

Gestión Flexible de Diálogos en el proyecto ADVICE

Luis Rodrigo Aguado¹, Ana García Serrano¹ y Paloma Martínez²

¹Grupo ISYS-Hermeneumática
Departamento de Inteligencia Artificial
Universidad Politécnica de Madrid
{agarcia,lrodrigo}@isys.dia.fi.upm.es

²Grupo de Bases de Datos Avanzadas
Departamento de Informática
Universidad Carlos III de Madrid
pmf@inf.uc3m.es

Resumen

En el presente artículo presentamos el trabajo realizado sobre la gestión del diálogo dentro del prototipo desarrollado en el proyecto ADVICE (EU Project IST 1999-11305).

Gestor de diálogos

El objetivo general del sistema es el diseño e implementación de un sistema de comercio electrónico capaz de asistir al comprador, pasando del enfoque actual basado en catálogos a una ayuda inteligente, tratando de emular de alguna manera la figura del vendedor humano. Para alcanzar este objetivo, se ofrecerá una selección de productos personalizada y consejos sobre los mismos por parte del asistente virtual, incluyendo sugerencias adaptadas para el usuario, explicaciones sobre los diferentes tipos de productos, características, alternativas y ofertas especiales. Con este objetivo, los principales elementos del sistema ADVICE incluyen: (A) un interfaz multimedia y basado en lenguaje natural que permita la interacción basada en el diálogo; y (B) una arquitectura basada en agentes, que incluye: un agente servidor, que se ocupa de acceder a las bases de datos de productos; un agente de interfaz, que maneje el diálogo multimedia; el agente de interacción que sea capaz de seguir la evolución de la conversación entre el usuario y el sistema y, por último un agente inteligente que contiene el conocimiento del dominio y permite la adaptación de los consejos por parte del sistema al usuario dependiendo de sus necesidades.

El diálogo es un acto en el que ambos participantes deben tener un contexto común para que se pueda llegar a buen puerto. Cuando los participantes coinciden sobre una misma intención, se dice que han llegado a un acuerdo. Las interfaces

clásicas de diálogo dejan toda la responsabilidad de la interacción al usuario, limitándose a responder las iniciativas de éste, en lugar de compartir la responsabilidad.

Para un diálogo flexible es necesario que exista al menos una meta común (para que los participantes se entiendan entre sí). Estas metas son las que provocan las aclaraciones y confirmaciones tan habituales en las conversaciones. Investigaciones para modelizar las metas comunes son los modelos teóricos de acción combinada [Cohen & Levesque, 91]. El reciente proyecto IST Trindi [Cooper & Larsson, 99] argumenta la necesidad de un contexto común entre el usuario y el sistema. El proyecto ADVICE se une a esta línea de investigación prototipando el modelo de hilos en un asistente virtual para el comercio electrónico.

Una gestión eficaz del diálogo requiere generalmente tanto información semántica (qué es lo que se está expresando) como pragmática (qué tipo de actos de habla se utilizan) [Grosz, Scott, y Kamp, en Cole 97]. De acuerdo a [Cohen, en Cole 97] también es necesario que el sistema participe de alguna manera en el diálogo.

Para el proyecto ADVICE distinguimos tres tipos distintos de actos de habla [Searle, 69], dependiendo de su función en el diálogo: una pieza de discurso (una expresión única), un discurso completo (un movimiento en el juego del diálogo), y el nivel de hilos (cada intención en cada subdiálogo). El conjunto de *speech acts* identificados se muestra en la figura 1:

<i>Courtesy acts</i>	
Salute: < c, a >	[c]: conventional (formal/informal)
Farewell: < c, a >	[a]: allowable (open/close)
Thank: < c, a >	<u>Examples:</u>
Disannoy: < c, a >	<i>Nice to see you [name]:</i> salute(i,c)
Empathetic: <c,a>	<i>Excuse me...:</i> disannoy(f,c)
Satisfy: < c, a >	<i>Don't worry...:</i> empathetic(i,c)
Wish: < c, a >	<i>Excellent:</i> satisfy(f,c)

Representative acts	
Inform: <t,m,s,c>	[t]: type (confirmation / data /...) [m]: matter (approve/deny/identity/...) [s]: subject (product/user/system/...) [c]: content (...) <u>Example:</u> <i>I'm AI: inform(data,identity,user,AI)</i>
Authoritative acts	
Authorize: <m,a>	[m]: matter (start / offer / task / ...) [a]: allowable (open / closed) <u>Example:</u> <i>Can I help you? authorize(task,open)</i>
Directive acts	
Request: <t,m,s,c>	[t]: type (choice/data/comparison/...) [m]: matter (approve/deny/identity/...) [s]: subject (user/system/...)
Command: <t,s,c>	[c]: content (values...) <u>Examples:</u> <i>Who are...?: request(data,identity,...)</i> <i>Show me some saws:</i> <i>command(search, system, product)</i>
Null Speech	
Null: < >	<i>Well,... : null()</i>

Figura 1: Conjunto de actos de habla del proyecto ADVICE

El sistema ADVICE contempla dos tipos de entrada de datos: en lenguaje natural (en inglés) a través del módulo intérprete de lenguaje, y a través del interfaz gráfico de usuario (navegando a través de las opciones de la página Web). El agente de interfaz se encarga de transformar las diferentes entradas del usuario a estructuras semánticas (cadenas de actos de habla) y se los pasa al agente de interacción (la parte del sistema encargada de la gestión del diálogo). Para presentar datos al usuario, existen tres posibilidades: el generador de LN, el interfaz gráfico de usuario y un avatar en 3D. Las estructuras semánticas de salida que genera el agente de interacción (actos de habla también) son la entrada para el agente de interfaz, que se encarga de construir una salida coordinada aprovechando los diferentes canales antes mencionados.

Las intervenciones en el diálogo se dividirán en 'piezas de discurso' (subconjunto del discurso con significado independiente), y cada uno de ellos es representado como un conjunto de actos de habla.

A lo largo de una intervención, hay que adaptar y actualizar el hilo o línea locutiva

y también producir una transición del estado de diálogo previo a el(los) siguiente(s). Finalmente, los actos de habla también contienen otro tipo de información: los detalles del diálogo (nombres, datos, características, etc.). Esta información es estática para toda la sesión y se almacena en el modelo de sesión.

El componente que se encarga de la gestión del diálogo tiene el control de la conversación (los aspectos formales a través de los estados del diálogo y los aspectos intencionales a través de los hilos). Después de procesar la información que contienen los actos de habla, el generador de discurso debería ser capaz de construir una respuesta del sistema. En este punto puede ser necesario cierto tipo de conocimiento para completar la respuesta. Este conocimiento puede encontrarse en el contexto del diálogo, o bien, si se trata de conocimiento nuevo el agente de interacción se lo pedirá al agente inteligente. La respuesta del agente inteligente es una oferta configurable modelizada como un árbol de decisión. El generador de discurso extraerá del árbol la información necesaria para alcanzar una solución, requiriendo ciertas explicaciones al usuario, si fuesen necesarias. Una vez se tiene toda la información necesaria se construye la estructura semántica, compuesta por una serie de actos de habla encadenados. El agente de interfaz será el encargado de traducir esos *speech acts* en frases, iconos, menús, fotografías, movimientos del avatar 3D, o una combinación de ellas.

Referencias

- Cohen, P.R., Levesque, H.J. (1991) *Confirmation and Joint Action*, Proceedings of International Joint Conf. on Artificial Intelligence, 1991
- Cole, R.A. ed. (1997) *Survey of the state of the art in Human Language Technology*.
- Cooper, R., Larsson, S., (1999). *Dialogue moves and information states*. In Proc. of the Third IWCS, Tilburg, 1999.
- Searle, J.R., (1969). *Speech Acts: an essay in the philosophy of language*. Cambridge Univ. Press.