

Visualización Basada en Semántica

Martín Larrea^{1,2,3}, Silvia Castro^{1,2,3}, Dana Urribarri^{1,2,3,4}, Sebastián Escarza^{1,2}

¹Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab)

²Departamento de Ciencias e Ingeniería en Ciencias de la Computación (DCIC)

³Instituto de Ciencias e Ingeniería de la Computación (ICIC),

⁴Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

Universidad Nacional del Sur, Av. Alem 1253, Bahía Blanca

{mll, smc, dku, se}@cs.uns.edu.ar

RESUMEN

Esta línea de investigación tiene como objetivo crear un modelo de visualización que considere la semántica de los datos, del contexto y de las etapas del “Modelo Unificado de Visualización” (MUV) para poder asistir al usuario en el seteo de los parámetros de la visualización logrando así una mejor representación visual. Este modelo deberá formalizar el modelo de referencia mediante una Ontología de Visualización, incorporar una Ontología de Datos e incorporar una capa de inferencia que le permita derivar nueva información, a partir de la ya conocida. Esta línea también busca introducir el uso de semántica y razonamiento semántico en diferentes aspectos del proceso de visualización.

Palabras clave: *Ontología, Representación Formal, Visualización basada en semántica, Visualización.*

CONTEXTO

Este trabajo se lleva a cabo 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. Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

Actualmente existe un gran número de modelos de referencia en Visualización ([1,2,3,4,5]) que, de diversas maneras, han identificado y delineado los principales componentes y procesos que sufren los datos para ser visualizados. En particular, en el seno del VyGLab hemos estado desarrollando el Modelo Unificado de Visualización (MUV) ([6]), un modelo que constituye un marco conceptual de referencia en términos del cual ubicar los procesos y estados intermedios de los datos y definir las interacciones explícitamente. Si bien estos esfuerzos por compilar, caracterizar y clasificar los aspectos más relevantes del área han sido y siguen siendo fructíferos, aún no existe un consenso definitivo que permita consolidar una teoría de base en Visualización. La incorporación de una semántica bien definida en el proceso de Visualización ya ha sido identificada como una necesidad ([7,8]) en el sentido que permitiría formalizar dicho proceso, estableciendo un vocabulario común que permita a los usuarios plantear sus requerimientos, y a los diseñadores de herramientas de visualización, expresar las transformaciones de los datos desde su obtención hasta la construcción de la vista, las operaciones que deben proveerse y las formas de interacción posibles entre los usuarios y la visualización. Para lograr una formalización consensuada del proceso de Visualización, es

preciso describir axiomáticamente el marco de referencia provisto por los modelos mencionados anteriormente. En este contexto, las ontologías surgen como una herramienta natural para estos fines. Un ejemplo de los beneficios de la formalización del proceso de visualización se encuentra en la formalización de los tipos de datos, es decir una Ontología de Datos. Contar con una ontología que clasifique los conjuntos de datos a visualizar es una guía que asiste a la hora de elegir la técnica de visualización apropiada para determinado conjunto de datos. En la literatura es posible encontrar taxonomías de datos, una versión menos formal de la ontología, pero no son presentadas desde un punto de vista de visualización. Por otro lado, el desafío de una visualización es encontrar una metáfora visual que permita entender y percibir en forma efectiva un conjunto de datos. Para este propósito es útil contar con cuantificaciones de diferentes aspectos de cada técnica de visualización. Una visualización debe proveer también un conjunto de interacciones a partir de las cuales el usuario explorará el conjunto de datos con una mínima carga cognitiva. La tecnología computacional actual permite la exploración de grandes conjuntos de información. Por un lado, esta situación es extremadamente útil pero la creciente cantidad de información genera una sobrecarga cognitiva. Mientras que el poder computacional ha crecido en forma exponencial, la habilidad para interactuar con dichos datos solo se ha incrementado en forma lineal. Hoy en día, una gran variedad de usuarios acceden, extraen y muestran información que está distribuida sobre diferentes puntos, con diferentes tipos, formas y contenidos. En muchos casos, el usuario debe tener un control activo sobre el proceso

de visualización pero, aún en este caso, es difícil obtener una visualización efectiva. Es común que la información que se desea representar no tenga una manifestación visual obvia, ante esta situación el proceso de mapeo del conjunto de datos a la vista puede llegar a ser no trivial ([9,10]).

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Este trabajo presenta dos ejes de investigación las cuales se entrelazan y tienen varios puntos en común. Tales ejes son:

1. “Formalización del proceso de Visualización mediante el uso de Ontologías”.
2. “Visualización Basada en Semántica”.

Formalización del proceso de Visualización mediante el uso de Ontologías. Este eje es una continuación directa de la línea de investigación presentada en ediciones anteriores de este Workshop ([11]), en la cual se plantea la necesidad de definir una Ontología de Visualización y de brindar soporte a dicha especificación formal a través de una plataforma de software que facilite la integración de los diversos componentes. Este trabajo integra el uso de ontologías en Visualización. Las ontologías, si bien se originaron hace tiempo en el área de Representación del Conocimiento, han cobrado nueva relevancia con el surgimiento de proyectos como la Web Semántica. A raíz de ello, numerosos estándares han sido definidos. Especificaciones como RDF, RDF Schema y OWL (el Lenguaje de Ontologías de la Web) permiten definir vocabularios estructurados que modelan jerarquías de herencia entre conceptos, relaciones, restricciones y reglas, que permiten un tratamiento más riguroso del conocimiento al

poseer una semántica formal asociada. En cuanto a la integración de semántica en visualización, aún no se cuenta con abordajes sistemáticos al problema y se identifican dos enfoques principales. Por un lado, se utilizan representaciones formales para mejorar la integración, consulta y descripción de los datos del usuario como una forma de enriquecer la visualización, y por el otro, trabajos más relacionados con el que proponemos utilizan descripciones semánticas, aunque en principio algo limitadas, del proceso de visualización en sí mismo para ayudar al usuario en la definición y configuración del mismo y en la selección de una técnica de visualización apropiada.

Visualización Basada en Semántica. Debido a que el objetivo de una visualización es lograr una representación que ayude al usuario a interpretar un conjunto de datos y comunicar su significado, es importante controlar el mapeo de las dimensiones físicas a las perceptuales; un usuario inexperto podría utilizar un mapeo incorrecto afectando negativamente la visualización resultante. Una estrategia para mejorar esta situación es guiar al usuario en la selección de los diferentes parámetros involucrados en la visualización a través de información semántica.

3. RESULTADOS OBTENIDOS/ESPERADOS

A nivel de la primera línea de investigación presentada, la transformación de los objetivos en resultados sigue las etapas delineadas en [11]. Al momento de esta publicación se cuenta con una definición parcial de la Ontología de Visualización. Concretamente, la arquitectura general de la ontología ya ha sido definida así como los principales elementos de la misma. Adicionalmente,

también se ha definido un conjunto de conceptos destinados a caracterizar el espacio de datos del usuario y la representación visual, y se continúa trabajando en la definición de los conceptos ontológicos necesarios para describir el proceso de visualización como una red de etapas interconectadas, así como en las definiciones necesarias para especificar las interacciones entre el usuario y la visualización. Sobre la segunda línea de investigación, se ha logrado incluir la semántica de los datos en el mapeo de datos a colores en una representación visual ([12,13,14]). Gracias a este trabajo, esta etapa del proceso de visualización ya no requiere de la participación del usuario; las decisiones de qué color usar para cada dato se determinan a partir de un razonador semántico. El próximo objetivo es integrar métricas al proceso de decisión semántica.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo relacionadas con las líneas de investigación presentadas:

- Tesis Doctoral. Sebastián Escarza. Tema: Ontologías de Visualización. Dirección: Dra. Silvia Castro.
- Tesis Magister. Georgina Inés Cerúsico. Tema: Taxonomías de Técnicas de Visualización para la Formalización del Proceso de Visualización. Dirección: Dra. Silvia Castro y Dr. Martín Larrea.
- Tesis Magister. Alejandra Elizabeth Herrera. Tema: Coordenadas Paralelas. Visualización e Interacciones. Dirección: Dra. Silvia Castro.
- Tesis Magister. César Escobal Blanco. Tema: La Reducción de Datos en las

Técnicas de Visualización. Dirección:
Dra. Silvia Castro.

5. BIBLIOGRAFÍA

- [1] E. H. Chi. A taxonomy of visualization techniques using the data state reference model. In Proceedings of the IEEE Symposium on Information Visualization (InfoVis'00), page 6975. IEEE Computer Society Press, 2000.
- [2] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. IEEE Symposium on Visual Languages, 0:336, 1996.
- [3] K. Brodlić and N. M. Noor. Visualization Notations, Models and Taxonomies. pages 207–212, Bangor, United Kingdom, 2007. Eurographics Association.
- [4] Brunetti, Josep Maria, et al. "Formal linked data visualization model." Proceedings of International Conference on Information Integration and Web-based Applications & Services. ACM, 2013.
- [5] Ma, Kwan-Liu. A Pervasive Parallel Processing Framework for Data Visualization and Analysis at Extreme Scale. No. DOE--UCD5373. Univ. of California, Davis, CA (United States), 2017.
- [6] Martig S., Castro S., Fillotrani P., Estévez E., Un Modelo Unificado de Visualización, Proceedings, pp. 881-892, 9º Congreso Argentino de Ciencias de la Computación. 6 al 10 de Octubre de 2003. La Plata. Argentina.
- [7] [DBD*05] D. J. Duke, K. W. Brodlić, D. A. Duce, and I. Herman. Do you see what I mean? IEEE. Computer Graphics and Applications, 25(3):6–9, 2005
- [8] M. Chen, D. Ebert, H. Hagen, R. S. Laramee, R. van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. IEEE Comput. Graph. Appl., 29(1):12–19, 2009.
- [9] Carpendale, M. S. T. 2001. Considering Visual Variables as a Basis for Information Visualization. Technical Report. University of Calgary, Department of Computer Science
- [10] Gu, Yi, et al. "Visualization and recommendation of large image collections toward effective sensemaking." Information Visualization 16.1 (2017): 21-47.
- [11] S. Escarza, S. Castro, S. Martig. Formalización del Proceso de Visualización mediante el uso de Ontologías. XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010) pp. 261-265. Universidad Nacional de la Patagonia San Juan Bosco. 5 y 6 de mayo de 2010. El Calafate. Santa Cruz. Argentina. ISBN: 978-950-34-0652-6
- [12] M. Larrea, S. Martig and S. Castro. Semantics-based Color Assignment in Visualization. Journal of Computer Science & Technology. Vol. 10 - No. 1 – April 2010 - ISSN 1666-6038.
- [13] M. Larrea, S. Martig and S. Castro. Formalización del Proceso de Visualización Basada en Semántica. XII Workshop de Investigadores en Ciencias de la Computación (WICC 2010) pp. 270-274. Universidad Nacional de la Patagonia San Juan Bosco. 5 y 6 de mayo de 2010. El Calafate. Santa Cruz. Argentina. ISBN: 978-950-34-0652-6.
- [14] Escarza, S., Larrea, M., Urribarri, D., Castro, S., Martig, S. "Integrating Semantics into the Visualization Process". Dagstuhl Follow-Up Seminar Book "Scientific Visualization: Interactions, Features Metaphors" (2011), pp. 92-102. ISBN: 978-3-939897-26-2.