

Diseño de Gestores de Diálogo

Rubén San Segundo Hernández

(5 Febrero 2004)



Servidores Vocales Interactivos

Las prestaciones conseguidas en Tecnología del Habla en los últimos años están permitiendo el desarrollo de sistemas que funcionan en Condiciones Reales para Dominios Restringidos

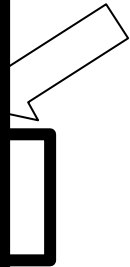
- ☞ **Información Meteorológica:** JÚPITER (Zue et al, 2000)
- ☞ **Información y Reservas de billetes de tren:** RailTel (Lamel et al, 1997) ARISE (Lamel et al, 2000; Baggia et al, 2000)
- ☞ **Información y Reservas de billetes de avión:** DARPA Communicator (Ward y Pellom, 1999; Pellom et al, 2000; Rudnicky et al, 2000)
- ☞ **Información Telefónica:** IDAS (Lehtinen et al, 2000; Córdoba et al, 2001) Philips (Schrâmm et al, 2000)
- ☞ **Información de Calificaciones:** STACC (Rubio et al, 1997) TADE (Casas, 1999)

Diagrama

Gestión del Diálogo



- Descripción de los Modelos de Diálogo utilizados en la actualidad y Ejemplos
- Ideas sobre Evaluación de Diálogo.
- Propuesta de una Metodología para el diseño de Gestores de Diálogo
- Aplicación de las Medidas de Confianza para el diseño eficiente de las estrategias de confirmación
- Modelado de usuario



Gestión del Diálogo (I)

Tipos de Gestores según el Agente que lleva la iniciativa:

📄 **Gestor de diálogo de Iniciativa del Sistema** (Sutton et al, 1998; San-Segundo et al, 1999; Baggia et al, 2000; Córdoba et al, 2000; San-Segundo et al, 2000c; Córdoba et al, 2001)

📄 **Gestor de diálogo de Iniciativa Mixta** (Zue, 1997b; Ward, 1999; Rudnicky, 1999; Rosset, 1999; Zue, 2000; Lamel, 2000; Schrämm, 2000; Pellom, 2000; Rudnicky, 2000)

- Flexibilidad en la secuencia de Datos.
- Flexibilidad en la secuencia de Objetivos.

📄 **Gestor de diálogo de Iniciativa Completa del Usuario** (No posible en la actualidad: la posibilidad de que existan ciertos errores obliga a que el sistema deba tomar la iniciativa en determinadas situaciones)

Gestión del Diálogo (II)

Estrategias de Diseño del Gestor de Diálogo:

- **Diseño por Intuición:** definición del diálogo en función de la experiencia del desarrollador.
- **Diseño por Observación:** definición del diálogo mediante el análisis de conversaciones usuario-operador (humano) en un servicio análogo al que se desea automatizar.
- **Diseño por Simulación (Mago de Oz):** simular parte de la funcionalidad del sistema que se desea automatizar para aprender el comportamiento que los usuarios tendrán ante este tipo de sistemas y poder hacer un mejor diseño.
- **Diseño por Prototipado y Mejora Iterativa:** desarrollo rápido de un primer prototipo (diseño por intuición) y proceso iterativo de prueba con usuarios reales y mejora del sistema.
- **Diseño por Aprendizaje sobre un Modelo Estocástico:** en este caso se pretende la definición de un modelo estocástico cuyos parámetros son estimados a partir de diálogo etiquetados.

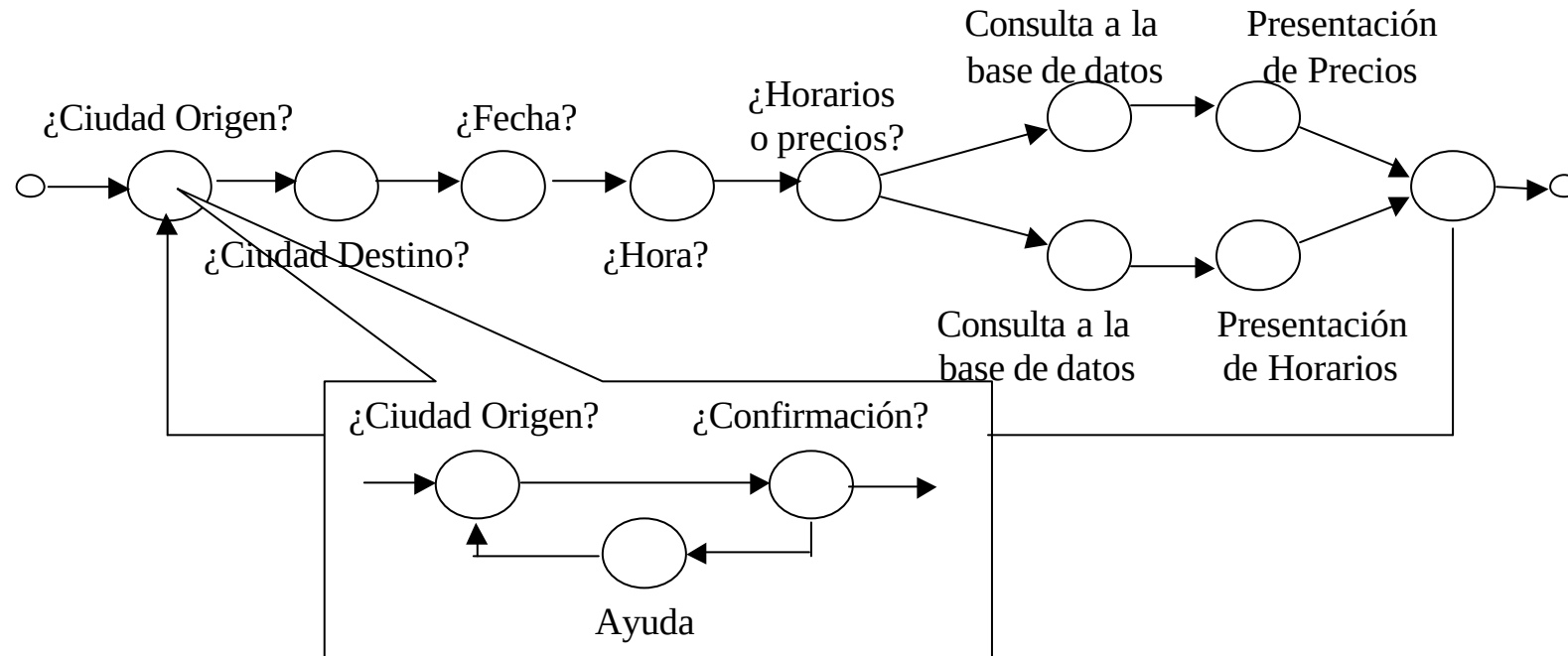
Modelo de Diálogo

Representación formal del proceso de diálogo entre usuario y sistema

Elementos del Modelo

- ➔ **Turnos de diálogo:** estructuras de datos en las que se almacena información relevante para cada turno del diálogo.
- ➔ **Historia y Estado del Diálogo:** estructura de datos en la que se almacena el estado actual del diálogo (datos obtenidos, datos confirmados, ...) y la historia del diálogo (secuencia de turnos ejecutados).
- ➔ **Algoritmo de ejecución de los turnos:** algoritmo que decide la secuencia de ejecución de los turnos que forman el diálogo.

Modelo basado en Autómatas Finitos



- ➔ **Turnos de diálogo:** estructuras en cada nodo del autómata.
- ➔ **Historia y Estado del Diálogo:** secuencia de nodos recorrida y nodo actual.
- ➔ **Algoritmo de ejecución de los turnos:** secuencia de nodos/turnos definida por el flujo del autómata.

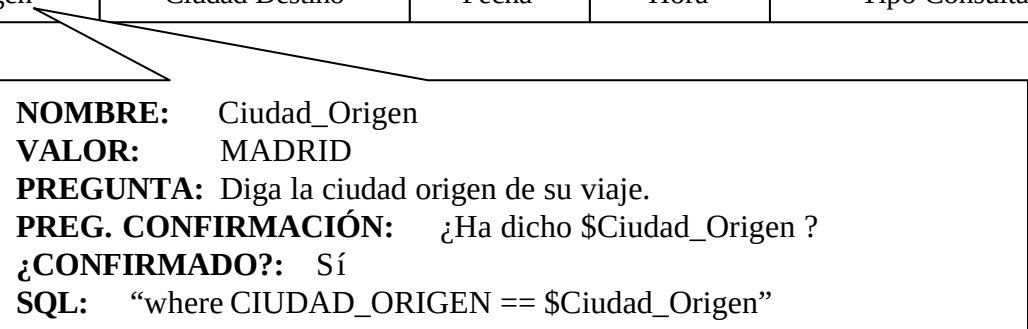
Modelo basado en Plantillas

OBJETIVO PRECIOS

Ciudad Origen	Ciudad Destino	Fecha	Hora	Tipo Consulta: PRECIOS
---------------	----------------	-------	------	------------------------

OBJETIVO HORARIOS

Ciudad Origen	Ciudad Destino	Fecha	Hora	Tipo Consulta: HORARIOS
---------------	----------------	-------	------	-------------------------



NOMBRE: Ciudad_Origen
VALOR: MADRID
PREGUNTA: Diga la ciudad origen de su viaje.
PREG. CONFIRMACIÓN: ¿Ha dicho \$Ciudad_Origen ?
¿CONFIRMADO?: Sí
SQL: “where CIUDAD_ORIGEN == \$Ciudad_Origen”

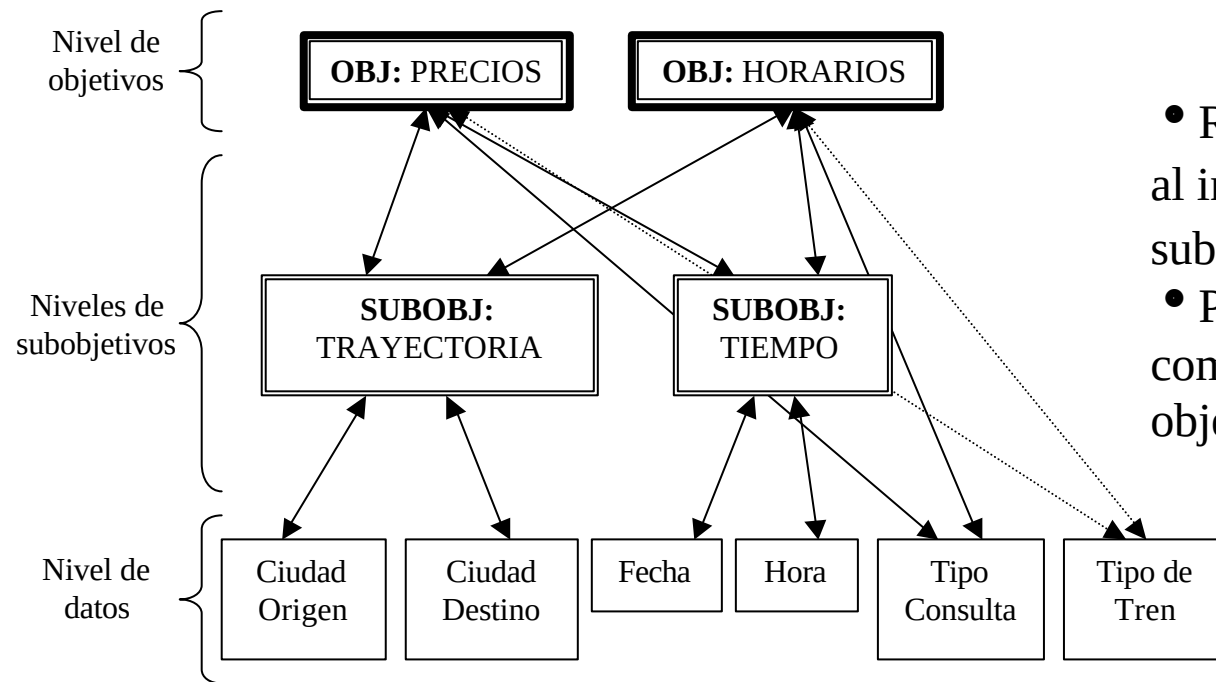
➔ **Turnos de diálogo:** se van generando de forma dinámica hasta completar los datos necesarios para satisfacer el objetivo/plantilla solicitado por el usuario.

➔ **Historia:** secuencia de turnos generados hasta ahora. **Estado del Diálogo:** queda reflejado según el contenido de los datos de cada plantilla.

➔ **Algoritmo:** los turnos se crean según las siguientes órdenes con esta prioridad:

- Clarificación o confirmación del último dato introducido.
- Si tenemos una plantilla rellena se consulta a la base de datos.
- Preguntar al usuario información que falta para completar la plantilla.

Modelo basado en Árboles de Objetivos (I)



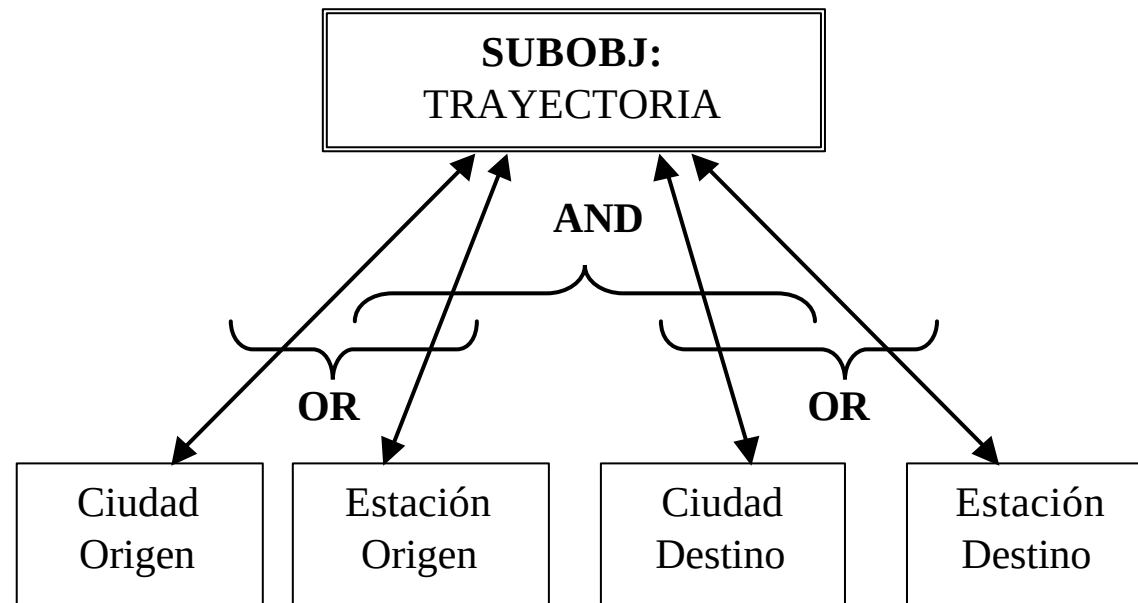
- Representación más flexible al introducir el nivel de subobjetivos.
- Permite modelar mejor la compartición de datos entre objetivos.

➔ **Turnos de diálogo:** se van generando de forma dinámica.

➔ **Historia:** secuencia de turnos generados hasta ahora. **Estado del Diálogo:** queda reflejado según el contenido de los datos del árbol.

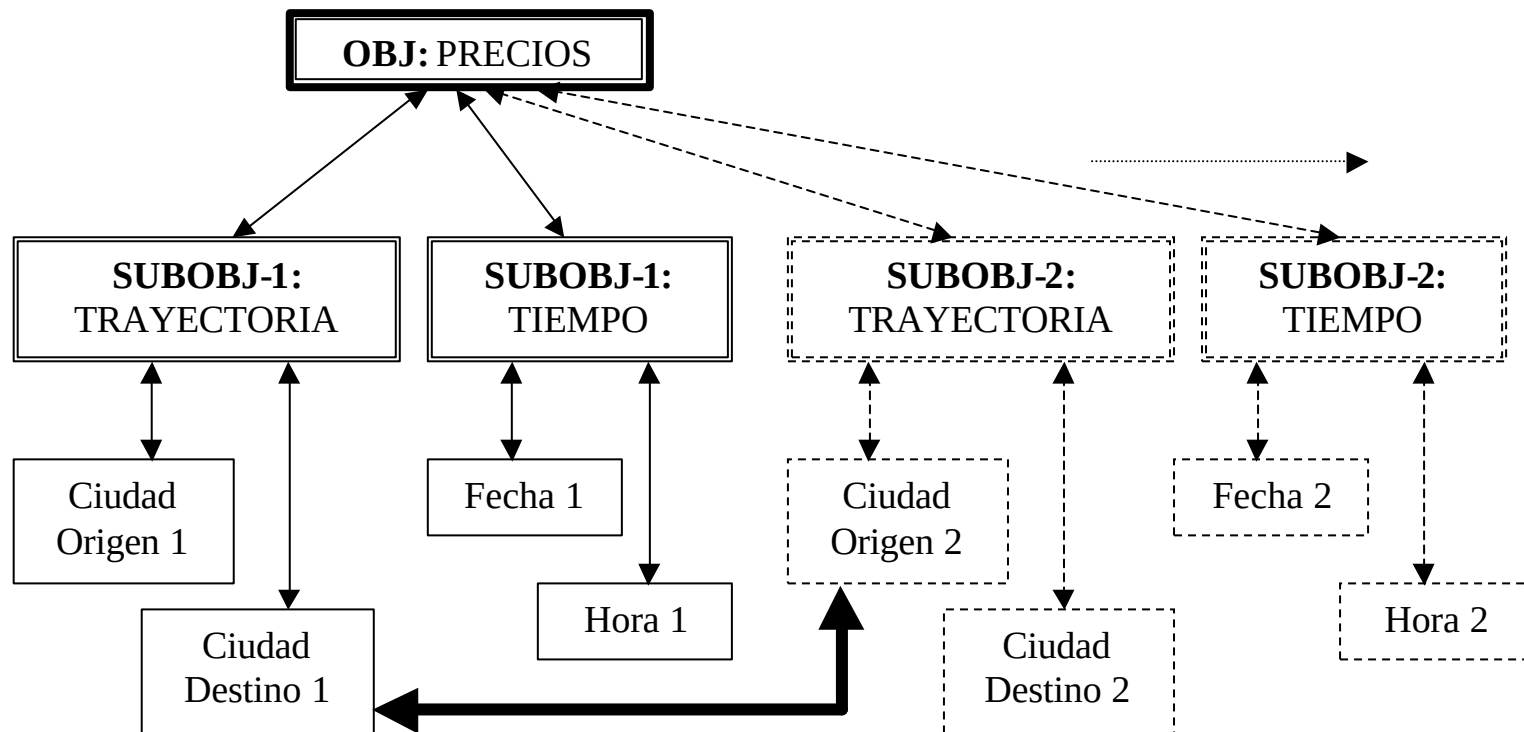
➔ **Algoritmo:** los turnos se crean de igual forma que en el modelo basado en plantillas.

Modelo basado en Árboles de Objetivos (II) (Opcionalidad)



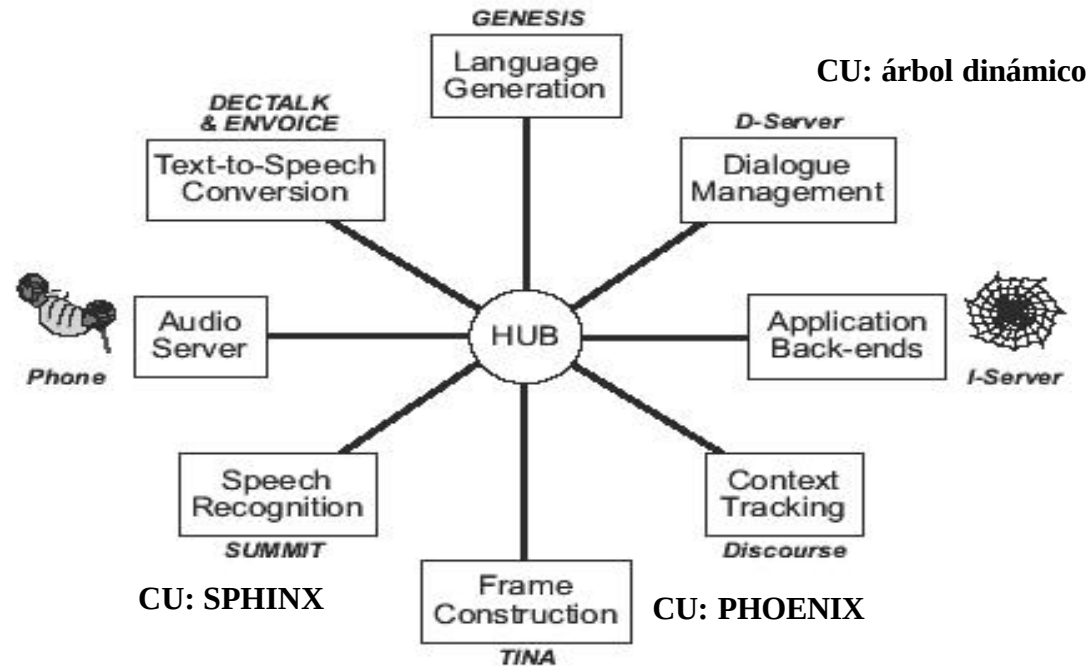
- ➔ Relación entre la información almacenada en los datos y en los subobjetivos.
- ➔ Estas relaciones permiten ampliar el criterio de opcionalidad.

Modelo de Árboles de Objetivos Dinámicos



- Pueden aparecer nuevos datos o subobjetivos, necesarios para completar un objetivo determinado.
- En otros casos, algunos de los datos se rellenan automáticamente y no hace falta preguntarles

EJEMPLOS (I): DARPA Communicator

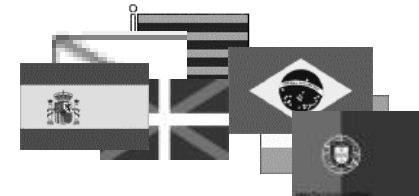


- <http://fofoca.mitre.org> .
- SISTEMA DE COLORADO: <http://communicator.colorado.edu>

EJEMPLOS (II): TELEFÓNICA

GESTIÓN DEL DIÁLOGO

- **VoiceXML o SALT.**
 - Mayor facilidad de desarrollo a costa de perder flexibilidad.
 - Intérprete VoiceXML propietario (*E-moción Voz*).
- **ÁGORA**
 - Gestión basada en la representación de la tarea: árbol dinámico
 - Iniciativa mixta.
 - Plataforma Multiservicio y Multilingüe.
 - Técnicas avanzadas de recogida de datos: Subdiálogos expertos: fechas, DNIs, Tarjeta de Crédito.
 - Técnicas de cambio de contexto o servicio.
 - Técnicas de recuperación de datos.
 - Generador de frases de salida según el estado del diálogo.
 - Servicios: 1003 automático, E-MATTER.



EJEMPLOS (III): TELEFÓNICA

SERVICIO 1003



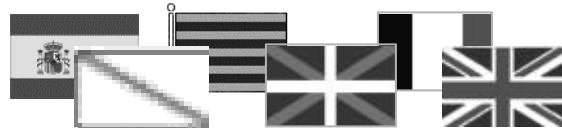
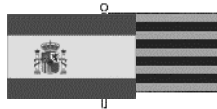
- **CARACTERÍSTICAS.**
 - Información de teléfonos Particulares y de Empresa.
 - Sistema ÁGORA.
 - Voz femenina corporativa.
 - Diseño Vocabularios. Cobertura vs. Tasa.
 - Servicio desde Cabinas Telefónicas.
- **PRINCIPALES PROBLEMAS**
 - Variedad de expresiones para las empresas.
 - Entorno Ruidoso.
 - Pitidos de las cabinas.
 - Variedad de Acentos del Castellano.

EJEMPLOS (IV): TELEFÓNICA

SERVICIO E-MATTER

- **CARACTERÍSTICAS.**

- Proyecto Europeo (IST-1999-21042).
- Acceso al correo electrónico a través de voz.
- Corrección de Texto.
- Identificación de idioma según el texto.
- Reconocimiento y Síntesis Multilingüe.



- Combinación de voces (masculina y femeninas) para la lectura de Mails (cabecera, fecha, hora y contenido).

- **PRINCIPALES PROBLEMAS.**

- Posibilidad de interrumpir los mails.

EJEMPLOS (V): TELEFÓNICA

SERVICIO EMOCIÓN-VOZ



EJEMPLOS (II): TADE (GTH)

Telephone Application Development Enviroment

TIPOS DE FUNCIONES

SECCION_VARIABLES:

n_max_duracion_llamada = 40; /*En minutos */

s_variable_de_tipo_string = "...";

SECCION_ERRORES:

TRATAMIENTO NO_RECONOCIDO sintetizar("Opción errónea"); reintentar;
FIN_TRATAMIENTO

TRATAMIENTO LONGITUD_CORTA sintetizar("Cadena corta."); reintentar;
FIN_TRATAMIENTO

TRATAMIENTO TIMEOUT sintetizar("Vuelva a repetir"); reintentar;

...

SECCION_SUBROUTINAS:

SUBROUTINA Bienvenida(s_frase_a_pronunciar);

sintetizar(s_frase_a_pronunciar);

retornar:

FIN_SUBROUTINA

SECCION_APLICACION:

INICIO:/* Comienzo de la aplicación */

esperar_llamada();

 s_variable_de_tipo_string = "hola a todos";

gosub Bienvenida(s_variable_de_tipo_string);

colgar();

FIN: /* Fin de la aplicación */

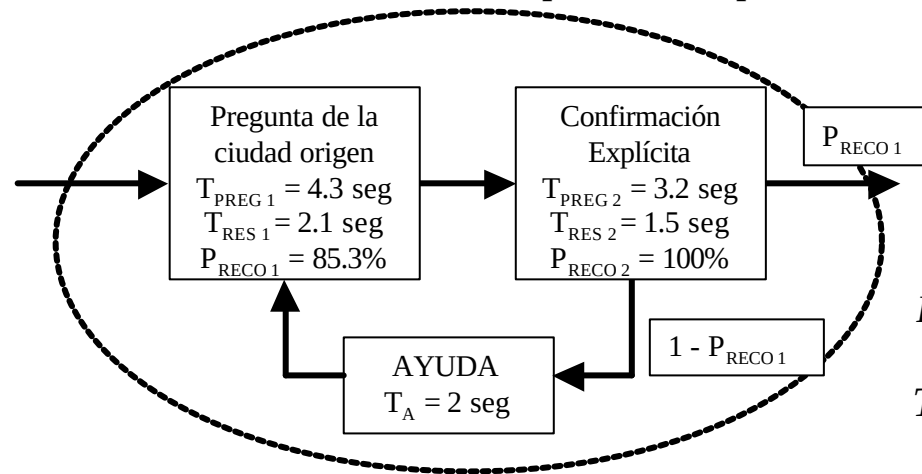
TADE: habla aislada

TADE: autómata de estados finitos

- **Gestión de la línea telefónica** (colgar, descolgar, marcar o esperar llamada)
- **Sentencias de voz** (reconocimiento, conversión texto a voz, reproducción y grabación).
- **Acceso a bases de datos** (abrir/cerrar bases de datos y ejecutar consultas).
- **Envío de correo electrónico**
- **Gestión de directorios y archivos**, el manejo de cadenas y operaciones aritméticas básicas.
- Intervención de un **operador humano**.

EVALUACIÓN (I)

1. Tiempo medio que se tarda en ejecutar ese estado. $T_{M.EST}$. Es el tiempo consumido en media para que la interacción usuario–sistema ejecute el estado correspondiente.
2. Probabilidad de fracaso del diálogo en ese estado. P_{ERROR} . Es la probabilidad de que el usuario no sea capaz de completar ese estado y acabe colgando.



$$P_{ERROR} = 1 - \sum_{n=1}^{n=N} \{(1 - P_{REC1})^{n-1} P_{REC1}\}$$

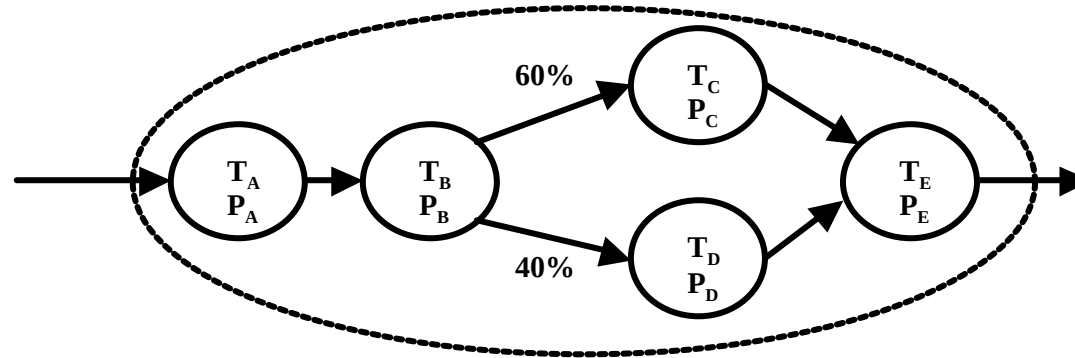
$$T_{MEST} = \sum_{n=1}^{n=N} \{(1 - P_{REC1})^{n-1} P_{REC1} [n(T_1 + T_2) + (n-1)(T_1 + T_2 + T_A)]\}$$

T_{PREG1} : Tiempo consumido para realizar la pregunta de la ciudad origen.
 T_{RES1} : Tiempo consumido en la respuesta de la ciudad origen.
 P_{RECO1} : Tasa de reconocimiento de la ciudad origen.
 T_{PREG2} : Tiempo consumido para realizar la pregunta de confirmación
 T_{RES2} : Tiempo consumido en la respuesta Sí/No.
 P_{RECO2} : Tasa de reconocimiento de Sí/No suponemos muy elevada (100%)
 T_A : Tiempo consumido en dar la ayuda.

$$T_1 = T_{PREG1} + T_{RES1}$$

$$T_2 = T_{PREG2} + T_{RES2}$$

EVALUACIÓN (II)



$$T_{MESTOTAL} = T_A + T_B + 0.6T_C + 0.4T_D + T_E$$

$$P_{ERROROTAL} = 1 - [(1 - P_A)(1 - P_B)\{0.6(1 - P_C) + 0.4(1 - P_D)\}(1 - P_E)]$$

EJEMPLOS:

1. Probabilidad 80% y tiempo 8 minutos.
2. Probabilidad 60% y tiempo 4 minutos.

Metodología de Diseño

- Etapas de la Metodología de Diseño:

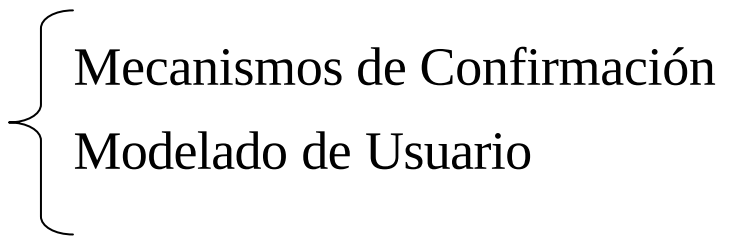
1. Análisis de la Base de Datos

2. Diseño por Intuición (Brain Storming)

3. Diseño por Observación

4. Diseño por Simulación

5. Diseño por Mejora Iterativa



Mecanismos de Confirmación
Modelado de Usuario

- Aplicación a un sistema de información ferroviaria

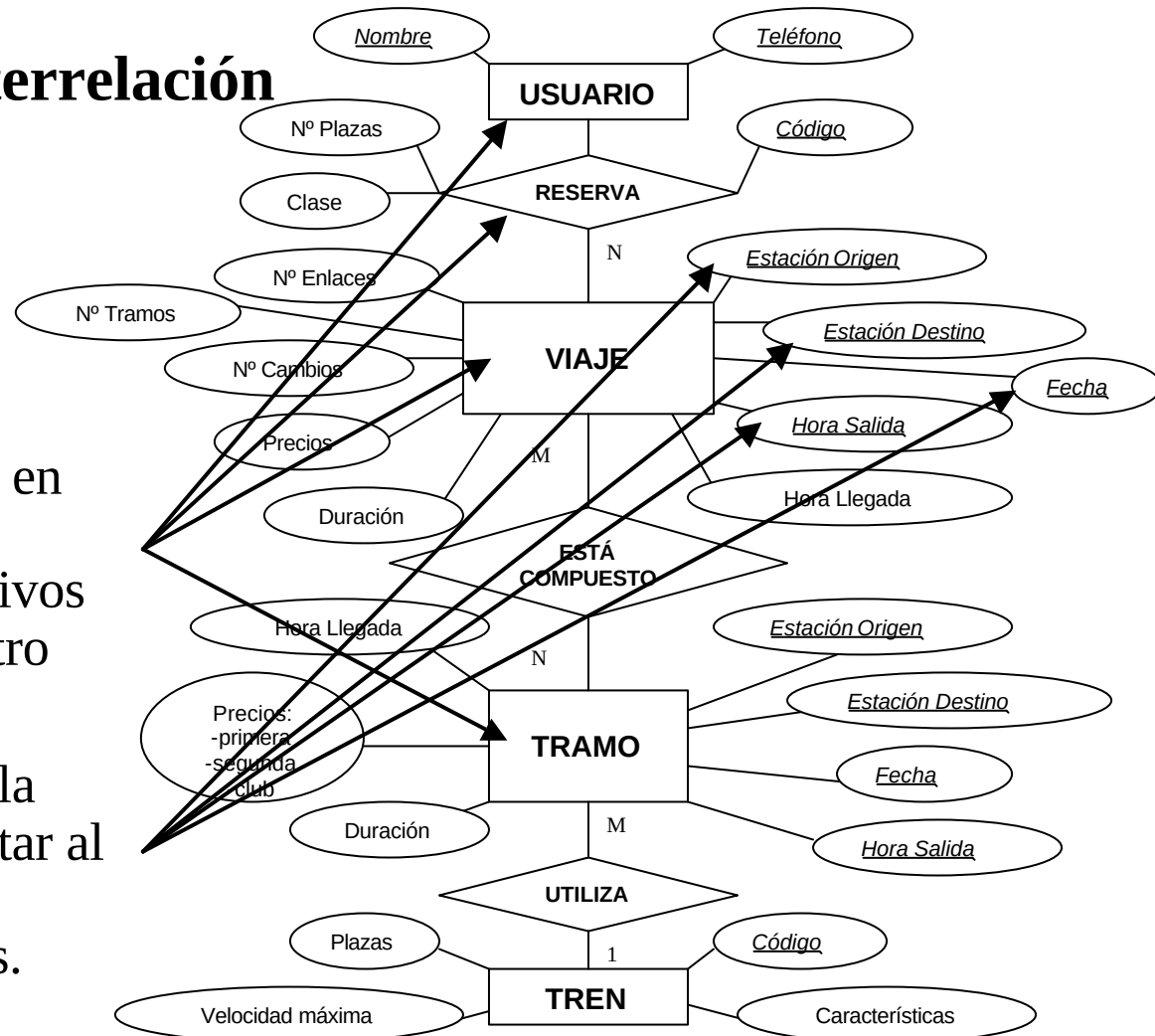
Análisis de la Base de Datos

Diagrama Entidad-Interrelación

- **Conjuntos Entidad.**
- **Atributos.**
- **Claves.**
- **Conjuntos Asociación.**

⇒ La información contenida en los **Conjuntos Entidad y Asociación** definirá los objetivos que podemos definir en nuestro servicio

⇒ Los atributos que forman la **Clave** son los datos a preguntar al usuario para acceder a la información de los Conjuntos.



Diseño por Intuición (Brain-Storming)

OBJETIVO: analizar el diagrama obtenido en la etapa anterior y proponer alternativas acerca de ...

“Los posibles objetivos que puede satisfacer el sistema”

“Secuencia de objetivos para formar el servicio”

“Datos necesarios para satisfacer cada objetivo”

“Formas de especificar un dato por parte del usuario”

“Información para satisfacer un objetivo”

RESULTADO: tabla con las diferentes alternativas que serán evaluadas en la siguiente fase: Diseño por Observación

Diseño por Observación (I)

OBJETIVO: análisis de la relevancia de cada una de las alternativas propuestas analizando diálogos reales usuario-operador en un servicio similar (100 llamadas)

Análisis de Objetivos

1. Los objetivos más solicitados por los usuarios y su secuencia

	%	Posición				
		1º	2º	3º	4º	5º
Horarios	64	57	6	1	-	-
Horarios Ida y Vuelta	20	-	14	5	1	-
Precios	46	6	30	10	-	-
Reserva	26	14	4	2	3	3
Frecuencia de trenes	2	2	-	-	-	-
Itinerario	14	8	4	1	1	-
Otros	12	8	2	-	2	-

2. La información que ofrece el operador para cubrir un objetivo

Diseño por Observación (II)

Análisis de Datos

1. Datos necesarios para cada objetivo y su secuencia
2. Clasificación de cada dato como Obligatorio u Opcional
 - **Obligatorio:** dato necesario para satisfacer el objetivo y que debe ser preguntado al usuario.
 - **Opcional:** dato no necesario pero si el usuario lo especifica permite personalizar la búsqueda.
3. Análisis de las diferentes formas de especificar un dato
4. Agrupación de datos
 - Definición de pasos del diálogo: ritmo del diálogo.
 - Posibilidad de incorporar confirmaciones conjuntas.

Diseño por Observación (III)

Análisis de la Negociación

Negociación: el usuario debe elegir o rechazar alguna de las alternativas que el sistema le ofrece, pudiendo cambiar alguna de las restricciones.

1. Información que más ayuda al usuario a decidir

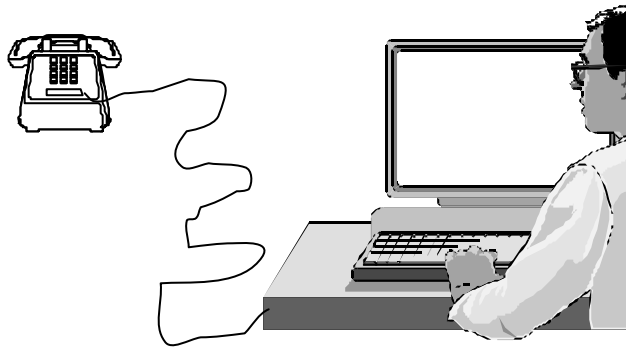
Criterio	%	Criterio	%
Hora salida	41	Nº de Enlaces	3
Hora llegada	14	Enlace	3
Est. De Salida	3	Clase	10
Precios	8	Duración	5
Tipo de tren	13		

2. Estrategias de negociación utilizada por el operador humano en su interacción con el usuario

- Presentar una opción al usuario y permitirle navegar.
- Presentar varias opciones y pedir que elija una.

Diseño por Simulación (I)

Herramienta de Mago de OZ



Se simula el comportamiento parcial del sistema mediante una persona, haciendo creer al usuario que está interactuando con un sistema completamente automático.

1. Es necesario desarrollar herramientas de ayuda al Mago: interfaz, etc...
2. Elementos de la herramienta del Mago de Oz:
 - **Aspectos Modificables por el Mago:** variables que debe decidir el Mago.
 - **Aspectos Fijos:** variables que no puede modificar el Mago.
 - **Plan de actuación:** reglas que guían la actuación del Mago.

El nº de variables Modificables por el Mago no debe ser excesivo. En el caso de que se quieran evaluar muchos aspectos, se deben hacer varios experimentos variando los aspectos fijos y modificables por el Mago

Diseño por Simulación (II)

OBJETIVO: simular y evaluar las mejores alternativas, así como analizar la respuesta de los usuarios frente a un sistema automático.

- Permite centrarnos en el diseño del **FLUJO DEL DIÁLOGO** dejando las técnicas de confirmación de datos para la etapa siguiente.
- Tipos de medidas de evaluación: obtenidas por el propio sistema (**sistema**) y las comentarios reflejados por los usuarios en los cuestionarios (**cuestionario**).

Análisis de Objetivos

1. Medidas para evaluar la secuencia de objetivos y su cobertura

Sistema:

- N° de veces que se solicita un objetivo
- Tiempo o n° de interacciones para satisfacerle

Cuestionario:

- Sugerencias de los usuarios sobre ampliaciones del servicio

Diseño por Simulación (III)

Análisis de Datos

1. Medidas para evaluar la inteligibilidad de las preguntas

Sistema: Grabación de las respuestas de los usuarios:

- Expresiones para definir Vocabularios y Modelos de Lenguaje
- N° de veces que el usuario permanece en silencio
- Tasa de reconocimiento (1ª versión del reconocedor)

2. Medidas para evaluar la secuencia de preguntas

Sistema: - Tasa de reconocimiento: multiplicación de las tasas de cada dato

Cuestionario: Preguntar a los usuarios sobre su preferencia

3. Evaluación de la gestión de datos obligatorios no especificados

Cuestionario: Preguntar a los usuarios sobre su preferencia:

- Fijar un valor por defecto
- Parametrizar la información que ofrezca el sistema

Diseño por Simulación (IV)

Análisis de la Negociación

Presentamos varias opciones de viaje a la vez

1. Evaluación del N° de opciones y de la información por opción

Sistema: Se prueban varias alternativas y se evalúa:

- Tiempo o n° de interacciones para elegir una opción

Cuestionario: Preguntar a los usuarios sobre su preferencia:

- Información que mejor le ayuda a elegir
- N° de opciones fáciles de recordar

Número de opciones				
1 en 1		2 en 2		3 en 3
21,4%		21,4%		57,2%
Criterio de Negociación				
T. Tren	Hora Sal.	Hora Lleg.	Precios	Duración
15,6%	37,5%	25,0%	15,6%	6,2%

Diseño por Mejora Iterativa (I)

Estrategias de Confirmación

1. Según el número de datos a confirmar

Un Dato: “¿Desea salir de Madrid?”

Varios Datos: “¿Desea viajar de Madrid a Sevilla?”

2. Según la confianza del dato a confirmar

Confirmación Explícita: “Entiendo que desea salir de Madrid ¿Es correcto?”

Confirmación Implícita: “He entendido que desea salir de Madrid ¿A dónde desea ir?”

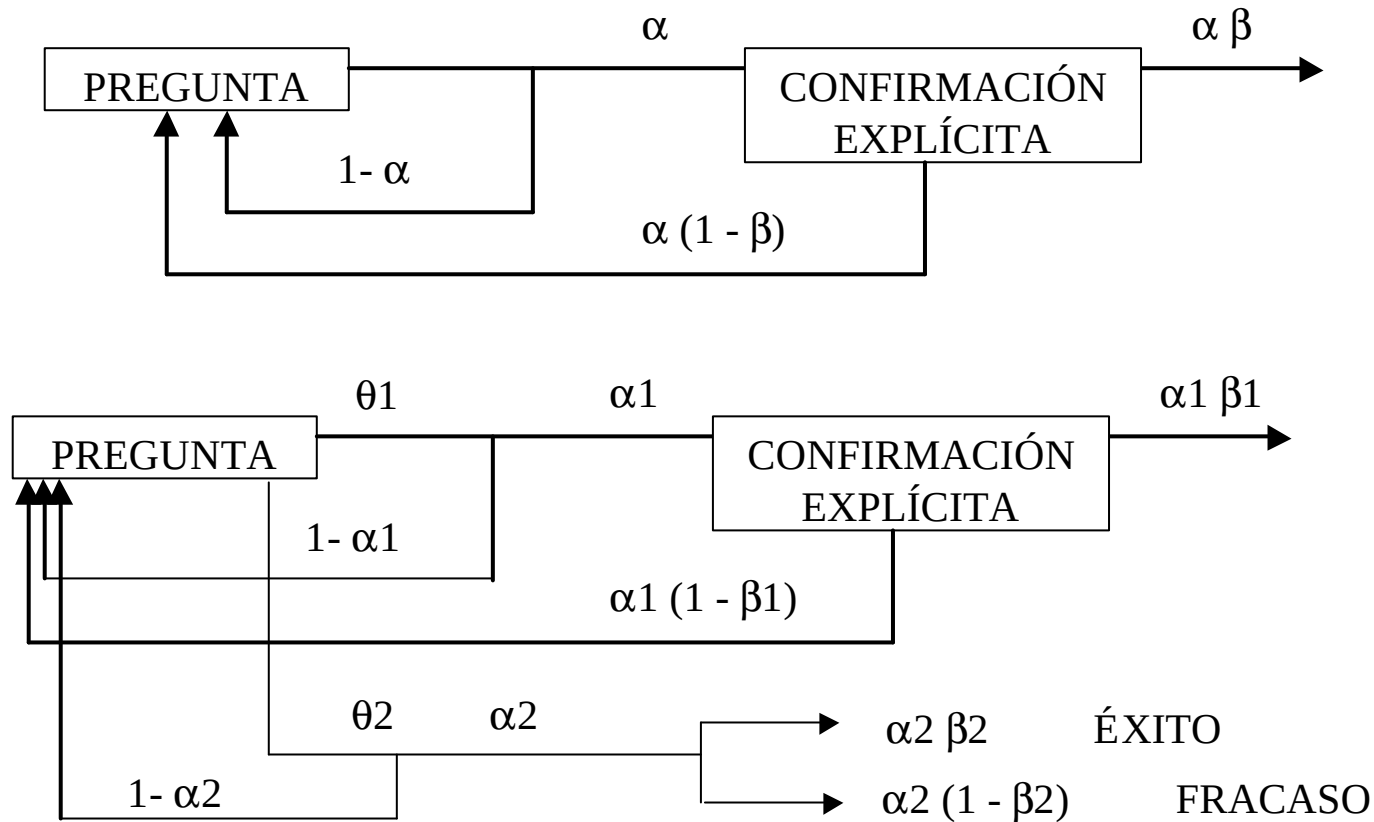
Confirmación Semi-Implícita: “He entendido que desea salir de Madrid. En caso de Error diga corregir si no diga su destino”

No Confirmación: El valor reconocido tiene tan buena confianza que no se confirma: Ej: preguntas de SÍ/NO

Rechazo: “Perdone, no he entendido muy bien, ¿Desde dónde desea salir?”

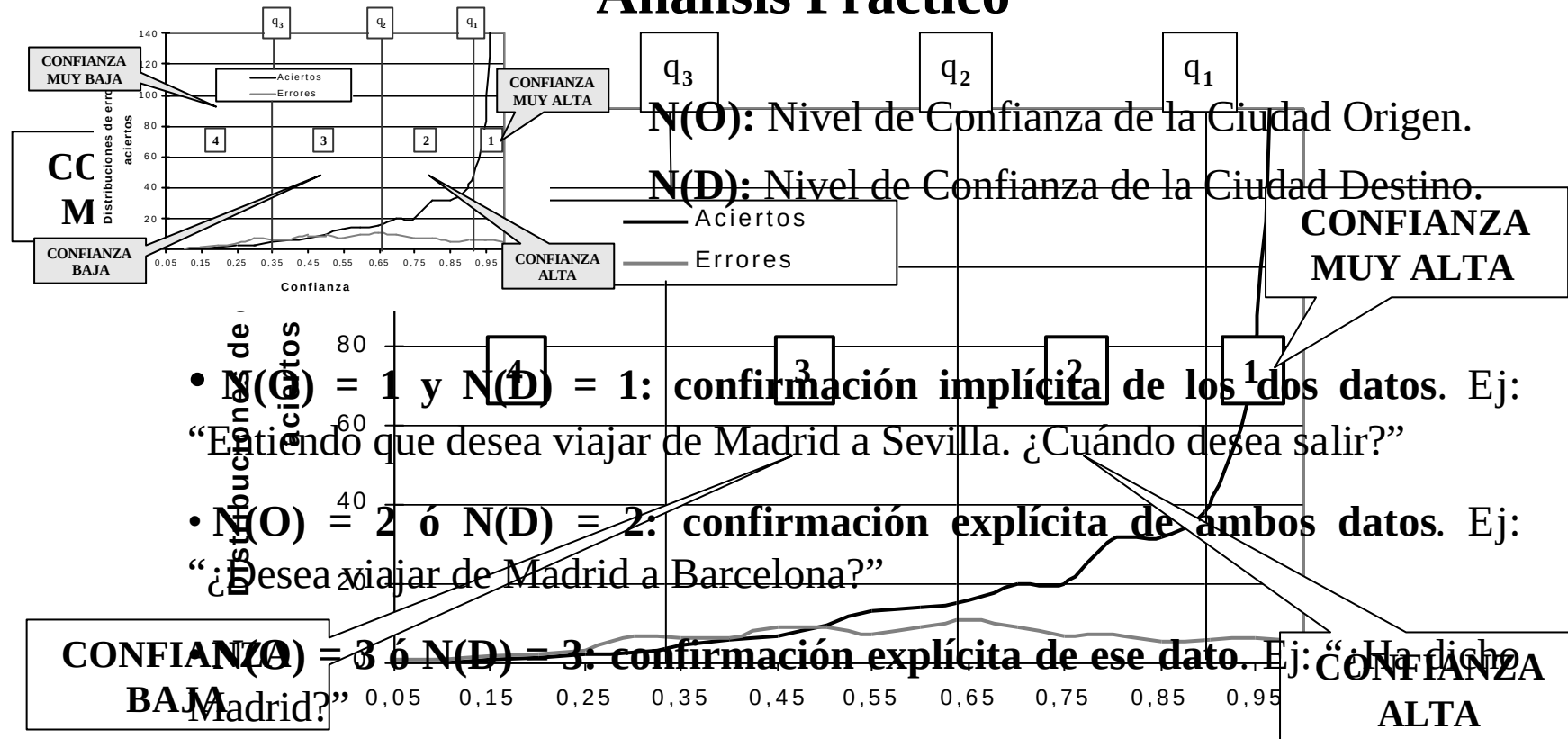
Diseño por Mejora Iterativa (III)

Análisis Teórico



Diseño por Mejora Iterativa (III)

Análisis Práctico



Diseño por Mejora Iterativa (v)

Modelo de Usuario

Modelo basado en la definición de varios niveles de destreza

1. Aspectos modificables del diálogo

- Preguntas que formula el sistema
 - **NIVEL 1:** Recuerde *hablar después de escuchar la señal*. Diga el *período del día* en el que desea viajar: *por la mañana, por la tarde o por la noche*.
 - **NIVEL 2:** Diga el *período del día* en el que desea viajar: *por la mañana, por la tarde o por la noche*.
 - **NIVEL 3:** Diga el *período del día* en el que desea viajar
 - **NIVEL 4:** ¿Cuándo desea viajar?
- Ayuda y comentarios ofrecidos por el sistema: mayor información cuanto menor destreza demostrada
- Mecanismos de confirmación: definición de diferentes umbrales según el nivel de destreza
- Definición de subdiálogos concretos: diálogos más o menos guiados

Diseño por Mejora Iterativa (VI)

Modelo de Usuario

2. Indicadores de la destreza del usuario

Positivos

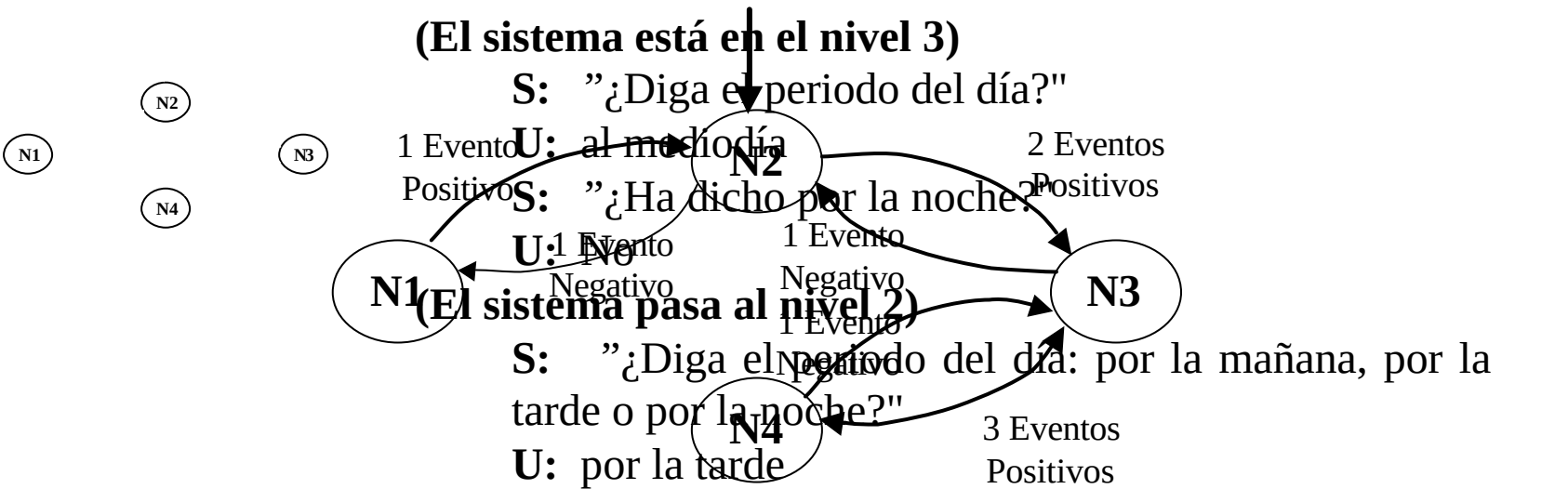
- N° de veces que un **usuario confirma** la propuesta del sistema.
- N° de veces que un **usuario corrige** un posible error del sistema.
- Porcentaje de confirmaciones implícitas frente a explícitas.

Negativos

- N° de veces que un **usuario no confirma** la propuesta del sistema.
- N° de veces que el **usuario no puede corregir** un error del sistema.
- N° de veces que el **usuario se queda callado** ante la pregunta del sistema.
- N° de veces que pide que el sistema le repita la información.
- N° de veces que solicita volver a empezar.
- Porcentaje de confirmaciones explícitas frente al de implícitas.

$$EventosTotalPositivos = \sum_{i=0}^N EventoPositivo_i \times Peso_i$$

$$EventosTotalNegativos = \sum_{i=0}^N EventoNegativo_i \times Peso_i$$



Diseño por Mejora Iterativa (VIII)

Aspectos GENERALES

1. Volver al comienzo desde cualquier punto: VOLVER A EMPEZAR
2. Prever mecanismos de CORRECCIÓN
3. Posibilidad de REPETIR: última pregunta o última información
4. Solicitar AYUDA en cualquier punto del diálogo
5. Suspensión momentánea de la llamada para atender alguna tarea urgente

Diseño por Mejora Iterativa (IX)

Ajustes Finales

1. Preguntar Fecha: ¿Desea viajar esta semana, la que viene o más adelante?
2. Comprobar la posibilidad de fechas y horas.
 - Fecha del viaje no puede ser anterior a la actual: en el caso del mismo día, considerar relación de horas.
 - Fecha de vuelta posterior a la fecha de llegada en la ida: en el caso del mismo día, considerar relación de horas.
3. Numeración absoluta de las opciones del TREN.

Evaluación (I)

- 105 usuarios completaron 4 escenarios de viaje
- 55 estudiantes y 50 empleados de la empresa RENFE

Características obtenidas del sistema

Medida	Valor
Duración Media de la consulta	205 seg
Nº Medio de preguntas realizadas por el Sistema	19,21
Nivel medio de destreza	1,91
% porcentaje de confirmaciones implícitas	70,0 %
Nº de veces que se vuelve a empezar en la mitad de la consulta	0,21
Nº de veces que se solicita la corrección de un dato	0,47
Duración de la negociación de la opción de un tren	57 seg
Nº de veces que se solicita la repetición de las opciones de tren	0,227

TASA DE
RECONOCIMIENTO

- **Vocab. Menores de 50 palabras: 95%** (con rechazo)
- **Ciudades (770): 88,1%** (con rechazo)

Evaluación (II)

Características obtenidas de los cuestionarios (Escala 1-5)

Medida	Valor
El sistema comprende lo que le dices	3,1
Las respuestas del sistema son claras y concisas	3,4
Entiendo lo que el sistema me dice	3,5
Se accede a la información de trenes rápidamente	2,9
El sistema es fácil de usar	3,5
Es fácil de aprender su comportamiento	3,7
El sistema me ayuda durante la conversación	3,1
En caso de error la corrección fue fácil	2,9
El sistema me hace las preguntas en un orden lógico	3,6
En general, es un buen sistema	3,0

35,2 % este sistema, 35,2% por WEB y 25,7% en la estación
(el resto 3,9% insistió en la opción de operadores humanos)