

Índice de figuras

Figura 1-1: Diagrama de bloques de un Servidor Vocal Interactivo.	1-1
Figura 3-1: Ejemplo de error de reconocimiento debido al efecto de co-articulación ..	3-3
Figura 3-2: Arquitectura en dos niveles: obtención de la secuencia de letras y comparación con los nombres del directorio	3-7
Figura 3-3: Árbol construido con los nombres del directorio	3-8
Figura 3-4: Arquitectura de hipótesis y verificación (San-Segundo et al, 2000b)	3-9
Figura 3-5: Modelo de Markov lineal	3-11
Figura 3-6: Diagrama de bloques del reconocedor de nombres deletreados	3-13
Figura 3-7: Topología del modelo de la letra A con silencios contextuales incorporados	3-16
Figura 3-8: Nuevas transiciones entre letras con los modelos contextuales	3-16
Figura 3-9: Modelos de ruido incorporados en el espacio de búsqueda para aplicar el modelo de lenguaje 3-gram (San-Segundo et al, 2002)	3-19
Figura 3-10: Variedad de errores producidos en la incorporación del modelo de lenguaje (ML) (San-Segundo et al, 2002)	3-22
Figura 3-11: Análisis de transiciones entre letras en el algoritmo de generación de las N mejores secuencias de letras	3-23
Figura 3-12: Grafo de letras para el nombre TORRA con COMPLEJIDAD_GRAFO=2	3-26
Figura 3-13: Incorporación en el grafo del modelo de lenguaje 3-gram	3-28
Figura 3-14: Evolución de la Tasa de Acierto (TAN) y de las probabilidades $P_H(M)$ y $P_V(M)$ según M (diccionario de 1.000 nombres) (San-Segundo et al, 2002)	3-32
Figura 3-15: Evolución de la Tasa de Acierto (TAN) y el Tiempo de Procesado según M (diccionario de 1.000 nombres)	3-32
Figura 3-16: Otros ejemplos de evolución de la Tasa de Acierto (TAN) según M	3-33
Figura 4-1: Diagrama del sistema de reconocimiento para fechas y horas	4-4
Figura 4-3: Transiciones para permitir variedad de pausas entre palabras	4-7
Figura 4-4: Transiciones entre palabras considerando ruidos intermedios	4-8
Figura 4-5: Modelo acústico de la palabra ‘lunes’	4-9
Figura 4-6: Cambio de características de los estados según la palabra predecesora	4-10
Figura 4-7: Estructuras de modelos acústicos utilizados con 3 y 5 estados respectivamente	4-11
Figura 4-8: Evolución de la Tasa de Error según el porcentaje de ficheros eliminados (San-Segundo, D., 2001)	4-13

Figura 4-9: Evolución de la Tasa de Error según el número de trifenemas considerado	4-14
Figura 4-10: Evolución de la Tasa de Error según el número de centroides considerado (San-Segundo, D., 2001)	4-15
Figura 4-11: Comparación de la evolución de la Tasa de Error para diferentes modelados semicontinuos según el número de centroides	4-16
Figura 4-12: Árbol inverso resultado del segundo paso en la generación del grafo de palabras (COMPLEJIDAD_GRAFO=2)	4-18
Figura 4-13: Grafo de palabras para el ejemplo de la figura 4-12	4-18
Figura 4-14: Utilización de la probabilidad del modelo de lenguaje considerando un modelo de ruido como modelo predecesor	4-22
Figura 4-15: Duplicación de nodos para la incorporación de un modelo 3-gram	4-23
Figura 4-16: Agrupación de varios modelos acústicos en un mismo nodo del grafo	4-24
Figura 4-17: Evolución de la tasa de error según el número de hipótesis	4-28
Figura 5-1: Etiquetación automática de conceptos como correctos o incorrectos	5-3
Figura 5-2: Ejemplo de reconocimiento erróneo pero con el mismo significado	5-6
Figura 5-3: Ejemplo de gestión del diálogo utilizando medidas de confianza. (S) frase del sistema, (U) frase del usuario	5-6
Figura 5-4: Ejemplo de representación del número de casos según el valor de confianza	5-11
Figura 5-5: Rechazo Correcto (RC) según el Rechazo Incorrecto (RI) para los parámetros del proceso de decodificación (PD) del modelo de lenguaje (ML) o todos combinados (PD + ML)	5-14
Figura 5-6: RC vs. RI para las confianzas medias al nivel de palabra (CMRC + CMRV), el resto de parámetros (NPR+NPV+HC+HCV+CML+PML) y combinándolos todos juntos	5-15
Figura 5-7: Ejemplo de frase fuera del dominio de aplicación cuyo reconocimiento genera una hipótesis con algún concepto válido en la tarea	5-17
Figura 5-8: Rechazo Correcto vs. Rechazo Incorrecto (RI) para la detección de frases fuera del dominio de aplicación	5-17
Figura 5-9: Rechazo Correcto vs. Rechazo Incorrecto (RI) para la detección de frases fuera del dominio de aplicación y frases sin ningún concepto correcto	5-19
Figura 5-10: Ejemplo de frase de referencia y de tres posibles hipótesis. Cada palabra se etiqueta con la confianza obtenida y las marcas de trama inicio y final	5-21
Figura 5-11: Grafo resultado del ejemplo presentado en la figura 5-10	5-21
Figura 5-12: Ejemplo de alineamiento entre hipótesis que no genera una unión de nodos por no cumplirse la segunda restricción	5-22
Figura 5-13: Grafo resultado del ejemplo presentado en la figura 5-12	5-23
Figura 5-14: Grafo obtenido con el algoritmo ROVER del ejemplo presentado en la figura 5-12	5-23

Figura 5-15: Rechazo Correcto vs. Rechazo Incorrecto (RI) para la detección de errores de reconocimiento considerando por separado parámetros de la etapa de hipótesis (H-1, H-2), de la etapa de verificación (V-3), y de ambas etapas (H-2 y V-3).	5-33
Figura 5-16: Rechazo Correcto (RC) vs. Rechazo Incorrecto (RI) para la detección de nombres fuera del diccionario, considerando por separado parámetros de la etapa de hipótesis (H-1, H-2), de la etapa de verificación (V-3), y de ambas etapas (H-2 y V-3).	5-35
Figura 5-17: Rechazo Correcto (RC) vs. Rechazo Incorrecto (RI) para la detección de errores en el reconocedor de fechas y horas para habla leída	5-40
Figura 5-18: Rechazo Correcto (RC) vs. Rechazo Incorrecto (RI) para la detección de errores en el reconocedor de fechas y horas en habla espontánea	5-41
Figura 6-1: Ejemplo de grafo para un servicio de información de horarios y precios de tren	6-5
Figura 6-2: Plantillas de los objetivos del diálogo presentado en la figura 6-1	6-8
Figura 6-3: Arbol de objetivos para el ejemplo de información de horarios y precios de tren	6-11
Figura 6-4: Ejemplo de regla de integridad para el subobjetivo TRAYECTORIA	6-13
Figura 6-5: Ejemplo de ampliación del árbol de objetivos según va evolucionando el diálogo	6-14
Figura 6-6: Ejemplo de reducción de objetivos según va evolucionando el diálogo	6-14
Figura 6-7: Cálculo de T_{MEST} y P_{ERROR} para un estado del diálogo formado por un sólo turno en el que se pide al usuario la ciudad de origen del viaje	6-16
Figura 6-8: Cálculo de T_{MEST} y P_{ERROR} para un estado del diálogo formado por más de un turno de diálogo: pregunta y confirmación del dato	6-17
Figura 6-9: Agrupación de estados para formar un estado mayor. También se muestra el cálculo de T_{MEST} y P_{ERROR} totales a partir de los parámetros de cada estado	6-18
Figura 6-10: Diagrama Entidad-Relación: información y reserva de viajes en tren	6-22
Figura 6-11: Diagrama de flujo para un diálogo genérico que podríamos implementar después de esta etapa	6-25
Figura 6-12: Rechazo Correcto vs. Rechazo Incorrecto (RI) para la detección de errores de reconocimiento en el primer candidato y en los dos primeros candidatos	6-49
Figura 6-13: Distribución de aciertos y errores en función del valor de confianza obtenido para el reconocedor de palabras y expresiones utilizado	6-50
Figura 6-14: Diagrama de flujo para el caso de utilizar las medidas de confianza para rechazar las hipótesis de baja confianza y volver a preguntar	6-50
Figura 6-15: Diagrama de flujo para el caso de introducir una pregunta de confirmación explícita	6-52
Figura 6-16: Distribución de aciertos y errores con los umbrales definidos manualmente para el diseño de los mecanismos de confirmación	6-54
Figura 6-17: Diagrama de niveles de destreza en la técnica de modelado de usuario ...	6-62
Figura A-1: Módulos de la arquitectura GALAXY-II (Seneff et al, 1998)	A-2
Figura A-2: Ejemplo de aplicación en el lenguaje del entorno TADE	A-10