuesto requiere que tanto la provisión, búsqueda y utilización de servicios en web sean realizados a través de protocolos estándar en los que todos los proveedores y usuarios de servicios concuerdan. Los servicios en web permiten entonces crear aplicaciones distribuidas a partir de componentes cuyo funcionamiento abstracto es anunciado por los creadores de los componentes y elegido de acuerdo a las necesidades del caso por los usuarios.

De la misma manera que la cantidad masiva de información disponible en internet y la correspondiente dificultad en procesarla eficientemente han llevado a la creación de tecnologías de web semántica (en las cuales anotaciones son provistas por encima de la información en si para permitir que las computadoras lleven a cabo dicho procesamiento eficiente), la potencial creación de numerosos servicios en web está acarreando similares problemas y soluciones [22,23,24,25,26]. El agregado de información semántica [4] por encima de los servicios en web posibilitaría un uso más eficaz, seguro y automático de los componentes de software provistos en web, permitiendo la creación de aplicaciones distribuidas en base a tales servicios.

Los métodos formales de ingeniería de software son herramientas lógico-matemáticas que toman la teoría de las ciencias de la computación y la enfocan en el problema de construir programas correctos. La obtención de especificaciones formales de los problemas a resolver, la derivación automática o semiautomática de programas a partir de tales especificaciones, y la verificación de algoritmos en cuanto a que posean propiedades deseadas son puntos centrales de los métodos formales (para nuestro caso, por ejemplo: [27]).

El uso de sistemas formales relacionales dentro de tales métodos tiene ya una larga historia y probado exito [3]. Desde su creación por Tarski en los 1940 como álgebras relacionales, estos sistemas matemáticos han evolucionado y diversificado en muchas clases de formalismos basados en la idea de relaciones matemáticas, permitiendo la descripción y análisis de muchos tópicos de las ciencias de la computación. En particular para nuestros objetivos, son de interés los tipos de álgebras relacionales que sirven como lógicas de programas, con las cuales la especificación, derivación y verificación formales de programas son llevadas a cabo.

Las clases de álgebras de mayor interés para los problemas específicos de servicios semánticos en web, son las álgebras relacionales propiamente dichas [9], las alegorías [2], y las álgebras de Kleene probabilísticas [7]. Las álgebras relacionales han sido usadas para expresar la estructura de sistemas de software, y en particular la composición de servicios semánticos en la web [12]. Las alegorías son una versión de estas álgebras en el estilo de la teoría de categorías, y son sumamente adecuadas para la especificación y derivación de programas, y por lo tanto para los servicios semánticos en web como clase particular de programas; estas alegorías se prestan particularmente al razonamiento en logica constructivista [8], como el autor encontró durante su trabajo previo [5,6]. Las álgebras de Kleene probabilísticas incorporan conceptos de paralelismo y comportamientos probabilísticos de los programas involucrados, siendo así adecuadas al análisis de colecciones de programas distribuidos y/o programas en los que respuestas correctas o con propiedades deseadas solo pueden garantizarse hasta cierto grado cuantificable.

Dada la complejidad de los componentes de software, como así también de las pruebas formales necesarias para verificar la corrección y otras propiedades relevantes de los mismos [17], es natural usar computadoras para asistir al usuario de métodos formales en estas tareas. Tanto en forma automatizada como interactiva, el uso de demostradores de teoremas, asistentes de demostración y verificadores de modelos será fundamental para los objetivos propuestos. Para las herramientas ya

existentes, la creación y desarrollo de bibliotecas de pruebas formales [1,16] referentes a servicios semánticos en web, usando sistemas relacionales como el medio de elección, será parte central del esfuerzo. Por otra parte, la creación y expansión de nuevas herramientas y prototipos de las mismas, en especial para sistemas relacionales probabilísticos, también formará parte central del trabajo a realizar.

4. Metodología y actividades a desarrollar

Las actividades a realizar consistirán en:

- identificación de las clases de sistemas relacionales que mejor se adecuen a métodos formales para servicios semánticos. Los candidatos de mayor interés serían variantes de las álgebras de Kleene probabilísticas.
- prueba de diferentes herramientas de software existentes para determinar las mejores en cuanto al uso de los sistemas formales relacionales elegidos. Esta es una lista no exhaustiva de herramientas de entre las que seleccionar:
  - Agda (asistente de demostración en teoría de tipos constructiva) http://appserv.cs.chalmers.se/users/ulfn/wiki/agda.php
  - Coq (asistente de demostración en teoría de tipos constructiva) http://coq.inria.fr/
  - Prover9 (demostrador de teoremas, admite lógica ecuacional de primer orden) http://www.cs.unm.edu/~mccune/prover9/
  - Isabelle (asistente de demostración para varias lógicas) http://www.cl.cam.ac.uk/research/hvg/Isabelle/
  - KAT-ML (asistente de demostración para álgebras de Kleene) http://www.cs.cornell.edu/Projects/KAT/
  - Yices (solucionador de satisfacción lógica con aritmética) http://yices.csl.sri.com/

  Es sumamente probable que combinaciones de dos o más herramientas sean necesarias para cubrir las diferentes necesidades de la formalización, y esta integración es otra tarea técnica a realizar en esta etapa.
- desarrollo de prototipos y herramientas de software para el uso de álgebras de relaciones probabilísticas en métodos formales para servicios semánticos. Se debe explorar la posibilidad de modificar una herramienta existente, por ejemplo de entre las estudiadas en la etapa previa. Sin emabargo, como se mencionó en la sección sobre logros previos, es probable que se necesite construir una herramienta de software completamente nueva para esta clase de álgebras.
- formalización y verificación de algoritmos distribuidos relevantes a los servicios semánticos en web. Ejemplos pertinentes serán tomados de la literatura del tema, con especial interés en aquellos que no posean una formalización por computadora al momento de ejecutarse esta etapa del proyecto.
- creación y uso de bibliotecas de formalizaciones y pruebas formales sobre servicios semánticos usando las herramientas de software ya existentes elegidas. La publicación en forma de hipertexto legible, disponible en internet, junto con el código de las formalizaciones en si, son los objetivos centrales de esta etapa.
- especificación, formalización y verificación de casos de interés en servicios semánticos, tanto a nivel de componentes como a nivel de su composición e interacción. Se considerarán casos tomados de estándares y productos de software cuyo uso esté difundido en el momento de abordar esta etapa, preferentemente a proyectos experimentales o de disponibilidad reducida.

La metodología a utilizar será esencialmente la prueba y evaluación de distintos formalismos y herramientas de software que los asistan, con la meta de identificar aquellos sistemas relacionales, asistentes de demostración, verificadores de modelos, demostradores de teoremas y similares, que mejor se presten a los objetivos antes mencionados. Habiendo establecido tales formalismos y herramientas como los deseados, la metodología para su uso y aplicación será la del practicante de métodos formales en ciencias de la computación: obtener prototipos de las formalizaciones de interés, para determinar así que se posee un fundamento correcto sobre el problema analizado, y luego expandir y detallar la formalización inicial hasta capturar todos los aspectos secundarios y casos particulares. Parte de este proceso incluirá en paralelo a la evaluación y prueba inicial de herramientas de software la construcción y mejora de prototipos de herramientas de software para álgebras relacionales probabilísticas, dada la escasez de tales herramientas en el presente momento.

5. Contexto y recursos humanos

Este proyecto tendrá lugar dentro del Dpto. de Cs. e Ing. de la Computación, Univ. Nac. Del Sur, Bahía Blanca. El autor formará parte del LISSI (Laboratorio de Invest. y Desarrollo en Ing. de Software y Sist. de Información) dentro de la mencionada unidad académica. Además de los recursos humanos y materiales allí disponibles para la realización del proyecto, el mismo ha sido enviado para su aprobación a la Agencia Nacional de Promoción Científica y Tecnológica como parte de los proyectos que dicha Agencia concedió a la UNS en la forma de PICT-PRHs, en los que el autor está enmarcado actualmente como docente e investigador del DCIC de la UNS.

Bibliografía

[1] H. Barendregt and H. Geuvers: Proof-assistants using Dependent Type Systems. In A. Robinson and A. Voronkov (eds.): *Handbook of Automated Reasoning*, Elsevier Science Publishers B.V., 2001, Volume II, chapter 18.

[2] R. Bird and O. de Moor: *Algebra of Programming*. International Series in Computer Science. Prentice Hall, 1997.

[3] C. Brink, W. Kahl and G. Schmidt (eds.): *Relational Methods in Computer Science*. Advances in Computer Science. Springer-Verlag, 1997.

[4] D. Fensel, H. Lausen, A. Polleres, J. de Bruijn, M. Stollberg, D. Roman, J. Domingue: *Enabling Semantic Web Services - The Web Service Modeling Ontology*. Springer, 2007.

[5] C. Gonzalia: *Relations in Dependent Type Theory*. PhD thesis. Technical Report no. 14D, Department of Computer Science and Engineering, Chalmers University of Technology and Göteborg University, 2006.

[6] C. Gonzalia: Towards a Formalisation of Relational Database Theory in Constructive Type Theory. En *RelMICS 7, Bad Malente, Germany, May 2003, Revised Selected Papers*, LNCS 3051, Springer-Verlag, 2004.

[7] A. McIver and C. Morgan: *Abstraction, Refinement and Proof for Probabilistic Systems*. Monographs in Computer Science, Springer, 2005.

[8] B. Nordström, K. Petersson and J.M. Smith: *Programming in Martin-Löf's Type Theory. An Introduction*. International Series of Monographs on Computer Science, No. 7. Oxford University Press, 1990.

[9] G. Schmidt and T. Ströhlein: *Relations and Graphs. Discrete Mathematics for Computer Scientists*. EATCS Monographs on Theoretical Computer Science. Springer-Verlag, 1993.

[10] R. Studer, S. Grimm, A. Abecker (Eds.): *Semantic Web Services - Concepts, Technologies, and Applications*. Springer, 2007.

[11] D. Berardi, D. Calvanese, G. De Giacomo, and M. Mecella: Composing Web Services with Nondeterministic Behavior. En *Proc. of the IEEE Int. Conf. on Web Services 2006 (ICWS 2006)*, IEEE Computer Society, 2006.
[12] P. Höfner, F. Lautenbacher: Algebraic Structure of Web Services. En *Proceedings of the 3rd International Workshop on Automated Specification and Verification of Web Systems (WWV 2007)*, Electronic Notes in Computer Science, Volume 200, Issue 3, 23 May 2008, Pages 171-187.
[13] C. Gonzalia and A. McIver: Automating Refinement Checking in Probabilistic System Design. En *ICFEM 2007, Boca Raton, FL, USA, November 14-15, 2007. Proceedings*. Lecture Notes in Computer Science, Volume 4789, Springer, 2007.
[14] A.K. McIver, C. Gonzalia, E. Cohen and C.C. Morgan: Using probabilistic Kleene algebra pKA for protocol verification. Journal of Logic and Algebraic Programming. Volume 76, Issue 1, May-June 2008, Pages 90-111.
[15] A. McIver, C. Morgan and C. Gonzalia: Proofs and refutations for probabilistic systems. En *FM 2008, Turku, Finland, May 26-30, 2008*. Lect. Notes in Comp. Sci., Vol. 5014, Springer, 2008.
[16] Hai H. Wang, Jin Song Dong, Jing Sun and Jun Sun: Reasoning support for Semantic Web ontology family languages using Alloy. Multiagent and Grid Systems - An International Journal, Vol. 2, pp. 455-471, IOS Press, 2006.
[17] Li Ye, Junliang Chen: Formal functional description of semantic web services: the logic description method. En *Proceedings of the ICSE 2006 international workshop on Service-oriented software engineering, Shanghai, China*, ACM Press, 2006.
[18] P. Höfner, G. Struth: On Automating the Calculus of Relations. En *Automated Reasoning - 4th International Joint Conference, IJCAR 2008 Sydney, Australia, August 12-15, 2008 - Proceedings*. Lecture Notes in Computer Science, Vol. 5195, Springer Verlag, 2008.
[19] P. Höfner, G. Struth: Automated Reasoning in Kleene Algebra. En *Automated Deduction - CADE-21 - 21st International Conference on Automated Deduction Bremen, Germany, July 17-20, 2007, Proceedings*. Lecture Notes in Computer Science, Vol. 4603, Springer Verlag, 2007.
[20] S. Agarwal, B. Sprick: Specification of Access Control and Certification Policies for Semantic Web Services. En *EC-Web 2005*, Lecture Notes in Computer Science, Vol. 3590, Springer Verlag, 2005.
[21] J. M. Garcia, D. Ruiz, A. Ruiz-Cortes, O. Martin-Diaz, M. Resinas: An Hybrid, QoS-Aware Discovery of Semantic Web Services Using Constraint Programming. En *Service-Oriented Computing - ICSOC 2007*, Lecture Notes in Computer Science, Vol. 4749, Springer Verlag, 2008.
[22] Jianwen Su, Tevfik Bultan, Xiang Fu and Xiangpeng Zhao: Towards a Theory of Web Service Choreographies. En *WS-FM 2007, Brisbane, Australia, September 28-29, 2007. Proceedings*. Lecture Notes in Computer Science, Vol. 4937, Springer Verlag, 2008.
[23] A. Friesen, E. Börger: A high-level specification for Semantic Web Service Discovery Services. En *Workshop proceedings of the sixth international conference on Web engineering (SMIWEP-MATeS'06)*, ACM International Conference Proceeding Series, Vol. 155, 2006.
[24] Wang, H. H., Gibbins, N., Payne, T., Saleh, A. and Sun, J.: A Formal Semantic Model of the Semantic Web Service Ontology (WSMO). En *Twelfth IEEE International Conference on Engineering of Complex Computer Systems, July 11 - 14, 2007, Auckland, New Zealand*.
[25] A. Bucchiarone, M. ter Beek and S. Gnesi: Formal Methods for Service Composition. En *3rd South-East European Workshop on Formal Methods (SEEFM'07) - Proceedings*. En prensa, SEERC.
[26] Montali, M., Pesic, M., van der Aalst, W., Chesani, F., Mello, P, and Storari, S: Declarative Specification and Verification of Service Choreographies. ACM Transactions on the Web, Volume 4, Issue 1, January 2010.
[27] Bhargavan, K., Fournet, C., Gordon, A.D.: Verifying Policy-Based Web Services Security. Microsoft Research, November 2007.