

Índice general

| | |
|---|------------|
| 1. Introducción | 27 |
| 1.1. Introducción. Sistemas de diálogo hablado | 27 |
| 1.1.1. Reconocimiento automático del habla | 32 |
| 1.1.2. Comprensión del habla | 34 |
| 1.1.2.1. Medidas de confianza | 34 |
| 1.1.3. Gestión de diálogo | 35 |
| 1.2. Objetivos y líneas de investigación tratadas en la tesis | 36 |
| 1.3. Estructura de la tesis | 37 |
| 2. Estado del arte | 41 |
| 2.1. Sistemas de diálogo hablado | 41 |
| 2.2. Gestión del diálogo | 72 |
| 2.2.1. Resumen de metodologías de gestión relevantes | 75 |
| 2.2.2. Teoría de los Estados de la Información | 81 |
| 2.2.3. Metodologías estadísticas. Aprendizaje Reforzado | 84 |
| 2.2.3.1. Modelado del diálogo mediante redes bayesianas | 94 |
| 2.3. Evaluación de sistemas de diálogo | 96 |
| 2.4. Simulación de usuarios | 104 |
| 2.5. Adaptación al usuario y/o a la tarea | 112 |
| 3. La tarea DIHANA | 119 |
| 3.1. Introducción | 119 |
| 3.2. Objetivos del proyecto DIHANA | 120 |
| 3.3. Adquisición del corpus DIHANA | 122 |
| 3.3.1. Tipos de escenarios definidos | 122 |
| 3.3.2. La plataforma de adquisición de diálogos | 124 |
| 3.3.3. Fases de la adquisición | 125 |
| 3.3.4. La estrategia del Mago de Oz | 126 |
| 3.3.5. Ficheros almacenados durante la adquisición | 127 |
| 3.3.6. Evaluación recibida de los usuarios | 128 |
| 3.3.7. Características del corpus DIHANA | 130 |
| 3.4. Etiquetado del corpus DIHANA | 130 |
| 3.4.1. Etiquetado de los turnos de usuario: representación semántica de la tarea | 130 |
| 3.4.2. Etiquetado de los turnos de sistema | 134 |

| | | |
|------------|--|------------|
| 7.4. | Evaluación del gestor de diálogo estadístico basado en redes neuronales . . . | 196 |
| 7.4.1. | Evaluación de la influencia de la talla del corpus | 197 |
| 7.4.2. | Evaluación de la influencia del sexo | 198 |
| 7.4.3. | Evaluación de la procedencia de los diálogos | 198 |
| 7.5. | Evaluación utilizando la voz y usuarios reales | 199 |
| 8. | Simulación de usuarios | 201 |
| 8.1. | Introducción | 201 |
| 8.2. | Simulador de usuarios estadístico | 202 |
| 8.2.1. | Codificación definida para el clasificador MLP | 204 |
| 8.2.2. | Implementación práctica del simulador estadístico | 206 |
| 8.2.3. | Evaluación del simulador | 210 |
| 8.3. | Definición de una técnica para la adquisición automática de diálogos | 213 |
| 8.3.1. | Adquisición de un corpus de diálogos | 214 |
| 8.3.2. | Ejemplos de diálogos adquiridos | 215 |
| 8.3.3. | Evaluación del corpus simulado | 217 |
| 8.4. | Experimentación con las metodologías descritas | 218 |
| 8.4.1. | Modelo estadístico tanto en el simulador de usuario como en el gestor de diálogo | 218 |
| 8.4.2. | Selección aleatoria de respuestas en el simulador de usuarios y modelo estadístico en el gestor de diálogo | 220 |
| 8.4.3. | Modelo estadístico en el simulador de usuarios y selección aleatoria de respuestas en el gestor de diálogo | 221 |
| 9. | Adaptación del gestor del diálogo al usuario y a nuevas tareas | 223 |
| 9.1. | Introducción | 223 |
| 9.2. | El proyecto EDECÁN | 223 |
| 9.3. | La tarea EDECÁN | 226 |
| 9.4. | Representación semántica de la tarea | 227 |
| 9.4.1. | Etiquetado de los turnos de usuario | 228 |
| 9.4.2. | Etiquetado de los turnos de sistema | 229 |
| 9.5. | Representación de las fuentes de información | 231 |
| 9.6. | Gestor de diálogo desarrollado | 232 |
| 10. | Conclusiones | 235 |
| A. | Adquisición Corpus DIHANA | 241 |
| A.1. | Ejemplo de ficha definida para la adquisición del corpus DIHANA | 241 |
| A.2. | Ficheros almacenados durante la adquisición de cada diálogo | 245 |
| A.2.1. | Fichero .fra.xml. Frases acústicamente balanceadas | 245 |
| A.2.2. | Fichero .rec.xml. Módulo de reconocimiento automático del habla | 246 |
| A.2.3. | Fichero .com.xml. Módulo de comprensión semántica | 247 |
| A.2.4. | Fichero .dia.xml. Módulo gestor de diálogo | 248 |
| A.2.5. | Fichero .que.xml. Consultas a la base de datos | 250 |
| A.2.6. | Fichero .piz.xml. Registro del diálogo | 251 |
| A.3. | Resultados de la encuestas | 253 |
| A.3.1. | Encuesta que rellenaron los solicitantes | 253 |

| | |
|---|------------|
| A.3.2. Estadísticas generales de las encuestas y evaluación de los escenarios | 255 |
| B. Sistema de diálogo DIHANA | 261 |
| B.1. Formato de los paquetes XML definidos | 261 |
| B.1.1. Módulo de entrada (RAH) | 261 |
| B.1.2. Módulo de comprensión (CH) | 261 |
| B.1.3. Gestor de diálogo (GD) | 262 |
| B.1.4. Módulo generador de respuestas (GR) | 263 |
| B.1.5. Módulo de síntesis de voz (CTV) | 264 |
| B.1.6. Módulo de base de datos (BD) | 264 |
| B.2. Categorías definidas en el corpus DIHANA | 264 |
| B.2.1. Ciudades | 264 |
| B.2.2. Servicios | 265 |
| B.2.3. Tipo-viaje | 265 |
| B.2.4. Tipo-tren | 265 |
| B.2.5. Precio | 265 |
| B.2.6. Número-relativo-orden | 265 |
| B.2.7. Dirección | 265 |
| B.2.8. Clase billete | 266 |
| B.3. Gestión del diálogo: Implementación en LUA de las redes neuronales | 266 |
| C. Generación de respuestas | 269 |
| C.1. Plantillas de respuestas definidas para la adquisición del corpus DIHANA | 269 |
| C.1.1. No entendido | 269 |
| C.1.2. No válido | 269 |
| C.1.3. Completar (confirmación/pregunta de concepto) | 269 |
| C.1.4. Confirmación explícita (de concepto y atributos) | 270 |
| C.1.5. Confirmación implícita | 271 |
| C.1.5.1. Duración | 271 |
| C.1.5.2. Horarios | 274 |
| C.1.5.3. Precio | 276 |
| C.1.5.4. Servicios | 283 |
| C.1.5.5. Tipo de tren | 283 |
| C.1.6. Confirmación mixta Seguro | 284 |
| C.1.6.1. Concepto Horario | 284 |
| C.1.6.2. Concepto Precio | 284 |
| C.1.6.3. Concepto Duración | 284 |
| C.1.6.4. Frame Tipo de tren | 285 |
| C.1.6.5. Concepto Servicios | 285 |
| C.1.6.6. Segunda parte de la confirmación mixta | 285 |
| C.1.7. Confirmación mixta Concepto Inseguro | 285 |
| C.1.7.1. Primera parte de la confirmación | 285 |
| C.1.7.2. Segunda parte de la confirmación | 285 |
| C.1.8. Concepto Seguro Hora, Precio, Duración, Tipo de tren o Servicios | 286 |
| C.1.8.1. Primera parte de la confirmación | 286 |
| C.1.8.2. Segunda parte de la confirmación | 286 |

| | |
|---|------------|
| C.1.9. Nueva Consulta | 286 |
| C.2. Plantillas de respuestas definidas tras el etiquetado del corpus. | 287 |
| C.2.1. Preguntas | 287 |
| C.2.2. Confirmaciones explícitas | 287 |
| C.2.2.1. Conceptos | 287 |
| C.2.2.2. Atributos | 287 |
| C.2.3. Confirmaciones mixtas | 288 |
| C.2.4. Respuestas a consultas a la base de datos | 289 |
| C.2.4.1. Respuesta-Hora_salida | 289 |
| C.2.4.2. Confirmación,Respuesta-[Hora_salida],Hora_salida | 290 |
| C.2.4.3. Confirmación,Respuesta-Hora_salida,Hora_salida | 290 |
| C.2.4.4. Respuesta-Hora_llegada | 290 |
| C.2.4.5. Confirmación,Respuesta-[Hora_llegada],Hora_llegada | 291 |
| C.2.4.6. Respuesta-Precio | 291 |
| C.2.4.7. Confirmación,Respuesta-Precio,Precio | 297 |
| C.2.4.8. Confirmación,Respuesta-[Precio],Precio | 298 |
| C.2.4.9. Respuesta-Tipo_tren | 298 |
| C.2.4.10. Confirmación,Respuesta-[Tipo_tren],Tipo_tren | 298 |
| C.2.4.11. Confirmación,Respuesta-Tipo_tren,Tipo_tren | 299 |
| C.2.4.12. Respuesta-Tiempo_recorrido | 299 |
| C.2.4.13. Respuesta-Servicio | 301 |
| C.2.4.14. Confirmación,Respuesta-[Servicio],Servicio | 301 |
| C.2.5. Otras | 301 |
| D. Desarrollo de sistemas de diálogo mediante VoiceXML | 303 |
| D.1. Introducción a VoiceXML | 303 |
| D.1.1. Orígenes de VoiceXML | 303 |
| D.1.2. Estructura y funcionamiento | 304 |
| D.1.3. Conceptos básicos | 306 |
| D.1.3.1. Diálogos y subdiálogos | 306 |
| D.2. Constructores de diálogo | 307 |
| D.2.1. Formularios | 307 |
| D.2.2. Menús | 308 |
| D.3. Modelado del lenguaje en VoiceXML | 308 |
| D.3.1. Gramáticas | 308 |
| D.3.1.1. Gramáticas para el reconocimiento de voz | 308 |
| D.3.1.2. Ámbito de las gramáticas | 309 |
| D.3.1.3. Activación de las gramáticas | 309 |
| D.4. Recursos disponibles | 310 |
| D.4.1. IBM Voice Toolkit para WebSphere Studio | 311 |
| D.4.1.1. Instalación del Voice Toolkit | 311 |
| D.4.1.2. Reconocimiento de voz | 312 |
| D.5. Aplicación práctica: Desarrollo de una centralita en castellano | 313 |
| D.6. Implementaciones de las especificaciones de interfaces vocales del W3C | 315 |
| D.6.1. VoiceXML 2.0 y 2.1 | 315 |
| D.6.2. Especificaciones para las gramáticas de reconocimiento de voz | 317 |

| | |
|--|-----|
| D.6.3. Interpretación semántica de gramáticas de reconocimiento de voz . | 318 |
| D.6.4. CCXML (Call Control XML) | 318 |

| | |
|---------------------|------------|
| Bibliografía | 319 |
|---------------------|------------|
