

MILES (Modelos de Interacción centrados en Lenguaje, Espacio y Semántica computacional)*

MILES (Models of Interaction centred on Language, spacE and computational Semantics)

Pablo Gervás
U. Complutense de Madrid
Madrid
pgervas@sip.ucm.es

Angélica de Antonio
U. Politécnica de Madrid
Madrid
angelica@fi.upm.es

Gabriel Amores
U. de Sevilla
Sevilla
jgabriel@us.es

Resumen: La finalidad principal del proyecto es desarrollar una arquitectura para sistemas de interacción que conjugue un motor de diálogo, un generador de lenguaje natural, y una representación semántica basada en ontologías que abarque tanto el espacio (real o virtual) como el usuario que en él se ubica.

Palabras clave: Generación de lenguaje natural, ontologías, sistemas de diálogo

Abstract: The main goal of this project is to develop an architecture for interactive systems that combines a dialogue engine, a natural language generator, and a semantic representation based on ontologies covering both the real (or virtual) physical space and the user located within it.

Keywords: Natural language generation, ontologies, dialogue systems

1. *Introducción*

En los últimos años ha habido grandes avances en las áreas de investigación de sistemas de diálogo, generación de lenguaje natural, y representación de conocimiento mediante ontologías. No obstante, a pesar de que estas tres áreas tienen en común el tratar interacciones entre personas y máquinas con componentes importantes de lenguaje natural, ha habido muy poco trabajo centrado en la conjunción de las tres.

En este proyecto coordinado, en el que participan grupos de investigación de las universidades Complutense de Madrid (UCM), Politécnica de Madrid (UPM) y Universidad de Sevilla (USE), se aplican soluciones tecnológicas de interacción para resolver dificultades de percepción o localización que un usuario real pueda tener con un entorno espacial concreto. Para los propósitos del proyecto, se tienen en cuenta usuarios con distinto tipo de diversidad funcional (distintos grados de facilidad en la percepción o en la movilidad). Se están desarrollando tecnologías para trabajar con entornos reales y con simulaciones virtuales de los mismos. El sistema resultante ofrecerá una funcionalidad de diálogo mediante la cual el usuario puede interactuar con el sistema, que comprende sus peticio-

nes y responderá coherentemente. Este tipo de sistema plantea la necesidad de resolver cuatro desafíos clave para hacer posibles las distintas funcionalidades contempladas. Primero, el sistema debe mantener un modelo del entorno. Segundo, debe mantener un modelo del propio usuario, sus capacidades y sus posibilidades de percepción. Tercero, debe ser capaz de generar mensajes en lenguaje natural para proporcionar al usuario descripciones e instrucciones acerca del entorno. Finalmente, debe ser capaz de mantener un diálogo multimodal con el usuario, conjugando todos estos ingredientes. Estos cuatro desafíos constituyen las claves de la investigación desarrollada.

El presente proyecto se basa en las siguientes hipótesis de partida. Para un sistema de diálogo es relevante el ser capaz de generar dinámicamente respuestas complejas en lenguaje natural integradas dentro de una arquitectura de presentación multimodal. Para un entorno virtual interactivo sería interesante establecer un diálogo fluido con el usuario. Finalmente, un sistema de generación de lenguaje natural debe ser capaz de enfrentarse a los retos específicos que puedan surgir de trabajar en contexto de diálogo y/o de aprovechar la riqueza de información disponible en entornos virtuales. Otra hipótesis importante para el proyecto es que las ontologías pueden constituir un nexo importante sobre

* Este proyecto de investigación ha sido financiado por el Ministerio de Innovación y Ciencia (TIN2009-14659-C03).

el que coordinar la interacción entre un sistema de diálogo, un generador de lenguaje natural, y un entorno virtual. La representación de conocimiento mediante ontologías ha sido objeto de investigación en tiempos recientes en distintos dominios de aplicación, pero su utilización en las áreas mencionadas se ha restringido hasta la fecha a prototipos exploratorios en cada área que validan la utilidad de la tecnología desde un punto de vista básico.

2. Trabajo previo

Como resultado de la investigación del grupo de investigación NIL (Natural Interaction based on Language) de la UCM en el contexto del proyecto GALANTE (TIN2006-14433-C02-01), financiado por el Plan Nacional de investigación, se desarrolló el framework TAP, diseñado para facilitar el desarrollo de aplicaciones de generación automática de lenguaje natural de forma que se pueda reutilizar el trabajo previamente desarrollado. Para ello se combinó trabajo previo sobre definición de arquitecturas genéricas para el desarrollo de aplicaciones de generación de lenguaje natural (Reiter y Dale, 2000; Cahill et al., 2001), con ideas de patrones de diseño (Gamma et al., 1995) y frameworks (Johnson y Foote, 1988). El framework TAP está capacitado para trabajar con representación de conocimiento en forma de ontologías. El grupo también ha participado con éxito en tareas de evaluación competitiva sobre generación de lenguaje natural relacionadas con la identificación de elementos observables a partir de descripciones textuales y con la generación de instrucciones para guiar a un usuario en un entorno virtual.

El grupo de investigación JULIETTA de la USE ha participado en numerosos proyectos de investigación y dispone de tecnología propia de sistemas de diálogo multimodal y procesamiento de lenguaje natural, desarrollada a lo largo de una trayectoria de más de 15 años de investigación. A lo largo de este tiempo ha desarrollado una serie de soluciones específicas para sistemas de diálogo que constituyen un punto de partida para la investigación que se está desarrollando en el proyecto MILES. Entre ellas se encuentran Episteme, un sistema de traducción automática que incluye en módulo completo de análisis léxico y gramatical junto con uno de generación basado en reglas, Delfos,

un sistema de diálogo unimodal inspirado en el ISU approach que desarrolla un gestor de diálogo basado en expectativas, y MIMUS (de Amores Carredano et al., 2010), una evolución sobre Delfos, dotándolo de mecanismos específicos de multimodalidad, como estrategias de fusión y presentación multimodal. En particular, MIMUS pivota su funcionamiento sobre fuentes de conocimiento externo, específicamente sobre ontologías tipo OWL. MIMUS fue integrado con el generador TAP desarrollado por el grupo de la UCM en el contexto del proyecto coordinado DIVA-GALAN (TIN2006-14433), dentro del subproyecto GILDA.

El grupo de investigación de la UPM disponía por su parte de varios resultados de interés para este proyecto, obtenidos durante los últimos 10 años como consecuencia de sus trabajos de investigación en otros proyectos como MAEVIF (TIC00-1346) o ENVIRA (TIN2006-15202-C03), y de varias tesis doctorales desarrolladas en el ámbito del grupo:

- Un modelo preliminar de ontología semántica para entornos virtuales, especialmente centrada en la representación e inferencia de información espacial acerca del entorno 3D y los objetos en él situados, capaz de inferir relaciones espaciales y estructurales de diversos tipos entre objetos
- Un algoritmo de generación de trayectorias para la navegación a través de un entorno tridimensional constituido por diversos sub-entornos, con evitación de obstáculos, con la posibilidad de considerar áreas de tránsito de mayor o menor coste asociado, y con algunos mecanismos para hacer la trayectoria lo más similar posible a la forma de navegación humana en entornos reales.
- Un modelo de percepción visual computerizado que puede ser utilizado por un agente virtual inteligente para extraer información visual de su entorno, así como para modelar las capacidades perceptivas de los usuarios y anticipar sus posibilidades de percepción
- Una ontología para el modelado de estudiantes y un método de diagnóstico cognitivo especialmente adaptado para su aplicación en entornos virtuales

Sobre estas bases previas se consideró que

se daban las circunstancias necesarias para que los tres grupos acometieran de forma coordinada la tarea de demostrar empíricamente la validez de la hipótesis de partida.

3. Objetivo y plan de trabajo

El proyecto se inició en enero de 2010, y finalizará en diciembre de 2012. El proyecto coordinado ha contemplado tanto el desarrollo de una arquitectura de sistema de gestión de interacciones como su validación en una serie de casos de estudio particulares.

El subproyecto NOVA (Navegación basada en Ontologías mediante la Verbalización de mensajes de Ayuda) se ha centrado en la tarea de guiar a un usuario por un espacio físico que desconoce mediante instrucciones y descripciones verbales. Se ha prestado especial atención a usuarios con diversidad funcional (percepción o movilidad restringidas).

El subproyecto SEPIA (Modelos Semánticos del Entorno y la Persona para una Interacción Adaptativa en entornos virtuales) ha prestado especial atención al uso de ontologías para representación y razonamiento sobre conceptos espaciales, así como para el modelado del usuario, sus conocimientos previos relevantes, sus capacidades perceptivas, y los conocimientos que éste va adquiriendo a través de la interacción con el sistema. Se ha investigado acerca de las posibles interacciones entre dichas ontologías y el proceso de cálculo de trayectorias, su interpretación semántica para facilitar la generación de instrucciones, y el seguimiento y supervisión del recorrido del usuario a través del entorno.

El subproyecto DIMMO (Dialogos Multi-Modales basados en Ontologías) se ha centrado en desarrollar soluciones avanzadas de gestión de diálogo que tengan en cuenta la selección de movimientos del diálogo y la presentación multimodal guiada por ontologías.

El grupo Julietta de Sevilla ha aportado un motor de diálogo avanzado, validado ya en aplicaciones prácticas reales (Amores, Manchón, y Pérez, 2010). El grupo de la UCM ha aportado el motor de generación TAP, validado en diversas tareas de evaluación competitiva (Gervás, Hervás, y León, 2008; Hervás y Gervás, 2009; Gervás, 2011) y cuya integración con el motor de diálogo del grupo Julietta ya fue el objetivo del proyecto DIVAGALAN (TIN2006-14433). De cara a la coordinación entre estos dos grupos, el objetivo parcial en este proyecto ha sido comprobar

que la integración realizada puede trasladarse a otros entornos de aplicación. El grupo de investigación de la UPM ha aportado su experiencia en el desarrollo de ontologías como modelos semánticos para guiar la interacción en entornos virtuales de simulación y aprendizaje. El motor de generación de la UCM ya se había integrado con las ontologías espaciales del grupo de la UPM en el marco del proyecto IVERNAO (CCG08-UCM/TIC-4300), financiado por la Comunidad de Madrid a lo largo del año 2009, que exploraba el papel de las ontologías de representación del espacio y el usuario a la hora de guiar a un usuario a través de un espacio físico que desconoce. La integración entre los módulos de UPM y USE se ha desarrollado mediante un interfaz basado en servicios web que hace accesible el diálogo verbal sobre el entorno virtual.

La arquitectura del sistema en desarrollo se basa en la propuesta realizada en la tesis doctoral (Méndez, 2008) de uno de los miembros del equipo UCM, Gonzalo Méndez, dirigida por la Investigadora principal del equipo UPM, Angélica de Antonio. En esta tesis se proponía una arquitectura que posibilitase la integración de distintos módulos de manera muy flexible, manteniendo la independencia entre ellos, lo que ha permitido que se puedan enchufar y desenchufar las aportaciones de cada grupo de manera transparente sin afectar a los módulos de los restantes grupos.

4. Conclusiones

El modelado del contexto de discurso interactivo que aporta un sistema de diálogo y la representación de relaciones espaciales complejas permitirán mejoras significativas en la naturalidad de las respuestas del generador de lenguaje natural que se contempla. La sustitución de los actuales textos enlatados utilizados en los entornos virtuales por texto generados dinámicamente permitirá ampliar la cobertura a mayor número de interacciones distintas con los usuarios. La posibilidad de generar respuestas con estructura narrativa, cuya información está extraída de una ontología que representa el entorno y su combinación con una presentación multimodal utilizando modelos estructurados como ontologías (de contexto, de usuario, etc.) constituirá una mejora significativa de la interacción para el usuario de un sistema de diálogo.

Es de esperar que el proyecto proporcione

resultados tanto en el campo de la generación de lenguaje natural, como en el campo de la representación de entornos virtuales o reales, como en el campo de los sistemas de diálogo. Dentro de la generación de lenguaje natural la exploración sistemática de los mecanismos de generación de indicaciones orientadas a la navegación basados en concepciones semánticas expresadas en forma de ontologías será una aportación valiosa. Para los sistemas de realidad virtual o aumentada, la construcción de representaciones semánticas, y no sólo geométricas, de entornos de trabajo, así como del usuario, sus características y capacidades, y también del contexto o situación y su evolución dinámica permitirá construir sistemas con un mayor grado de inteligencia y capacidad de adaptación. Su integración con mecanismos de generación de lenguaje natural y gestión de diálogo permitirá plasmar las nuevas posibilidades de interacción adaptativa del sistema de una forma muy natural para el usuario. En el contexto del proyecto MILES este objetivo se concreta en dar apoyo al usuario para la navegación a través del entorno, pero estos mecanismos podrían extenderse a otras tareas como la búsqueda de objetos, o la manipulación de elementos del entorno.

Para los sistemas de diálogo se explorará una arquitectura de presentación multimodal que incluya un proceso de decisión que escoja la presentación óptima utilizando modelos expresados mediante ontologías. Además se analizarán las características de alto y bajo nivel que influyen en la elección de movimientos de diálogo y se desarrollará un algoritmo de selección que mejore las tasas de acierto a nivel de análisis dentro de los sistemas de diálogo.

Se espera obtener resultados con posibilidad de transferencia a medio plazo, ya que se está en proceso de integrar las diferentes técnicas y tecnologías desarrolladas en una arquitectura software modular que permita diversos tipos de aplicaciones.

Bibliografía

Amores, J. Gabriel, Pilar Manchón, y Guillermo Pérez. 2010. Humanizing conversational agents: Indisys practical case study in ehealth. En Diana Perez-Marin y Ismael Pascual-Nieto, editores, *Conversational Agents and Natural Language In-*

teraction: Techniques and Effective Practices.

- Cahill, L., R. Evans, C. Mellish, D. Paiva, M. Reape, y D. Scott. 2001. The RAGS reference manual. Informe Técnico ITRI-01-08, ITRI, University of Brighton.
- de Amores Carredano, José Gabriel, Guillermo Pérez García, Pilar Manchón Portillo, Carmen del Solar, Jesús González Martí, David Ávila Membrives, Antonio Ávila Membrives, y David Moral Alcázar. 2010. MIMUS: Gestión de diálogo multilingüe y multimodal. Informe técnico, Universidad de Sevilla.
- Gamma, Erich, Richard Helm, Ralph Johnson, y John Vlissides. 1995. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley.
- Gervás, Pablo. 2011. UCM Submission to the Surface Realization Challenge. En *Generation Challenges 2011 Session at 13th European Workshop on Natural Language Generation (ENLG 2011)*.
- Gervás, Pablo, Raquel Hervás, y Carlos León. 2008. NIL-UCM: Most-Frequent-Value-First Attribute Selection and Best-Scoring-Choice Realization. En *Referring Expression Generation Challenge 2008, Proc. of the 5th International Natural Language Generation Conference (INLG'08)*, Ohio, USA, 06/2008.
- Hervás, Raquel y Pablo Gervás. 2009. Evolutionary and Case-Based Approaches to REG: NIL-UCM-EvoTAP, NIL-UCM-ValuesCBR and NIL-UCM-EvoCBR. Athens, Greece, 04/2009.
- Johnson, R. E. y B. Foote. 1988. Designing reusable classes. *Journal of Object-Oriented Programming*, 1(2):22–35, June/July.
- Méndez, Gonzalo. 2008. *Una Arquitectura Software Basada en Agentes y Recomendaciones Metodológicas para el Desarrollo de Entornos Virtuales de Entrenamiento con Tutoría Inteligente*. Ph.D. thesis, Facultad de Informática, Universidad Politécnica de Madrid.
- Reiter, Ehud y Robert Dale. 2000. *Building Natural Language Generation Systems*. Cambridge University Press.