

MODELADO ESPECÍFICO DE DOMINIO (DSM) PARA UNA FAMILIA DE APLICACIONES DE SOFTWARE DE SIMULACIÓN DE ECOSISTEMAS NATURALES

Urciuolo Adriana, Iturraspe Rodolfo, Moyano Ezequiel, Gel Matías
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Onas 450, (9410) Ushuaia. Tierra del Fuego. Tel: ++54-2901-443533

aurciuolo@untdf.edu.ar, riturraspe@untdf.edu.ar, emoyano@untdf.edu.ar, mgel@untdf.edu.ar

Resumen

El Dominio de aplicaciones de simulación de ecosistemas naturales (DSEN) presenta características propias que requieren de un manejo específico, tales como la complejidad inherente a la información ambiental y a la aplicación de modelos matemáticos para simular los fenómenos naturales. La evolución del software del dominio y los requerimientos del mismo: manejo de complejidad, facilidades de integración y variabilidad en la configuración de aplicaciones particulares hacen necesaria la utilización de técnicas apropiadas de ingeniería de dominio, que faciliten reuso y permitan elevar el nivel de abstracción de los lenguajes utilizados acercando al desarrollador a los conceptos del dominio.

El modelado específico del dominio (DSM) propone elevar el nivel de abstracción, expresando las soluciones directamente en términos del dominio. Utiliza lenguajes específicos de dominio en combinación con generadores de código y frameworks de dominio, pasando de especificaciones de muy alto nivel a código ejecutable. Provee verdaderos beneficios cuando el dominio es bien acotado y las posibles variaciones están previstas.

En el marco de la presente línea de investigación se propone utilizar DSM para el desarrollo de software de una familia de aplicaciones DSEN definida mediante un enfoque de reuso sistemático.

Palabras clave: reuso, dominio específico, familia aplicaciones

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la UNTDF, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática de Sistemas de Información Ambiental – Hidroinformática desde hace más de 10 años (acreditados por la UNPSJB previamente a la creación de la UNTDF). Actualmente esta línea se enmarca en el Proyecto recientemente iniciado: “Red para la conservación de Ecosistemas fluviales de Patagonia - CONICET (Feb/2013-Feb/2018 - coordinado por el CENPAT-CONICET, con nodos en Chubut, Neuquén y Tierra del Fuego) siendo coordinado el Nodo Tierra del Fuego por la UNTDF. El Proyecto se lleva adelante con un enfoque multidisciplinario por parte de organismos de investigación y gestión en la Patagonia y consta de diversos componentes. El grupo responsable de la presente línea de investigación desarrolla sus actividades en el Nodo Tierra del Fuego de la Red Ecofluvial, en los siguientes componentes del Proyecto: Sistema integrado de información de cuencas y Modelado de ecosistemas en cuencas patagónicas.

Introducción

El estudio de los sistemas de información ambiental adquiere cada vez mayor importancia, dada la función esencial que los mismos cumplen en la toma de decisión para el manejo y aprovechamiento sustentable de los recursos naturales y el medio ambiente [5]. Las particularidades de este dominio hacen necesario un permanente estudio, análisis y reflexión acerca de metodologías de desarrollo de software adecuadas para modelar los ecosistemas naturales [8]. Aplicaciones típicas del dominio son las relativas a la modelación de procesos físicos (por ej. crecidas de ríos), como es el caso del software de modelos de simulación hidrológica, climática, calidad de aguas, etc. Existen varias generaciones de estos modelos, los cuales proveen medios de extrapolación cuantitativa y predicción de fenómenos naturales a partir de variables físicas y parámetros (por ej. modelos lluvia-caudal). Uno de los requerimientos básicos de estas aplicaciones es el *manejo de la complejidad*, debida principalmente a ciertas características de la información ambiental (datos espacio-temporales) y a la definición de escenarios de simulación con diferentes procesos naturales y métodos intercambiables. La integración de información proveniente de diferentes subdominios y la simulación de procesos físicos de la naturaleza a través de modelos matemáticos constituyen problemas que caracterizan estas aplicaciones y deben resolverse repetidas veces ante las diferentes situaciones naturales a simular. El software en muchos casos ha sido desarrollado y utilizado principalmente por expertos del dominio, con grandes dificultades para su extensión, su ajuste y/o adaptación a distintos escenarios, así como para la incorporación de nuevas o diferentes formas de simular un mismo proceso, sin posibilidad de reuso del esfuerzo realizado. Por otra parte los expertos no cuentan con facilidades para trabajar en un nivel de abstracción apropiado.

No obstante los problemas mencionados, el desarrollo de modelos reusables utilizando medios de especificación y lenguajes apropiados a las particularidades del software en el dominio de simulación de ecosistemas naturales (DSEN), aún es incipiente. La evolución del software del dominio y los requerimientos del mismo: manejo de complejidad, facilidades de integración y variabilidad en la configuración de aplicaciones particulares, hacen necesaria la utilización de técnicas apropiadas de ingeniería de dominio, que permitan la construcción de una familia de productos, los cuales compartiendo ciertos artefactos de software, puedan ser adaptados a diferentes configuraciones. Asimismo resulta necesario contar con lenguajes específicos del dominio que permitan a los desarrolladores del software trabajar con mayor nivel de abstracción.

Existen numerosas líneas de investigación en la actualidad en las cuales se analiza el problema de la complejidad de las aplicaciones ambientales, abordando la cuestión del reuso y modelado específico de dominio. En el pasado se han desarrollado frameworks, modelos conceptuales y microarquitecturas para el problema de la representación de objetos geográficos y campos continuos, así como modelos conceptuales y microarquitecturas OO para modelos de simulación hidrológica [10] [11] [12][13] que abordaban el problema de reuso de diseño. En cuanto al modelado apropiado al dominio se ha propuesto un Perfil UML 2 para aplicaciones de monitoreo ambiental [14]. No obstante, aún no se ha trabajado bajo el criterio de reuso sistemático [3] y son escasos los DSL para la modelación del comportamiento de ecosistemas naturales [1]

Los DSL (Lenguajes específicos de dominio) se han desarrollado en diversas áreas para facilitar la construcción de modelos en un nivel más cercano al conceptual a los expertos del dominio. El modelado de software con adecuadas herramientas para la definición de

los conceptos específicos del dominio es muy conveniente para el desarrollo de sistemas complejos. En DSM [4] los modelos se construyen usando conceptos que representan objetos del dominio de aplicación en lugar de conceptos de un lenguaje de programación determinado. El lenguaje de modelado sigue las abstracciones y semántica del dominio, permitiendo que los desarrolladores se perciban trabajando directamente con los conceptos del dominio. Para estos fines, DSM utiliza lenguajes específicos de dominio en combinación con generadores de código y frameworks de dominio, pasando de especificaciones de muy alto nivel a código ejecutable.

El desarrollo de una plataforma DSM requiere una considerable inversión de recursos, la necesidad de expertos del dominio, así como desarrolladores experimentados. Este trabajo se ve facilitado en la actualidad por cuanto existen entornos especializados que permiten desarrollar los lenguajes y generadores específicos de la solución DSM sobre la arquitectura que ofrecen (“Language Workbenches” s/Fowler)

Mientras un lenguaje de modelado tradicional tiende a ser general, un lenguaje DSM resulta lo más específico posible, elevando el nivel de abstracción de los modelos. Mientras más específico sea el dominio, mayor es el beneficio en productividad de aplicar DSM.

Al analizar la viabilidad de una solución DSM se debe considerar el beneficio devuelto durante su uso: cuántas variaciones de un producto permite y el número de características en que varían los productos permitidos. En el dominio DSEN, si bien pueden realizarse numerosas variaciones de un mismo producto, la gran dispersión de modelos de simulación de ecosistemas naturales existentes, puede disminuir la utilidad del DSL y las posibilidades de reuso. De allí la importancia de definir claramente una Familia de aplicaciones con un enfoque de reuso

sistemático. Las Líneas de Productos de Software (LPS) proveen reuso sistemático y un modo conveniente de diseñar e implementar una variedad de sistemas fuertemente relacionados. [3]. En el marco de este enfoque es posible trabajar en familias de aplicaciones que comparten características comunes, maximizando los beneficios de DSM [6].

El diseño sistemático de una arquitectura y un conjunto de recursos (assets) reusables constituyen la base para construir una familia de aplicaciones relacionadas [2]. Esta definición es importante para la posibilidad de acotar el trabajo en un dominio a un subconjunto del mismo, llamado “familia de aplicaciones”

Se debe tener en cuenta al adoptar este enfoque, que no es necesario construir todos los productos de una Línea de Productos de software, sino que una LPS puede pensarse como la definición de un conjunto virtual de productos, que comparten características comunes a ser construidos a partir del mismo conjunto de recursos centrales. Solo aquellos productos que satisfagan cierta necesidad serán construidos, mientras que los otros representan la capacidad potencial de construcción de productos. Así se espera desarrollar la arquitectura y recursos para una familia DSEN, implementando sólo aquellos componentes de interés al caso de estudio del Proyecto.

Los productos en una familia de aplicaciones DSEN pueden variar en términos de su comportamiento, configuración física, factores de escala, tipo de parámetros, etc.

En el marco de esta línea de investigación se espera avanzar en el desarrollo de los recursos centrales (arquitectura y componentes básicos) necesarios para la construcción de software en una familia de aplicaciones DSEN, generando además un Framework de aplicación y un lenguaje específico para el dominio que facilite a los expertos el desarrollo de nuevas aplicaciones.

Líneas de Investigación y Desarrollo

La línea de investigación actual se enfoca en DSM para familias de aplicaciones DSEN. De acuerdo a lo expuesto, y a los fines de lograr los objetivos del proyecto, se hace necesario el estudio de diferentes tópicos vinculados al tema Ingeniería de Dominio, principalmente en materia de reuso sistemático y DSM. Se describen las líneas de investigación y desarrollo del grupo en estos dos aspectos:

1) Reuso sistemático: En relación e este tema, se plantea el estudio de métodos para las distintas etapas del Desarrollo de Líneas de Productos de Software [3]: Análisis de dominio, Modelado de la variabilidad, Desarrollo de arquitecturas referenciales para dominios específicos, Definición de recursos centrales, Definición familias de aplicaciones (dado que se trabajará con una familia específica)

En cuanto a desarrollo, se obtendrá en esta línea una arquitectura referencial (a partir del análisis de dominio), donde desarrollar los componentes principales de un modelo de simulación de cuenca básico (incluyendo simulación hidrológica), teniendo previstas distintas variaciones para futuros trabajos.

La arquitectura definida deberá entonces proveer facilidades para ser instanciada en aplicaciones de la familia DSEN seleccionada

2) DSM (Modelado específico de dominio)

Los temas centrales de estudio en este aspecto, para dar soporte a un alto nivel de abstracción, son: DSLs (lenguajes específicos de dominio), frameworks específicos de dominio y generadores de código [9].

En cuanto a desarrollo se prevé la creación de un ambiente integrado de desarrollo y de un DSL específicos para la familia de aplicaciones del dominio DSEN seleccionada, permitiendo se eleve el nivel de abstracción, ocultando a los desarrolladores la complejidad en implementación.

Se trabaja con expertos del dominio, para obtener el conocimiento necesario acerca de

métodos de simulación de procesos físicos y la validación continua de resultados.

Resultados y Objetivos

Si bien la línea de investigación en el marco del Proyecto de Red aprobado por CONICET (febrero/2013) es incipiente, el grupo de investigación ya ha obtenido resultados de proyectos anteriores que constituyen un insumo para las actuales investigaciones. Entre ellos pueden citarse:

- Microarquitecturas de diseño OO para sistemas de modelación hidrológica [1].
- Patrones conceptuales para sistemas de información hídrica [2]
- Perfil UML2 para aplicaciones de Monitoreo ambiental [14]
- Análisis de dominio para sistemas complejos (Tesis de grado Lic. Ezequiel Moyano)
- Framework de soporte a aplicaciones de Glaciología (Tesina de grado en curso Matías Gel)

Los resultados mencionados constituyen aportes al conocimiento del dominio, sus características y avances en el desarrollo de una arquitectura referencial.

Objetivo general

Contribuir al desarrollo de modelos y lenguajes específicos de dominio para una familia de aplicaciones DSEN que permitan incrementar productividad y reuso, así como elevar el nivel de abstracción de lenguajes utilizados.

Objetivos específicos

- ✓ Adaptación de técnicas de análisis de dominio apropiadas para el modelado de sistemas que simulan el comportamiento de ecosistemas de la naturaleza.
- ✓ Definición de características comunes, ámbito y variabilidad para aplicaciones de simulación de ecosistemas naturales.
- ✓ Definición de la familia de aplicaciones de software DSEN y sus recursos centrales (core assets).

- ✓ Definición de una arquitectura referencial para aplicaciones DSEN, que considere el conjunto de variaciones permitidas para la familia de productos seleccionada.
- ✓ Desarrollo de componentes centrales para la implementación de productos de la familia de aplicaciones
- ✓ Definición de un lenguaje específico de dominio y de un Framework de aplicación para expertos del dominio.

La validación de resultados será realizada en la construcción de un modelo de simulación para la cuenca del río Grande, correspondiente al caso del estudio del Nudo Tierra del Fuego de la Red Ecofluvial.

Formación de Recursos Humanos

En el marco de esta línea de investigación se han concluido dos tesis de grado de Licenciatura en Informática de integrantes del Grupo de investigación (Lic. Ezequiel Moyano, APU Matías Gel) y se está formulando la propuesta para la tesis del posgrado (Magister en Ing. de Software – UNLP) (Lic Ezequiel Moyano). Asimismo se encuentra en desarrollo la tesis de Doctorado en Ciencias Informáticas (UNLP) sobre Componentes de Dominio para aplicaciones ambientales

El grupo de investigación es multidisciplinario, por cuanto, si bien se consideran fundamentales las actividades relativas al análisis y diseño de sistemas, es necesario el aporte de expertos del dominio, a los fines obtener conocimiento del mismo.

Referencias

- [1] Andrew Fall, Joseph Fall *A domain-specific language for models of landscape dynamics* Original Research Article *Ecological Modelling, Volume 141, Issues 1–3, 1 July 2001, Pages 1-18*
- [2] Aurangzeb Khan, Farooque Azam and Jahanzaib Khan. *Architecture Centric Development in Software Product Lines. International Journal of Computer Applications* 30(2):6-12, September 2011.

[3] Clements P. and Northrop. *Software Product Lines: Practices and Patterns*. Addison-Wesley, Boston, MA, 2001.

[4] Fowler M. *Domain-Specific Languages – USA* - Addison-Wesley Professional - 2011.

[5] Green D., Klomp N. *Environmental informatics - a new paradigm for coping with complexity in nature*. Complexity International Vol 6, 1998.

[6] Ilker Altintas N., Semih Cetin, Ali H. Dogru, Halit Oguztuzun, "Modeling Product Line Software Assets Using Domain-Specific Kits," *IEEE Transactions on Software Engineering*, vol. 38, no. 6, pp. 1376-1402, Nov.-Dec., 2012.

[8] Parent C., S. Spaccapietra E. Zimányi, *Spatio-temporal conceptual models: Data structures + space + time*. In C. B. Medeiros, editor, Proc. ACM-GIS, pp 26-33, 1999.

[9] Steven K. y Juha-Pekka, Tolvanen - *Domain-Specific Modeling Enabling Full Code Generation* - USA - IEEE Computer Society / John Wiley & Sons – 2008.

[10] Urciuolo Adriana, Iturraspe Rodolfo, Villarreal Martín. "Microarquitecturas de Diseño OO para Sistemas de Modelación Hidrológica". La Plata, Octubre de 2003. Publicado en *Proceedings del Congreso Argentino de Ciencias de la Computación 2003 – CACIC 03 (12 páginas)*.

[11] Urciuolo Adriana, Iturraspe Rodolfo. 2002. "Conceptual Patterns for Water Resources Information Systems". *Journal of Computer Science and Technology* Vol. 3 - No. 1 - April 2003 pp 20-26.

[12] Urciuolo A., R. Iturraspe, A. Parsón, "Conceptual Microarchitectures for Hydrologic Simulation Models", En: *CLEI Electronic Journal*, ISSN 0717-5000, Cecilia Bastarrica ed. <http://www.clei.cl>. Vol 7, 1, Jun 2004, paper 6, 18 pp.

[13] Urciuolo A., R. Iturraspe, "Conceptual Patterns for Water Resources Information Systems", En: *Journal of Computer Science and Technology* Vol. 3 - No. 1 - April 2003 - ISSN: 1666-6038, pp 20-26.

[14] Urciuolo Adriana, Iturraspe Rodolfo, Moyano Ezequiel. *Perfil UML 2.0 para Aplicaciones de Monitoreo Ambiental*. Actas de las VI JIISIC'07, 31 de Enero al 2 de Febrero del 2007, Lima, Perú. Facultad de Ciencias e Ingeniería and Departamento de Ingeniería, Pontificia Universidad Católica del Perú 2007, ISBN 978-9972-2885-1-7. pp. 393-401.