

Planteamiento semántico y pragmático para gestión de diálogos en asistentes virtuales

Luis Rodrigo Aguado
lrodrigo@isys.dia.fi.upm.es
Laboratorio de Sistemas
Inteligentes
Facultad de Informática
Universidad Politécnica de
Madrid

Ana M. García Serrano
agarcia@isys.dia.fi.upm.es
Laboratorio de Sistemas
Inteligentes
Facultad de Informática
Universidad Politécnica de
Madrid

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

Resumen: La gestión de diálogos y el tratamiento del lenguaje aparecen como piezas clave en los asistentes virtuales para tiendas electrónicas, al ser su principal objetivo el asistir al usuario tanto en la navegación por las páginas de productos de la tienda como en la búsqueda del producto más adecuado a los requisitos del usuario. En este trabajo se presentan los detalles del planteamiento semántico y pragmático que se ha realizado en el marco del proyecto ADVICE (IST-1999-11305) para una tienda de herramientas de bricolaje.

Palabras clave: gestión de diálogo, gramáticas semánticas, asistente virtual.

Abstract: Dialogue management and language processing appear as key points in virtual assistants for e-shops, as their main goal is to assist the user both in his navigation through the shop product pages and in his search for the most appropriate product. In this paper we present the details of the semantic and pragmatic approach done in the ADVICE project (IST-1999-11305) for a bricolage tools shop.

Key words: dialogue management, semantic grammars, virtual assistant.

1 Introducción

La vertiginosa popularización de Internet, ha provocado que el número de personas que necesitan interactuar con ordenadores haya crecido de manera espectacular. Normalmente, son las personas las que adquieren la preparación necesaria para que esa interacción se pueda llevar a cabo en términos satisfactorios, aunque sería deseable reducir esa distancia comunicativa al mínimo. Este objetivo hasta el momento no se ha conseguido de manera genérica y se plantea como necesario, por un lado, el desarrollo de recursos intermedios que faciliten el procesamiento del lenguaje natural y, por otro lado, probar nuevas técnicas en dominios restringidos que permitan una forma de comunicación satisfactoria con los recursos disponibles.

El trabajo descrito en el presente artículo se enmarca dentro de esta segunda tendencia, ya que se describe cómo un vendedor virtual asociado a una tienda electrónica puede

comunicarse con los usuarios o potenciales compradores (humano). En este entorno ocurre de manera natural que el conocimiento involucrado está limitado al dominio de productos y aplicaciones de la tienda electrónica y el sistema debe tratar solo con el sublenguaje propio de una relación cliente-vendedor.

El asistente virtual desarrollado para esta aplicación (Rodrigo, García Serrano y Martínez, 2001) es capaz de interactuar con el usuario en lenguaje natural, adaptar sus recomendaciones a los requisitos y características del usuario y proporcionar explicaciones sobre los productos que ofrece, así como productos alternativos.

El diálogo se centra en llevar a cabo dos tareas específicas, una es vender productos en una tienda de electrónica de acuerdo con los requerimientos del cliente y otra ofrecer consejos sobre estos productos. La complejidad reside en la gran variedad de productos, con gran especialización tanto en las herramientas

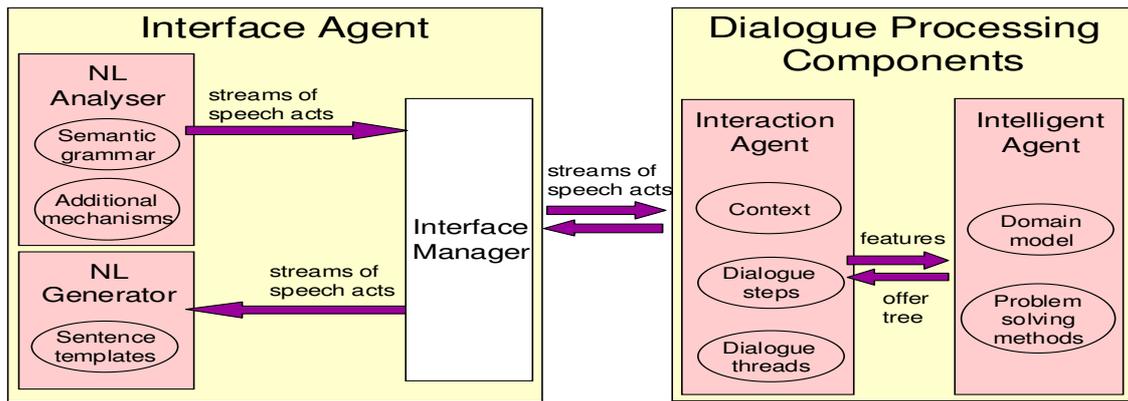


Figura 1: Arquitectura general del sistema

como en las aplicaciones que pueden tener estas herramientas.

En los siguientes capítulos se detallarán los componentes de información necesarios para el dialogo, haciendo especial énfasis en los de análisis del lenguaje natural, aunque contando brevemente aspectos de la gestión del diálogo y de la generación del lenguaje.

2 Asistentes virtuales

Aunque la idea de integrar un agente virtual con capacidad para comunicarse en lenguaje natural con los usuarios de una tienda electrónica es relativamente reciente, existen varias páginas que ya cuentan con este tipo de asistentes con dos diferentes facilidades: la cumplimentación de formularios para (a) acortar la navegación por las páginas de productos, o (b) recuperar páginas de tiendas electrónicas con determinados productos.

Desde el punto de vista del procesamiento del lenguaje, las aplicaciones web del tipo "chatter bots", como ocurre en los casos de ALICE, Hal, Max y NICOLE, si bien son capaces de aceptar la entrada del usuario y dar alguna respuesta que parece inteligente, en realidad no realizan ningún tipo de procesamiento del lenguaje. Para generar estas respuestas, disponen de una gran cantidad de patrones de frases pregunta (del orden de decenas de miles) con respuestas predefinidas asociadas, pero no suelen incorporar gestión del contexto ni conocimiento relacionado con el tema del dialogo. Esto hace que no sean capaces de mantener un diálogo con un objetivo concreto, ni sean capaces de utilizar referencias a puntos anteriores de la conversación ya que no manejan el hilo del diálogo. En los casos

más avanzados estos bot acceden a bases de datos de la compañía y son capaces de incluir cierta información dentro de sus respuestas.

Sin embargo, se desearía que el agente virtual supiera responder las peticiones del comprador de la manera más cercana a como lo haría un vendedor humano, lo que implica, por ejemplo, recordar los detalles que vaya proporcionando el usuario para no tener que preguntar más tarde los mismos, ser capaz de anticiparse a alguna de las preguntas del cliente o bien aportar las explicaciones necesarias en cada momento sobre los productos para que el usuario no tenga que visitar las páginas que se le ofrecen y así hacer más eficiente su navegación.

Para alcanzar este tipo de interacción es necesario:

1. Durante el procesamiento de las frases del usuario (a) extraer la información que el usuario codifica, (b) las intenciones que tiene al decirlas y (c) las informaciones puntuales que aporta.
2. Gestionar el desarrollo del diálogo, manteniendo un hilo argumental lógico y de acuerdo a las intenciones del usuario y, por último,
3. un conocimiento exhaustivo acerca del dominio de la tienda electrónica, para poder aportar contenido a la conversación.

De acuerdo con los requisitos anteriores el diseño general del sistema (Figura 1) se ha realizado sobre la base de dos componentes: el agente interfaz que incluye los módulos necesarios para la interpretación de las intervenciones del usuario en lenguaje natural o utilizando el interfaz gráfico y el componente de proceso del dialogo que incluye al agente de interacción que gestiona el dialogo y al agente

inteligente con conocimiento sobre la tienda electrónica.

Cuando llega una intervención por parte del usuario a través del agente de interfaz, mediante el interfaz gráfico o de lenguaje natural, se extraen la intención y el contenido que se envían al agente de interacción. Este generará una intención de respuesta, que, en caso de necesitar consultar o incluir algún dato relativo al dominio, se lo preguntará al agente inteligente, quien posee una descripción detallada de los productos del dominio y sus aplicaciones. Finalmente, esta intención de respuesta genera una respuesta concreta en el agente de interfaz, que se lleva a cabo coordinando al interfaz gráfico y el texto en lenguaje natural con el avatar que representa al vendedor virtual.

3 *Análisis intencional del lenguaje*

El objetivo principal de un sistema de procesamiento de lenguaje natural que forme parte de un sistema de diálogo es identificar y transmitir el contenido y la intencionalidad de la intervención del usuario, para que la gestión del diálogo se realice aprovechando toda la información que el usuario haya aportado.

Se presentan a continuación los procesos de análisis semántico-pragmáticos, tras una breve descripción del análisis del corpus realizado.

3.1 *Análisis del corpus*

Dado que el sistema debe tratar con un sublenguaje concreto, se hizo un estudio de corpus de diálogos del dominio. Este corpus fue creado artificialmente para esta tarea y contiene 202 consultas de usuarios al sistema y 10 diálogos completos de compra de productos, un tamaño suficiente para obtener conclusiones de cara al diseño del sistema. En primer lugar, se identificó la terminología propia del dominio y las relaciones que existían entre estos términos.

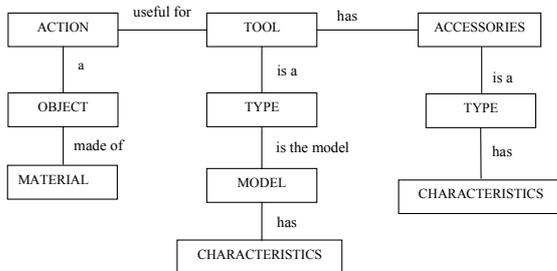


Figura 2. Ontología del dominio.

A partir de esos datos y con un paso de abstracción de los términos concretos a conceptos de más alto nivel, se obtuvo un modelo conceptual del dominio con las relaciones que existen entre los diferentes conceptos y la terminología asociada a cada concepto, es decir, una ontología del dominio (Figura 2).

Basándonos en la ontología se realizó una nueva etapa de estudio de las frases del corpus, que confirmaban patrones de frases que incluyen terminología (Figura 3).

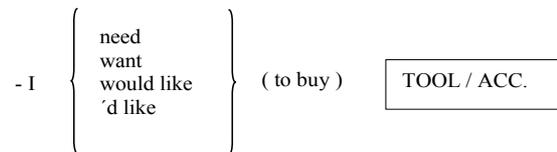


Figura 3. Ejemplo de patrón de frase

3.2 *Análisis semántico-pragmático*

De las intervenciones del usuario es necesario identificar su intención y extraer los datos concretos que estén incluidos en las frases, los detalles de la conversación. Para codificar estos dos tipos de información conjuntamente se ha creado un conjunto de actos comunicativos a partir de los propuestos por Searle, (1969). Los siete tipos diferentes de actos comunicativos se presentan a continuación.

3.2.1 *Aportando información: inform*

Este acto de habla se usa cuando uno de los interlocutores aporta una información (no necesariamente nueva en la conversación) dentro de una frase. Tiene cuatro parámetros:

- Type: determina el tipo de información de la que se trata. El valor más usual es el de "data", que indica que se proporciona un dato. Los otros valores posibles hasta el momento son "confirmation" y "suggest".
- Matter: Especifica el porqué de la información. Si, por ejemplo, el tipo de la información es "data", a qué corresponde ese dato. Si es de tipo "confirmation", el carácter de la confirmación, etc. Este parámetro especializa de alguna manera la información proporcionada con el primero ("Type").
- Subject: A qué se refiere la información.
- Content: La información en sí misma. Mientras los posibles valores de los otros campos son fijos, este campo es completamente

abierto. Puede contener nombres propios, números, identificadores de productos, en resumen, los datos específicos que se proporcionan en una frase.

Por ejemplo, una frase como "I am John Smith", se codificaría como:

INFORM	{	Type: data Matter: identity Subject: user Content: John Smith
--------	---	--

ya que la única intención de la frase es proporcionar una información y toda esa información (el nombre del hablante) está contenida en el acto de habla.

3.2.2 Solicitando información: request

Este acto de habla se usa para representar la situación en la que uno de los hablantes solicita al otro que aporte algún tipo de información. Tiene cuatro parámetros análogos a los del acto de habla "inform".

- Type: representa el tipo de petición que se realiza. Los valores posibles para este campo detectados hasta el momento, son "data", "comparison", "choice" y "confirmation", según lo que se solicita es un dato, una comparación, una elección, o una confirmación, respectivamente.

- Matter: especifica exactamente qué se demanda del interlocutor. Como en el caso anterior, este campo complementa el parámetro "type", para expresar exactamente qué se está demandando. Por ejemplo, si el campo "type" tiene el valor "data" (que al igual que en el caso anterior es el valor más habitual para este campo) y este campo tiene el valor "identity", el hablante está solicitando información para identificar algo (posiblemente el nombre sea lo más adecuado).

- Subject: Dice a qué se refiere la información que se solicita. Puede ser un producto, un accesorio, un material, o cualquier elemento relacionado con el dominio.

- Content: Este campo tiene valor nulo en todos los casos excepto cuando el campo "type" tiene valor "choice", en cuyo caso se especifican aquí la lista de valores entre los que se debe hacer la elección.

Un ejemplo de este tipo de acto de habla se puede observar en la frase "What is a

systainer?", que se codificaría como:

REQUEST	{	Type: data Matter: identity Subject: accessory Content: null
---------	---	---

INFORM	{	Type: data Matter: identity Subject: accessory Content: systainer
--------	---	--

En este caso no se puede incluir toda la información de la frase en un solo acto de habla, que represente tanto la intencionalidad (preguntar un dato), como la información asociada ("systainer"), así que se utilizan varios para codificar la intervención.

3.2.3 Cambiando la iniciativa: authorize

Este es el acto de habla probablemente menos intuitivo. Se usa para representar cuando uno de los interlocutores le cede la iniciativa en el diálogo al otro. Por ejemplo, cuando un usuario se conecta, es el sistema el que coge la iniciativa en el diálogo saludándole, pero inmediatamente se la cede al usuario preguntándole "¿Te puedo ayudar en algo?". Esto es exactamente lo que representa este acto de habla.

Tiene tres parámetros:

- Matter: Representa la razón de la autorización. Puede ser simplemente un nuevo diálogo, o una oferta que tenga que ser aceptada o rechazada.

- Allowable: Los valores posibles son "open" y "closed". Se usa para diferenciar si se espera una respuesta explícita del usuario (valor "open") o no (valor "closed"). Es puramente intuitivo, y de hecho se puede encontrar una respuesta incluso si el parámetro tenía el valor "closed".

- Subject: A qué se refiere la nueva iniciativa. Puede ser un nuevo producto, o una nueva tarea dentro del diálogo.

Un ejemplo de frase que necesita este acto de habla sería, por ejemplo, "Can I help you?", que se representaría como:

AUTHORIZE	{	Matter: start Allowable: open Subject: task
-----------	---	---

3.2.4 Dando órdenes: command

El acto de habla "command" se usa cuando uno de los hablantes le manda hacer algo al otro. Debido a la limitación del dominio, este acto de habla queda bastante simplificado, ya que el único que lo utiliza es el usuario, para mandar al sistema que haga una búsqueda de productos, o le ayude en la decisión de la compra de un producto.

Tiene tres parámetros:

- Matter: especifica cuál es la orden. En este momento solo se han detectado como posibles valores "purchase" y "search", que se asocian a las órdenes de asistir en la compra de un producto y de realizar una búsqueda, respectivamente.

- Subject: representa a quién se dirige la orden. Por la razón que ya se comentó, en este momento solo es posible el valor "system" aquí, ya que las órdenes siempre van dirigidas al sistema.

- Object: aclara a qué se refiere la orden.

La frase "I want a hammer" tiene, entre otros un acto de habla asociado de este tipo, ya que aunque a nivel locutivo solo expresa el deseo de poseer un martillo, a nivel ilocutivo lleva asociada un mandato para que el sistema le ayude en la compra de un martillo. El conjunto de actos de habla que codifican esta frase sería:

COMMAND	{	Matter: purchase
		Subject: system
		Object: product
INFORM	{	Type: data
		Matter: identity
		Subject: product
		Content: hammer

3.2.5 Siendo educados: actos de cortesía

Dentro de esta categoría se encuadran un conjunto de actos de habla que no son estrictamente necesarios en una conversación en términos informacionales, sino que aportan algo de carácter emocional, la única información de este tipo con la que trata el sistema. Este conjunto de actos de habla comparte los mismos dos parámetros:

- Allowable: Este campo tiene el mismo objetivo que el campo del mismo nombre incluido en el acto de habla "allowable".

- Conventional: este parámetro se refiere a la formalidad en la expresión usada. Tiene dos

posibles valores, "formal", si se usa una expresión más refinada, o "informal", para una expresión más coloquial.

Se han encontrado siete tipos de actos de habla que pertenecen a esta categoría. Estos actos de habla son muy simples, por lo que se proporciona simplemente un ejemplo de su uso:

Salute : "Hello, Mr. Smith."

Farewell (despedidas): "See you soon."

Thank: "Sincerely grateful"

Disannoyance: "Wait a minute, please."

Empathetic: "Don't worry, Mr. Jones."

Satisfy (satisfacción): "Excellent!"

Wish: "I hope that was of use for you."

3.2.6 Nada que decir: null

Finalmente, hay veces que algunas palabras de una persona no tienen una intención concreta. Algunas palabras son simplemente convenciones o clichés. Este tipo de expresiones carentes de significado son las que se simbolizan con este acto de habla.

Este acto de habla tiene un parámetro asociado para diferenciar la situación en la que se usa: el parámetro "Type".

3.3 Gramática semántica

Una vez hemos generalizado las frases del corpus a patrones y, además, disponemos de una manera de anotar el conocimiento que lleva incluido cada patrón, se puede etiquetar cada patrón con los actos comunicativos que se deberían obtener de su análisis, con lo que, si el corpus elegido es suficientemente representativo, y asumimos que la abstracción desde las frases a los patrones ha sido suficientemente genérica, obtendremos el resultado deseado. Estos patrones de frases y sus actos comunicativos asociados se representan de manera natural mediante una gramática semántica.

Este tipo de análisis, si bien resulta muy intuitivo y muy directo, tiene un grave problema, y es que no importa cómo de extenso y completo sea el corpus analizado que, debido a la riqueza propia del lenguaje, siempre nos encontraremos con frases no conocidas, que no encajen con ninguno de los patrones que tenemos, por lo que la cobertura tiende a ser bastante baja. Por ello, se complementó el análisis por medio de esta gramática con dos mecanismos que mejoran la cobertura del análisis a costa de sacrificar algo de precisión, que en el caso del análisis directo con la gramática semántica es prácticamente perfecta.

El primero de ellos es una técnica de relajación de las reglas de la gramática, de manera que no es necesario que la frase del usuario encaje perfectamente con alguno de los patrones que hay en la gramática, sino que se permite que haya partes de la frase que el mecanismo de análisis se pueda saltar, si con ello se consigue que el resto de la frase encaje entera con el patrón. Esto hace que haya muchas más frases que se puedan analizar, si bien en las partes que se saltan de la frase de entrada podemos perder algo de información, y con ello precisión en el resultado.

El segundo de los mecanismos añadidos deriva de la ontología que surgió fruto del estudio del corpus, y que identifica los conceptos principales del dominio junto a la terminología que lexicaliza cada uno de esos conceptos. En un sistema de este tipo es esencial que el usuario confíe en que el sistema va a ser capaz de entenderle, ya que si pierde esa confianza dejará de hacer uso del sistema. Por tanto, el conjunto de conceptos marcados por la ontología siempre deberían ser entendidos por el sistema, se podría decir que forman la línea base del análisis. Se ha implementado un análisis que siempre que aparezca una palabra que representa alguno de los conceptos de la ontología del dominio es capaz de devolver la información asociada, aunque no es capaz de devolver otra cosa que no esté asociada a alguna de estas palabras. La cobertura de este tipo de análisis es muy alta, ya que una gran mayoría de las frases del usuario contiene alguna de estas palabras, aunque la precisión es baja ya que no se aprovecha la información del resto de la frase.

Finalmente, hay que escoger entre los tres resultados (posiblemente diferentes) que obtenemos de los tres tipos de análisis cuál será el resultado final. En primer lugar, si el análisis directo usando la gramática semántica obtiene un resultado, lo escogemos como resultado del sistema, ya que es el tipo de análisis que consigue una mejor precisión. Si este tipo de análisis no tiene éxito, la estrategia que se ha decidido es la de combinar los resultados obtenidos por los otros dos mecanismos de análisis, ya que el primero de ellos tiene más en cuenta la estructura general de la frase, aunque se puede "saltar" informaciones importantes, mientras que el segundo recupera la información puntual básica sin atender a la estructura de la frase en la que se encuentra.

3.4 Gestión del diálogo

La información proveniente del análisis de la intervención del usuario se trata de tres maneras distintas, (García Serrano et al., 2001).

En primer lugar, se extrae la información estática, los detalles de la conversación. Si esta información se limita al ámbito de la conversación, se almacena como contexto. Por ejemplo si el usuario nos dijera que necesita una herramienta para cortar madera, este dato sería parte del contexto de la conversación, y cuando acaba la conversación deja de tener valor. En contraposición a este tipo de información estática está la información (también estática) que perdura más allá del ámbito de la conversación como sería si, por ejemplo, el usuario nos informase de que es carpintero. Esta información es susceptible de formar parte de un modelo de usuario, sin embargo, el desarrollo de un modelo de usuario es una tarea compleja que queda fuera del ámbito de este trabajo, y actualmente toda la información estática se trata como información contextual y se desecha cuando acaba el diálogo.

Los otros dos tipos de información que se extraen del mensaje del usuario son de carácter dinámico, información que indica cómo se está desarrollando el diálogo. En primer lugar están los pasos de diálogo, (Amores y Quesada, 2001). Un paso de diálogo limita los posibles pasos siguientes que puede tomar el interlocutor. Esta información se ha modelizado por medio de pares de adyacencia que se han extraído del corpus de diálogos. Por otro lado se tiene en cuenta el hilo de la conversación. Para que dos interlocutores se entiendan no es suficiente (aunque sí necesario) con que respeten los pasos de diálogo mencionados anteriormente. También es obligado que estén hablando de lo mismo, que compartan el hilo de la conversación. Ese es el tercer (y último) tipo de información que se maneja. Se trata de intentar averiguar cuál es la intención que subyace a lo largo de las intervenciones del usuario, para que el sistema pueda responder adecuadamente, de manera que la comunicación sea exitosa. Estos hilos se organizan en forma de árbol, de manera que las conversaciones se pueden anidar entre sí. Por ejemplo, en un momento de la conversación en el que el sistema está vendiendo un producto al usuario, éste puede preguntar sobre los accesorios disponibles para el producto, iniciando así un sub-diálogo. Cuando este sub-

diálogo se acaba, el sistema retoma la conversación anterior e intenta continuar con la venta del producto.

4 Generación del lenguaje

Finalmente, hay que responder al usuario con una salida acorde con la entrada, por lo que el agente de interacción analiza cuales son los posibles pasos de diálogo que se pueden tomar y cuál es el hilo del diálogo activo para generar un conjunto de actos comunicativos de respuesta. El contenido que completa esos actos comunicativos puede provenir bien del contexto si son datos que ya se han mencionado durante el diálogo o bien del agente inteligente si es nuevo conocimiento acerca del dominio el que se aporta. Esta información se remite al agente de interfaz, que se encarga de presentar al usuario la información que representan.

Para presentar una salida en lenguaje natural, dado que el conjunto de intenciones que puede mostrar el sistema está acotado y bien definido, existen unas plantillas de frases que se completan con la información estática que incluyen los actos comunicativos recibidos. Existe suficiente variedad de plantillas de frases preparadas para cada posible intención de manera que la interacción no sea repetitiva, incluso aunque algunas intenciones se puedan repetir a lo largo del diálogo.

5 Evaluación

Durante el desarrollo del sistema, se llevaron a cabo dos procesos de evaluación, el primero tras

la fase de desarrollo de la gramática semántica, y el segundo cuando se completó la fase de mejora de la cobertura con los dos mecanismos de análisis comentados. Los resultados ratifican las afirmaciones realizadas hasta el momento sobre la escasa cobertura de la gramática semántica y la mejora de la misma con los mecanismos añadidos. Para poder comparar los resultados obtenidos en ambas fases, el conjunto de frases que se utilizó para evaluar el sistema fue el mismo, esto es con la misma gramática semántica y con el análisis usando reglas relajadas y el basado en elementos aislados en el segundo proceso de evaluación.

Con el objetivo de la evaluación, se creó un nuevo conjunto de diez diálogos. Estos diez diálogos suman un total de 68 intervenciones del usuario pudiendo contener cada intervención una o más frases, con un máximo de 14 intervenciones en el diálogo más largo y un mínimo de 3 intervenciones en el más corto.

Los resultados del sistema al analizar cada una de las frases se clasificaron en cuatro categorías:

- P: el resultado ofrecido por el sistema es perfecto.
- Par: el resultado devuelto por el sistema es correcto, pero parcial, no es completo.
- N: El sistema no es capaz de analizar la frase de entrada y por lo tanto no es capaz de devolver ningún resultado.
- W: El sistema analiza mal la frase de entrada y devuelve un resultado erróneo.

El hecho de obtener como resultado "W" debe considerarse un error grave, ya que indica

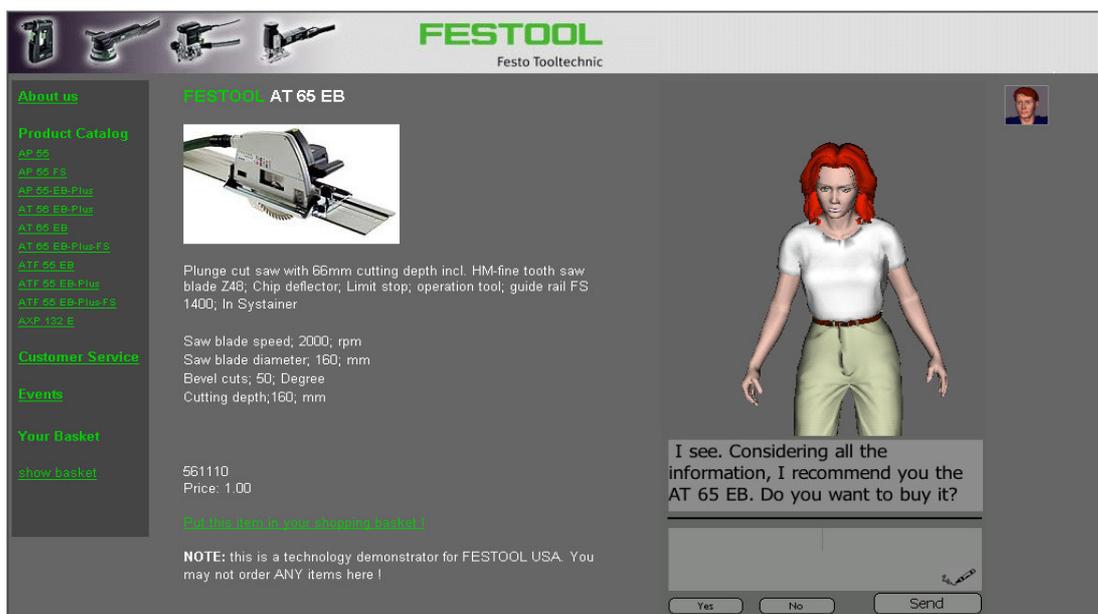


Figura 4. Interfaz del sistema completo

un fallo en el diseño del sistema en algún punto. Sin embargo, el obtener bien "N" o bien "Par" como resultado solo indica que hay una falta del conocimiento necesario para analizar la frase. Dado que el sistema se ha diseñado teniendo en cuenta que sea fácilmente ampliable, añadir el conocimiento preciso para convertir estos resultados en "P" es abordable. Finalmente, obtener "P" es lo deseable, ya que significa que se ha conseguido extraer toda la información relevante de la intervención del usuario.

Se observa en la Tabla 1 que se ha pasado de conseguir recuperar información correcta de las frases (suma de los resultados "P" y "Par") en un 60.29% en la primera evaluación a un 77.93% en la segunda, con lo que se ha mejorado la cobertura del sistema. Además, dado que la información que devuelve el sistema completa y correcta (resultado "P") también ha aumentado, podemos concluir que la precisión también se ha mejorado con en el segundo proceso.

Resultado	Eval. 1	Eval. 2	Diferencia
P	30.88 % (21)	41.17 % (28)	+ 10.29 %
Par	29.41 % (20)	36.76 % (25)	+ 7.35 %
N	38.23 % (26)	19.11 % (13)	- 19.12 %
W	4.41 % (3)	2.94 % (2)	- 1.47 %

Tabla 1. Resumen de los datos de las evaluaciones.

En términos genéricos, se puede considerar satisfactorio el rendimiento del sistema, ya que se acerca a un 80% de intervenciones en las que se es capaz de extraer información, de manera que el diálogo puede continuar sin interrupciones. Consideramos negativo, a la hora de juzgar estos resultados el hecho de que no se dispuso de un corpus real, sino que se creó artificialmente por no expertos en el tema, y consideramos que caso de disponer de un corpus real los resultados a los que se puede aspirar serían mejores.

6 Conclusiones

Como conclusiones del trabajo descrito cabe destacar que se ha conseguido un nivel mínimo de gestión de dialogo para que un usuario pueda "confiar" en el asistente virtual de la tienda.

Además, el manejo uniforme de la información del diálogo sobre la base de los actos comunicativos permite una interacción multimodal, ya que no es necesario que la

entrada provenga exclusivamente en lenguaje natural, sino que puede provenir de cualquier medio que se pueda traducir en actos comunicativos, como en el caso más habitual del interfaz gráfico de usuario (por medio de clicks de ratón).

Se puede observar un ejemplo de esta integración en el piloto desarrollado en la Figura 4, en la que lenguaje natural, interfaz gráfico de usuario y avatar se coordinan para responder al usuario. No obstante, el sistema es mejorable en muchos aspectos. En primer lugar, es necesario trabajar con un corpus real y de amplia cobertura, que haga lo más realista posible las conclusiones extraídas de su estudio.

En cualquier caso, el modelo de análisis expuesto está limitado por su propio diseño, y en los trabajos futuros se plantea conseguir un análisis más robusto, incluyendo los niveles morfológico y sintáctico a partir de la adaptación de recursos lingüísticos existentes (García Serrano, Martínez y Rodrigo, 2001).

Bibliografía

- ALICE (<http://www.alicebot.org>)
 Amores J. Gabriel, J.F.Quesada. 2001. Dialogue moves for Natural Language Command Languages, *Revista SEPLN*, Volumen: nº 27, pp. 81-88.
 Dale R., H. Moisi, H. Somers, editores. 2001 Handbook of Natural Language Processing. Marcel Dekker Inc.
 García-Serrano, A., P. Martínez, L. Rodrigo. 2001. Adapting and extending lexical resources in a dialogue system. *Proceedings of the Annual Meeting of the Association for Computer Linguistics*, Toulouse, Francia.
 Garcia-Serrano, A., et al. 2001. A knowledge based design of a NL-Interaction for intelligent assistance in the e-commerce scenario. *IEEE International Workshop on Natural Language processing and Knowledge Engineering*, Tucson, EE.UU.
 Hal (<http://www.zabaware.com>)
 Max (<http://www.extempo.com/tours/>)
 NICOLE (<http://www.neuromedia.com>)
 Rodrigo Aguado, L., A. García Serrano, P. Martínez. 2001. Demostración de prototipo: Gestión flexible de diálogos en el proyecto ADVICE. *XVII Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural*, Jaén, España.
 Searle J. R.. 1969. *Speech Acts*. Cambridge University Press, Cambridge, Reino Unido.