

Hacia la interacción en lenguaje natural

Towards Natural Language Interaction

Maria Fuentes y Meritxell González

{mfuentes,mgonzalez}@lsi.upc.edu

Centro de Investigación TALP

Universitat Politècnica de Catalunya

Barcelona, España

Resumen: En éste documento se presenta la investigación que está siendo llevada a cabo en el Grupo de Procesamiento de Lenguaje Natural (GPLN) de la Universidad Politécnica de Cataluña (UPC). En concreto, hemos articulado la presentación de las diferentes líneas de trabajo tomando como referencia su aplicación en un asistente virtual. Creemos que su uso y implantación irá en aumento en los próximos diez años, de ahí la importancia del estado de las tecnologías del lenguaje natural y, aún más, de los nuevos retos que este tipo de aplicaciones nos plantean.

Palabras clave: Lenguaje Natural, Comprensión, Interacción Hombre-Máquina, Asistente Virtual

1. *Introducción*

Después de décadas de investigación, la comprensión del lenguaje natural sigue siendo una de las muchas áreas de la Inteligencia Artificial (IA) que, siendo relativamente fácil para los humanos, resulta inalcanzable para los ordenadores. Para comprender expresiones en lenguaje natural se requiere conocimiento de muchos tipos diferentes y capacidad para razonar de forma eficiente. La gran cantidad de conocimiento heterogéneo involucrada nos lleva a una gran cantidad de posibles interpretaciones para cada expresión. De forma análoga, una sola idea puede expresarse de diversas formas.

Existen diferentes alternativas para abordar la complejidad de comprensión del lenguaje natural. En algunos casos se aplican modelos cognitivos sobre los procesos humanos de procesamiento del lenguaje. En otros casos se aplican aproximaciones mucho más superficiales, que atacan solamente problemas específicos para alcanzar un objetivo práctico muy concreto. En ambos casos se trata el lenguaje de forma acotada, ya sea porque solamente se tratan algunos tipos de expresiones (primer caso) o bien porque se dejan de lado los fenómenos que no están directamente relacionados con el problema que se quiere solucionar (segundo caso).

Los sistemas de diálogo hombre-máquina son un punto de encuentro para estas

dos aproximaciones. Por un lado, se requiere de un modelo que guíe la gestión del diálogo, integre la información obtenida en una base de conocimiento y haga las inferencias y acciones requeridas para llevar a buen término el diálogo. Este modelo suele estar diseñado e implementado por expertos en diálogo, basándose en teorías sobre el funcionamiento del diálogo entre humanos. Por otro lado, el tratamiento de las expresiones de los usuarios al interactuar con la máquina suele aplicar métodos robustos, que buscan identificar las piezas claves de información, incluso de forma superficial, dejando de lado información que no parece crucial para la tarea. Además, dado que muchas veces los sistemas de diálogo interactúan oralmente con los usuarios, se integran en esta aplicación las técnicas de tratamiento del habla y procesamiento del lenguaje natural (PLN), tradicionalmente separadas.

1.1. **Interés de los asistentes virtuales**

Los asistentes virtuales son un campo de aplicación privilegiado para los sistemas de interacción hombre-máquina, con cada vez más presencia entre las aplicaciones habituales. Las propiedades más relevantes de los asistentes se presentan a continuación.

En primer lugar, resultan un campo de prácticas asequible y adecuado para la inves-

tigación y el desarrollo en el área de la interacción hombre-máquina, e incluso pueden ser útiles para la investigación en teorías de comunicación entre personas.

En segundo lugar, facilitan el acceso a sistemas complejos. El objetivo principal de estos sistemas es ofrecer ayuda a los usuarios mientras realizan una tarea concreta; y el habla es la forma más natural de comunicarse, especialmente si el asistente se refuerza con la presencia de un avatar. Éste tipo de sistemas suelen estar dotados de mecanismos que adaptan la interacción con el usuario a sus características particulares. De esta forma facilitan el acceso no sólo a usuarios expertos, sino, más importante, a usuarios noveles o con dificultades de acceso, los cuáles han sido históricamente excluidos, contribuyendo así a reducir la brecha digital. De esta forma, los asistentes virtuales pueden constituirse en la interfaz natural para realizar consultas a entidades y servicios, para acceder y personalizar servicios web¹; o en la toma de decisiones²; para hacer reaccionar a una sala inteligente³; para interactuar con otros dispositivos⁴, como un GPS o el TDT; o para establecer una relación para obtener por ejemplo soporte emocional⁵.

En último lugar, los asistentes virtuales requieren de la mayor parte de tecnologías de PLN existentes, ya que requieren una gran dosis de comprensión del lenguaje. En este sentido, los asistentes virtuales nos parece una buena plataforma para diagnosticar la viabilidad o utilidad de las técnicas y métodos que se desarrollan de forma más teórica.

Por estas razones, vamos a tomar los asistentes virtuales como eje para concretar la investigación relacionada con el lenguaje natural, según la visión de nuestro grupo de investigación, para los próximos diez años.

2. *El grupo TALP en la UPC*

GPLN inicia su actividad en el departamento de Lenguajes y Sistemas Informáticos de la UPC en el año 1988. Desde sus orígenes, ha sido un grupo interdisciplinario, con una participación activa y continuada de lingüistas. En el año 1999, el GPLN se unió al Grupo de Procesado del Habla del Departamento de

Teoría de la Señal y Comunicaciones (TSC) de la UPC creando el Centro Específico de Investigación TALP. En el 2004, en colaboración con un tercer grupo de Procesado de Vídeo e Imagen del TSC, se crea en la UPC la primera sala inteligente. Éste entorno permite estudiar el habla oral en contextos más o menos naturales. Estudios recientes demuestran que en el proceso de comprensión se tiene en cuenta tanto información lingüística como no lingüística. Esto supone integrar una gran variedad de fuentes de conocimiento, incluyendo conocimiento del mundo o del contexto, conocimiento del hablante y/o el tema, frecuencia léxica, uso previo de una palabra o un tema semánticamente relacionado, expresiones faciales, posición del hablante/oyentes, prosodia y/o tono.

GPLN ha desarrollado una prolífica actividad en diversas áreas de investigación dentro del PLN y la IA. En procesamiento básico de la lengua (normalmente inglés, castellano, catalán y árabe) destacan los campos de desambiguación y anotación morfosintáctica y semántica (Carreras, 2005), la desambiguación semántica de sentidos (Escudero, 2006) y la aplicación de métodos estadísticos al procesamiento básico de la lengua (Padró, 2008), así como la aplicación de técnicas de aprendizaje automático a todo tipo de tareas y problemas de PLN. La solución no suele ser una aproximación estadística contraria a un enfoque lingüístico o viceversa. Los sistemas pueden utilizar una aproximación híbrida, combinando técnicas propias de ambos enfoques.

También se investiga en procesamiento de diccionarios, corpus textuales y recursos lingüísticos en general, con particular atención a los recursos léxicos, y la adquisición de conocimiento a partir de la comprensión semántica (Atserias, 2006) de documentos. Con ello se quiere adquirir (Farreras, 2005), enlazar (Daudé, 2005) y enriquecer ontologías y definir una anotación semántica del documento. Ésta anotación permitirá representar los documentos para más tarde aplicar razonamiento sobre ellos. Para llegar a obtener ésta representación se requiere investigar en entendimiento temporal, espacial, referencias, relaciones de causalidad (implicación textual), etc. Consideramos que una buena comprensión del texto es clave para la mejora de los procesos de análisis subyacentes a la gran mayoría de aplicaciones

¹proyecto HOPS, <http://www.bcn.es/hops>

²proyecto TRIPS, <http://trips.uic.org>

³proyecto CHIL, <http://chil.server.de>

⁴proyecto TALK, <http://www.talk-project.org>

⁵proyecto COMPANION, <http://www.companions-project.org>

de lenguaje natural, especialmente aquellas que requieren una interpretación precisa del significado del texto como: traducción automática (Giménez, 2008), extracción de información, respuesta a preguntas, resumen automático (Fuentes, 2008) y sistemas de diálogo (González, 2010), en las cuales el grupo también tiene líneas de investigación.

3. *Aplicación de las técnicas de PLN en los asistentes virtuales*

Dado un ámbito específico (turismo, compra/venta, blogs de opinión), mediante la aplicación de técnicas de minería de textos podemos generar bases de conocimiento de forma automática. La minería de textos es el proceso por el cual un sistema informático descubre nueva información a partir de diversos recursos escritos. El área de la minería de textos abarca extracción de información, sistemas fundamentales de pregunta-respuesta y clustering de documentos, entre otros.

A diferencia de la minería de datos, en la minería de textos encontramos datos no estructurados, por lo que es necesario un pre-proceso que nos ayude a identificar las entidades y sus relaciones. Uno de los desafíos que presenta la minería de textos es que los ejemplos no están clasificados ni etiquetados; y los recursos tampoco están conectados con otros tipos de conocimiento que faciliten la interpretación. Por otra parte, una de las mayores aportaciones de las técnicas de minería de textos es la generación de bases de conocimiento estructurado a partir del cual es posible inferir nuevo conocimiento. Estas bases de conocimiento pueden ser accedidas mediante interfaces de usuario más o menos complejas: sistemas de pregunta-respuesta, sistemas de diálogo o asistentes virtuales más completos, como los avatares.

4. *Caso práctico: Turismo interactivo*

Consideremos el caso de dos personas, María y Johan, que asisten a la SEPLN, que tiene lugar en Valencia. María realiza un doctorado en diacronía. Entre sus actividades se encuentra realizar una visita al Archivo del Reino de Valencia. Johan, por su parte, acompaña a María, es biólogo y quiere aprovechar para visitar la Albufera.

Para organizar una agenda a María se necesita, por una parte, obtener información sobre el programa de la SEPLN y los horarios

de las sesiones en las que María está interesada; y por otra parte se debe obtener información sobre el horario del Archivo. Además, para poder desplazarse desde el lugar de celebración de la SEPLN hasta el Archivo se necesita conocimiento geográfico así como de transporte público, horarios y tiempo de desplazamiento. La información se debería adquirir a partir de todos los recursos disponibles, mayoritariamente texto escrito y estructurado en bases de datos y/o servicios web. El conocimiento adquirido se estructura en forma de relaciones entre entidades y sus atributos. Por ejemplo, las relaciones entre estaciones de autobús y lugares específicos. Para realizar esta tarea se necesitan varias técnicas de minería de textos, desde las más simples hasta las más complejas, así como fuentes de conocimiento multilingüe y técnicas de traducción automática. Es decir, debemos tener en cuenta la interlingualidad de los recursos. A continuación, se procesan las fuentes usando técnicas de diferentes niveles de abstracción. Por ejemplo reconocimiento de nombres de entidades y expresiones temporales, o extracción y selección de patrones. Finalmente obtendremos el conocimiento sobre la tarea en forma estructurada, lista para ser usada por el asistente virtual.

Cuando María visita el Archivo en busca de información sobre documentos antiguos, el asistente accede a fuentes documentales que pueden estar escaneadas y transcritas automáticamente usando un OCR o manualmente en el idioma original o en otro. La mayoría de originales estarán escritos en latín, castellano o catalán antiguo. Para recuperar los diferentes documentos de interés para María se necesita extraer la información asociada a su perfil particular, teniendo en cuenta la multilingualidad de los recursos.

En el caso de Johan, el asistente le ayuda a llegar a la Albufera y accede a diferentes fuentes de información sobre éste lugar: desde una guía de ámbito general en inglés y la información disponible en las wikipedias, hasta guías más específicas, como la guía que proporciona la oficina de turismo de Valencia, escrita en castellano o valenciano. Asimismo, puede acceder a reportajes específicos publicados tanto en revistas científicas del área (p.e. Nature, National Geographic) como en videos documentales, tanto científicos (p.e. BBC) como de divulgación (p.e. Canal 9). Así, en este caso, la mayor parte de la

información que el asistente debe adquirir no se encuentra solo disponible en texto escrito o estructurado, sino también en prensa escrita y otros medios audiovisuales o no estructurados, como podcast, radio, televisión. Lo que conlleva recuperación, extracción y comprensión de contenidos multilingüe y multimedia.

Otra aplicación práctica de los asistentes virtuales es la búsqueda, clasificación y selección de opiniones. Esto implica abordar distintos subproblemas, p.e. reconocimiento y clasificación de entidades nominales, clustering (descubrir categorías y propiedades relevantes), así como la clasificación de weblogs en varios grados de subjetividad en diferentes dimensiones, p.e. influencia y/o sentimiento.

Por ejemplo, imaginemos que Johan y María buscan un restaurante donde ir a comer un arroz en Valencia. El restaurante debe cumplir una serie de requisitos, como que sirvan comida para vegetarianos y que se pueda llegar con transporte público. Sería deseable poder seguir recomendaciones de especialistas, como por ejemplo los de la guía gastronómica Michelin, o tener opiniones positivas de los usuarios en diferentes foros o blogs.

Para mejorar la calidad de los resultados el asistente necesita interactuar con el usuario para desambiguar información, seleccionar entre diferentes opciones o refinar una búsqueda. Cada interacción se divide en tres etapas: comprensión del usuario, control de la interacción y generación de una respuesta.

Comprender al usuario consiste en analizar diferentes niveles de su intervención, p.e. análisis sintáctico-semántico, análisis del discurso y reconocimiento automático del habla en el caso de comunicación oral.

La información obtenida del usuario se incorpora, durante la etapa de control de la interacción, al conocimiento específico sobre la misma. Durante esta etapa el asistente decide si necesita obtener más información por parte del usuario, p.e. cuál es su ubicación actual para buscar las estaciones de autobús más cercanas (esta información se podría inferir automáticamente a partir de geolocalizadores). Controlar la interacción es un proceso complejo que puede implicar varios turnos de clarificaciones por parte tanto del usuario como del sistema, hasta que éste último llega a comprender el objetivo del usuario y puede ofrecer una respuesta satisfactoria.

En la etapa de generación de respuesta de cada interacción el asistente establece el con-

tenido de la respuesta: resultados parciales en forma de resumen de documentos encontrados en el Archivo, información específica sobre alguno de los autores de los documentos, o recomendaciones adicionales, como visitar la Catedral de Santa María, donde se encuentra enterrado Ausiàs March, uno de los autores de los documentos obtenidos. Algunos de los métodos de PLN que intervienen durante esta etapa són la planificación del discurso, que da más naturalidad y coherencia a la interacción y la generación del lenguaje, donde se pueden aplicar técnicas de detección/selección del registro. La lengua generada puede ser distinta a la de la fuente de donde se extrae la información requerida. Por otro lado el asistente a parte de mostrar información textual podría sintetizar voz o extraer partes de documentos de video o audio.

5. Conclusiones

Éste artículo presenta a los asistentes virtuales como una aplicación que aborda distintas tareas de PLN, y a la vez plantea un nuevo horizonte de retos que afrontar en el futuro próximo, con el objetivo de conseguir una interacción natural con la tecnología que nos envuelve. El aumento de capacidad de los ordenadores no es suficiente, es necesario seguir trabajando en nuevos enfoques para superar las deficiencias todavía existentes.

Bibliografía

- Atserias, J. 2006. *Towards Robustness in Natural Language Understanding*. Ph.D. tesis, EHU.
- Carreras, X. 2005. *Learning and Inference in Phrase Recognition: A Filtering-Ranking Architecture Using Perceptron*. Ph.D. tesis, UPC.
- Daudé, J. 2005. *Enlace de Jerarquías Usando Etiquetado por Relajación*. Ph.D. tesis, UPC.
- Escudero, G. 2006. *Machine Learning Techniques for Word Sense Disambiguation*. Ph.D. tesis, UPC.
- Farreras, J. 2005. *Automatic Construction of Wide-Coverage Domain-Independent Lexico-Conceptual Ontologies*. Ph.D. tesis, UPC.
- Fuentes, M. 2008. *A Flexible Multitask Summarizer for Documents from Different Media, Domain, and Language*. Ph.D. tesis, UPC.
- Giménez, J. 2008. *Empirical Machine Translation and its Evaluation*. Ph.D. tesis, UPC.
- González, M. 2010. *DIGUI: A flexible dialogue system for accessing web services*. Ph.D. tesis.
- Padró, Muntsa. 2008. *Applying Causal State Splitting Reconstruction Algorithm to Natural Language Processing Tasks*. Ph.D. tesis, UPC.