

Ontologías de Visualización

Sebastián Escarza Silvia Castro Sergio Martig

Laboratorio de Investigación en Visualización y Computación Gráfica
Departamento de Ciencias e Ingeniería de la Computación - Universidad Nacional del Sur
Av. Alem 1253 – 8000 – Bahía Blanca – Buenos Aires – Argentina
Tel: 4595101 - Interno: 2618
{se, smc, srm}@cs.uns.edu.ar

Resumen

La Visualización de Información ha alcanzado un estado de desarrollo en el cuál se han identificado los componentes y procesos esenciales y comunes a toda Visualización. En este contexto se han planteado muchos modelos de referencia para contribuir al desarrollo más sistemático de las herramientas del área. La incorporación de una semántica bien definida en el proceso de Visualización, permitiría formalizarlo estableciendo un vocabulario común en el cuál los usuarios planteen sus requerimientos y los diseñadores de herramientas de visualización puedan expresar las transformaciones de los datos desde su obtención hasta la obtención la vista, las operaciones que deben proveerse y las formas de interacción posibles entre los usuarios y la visualización. La incorporación de información acerca del propio proceso de Visualización en las herramientas permitiría, además, lograr una asistencia al usuario más efectiva durante la construcción de la representación visual. Para dotar de semántica al proceso de visualización se hace necesario incorporar ontologías. En este trabajo se propone la construcción de una ontología de visualización que contribuya a definir formalmente los conceptos que conforman la teoría de base, que provea un vocabulario común y bien definido para el desarrollo de visualizaciones y que haga posible al software utilizar sus conceptos y relaciones para asistir al usuario durante el proceso de construcción de la representación visual.

Palabras Clave: Visualización – Visualización de Información – Ontología

Introducción

La Visualización de Información ha venido sufriendo un gran desarrollo en los últimos años motivada principalmente por la necesidad creciente de contar con técnicas que permitan abordar la problemática relativa al análisis y manipulación efectiva de volúmenes de información cada vez mayores. Detrás del objetivo de contar con metodologías más sistemáticas para el desarrollo de visualizaciones, mucho camino se ha recorrido desde los primeros intentos aislados de herramientas. A partir de eso, ha sido posible comenzar a identificar tareas propias del proceso de visualización que son comunes a todas las herramientas del área. Estos esfuerzos son el paso previo para el desarrollo de marcos de trabajo comunes que permitan caracterizar las transformaciones que los datos deben sufrir desde su obtención hasta la generación de la vista. Los modelos proporcionados por estos marcos de trabajo, brindarán un vocabulario común en el que los usuarios de los diversos dominios de aplicación expresen sus necesidades y en el contexto del cual los diseñadores de herramientas de visualización puedan expresar las transformaciones del proceso.

Paralelamente a los avances en materia de Visualización, y motivado en gran medida por el proyecto conocido como la Web Semántica [W3C], han surgido un conjunto de tecnologías destinadas a dotar de semántica a la información. Si bien, toda la información posee semántica asociada, se ha avanzado mucho en el desarrollo de tecnologías que permitan expresar explícitamente la misma con el objetivo último de lograr que la información sea útil en igual medida para humanos y agentes inteligentes. Si se estructura la información de manera que pueda ser accedida de manera uniforme y se provee un marco para definir recursos y sus relaciones, sobre

ello es posible definir ontologías y se tiene la base necesaria para especificar formalmente el significado de la información.

Las ontologías constituyen bases de conocimiento estructuradas en las cuales se modelan conceptos, sus características (atributos) y relaciones. Son una especificación formal de contenido (semántica), un conjunto de conceptos que puede ser utilizado por agentes para dialogar en un lenguaje común ya que modelan un determinado dominio de manera rigurosa. De esta manera las ontologías cobran un papel importante a la hora de dotar a un sistema de meta-información. Se pueden utilizar para especificar información sobre la semántica de las entidades de un sistema de manera que el mismo sistema pueda utilizarla.

La introducción de ontologías en Visualización de Información puede abordarse en dos frentes. Por un lado, el uso de ontologías daría a los usuarios la posibilidad de proveer los datos de su dominio de aplicación con semántica asociada. Dicha semántica, además de ser potencialmente visualizable, podría ser utilizada para conducir el proceso de visualización hacia representaciones visuales en las cuales se explote el significado de los datos y se refleje de manera más adecuada el espacio de información a explorar.

Por otro lado, la introducción de ontologías en el modelo de visualización mismo, permitiría a un agente conducir el proceso de visualización al contar con una semántica clara acerca de las decisiones que puede tomar y se lograría de esta manera brindar asistencia al usuario durante la construcción de la representación visual. El contar con una semántica rigurosa para el modelo de datos de la visualización, las transformaciones, estados intermedios y las diferentes técnicas aplicables también permitiría definir el vocabulario común a usuarios y diseñadores antes mencionado, logrando el doble beneficio de, por un lado, establecer una vía de comunicación inteligible entre quienes plantean los requerimientos y quienes diseñan las visualizaciones y, por el otro, avanzar hacia la definición de metodologías de construcción de visualizaciones más sistemáticas.

Antecedentes

Si bien actualmente no se cuenta con resultados consensuados con respecto a una Ontología de Visualización, la comunidad científica en general reconoce la necesidad de contar ésta. Los primeros trabajos en la búsqueda de dicha ontología fueron realizados por Ken Brodli et. al. [VO04]. Estos trabajos constituyen el primer intento serio y pueden considerarse un importante punto de partida para nuestro desarrollo.

Ontologías y OWL

Tom Gruber define a las ontologías como “una especificación explícita de una conceptualización.”. Las ontologías constituyen bases de conocimiento estructuradas en las cuales se modelan conceptos, instancias, sus atributos y relaciones. Dichos conceptos, instancias, atributos y relaciones tienen contrapartes directas en el dominio que se intenta modelar, las cuales conforman la conceptualización antes mencionada. Las ontologías son una especificación formal de contenido (semántica), un conjunto de conceptos que puede ser utilizado por agentes para dialogar en un lenguaje común ya que modelan un dominio de manera rigurosa. La definición de la semántica de una ontología se sustenta sobre formalismos basados en la lógica quienes brindan los axiomas y las reglas de inferencia necesarios para derivar información a partir de la ontología.

La definición de una ontología puede realizarse de diversas maneras debido a la existencia de varios lenguajes de ontologías. Ante este abanico de posibilidades la adopción de estándares maduros bien establecidos y aceptados, resulta importante. En este sentido, la utilización de OWL como lenguaje para la especificación de la ontología resulta muy conveniente debido a la gran disponibilidad de herramientas y utilidades existentes y en desarrollo hoy en día.

OWL es el lenguaje de ontologías propuesto por el W3C [W3C] y constituye un pilar fundamental de la Web Semántica, un proyecto para proveer de semántica a la Web actual. OWL descansa sobre otras tecnologías como XML, XML Schema, RDF y RDF Schema, y juntos permiten dotar de estructura a los documentos y definir recursos interrelacionados con semántica explícita. Todas estas tecnologías constituyen estándares abiertos, ampliamente probados y aceptados.

Sin duda, el principal atractivo de la utilización de una ontología es la posibilidad de razonar y derivar conclusiones a partir de ella. Para ello, es necesario contar con un mecanismo de razonamiento capaz de procesar la ontología, validarla (chequear potenciales inconsistencias) y responder consultas en base a la semántica de la misma. En este sentido se definen tres sublenguajes de OWL: OWL Lite, OWL DL y OWL Full de acuerdo a la complejidad computacional que tiene la construcción de un razonador completo para ontologías. Debido a que brinda la mejor combinación entre poder expresivo y complejidad de razonamiento, OWL DL es el lenguaje a adoptar en este trabajo.

Contexto de la Investigación

El trabajo se llevará a cabo en el contexto de los trabajos realizados en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab) del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur. Actualmente, el laboratorio lleva a cabo investigaciones en diferentes técnicas de Visualización de Información entre las que podemos citar Visualización de Grafos, Scatterplots y Coordenadas paralelas; aplicadas a diversos dominios como por ejemplo Visualización de Argumentaciones Rebatibles, Visualización de Bases de Datos y Sistemas de Información Geográfica, entre otros. Este contexto de desarrollo provee dominios de aplicación variados en el cual validar el desarrollo de una Ontología de Visualización que sea lo más general posible.

Objetivo de la Investigación

El objetivo de este trabajo consiste en la definición de una ontología inicial de visualización. Dicha definición se basará en el Modelo Unificado de Visualización (MUV) [MCFE03], un modelo conceptual que provee el marco necesario para el diseño y la construcción de visualizaciones.

En una primera etapa, se planea realizar la definición de los conceptos de más alto nivel. Los estados intermedios y transformaciones del modelo, así como los operandos y operadores involucrados en el proceso de visualización, deberán mapearse a conceptos y relaciones en la ontología, constituyendo un modelo de datos en término del cual será posible implementar visualizaciones y agentes que utilicen dicha información de manera provechosa. También en relación con el MUV, deberá plantearse como parte de la ontología, un conjunto de conceptos y relaciones destinado a modelar las interacciones que la visualización podría proveerle al usuario.

En una segunda instancia, se incrementará el nivel de detalle de la ontología, incluyendo los conceptos y relaciones propias de cada etapa. Se definirá un modelo de datos ontológico para las etapas iniciales, se especificarán las diferentes formas de mapear visualmente los diversos tipos de datos del usuario y se modelarán las principales técnicas de visualización desde una óptica lo más general y uniforme posible pero sin perder las características distintivas de cada una de ellas.

Para la definición de la ontología es importante seguir un enfoque iterativo en el cual a medida que se avance en el desarrollo se refinen los conceptos antes planteados. También es fundamental validar los conceptos en la ontología a medida que se introducen con el propósito de no arribar a ontologías inconsistentes impidan la posibilidad de procesarlas automáticamente y razonar sobre ellas.

Durante el proceso de desarrollo de la ontología es fundamental analizar otros trabajos relacionados a fin de evaluarlos y compararlos con el nuestro. Resulta importante fortalecer las similitudes y analizar las diferencias de manera de lograr consenso en la comunidad científica y orientar los esfuerzos en una dirección común. En este sentido el análisis de trabajos similares, de diferentes modelos de datos y su utilización y aceptación generalizada son factores clave a tener en cuenta.

Adicionalmente, se contempla la implementación de prototipos de herramientas de visualización con el objeto de evaluar de manera pragmática las definiciones planteadas de manera de realizar los ajustes necesarios en etapas tempranas y en concordancia con el enfoque iterativo mencionado. Para el desarrollo del software se tendrá especial cuidado en priorizar enfoques modulares que faciliten la incorporación e interoperabilidad de componentes de software. Esto tiene vital importancia en un entorno de experimentación y evaluación como el que se pretende lograr.

Agradecimientos

Este desarrollo es parcialmente financiado por los PGI 24/N015 y 24/Zn12, Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur, Bahía Blanca, Argentina y el PICT 2003 N° 15043, ANPCyT, SeCyT, Argentina.

Referencias

- [DB05] D. Duke, K. Brodlie, *Visualization Viewpoints*, IEEE Computer Graphics and Visualization journal, May/June 2005, Volume 25, Number 3, pp. 6-9.
- [DBD] D. Duke, K. Brodlie, D. Duce, *Building an Ontology of Visualization*. <http://ieeexplore.ieee.org/iel5/9449/29999/01372256.pdf>.
- [GC02] V. Geroimenko, C. Chen (Eds). *Visualizing the Semantic Web – XML-based Internet and Information Visualization*, Springer Verlag, 2003, Second Print.
- [MCFE03] S. Martig, S. Castro, P. Fillotrani, E. Estévez, *Un Modelo Unificado de Visualización*, Proceedings, pp. 881-892, 9° Congreso Argentino de Ciencias de la Computación. 6 al 10 de Octubre de 2003. La Plata. Argentina.
- [VO04] Reporte del *Visualisation Ontology Workshop* realizado en Nacional e-Science Centre el 7 y 8 de Abril de 2004. <http://www.nesc.ac.uk/esi/events/393/>
- [W3C] Documentos y recomendaciones del sitio <http://www.w3c.org> (World Wide Web Consortium).