

Dos aproximaciones basadas en reglas para la gestión del diálogo*

David Griol, Lluís F. Hurtado, Emilio Sanchis, Encarna Segarra

Departament de Sistemes Informàtics i Computació
Universitat Politècnica de València. E-46022 València, Spain
{dgriol, lhurtado, esanchis, esegarra}@dsic.upv.es

Resumen: El objetivo principal del artículo es presentar dos modelos de gestión del diálogo basados en reglas, diseñados a partir de la estrategia utilizada para adquirir un corpus mediante la técnica del Mago de Oz y fundamentados en la utilización de medidas de confianza para la detección y corrección de errores. Este trabajo se enmarca dentro del Proyecto DIHANA, que tiene como objetivo el estudio y desarrollo de un sistema de diálogo robusto para el acceso a sistemas de información. **Palabras clave:** Sistemas de diálogo, reconocimiento automático del habla espontánea, comprensión del habla, procesamiento del lenguaje natural, gestión del diálogo.

Abstract: This article presents two strategies for the development of two dialogue managers based on rules. These rules are defined from a corpus obtained by the technique of Wizard of Oz. Confidence measures are used for error detection and recovery. This work is included in the DIHANA Project, whose goal is the design and development of a dialogue system for the access to an information system using spontaneous speech.

Keywords: Dialogue systems, automatic spontaneous speech recognition, language understanding, natural language processing, dialogue management.

1. Introducción

Un sistema de diálogo es una interfaz hombre-máquina capaz de reconocer y comprender una entrada hablada y producir una salida oral como respuesta. Para llevar a cabo este proceso intervienen sistemas de distinta complejidad que: deben reconocer las palabras pronunciadas, comprender su significado, gestionar el diálogo (incluyendo la información contextual, manejo de errores y acceso a la aplicación final) y generar la respuesta oral. Una descripción de sistemas de diálogo disponibles en la actualidad puede encontrarse en (Rudnicky et al., 1999), (Zue et al., 2000) y (Lamel et al., 2000).

El trabajo que presentamos en esta comunicación describe dos estrategias basadas en reglas para el diseño de un gestor de diálogo para una aplicación de acceso a un sistema de información mediante habla espontánea.

La primera estrategia que se describe, y que hemos denominado gestión *Completa-Confirma*, se basa en la solicitud de los atributos mínimos para la realización de una consulta y la posterior confirmación de los

valores que el sistema tenga anotados como poco fiables. El segundo gestor, denominado *Dihana-Oz*, implementa la estrategia definida para la adquisición de un corpus de diálogo mediante la técnica de Mago de Oz, en la que un humano imita el comportamiento del sistema. Esta estrategia se basa en la confirmación de los valores suministrados por el usuario cada vez que el gestor considera que su fiabilidad no es suficiente, utilizándose diferentes tipos de confirmación.

En ambos casos se utilizan medidas de confianza, proporcionadas por el módulo de comprensión, para determinar el nivel de fiabilidad de los datos. El uso de medidas de confianza para el tratamiento de errores en la gestión del diálogo ha sido propuesto por otros autores (San-Segundo et al., 2001), (Torres et al., 2005).

Las dos estrategias se han evaluado a partir de la realización de 120 diálogos y la posterior obtención de una serie de medidas comparativas.

Ambos gestores se engloban en un sistema de diálogo completo, en el que se ha definido un sistema de comunicación entre los diversos módulos. A continuación se mencionan algunas características básicas del resto

* Este trabajo se ha desarrollado en el marco del proyecto DIHANA subvencionado por la CICYT número TIC2002-04103-C03-03.

de módulos relevantes del sistema:

- Módulo de Reconocimiento Automático del Habla: Se ha realizado la integración del reconocedor del habla Sphinx2 (Website Sphinx2, 2001) y el entrenamiento de los modelos acústicos con Sphinx-Train, ambos desarrollados por la Carnegie Mellon University.
- Módulo de comprensión: Para la definición de la semántica de la tarea se utiliza el concepto de frame: cada intervención del usuario genera frames que representan el significado de la intervención (Segarra et al., 2002).
- Módulo de acceso a la base de datos de la aplicación: El sistema utiliza una base de datos relacional en PostGreSQL.
- Módulo de generación de respuestas: La estrategia que sigue el sistema consiste en la utilización de plantillas.
- Módulo de síntesis texto-voz: Se ha realizado la integración de un sintetizador desarrollado expresamente para el proyecto.

El proyecto DIHANA tiene como uno de sus principales propósitos el estudio de la metodología de diseño y el desarrollo de un sistema de diálogo adecuado para el acceso a la información mediante habla espontánea en diferentes entornos. La tarea del proyecto DIHANA es la consulta en castellano a un sistema de información sobre horarios y precios de trenes de largo recorrido. En (Benedí et al., 2004) puede encontrarse información referente al corpus DIHANA, en cuanto a estructura y resto de características relevantes.

2. Descripción de las estrategias de gestión del diálogo

La idea fundamental en la que se basan los modelos de gestión de diálogo desarrollados consiste en la conveniencia de elaborar gestores estándares, que sean independientes de la tarea, portables y con la capacidad de poder incorporar nuevos idiomas de forma sencilla. La utilización de modelos basados en reglas permite desarrollar una solución sencilla y efectiva al problema de la gestión del diálogo, ofreciendo fiabilidad y robustez en la consecución del objetivo.

La implementación práctica de estas ideas se materializa mediante la utilización de

diccionarios y ficheros de definición de la semántica de la tarea, en los que se almacenan los parámetros fundamentales del sistema (frames definidos en la tarea, listado de atributos, atributos obligatorios para cada tipo de consulta...), y en la determinación de un formato estándar para la salida del gestor del diálogo. Todo ello viene apoyado por la definición de un protocolo de comunicación entre los módulos del sistema.

El diseño de los gestores es compatible con algunas de las recomendaciones definidas en el proyecto TRINDI (Website TRINDI, 1998), como la simplificación de las estrategias de gestión, portabilidad, utilización de estructuras de datos comunes, etc.

2.1. Gestor del Diálogo COMPLETA-CONFIRMA

El primer modelo desarrollado para la gestión de diálogo sigue un proceso iterativo en base al siguiente algoritmo:

1. Lee la salida del módulo de comprensión (representación semántica del turno de usuario en base a frames).
2. Realiza la extracción de frames y atributos a partir de la salida de comprensión.
3. Actualiza el historial del diálogo con los datos proporcionados por el usuario.
4. Actualiza el historial del diálogo cuando se realiza una consulta a la base de datos. La totalidad de los resultados obtenidos al realizar una consulta a la base de datos se almacenan en el historial del diálogo.
5. Genera la salida, siguiendo un formato definido, y la suministra al generador de respuestas, para que construya la respuesta asociada.

Como puede observarse en el algoritmo anterior, el historial del diálogo es el elemento fundamental para establecer la estrategia del diálogo y decidir las respuestas del sistema en cada momento, interviniendo de un modo fundamental las medidas de confianza almacenadas junto a los atributos y frames.

La utilización de medidas de confianza surge como solución al problema de la propagación de errores entre los diferentes módulos que componen el sistema. Existen diferentes

estrategias para realizar el tratamiento de estos posibles errores en la información de entrada de los módulos. La técnica que utiliza el gestor se basa en la utilización de confirmaciones explícitas únicamente para aquella información que se ha detectado que tiene poca fiabilidad en el proceso de comprensión.

2.1.1. Historial del Diálogo

El historial de diálogo estructura la información en base a tres bloques diferenciados:

- **Frame actual.** Por frame actual, entendemos el concepto sobre el cual el usuario realiza la consulta. En el caso de detectar más de un frame en la respuesta del módulo de comprensión, se aplican técnicas de simplificación de frames, que seguidamente se detallarán.
- **Atributos.** Cada uno de los frames dispone de un conjunto de atributos obligatorios, considerándose el resto como modificadores de la consulta básica establecida para cada tipo de frame.
- **Medidas de confianza asociadas.** Se almacenan las confianzas asociadas a frames y atributos, incluidas en la representación semántica generada por el módulo de comprensión.

El historial del diálogo permite informar al usuario sobre el estado del diálogo en cualquier instante del mismo, generándose dos elementos adicionales para realizar esta funcionalidad. Estos dos componentes visuales son el historial de turnos del diálogo y la pizarra del diálogo.

El historial de turnos de diálogo informa sobre la totalidad de frases reconocidas del usuario y las correspondientes respuestas generadas por el sistema. La pizarra del diálogo muestra el frame actual y el conjunto de atributos mencionados a lo largo de los diferentes turnos de usuario.

2.1.2. Petición de atributos obligatorios: Iniciativa mixta

Para cada uno de los frames definidos para la tarea, se ha detallado un conjunto de atributos obligatorios, necesarios para poder realizar la consulta a la base de datos. Dicho conjunto aparece en la Figura 1.

Cabe mencionar, que la solicitud de atributos obligatorios contiene la totalidad de atributos mínimos restantes, pudiéndose completar el valor de todos ellos en un único

FRAMES	ATRIBUTOS MÍNIMOS
<i>HORA-SALIDA</i> <i>HORA-LLEGADA</i>	<i>CIUDAD-ORIGEN</i> <i>CIUDAD-DESTINO</i> <i>FECHA-SALIDA</i>
<i>PRECIO</i>	<i>CIUDAD-ORIGEN</i> <i>CIUDAD-DESTINO</i> <i>FECHA-SALIDA</i> <i>CLASE-BILLETE</i>
<i>TIEMPO-RECORRIDO</i>	<i>CIUDAD-ORIGEN</i> <i>CIUDAD-DESTINO</i>
<i>TIPO-TREN</i>	<i>CIUDAD-ORIGEN</i> <i>CIUDAD-DESTINO</i>
<i>SERVICIOS</i>	<i>CLASE-BILLETE</i> <i>TIPO-TREN</i>

Figura 1: Atributos obligatorios definidos para cada uno de los frames del sistema.

turno de usuario. Por último, merece destacarse que la iniciativa del diálogo es mixta, teniéndose en cuenta la información aportada por el usuario independientemente de la pregunta realizada por el sistema.

2.1.3. Simplificación de frames

Dado que el sistema está orientado a completar los datos mínimos para hacer una única consulta a la base de datos, debe tratarse la situación en la que el usuario realiza más de una consulta en el mismo turno. En este caso, se aplica un preprocesado de la representación semántica del turno de usuario, de forma que se provea una entrada más flexible al gestor. Las técnicas de simplificación de frames utilizadas conllevan la pérdida de información con respecto a la secuencia semántica suministrada por el módulo de comprensión. El sistema se vale de dos técnicas:

- **Selección jerárquica:** La secuencia de unidades semánticas suministradas por comprensión se simplifica teniendo en cuenta una jerarquía en la ordenación de los frames y seleccionando aquel que aparezca en un nivel superior. Para realizar la simplificación del número de frames se tiene en cuenta la aparición o no del frame actual en la secuencia de frames detectada en la salida de comprensión. En caso de aparición, se podan el resto de frames de la secuencia.
- **Selección condicionada a las confianzas:** Al igual que en el método anterior, se utiliza la poda de frames para simplificar la representación semántica original. La técnica que se sigue a la

hora de realizar esta poda consiste en ordenar los frames de acuerdo con sus medidas de confianza y descartar los menos fiables. Para ello, se tiene en cuenta que cada uno de los frames puede disponer de más de una medida de confianza asociada, calculándose la media de las medidas suministradas y permaneciendo aquel frame que disponga de una fiabilidad mayor.

La Figura 2 muestra un ejemplo para los tipos de simplificación de frames descritos. Entre corchetes aparecen las dos medidas de confianza utilizadas para cada ítem (medidas sintáctica y semántica proporcionadas por el módulo de comprensión), calculándose la media a la hora de operar con las mismas. En (García et al., 2003) puede encontrarse información sobre el cálculo y significado de las medidas de confianza utilizadas.

Frase de entrada en lenguaje natural
Hola, podría decirme horarios y precios para ir a Bilbao.
Representación semántica en base a frames y atributos
(HORA-SALIDA) [0.53, 0.67] CIUDAD-DESTINO:bilbao [0.77, 0.51] (PRECIO) [0.35, 0.42] CIUDAD-DESTINO:bilbao [0.77, 0.51]
Selección jerárquica (Nota: El frame actual es (PRECIO)).
(PRECIO) [0.35, 0.42] CIUDAD-DESTINO:bilbao [0.77, 0.51]
Selección condicionada a las confianzas
(HORA-SALIDA) [0.53, 0.67] CIUDAD-DESTINO:bilbao [0.77, 0.51]

Figura 2: Simplificación de frames.

2.1.4. Ejemplo de Diálogo

Las siguientes figuras muestran un ejemplo de diálogo adquirido con la herramienta. Las siglas utilizadas se corresponden con *U* (Entrada de Usuario), *S* (Salida del Generador de Respuestas), *P* (Pizarra del Diálogo) y *R* (Representación semántica generada por el módulo de comprensión).

En primer lugar, el diálogo se inicializa con un *prompt* de presentación. El único atributo disponible en la pizarra es *CIUDAD-ORIGEN*, como se observa en la Figura 3.

P: CIUDAD-ORIGEN:valencia S: Bienvenido al sistema de información de trenes. ¿En qué puedo ayudarle?

Figura 3: Inicialización del diálogo.

Tras cada intervención del usuario, se extraen los frames y atributos de la secuencia semántica suministrada por comprensión. En caso de no detectar ningún frame, el generador de respuestas devuelve una respuesta estándar (Figura 4).

U: Quiero ir a Bilbao. R: () [0.81, 0.77] CIUDAD-DESTINO:bilbao [0.33, 0.36] P: CIUDAD-ORIGEN:valencia CIUDAD-DESTINO:bilbao S: Por favor, indíqueme que tipo de información necesita.
--

Figura 4: Solicitud del tipo de consulta.

Una vez el gestor ha detectado un frame en la interpretación semántica del turno de usuario, se solicitan los atributos obligatorios para dicho frame, y de cuyo valor no se disponga en el estado actual del diálogo, como muestra la Figura 5.

U: Quisiera horarios. R: (HORA-SALIDA) [0.71, 0.67] P: CIUDAD-ORIGEN:valencia CIUDAD-DESTINO:bilbao S: Dígame la fecha de salida.

Figura 5: Solicitud de atributos obligatorios.

Tras completar la totalidad de atributos obligatorios, se realiza una confirmación explícita, aplicable únicamente a aquellos atributos cuya medida de confianza asociada sea inferior al umbral predeterminado. Para el ejemplo citado, se observa que el valor de la media de las medidas de confianza de los atributos *CIUDAD-DESTINO* y *FECHA-SALIDA* es inferior al umbral 0.5, fijado para el ejemplo (Figura 6).

En caso de recibir una confirmación positiva por parte del usuario, se realiza la consulta a la base de datos, actualizando el historial del diálogo y mostrando el resultado de la consulta conjuntamente con una respuesta estándar, como se observa en la Figura 7.

U: Me gustaría viajar el lunes.
 R: () [0.11, 0.45]
 FECHA-SALIDA:18-04-2005/18-04-2005
 [0.43, 0.32]
 P: CIUDAD-ORIGEN:valencia
 CIUDAD-DESTINO:bilbao
 FECHA-SALIDA:18-04-2005/18-04-2005
 S: ¿Desea horarios a Bilbao saliendo el día
 18-05-2005?

Figura 6: Confirmación explícita basada en las medidas de confianza.

U: Sí.
 R: (AFIRMACIÓN) [0.61, 0.85]
 P: CIUDAD-ORIGEN:valencia
 CIUDAD-DESTINO:bilbao
 FECHA-SALIDA:18-04-2005/18-04-2005
 S: Hay varios trenes. El primero sale a las 08:54
 y el último a las 23:45 . ¿Desea algo más?

Figura 7: Consulta a la base de datos.

El diálogo finaliza tras detectar un frame *CIERRE* en la representación semántica generada por comprensión (Figura 8).

U: No gracias. Adiós.
 R: (CIERRE) [0.61, 0.55]
 P: CIUDAD-ORIGEN:valencia
 CIUDAD-DESTINO:bilbao
 FECHA-SALIDA:18-04-2005/18-04-2005
 S: Gracias por utilizar el sistema de información de trenes. Adiós.

Figura 8: Finalización del diálogo.

2.2. Gestor del Diálogo DIHANA-OZ

La adquisición de un corpus específico de diálogo usuario-sistema plantea una gran dificultad; ya que, para que esta adquisición se realice de una manera natural se precisa un sistema de diálogo que funcione eficientemente, pero para desarrollar un sistema de diálogo eficiente es necesario una gran cantidad de datos (diálogos naturales) para el entrenamiento de sus modelos. De ahí, la utilización de la técnica de Mago de Oz, en la que una persona asume el papel del gestor de diálogo y ayuda al usuario a obtener respuestas a sus consultas siguiendo una estrategia definida.

En este trabajo se ha desarrollado un segundo gestor, Dihana-Oz, que implementa la estrategia que se elaboró para la adquisición con la técnica de Mago de Oz del corpus DIHANA. En esta estrategia, dados el estado de la pizarra y la información proporcionada por el usuario en el turno correspondiente, el gestor interacciona con el usuario en base a los niveles de confianza suministrados por el sistema.

Si todos los datos de la pizarra disponen de una medida de confianza asociada mayor que el umbral fijado (estado seguro), el gestor elige una de las tres interacciones siguientes:

- Confirmación Implícita y Consulta a la base de datos si dispone de frame y, al menos, de los valores de sus atributos mínimos. (Ej. *Le consulto horarios de trenes con salida en Bilbao destino Santander en primera clase.*)
- Completar si no dispone de valor para *FRAME_ACTUAL* y/o alguno de sus atributos mínimos (sin valor por defecto).
- Confirmación Mixta. (Ej. *Quiere horarios de trenes a Vigo , ¿desde Pontevedra?*). En las confirmaciones mixtas se incorporan referencias no únicamente a valores marcados como poco fiables (Vigo en el ejemplo anterior), sino también a uno o más conceptos con suficiente fiabilidad (solicitud de horarios y Pontevedra), favoreciéndose una mayor naturalidad en el diálogo. Se realiza sobre el 30 % de turnos seguros en lugar de una Confirmación Implícita-Consulta.

Si el estado es inseguro (aquel en el que uno o más datos de la pizarra poseen una confianza menor que el umbral), el gestor selecciona una de las dos interacciones siguientes:

- Confirmación Explícita del primero de los ítem inciertos que aparezca en la pizarra. (Ej. *¿Quiere viajar a Madrid?*)
- Confirmación Mixta para darle naturalidad al diálogo. Se realiza sobre el 30 % de turnos de diálogo inseguros en lugar de una Confirmación Explícita.

En esta estrategia cumplen un papel esencial la pizarra del diálogo y los métodos de simplificación de frames, descritos previamente. Asimismo, el almacenamiento de la

última respuesta del sistema permite resolver las confirmaciones de frames y atributos.

2.2.1. Ejemplo de Diálogo

Las siguientes figuras muestran un ejemplo de diálogo adquirido con la herramienta. Las siglas utilizadas se corresponden con las citadas para el modelo de gestión anterior.

En primer lugar, el diálogo se inicializa con un *prompt* de presentación. El único atributo disponible en la pizarra es *CIUDAD-ORIGEN*, como se observa en la Figura 9.

```
P: CIUDAD-ORIGEN:valencia
S: Bienvenido al sistema de información de trenes. ¿En qué puedo ayudarle?
```

Figura 9: Inicialización del diálogo.

En caso de no detectar ningún frame tras la intervención del usuario, el gestor de diálogo realiza la petición del tipo de consulta (Figura 10).

```
U: Quiero ir a Bilbao.
R: ( ) [0.81, 0.77]
   CIUDAD-DESTINO:bilbao [0.33, 0.36]
P: CIUDAD-ORIGEN:valencia
   CIUDAD-DESTINO:bilbao
S: ¿Quiere saber horarios?
```

Figura 10: Solicitud del tipo de consulta.

Una vez el gestor confirma un frame en la interpretación semántica del turno de usuario, comprueba si existen en la pizarra valores con una confianza inferior al umbral prefijado. Para el ejemplo, se observa que el valor de la media de las medidas de confianza del atributo *CIUDAD-DESTINO* es inferior al umbral 0.5, generándose una confirmación del valor correspondiente (Figura 11).

```
U: Sí.
R: (AFIRMACION) [0.61,0.85]
P: CIUDAD-ORIGEN:valencia
   CIUDAD-DESTINO:bilbao
S: ¿Desea ir a bilbao?
```

Figura 11: Confirmación de los atributos.

Tras recibir una confirmación positiva, y comprobar que no existen valores en la pizarra con baja fiabilidad, se solicitan los atributos obligatorios para completar la consulta solicitada, como muestra la Figura 12.

```
U: Sí.
R: (HORA-SALIDA) [0.71, 0.67]
P: CIUDAD-ORIGEN:valencia
   CIUDAD-DESTINO:bilbao
S: ¿Qué día desea salir?
```

Figura 12: Solicitud de atributos obligatorios.

Tras incorporar a la pizarra atributos cuya medida de confianza asociada sea inferior al umbral predeterminado, el gestor selecciona de nuevo uno de los posibles estados de confirmación (Figura 13).

```
U: Me gustaría viajar el cuatro de mayo.
R: ( ) [0.34,0.43]
   FECHA-SALIDA:04-05-2005/04-05-2005
   [0.33, 0.46]
P: CIUDAD-ORIGEN:valencia
   CIUDAD-DESTINO:bilbao
   FECHA-SALIDA:04-05-2005/04-05-2005
S: ¿Desea salir el 04-05-2005?
```

Figura 13: Confirmación de los atributos.

En caso de recibir una confirmación positiva por parte del usuario, se realiza la consulta a la base de datos y se genera una respuesta que muestre el resultado de la misma, tal y como se observa en la Figura 14.

```
U: Sí.
R: (AFIRMACIÓN) [0.61, 0.85]
P: CIUDAD-ORIGEN:valencia
   CIUDAD-DESTINO:bilbao
   FECHA-SALIDA:04-05-2005/04-05-2005
S: Un momento por favor. Hay varios trenes.
   El primero sale a las 08:54 y el último a las
   23:45 . ¿Desea algo más?
```

Figura 14: Consulta a la base de datos.

El diálogo finaliza tras detectar un frame *CIERRE* en la representación semántica generada por comprensión (Figura 15).

La observación de los ejemplos expuestos para las dos estrategias de diálogo muestra la importancia de conocer el tipo de consulta que desea realizar el usuario antes de realizar otras operaciones. Sin embargo, mientras la primera de las técnicas espera a completar el conjunto de atributos mínimos antes de confirmar valores de la pizarra con baja confianza, la segunda de ellas confirma dichos

U: No gracias. Adiós. R: (CIERRE) [0.61, 0.55] P: CIUDAD-ORIGEN:valencia CIUDAD-DESTINO:bilbao FECHA-SALIDA:04-03-2005/04-03-2005 S: Gracias por utilizar el sistema de información de trenes. Adiós.
--

Figura 15: Finalización del diálogo.

valores nada más realizar su detección. Este hecho favorece que el número medio de turnos de sistema y de confirmaciones explícitas sea mayor en la segunda estrategia.

3. Técnicas de generación de respuestas

El generador de respuestas realiza la traducción de las representaciones semánticas de los turnos de usuario a frases en castellano, basándose en plantillas y combinando reglas.

Al igual que en el módulo de comprensión, la entrada del generador de respuesta está compuesta por frames y atributos, con medidas de confianza asociadas, lo que permite la generación de respuestas en lenguaje natural detalladas, en las que los atributos se mencionan o no dependiendo de su confianza correspondiente.

La técnica utilizada consiste en disponer de una serie de plantillas asignadas a cada uno de los diferentes tipos de frame, en las que aparecen reflejados los nombres de los atributos. Éstos se sustituyen por los valores de los mismos, obtenidos del historial del diálogo, a la hora de mostrar la respuesta al usuario.

4. Evaluación de los modelos de gestión del diálogo

Se ha evaluado el comportamiento de los gestores de diálogo presentados fijando un umbral de confianza (0.5) para valorar la fiabilidad de la información y utilizando un corpus de 15 escenarios sobre consultas de horarios y precios en viajes de ida o de ida y vuelta. Se llevaron a cabo un total de 120 diálogos, 60 para cada una de las estrategias, desarrollados por 6 usuarios familiarizados con el sistema. Para la evaluación de las estrategias se ha tenido en cuenta las siguientes medidas:

1. Porcentaje de diálogos en los que se alcanzó el objetivo (% éxito). Para cada

escenario el usuario debe obtener información correspondiente a una o varias consultas, dependiendo del éxito del sistema del hecho que el gestor suministre correctamente la totalidad de la información.

2. Número medio de turnos (nT) de sistema por diálogo.
3. Ratio de confirmaciones (% confirm). Este valor se obtiene contando el número de turnos correspondientes a confirmaciones explícitas, nCT, con respecto al total de turnos de sistema, es decir, nCT/nT .
4. Número medio de errores corregidos por diálogo (nCE). Se trata de la media de errores detectados y corregidos por el gestor, habiéndose contemplado como errores únicamente aquellos que modifican el valor de los atributos (y que pudieran causar el fallo del diálogo).
5. Número medio de errores no corregidos por diálogo (nNCE). Se consideran únicamente los errores que modifican los valores de los atributos.
6. Tasa de corrección de errores (% correct). Se trata del porcentaje de corrección de errores, es decir, $nCE/(nCE + nNCE)$.

Los resultados de la evaluación (Figura 16) muestran en primer lugar la eficacia de los gestores de diálogo desarrollados, alcanzando un porcentaje de éxito del 93,45% para la primera de las estrategias y un 97,33% para la segunda de ellas.

La estrategia Completa-Confirma, basada en la petición de los atributos mínimos para realizar una consulta determinada y en la confirmación de los atributos con menor fiabilidad una vez están disponibles todos ellos, posee una media de turnos de sistema inferior a la estrategia Dihana-Oz, basada en la confirmación de los atributos en el instante en el que se incorporan a la pizarra del diálogo. Este es el coste a pagar de cara a obtener una mayor tasa de detección y corrección de los errores introducidos durante el diálogo.

Estos factores posibilitan que estrategia Dihana-Oz, a costa de incrementar la media de turnos de sistema y de confirmaciones explícitas por diálogo, logre un porcentaje de

	% éxito	nT	% confirm	nCE	nNCE	% correct
Gestor Completa-Confirma	93,45	6,61	21,39	0,62	0,46	57
Gestor Dihana-Oz	97,33	9,03	35,04	0,86	0,11	89

Figura 16: Resultados de la evaluación realizada para los modelos de gestión desarrollados.

éxito y de corrección de errores mayor que la estrategia Completa-Confirma.

Este hecho se ve favorecido por la utilización de confirmaciones mixtas, que además de dotar de mayor naturalidad al diálogo, posibilitan al usuario un mayor grado de conocimiento de los atributos y valores que se introducirán en la sentencia de consulta a la base de datos, y que marcarán la validez de la información obtenida con respecto a la consulta requerida por el usuario. La utilización de confirmaciones mixtas, incluso cuando la fiabilidad de la información supera el umbral prefijado, permite eliminar o rectificar de la consulta aquellos valores marcados como fiables de forma errónea, y que son los principales causantes del fracaso del diálogo para estrategia Completa-Confirma.

5. Conclusiones

En este trabajo se han presentado dos estrategias para realizar la gestión de diálogo basadas en reglas. Ambas comparten similitudes en la forma de operar, destacando la utilización de medidas de confianza para el tratamiento y detección de errores, la definición de diferentes tipos de confirmaciones y el empleo del historial del diálogo como elemento fundamental para la toma de decisiones. La principal diferencia radica en las técnicas utilizadas para la confirmación de los valores con baja fiabilidad. La evaluación llevada a cabo muestra la eficacia de los modelos desarrollados de cara a obtener los objetivos marcados en el diálogo.

Paralelamente, cabe destacar los esfuerzos realizados para dotar a ambas estrategias de un grado de estandarización que permita su validez independientemente de la tarea del sistema.

Bibliografía

- Benedí, J.M., Varona A., Lleida E., y alt. 2004. DIHANA: Sistema de diálogo para el acceso a la información en habla espontánea en diferentes entornos. *Actas de las III Jornadas en Tecnología del Habla, Valencia (España)*, páginas 141–146.
- García, F., Hurtado L.F., Sanchis E., y Segarra E. 2003. The incorporation of Confidence Measures to Language Understanding. *International Conference on Text Speech and Dialogue (TSD 2003), Ceské Budejovice (Czech Republic)*, Lecture Notes in Artificial Intelligence series 2807:165–172.
- Lamel, L., Rosset S., Gauvain J.L., Bennacef S., M. Garnier-Rizet, y Prouts B. 2000. The LIMSI ARISE System. *Speech Communication*, 4(31):339–353.
- Rudnický, A., Thayer E., Constantinides P., Tchou C., Shern R., Lenzo K., Xu W., y Oh A. 1999. Creating natural dialogs in the Carnegie Mellon Communicator system. *Proceedings of Eurospeech*, 1(4):1531–1534.
- San-Segundo, R., Pellom B., Hacıoglu K., Ward W., y Pardo J. 2001. Confidence measures for spoken dialogue systems. *Proc. ICASSP. Salt Lake City, USA*.
- Segarra, E., Sanchis E., García F., y Hurtado L.F. 2002. Extracting semantic information through automatic learning techniques. *International Journal of Pattern Recognition and Artificial Intelligence*, 16(3):301–307.
- Torres, F., Hurtado L.F., García F., Sanchis E., y Segarra E. 2005. Error handling in a stochastic dialog system through confidence measures. *Speech Communication*, 45:211–229.
- Website Sphinx2. 2001. Disponible en: www.speech.cs.cmu.edu/sphinx/doc/sphinx2.html .
- Website TRINDI. 1998. (Task Oriented Instructional Dialogue). Disponible en: www.ling.gu.se/projekt/trindi/ .
- Zue, V., Seneff S., Glass J., Polifroni J., Pao C., Hazen T.J., y Hetherington L. 2000. JUPITER: A Telephone-Based Conversational Interface for Weather Information. *IEEE Transactions on Speech and Audio Processing*, 8(1), January.