

Plataforma para Creación de Agentes de Diálogo Inteligentes

Andrés T. Hohendahl^{1,2}
andres.hohendahl@fi.uba.ar

1 Laboratorio de Estereología y Mecánica Inteligente, Facultad de Ingeniería, UBA.

2 IIBM (Instituto de Ingeniería Bio Médica), Facultad de Ingeniería, UBA.

Personal-Blog: <http://web.fi.uba.ar/~ahohenda>

INTRODUCCIÓN

Sin duda el dialogar con una máquina, algo presente hace tiempo en el imaginario colectivo, enunciado por autores de ciencia a ficción, es el desafío inexorable de esta próxima década. Siguiendo este razonamiento sistemas inteligentes de la capacidad de la famosa computadora HAL9000 de “*2001 a Space Odissey*” del director *Stanley Kubrik* pronto serán realidad, la consigna es pavimentar el camino.

La necesidad de agentes de diálogo se puede ver en especial por el paradigma actual de acceso a la información, hoy mismo todo el acceso para búsqueda de información se realiza ya sea por menús estructurados en aplicaciones de escritorio y/o web, siendo preferente y muy frecuente la búsqueda por coincidencia de palabras en buscadores masivos como Google, por lote de palabras clave, el cual dista mucho del modo en el cual los humanos accedemos nuestro conocimiento al realizar consultas a humanos, nosotros accedemos preferentemente en forma verbal (hablada o escrita) con un fuerte contenido semántico y con una intención determinada, elaboración lógica y razonamiento sobre la mera consulta. Realizamos un intercambio previo con el interlocutor para asegurarnos de que hayamos entendido exactamente la pragmática de la consulta, y luego procedemos a realizar los pasos necesarios en algún sistema de información que conocemos a la perfección para brindarle las respuestas o resultados a quien necesitaba o consultó la información.

Esto pasa a diario y multiplicado por miles, en los llamados “*callcenters*” o centros de consulta telefónica, consumiendo infinidad de recursos humanos dilapidados en consultas en su mayoría muy simples, consultas que una máquina podría resolver sin inconvenientes, con un razonamiento ultra sencillo si puede manejar lenguaje natural.

ANTECEDENTES

Sistemas de diálogo han sido presagiados por todos los futuristas y escritores de ciencia ficción, aunque también han sido programados ciertos sistemas con esta intención, durante las fases muy iniciales de la computación. El primer sistema conocido de este tipo fue conocido como *Eliza* [1] luego hubo varios proyectos usando diálogo en lenguaje natural como el famoso *SHRDLU* [3] interactuando en el contexto de un ambiente acotado con objetos geométricos simples; en la universidad de Rochester con fondos militares y públicos se hizo un proyecto muy importante llamado *TRAINS* [2] que permiten interactuar mediante diálogo en un contexto acotado para reserva de trenes. Otros proyectos de este tipo son *TRIPS* [4] que agregan el concepto abstracto y operativo de planificación de acciones. Todos ellos pertenecen a laboratorios universitarios y son subsidiados para promover el estado del arte y experimentar, y difícilmente dejen el laboratorio pues no son concebidos para la industria práctica.

Si bien con tanto antecedente de trabajos, pareciera que fabricar un sistema de diálogo es algo trivial, en realidad no lo es en lo absoluto, puesto que la cantidad de tareas interdependientes a resolver es importante, además de la dificultad de realizar cada una de ellas. Otro problema radica en que cada bloque de proceso suele ser de diferente autor, fuente o teoría y la compatibilidad entre las mismas no ha sido nunca una meta en sus diseños individuales, por lo cual la interfase entre estos sistemas es un tema en sí mismo bastante espinoso y hasta sin solución pues a veces dos bloques refieren a teorías subyacentes que no están encontradas, esto pasa muy a menudo en lingüística.

DESCRIPCIÓN y PROPÓSITO

Se ha creado una plataforma de desarrollo orientada a poder generar un sistema cuya exteriorización es la de un agente inteligente, capaz de conversar con una persona en forma natural, mediante texto plano, en forma racional y lógica, respondiendo a preguntas, siguiendo hilos de conversaciones, presentando declaraciones, propuestas, sugerencias y mucho más; en definitiva entablando un diálogo, tal como lo haría un humano real. La finalidad, dominio, extensión y transcurso del diálogo es completamente dependiente de cada necesidad y dominio.

La plataforma permite crear servicios de conversación y diálogo en cualquier idioma alfabético aglutinante del tipo hindo-europeo del cual se posea un diccionario flexivo y un analizador morfológico, que entregue etiquetas compatibles con *EAGLES* [7], el sistema actualmente posee recursos para el castellano e inglés.

La idea es que se puedan hacer con esta plataforma, sistemas de diálogo, permitiendo un diálogo coloquial para su uso en aplicaciones como: soporte técnico, promociones, ventas, encuestas, asesoramiento, asistencia técnica, entretenimiento, consultas y transacciones en sitios web, entre tantos otros. Podrá contestar con facilidad una enorme cantidad de información pues posee la posibilidad de operar como puente entre el lenguaje natural y la información en bases de datos y la web, permitiendo acceder desde el lenguaje a todo un cúmulo de información específica para evacuar dudas y preguntas.

REALIZACIÓN

El tipo de trabajo necesario para crear un agente de diálogo inteligente, sin duda es altamente interdisciplinario, en especial porque la inteligencia no es algo que se pueda definir fácilmente, por lo cual enfocamos el problema desde un punto de vista práctico: sólo nos interesa poder crear un agente que pueda entender texto natural, extrayendo la pragmática y retórica de un cierto número limitado de situaciones, en un formato lingüístico acotado y simple, en el cual las acciones estarán igualmente limitadas.

Para poder hacer realidad esto hace falta el trabajo de programadores con conocimiento profundo de lengua y científicos con capacidad de crear sistemas nuevos, que tengan actitudes similares a las de un humano, lingüistas computacionales que sepan y puedan programar los mecanismos de reconocimiento del lenguaje natural, incluyendo profesionales capaces de unir todo este trabajo logrando que funcione todo perfectamente y en forma coordinada.

Todo esto debe estar unido a un número importante de recursos computacionales puesto que el poder reconocer una palabra ya plantea un fenomenal problema de computación,

en especial cuando el léxico no está completamente acotado y hay presunción de errores de ingreso como ser los producidos por ortografía, tipeo, olvidos, formato, palabras extranjeras y otras fuera de vocabulario o parasintéticas, se han usado un número de sistemas previamente desarrollados para tal fin descritos en [5] y [6].

Debido a todo esto se ha decidido dividir el problema en programación (algoritmos, etc.) y manejo de diálogo humano (qué decir y cuándo, cómo contestar, etc.).

Esta idea se cristalizó en un sistema de desarrollo (*IDE Integrated Development Environment*) que permite editar y probar un nuevo lenguaje de definición de diálogo (DDL) convirtiéndolo en un producto funcionando, inyectando la programación necesaria automáticamente mediante compilación activa y dinámica.

Se ha definido para esto un nuevo lenguaje (*DDL*) capaz de poder expresar y contener conocimiento latente, reglas complejas y plasmar todo eso en un sistema cuya exteriorización es la de un agente inteligente. Este lenguaje *DDL*, no es un lenguaje de computación nuevo sino un lenguaje capaz de ser manejado y creado por un humano sin conocimiento de computación alguno.

El *DDL* contiene lógica de diálogo tipo novela/guión, combinable con operadores lógico, matemáticos y lingüísticos simples a nivel intuitivo desde concatenación, operaciones del tipo de parecido fonético o semántico de palabras, partes de la oración, sintagmas, hallar raíces ontológicas, permite reconocer el idioma de una palabra o frase, reconocer palabras raras y fórmulas, entre otros muchos fenómenos de texto. Todas partes de las tareas cognitivas simples que alguien pueda reconocer algo escrito, decir algo o responder o una pregunta.

El sistema *IDE* creado se opera como una aplicación accesible desde la web, es modular y permite agregar funciones nuevas y externas como acceso a bases de datos, buscar en Google, consultar Wikipedia, etc.

El motor de diálogo (*MD*) posee un planificador de alto nivel que analiza las secuencias de diálogos bajo un concepto de satisfacción de turnos de conversación, basado en lógica pura de conversación humana, analizando entre otras cosas: la deixis usada, las coreferencias y anáforas, los tiempos de verbos empleados, el tipo de palabras usadas midiendo hasta el nivel de cultura del interlocutor, el número de errores de ortografía e interpretación y los modismos usados como: gestos, emoticones y palabras expresas o comentarios que expresan sentimientos como agrado, conformidad o disgusto.

Todo esto lo realiza el *MD* haciendo un análisis agudo y adaptativo para poder guiar la conversación hacia la maximización de la función de diálogo deseada como ser brindar respuestas y/o resolución de problemas planteados o cualquiera sea el objetivo que se le haya fijado con reglas claras.

El *MD* puede dar instrucciones al interlocutor, solicitando aclaraciones, o instando a conversar de un cierto modo, proveer información de ayuda para encauzar la conversación. Finalmente, basado en la decisión del sistema de supervisión con apoyo emocional y temático, el sistema hará un último esfuerzo (si hay suficientes elementos) en buscar temáticamente y por aproximación usando un conjunto grande de técnicas desde refinadas hasta muy ambiciosas, buscando en una base de conocimiento que el

cliente puede proporcionar, a fin de brindar las soluciones mas frecuentes o los planteos mas frecuentes de los clientes (FAQs), las preguntas frecuentes o los manuales mismos de los productos o servicios. En otras palabras, el *MD* ante cosas coherentes pero cuyo planteo no es entendible fácilmente, decidirá brindar una aproximación a lo conversado, devolviendo algo ‘parecido’ o que ‘suena similar a algo como lo que preguntó’, se puede indicar para que busque partes similares al texto dicho por el usuario y los elementos importantes de la conversación en curso, en una base de datos de conocimientos con ejemplos, manuales de uso del sistema o producto, preguntas frecuentes con las correspondientes respuestas, ver en foros internos, etc.

Todo esto apunta a poder siempre responder lo que parece más alineado con el tema de la consulta o conversación, por más que no se la comprenda. Lo buscado y priorizado es maximizar la satisfacción del usuario

CARACTERÍSTICAS DESTACADAS

RECONOCIMIENTO

- Reconocedor de Lenguaje de cada texto (a nivel palabra y oración)
- Diccionario Morfológico Extensible con Reconocimiento de >2.5 millones de palabras (español) y 1.2 M (inglés)
- Corrección Inteligente y Robusta de Errores de Ortografía (falta/omisión/teclado/fonético/etc.)
- Reconocimiento de Palabras, Locuciones y Entidades Nombradas
 - Análisis Morfológico (verbos conjugados y palabras flexionadas)
 - Reconstrucción Fonética de palabras mal escritas (por similitud)
 - Análisis Gramatical de Partes de la Oración (POS)
 - Análisis Semántico (*razgos + vinculación con synsets de WordNet*)
 - Palabras Fuera de Vocabulario (Parasintéticas)
 - Unidades Internacionales (~180 entre SI, MKS, CGS: es, en)
 - Cifras (numéricas, coloquiales, romanos, hexadecimales, etc.)
 - Fechas y Tiempos (Calendario y Horarios)
 - Locuciones (es:4500, en:1200)
 - Thesauro de sinónimos (inglés:EN y español:ES con > 250k relaciones)
 - Traducción de palabras (EN-ES-EN)
 - Entidades y Nombres Propios (es:18.500, en:8.500)
 - Distinción entre palabras locales, extranjeras e impronunciables (basura)
 - Direcciones web y de e-mail (incluso puede validarlas on-line)
- Posibilidad de extracción de información a partir de lo que ‘dice’ el usuario.
- Conversación basada en Patrones de coincidencia en cascada
 - Simples (textuales), con posibilidad de partes opcionales.
 - Sofisticados (grupos muy largos de palabras)
 - Condicionales basados en rasgos y datos.
 - Morfológicos (Etiquetas EAGLES2.0 ampliadas)
 - Gramaticales (Sintgmas, Entidades Nombradas, Segmentos de la oración)
 - Plug-ins opcionales (a medida):
 - Desambiguación: WSD Word Sense Desambiguation
 - Parser / Chunker Gramatical: GLR
 - Extracción Semántica Superficial (Shallow Semantic Extraction)

PROCESO

- Lenguaje de procesamiento de alto nivel de abstracción, capaz de operar con ‘palabras’ partes de la oración y conceptos ontológicos extractados de la conversación o introducidos para producir resultados de texto en forma simple e intuitiva mediante un arsenal de más de 70 potentes funciones provistas, incluyendo operadores lingüísticos proveyendo operaciones semánticas, gramaticales, similitud fonética, etc.
- No es necesario saber programar el sistema es muy intuitivo y simple.
- Adicionalmente incorpora la posibilidad de crear nuevas funciones por parte del usuario y/o proveerlas externamente con librerías tipo plug-in, en lenguaje s .NET permitiendo conexión con bases de datos, servicios xml, sitios web, etc.

MECANISMO de DIÁLOGO

- Sistema inteligente: “casi humano” para seguimiento del hilo de una conversación, creando verdaderas Cadenas de Razonamiento Temáticas

Ej: si contestó que sí, entonces preguntar por esto otro.. si no decirle aquello.. o preguntarle esto otro

- Generación de Respuestas basadas en datos y plantillas en la generación del diálogo de respuesta con
 - Funciones de concordancia sintáctica.
 - Capacidad de Flexión (Conjugación)
 - Manejo simple de palabras, listas, concatenaciones, etc.
- Persistencia de Información extraída del diálogo para con cada usuario.
 - El sistema puede recordar selectivamente partes de lo que el usuario dijo.
 - La capacidad de Memoria asignable a esto es completamente configurable
 - Hay posibilidad de hacer que de vinculen los usuarios entre sí para servicios de chat en grupo y redes sociales.

PERFORMANCE y DESEMPEÑO

- Velocidad de Análisis y Respuesta (>20.000 cláusulas/patrones por segundo) (*Server HP tipo ProLiant Xeon X3500/2GHz 4Gbytes RAM*)
- Posibilidad de sostener conversaciones simultáneas (>500) con 1 Gb de RAM
- Tiempo de respuesta promedio con 1800 patrones complejos: 0.2 s (< 1 seg.) (no incluye consultas a bases de datos externas al sistema cuya respuesta agregará demoras)
- Tráfico típico (en un Server chico) hasta 5 preguntas por segundo simultáneas, con hasta 50 conversaciones activas (usuarios simultáneos)

DIAGRAMA GENERAL

Se muestra un diagrama general del sistema, puesto que una descripción detallada del sistema excede la extensión del trabajo, por ejemplo el Manual del lenguaje DDL tiene 120 páginas A4 en cuerpo 10, y el tamaño de los módulos de software es de 500 mil líneas de código en 880 archivos y 2500 clases e interfases en lenguaje C#.

Se muestra en la **Tabla 1** un diagrama simple en bloques del sistema.

Tabla 1



RESULTADOS Y APLICACIONES

El sistema resultante está programado en C#, posee una gran cantidad de módulos como se puede apreciar en la **Tabla 1** y entre sus diversos módulos hay un servidor léxico multilingüe robusto, descrito en un trabajo en este mismo congreso ProLen.

El MD resultante fue efectivamente aplicado en la creación de agentes inteligentes de diálogo en lenguaje natural y asistentes matemáticos que ya funcionan en la Argentina y en Uruguay por MSN y por SMS (mensajes de texto) contra números cortos móviles.

Un uso importante del mismo fue para crear un agente inteligente para el Ministerio de Turismo de Uruguay en dic. del 2009, llamado Yorugua, quien chateando por MSN asesoraba e invitaba a veranear en el vecino país a turistas de toda Latinoamérica, en apenas unas 3 semanas, el sistema tuvo aprox. 20 mil usuarios únicos.

También fue usado anteriormente en un sistema de diálogo por MSN que imitaba un chef experto en Arroz SAMAN (una conocida marca Uruguaya), llamado Maestro SAMAN el cual ya sostuvo varios millones de turnos de conversación durante la segunda mitad del año 2009.

BIBLIOGRAFIA

- [1] *Joseph Weizenbaum*, MIT. Program: "ELIZA - Un Programa informático para el estudio del lenguaje natural en la comunicación Hombre-Máquina", *Communications of the Association for Computing Machinery* 9 (1966): 36-4
- [2] TRAINS <http://www.cs.rochester.edu/research/cisd/projects/trains/>
- [3] *Terry Winograd*, MIT. Program: SHRDLU, AI Technical Report 235, February 1971 "Procedures as a Representation for Data in a Computer Program for Understanding Natural Language", *Journal of Cognitive Psychology* Vol. 3 No 1, 1972; book: *Understanding Natural Language* (MIT Academic Press, 1972).
- [4] TRIPS "The Rochester Interactive Planning System"
<http://www.cs.rochester.edu/research/cisd/projects/trips/>
- [5] "Algoritmos eficientes para detección temprana de errores y clasificación idiomática para uso en procesamiento de lenguaje natural y texto", Hohendahl, A.T.; Zelasco J. F., WICC2006 - ISBN 950-9474-35-5
- [6] "Desarrollo de un algoritmo para la medición del grado de similitud fonológica entre formas escritas", Hohendahl, A.T.; Zanutto, B. S.; Wainselboim, A. J.; SLAN2007. X Congreso Latinoamericano de Neuropsicología 2007, Buenos Aires, Argentina
- [7] "Etiquetas EAGLES", www.lsi.upc.es/~nlp/tools/parole-sp.html, 6/2011