

**UNIVERSIDAD DE BUENOS AIRES**

**FACULTAD DE CIENCIAS EXACTAS Y NATURALES**

1. DEPARTAMENTO: Computación
2. CUATRIMESTRE: Segundo de 2002
3. ASIGNATURA: Introducción a las Tecnologías del Habla
4. CARRERA: Licenciatura en Ciencias de la Computación
5. CARACTER DE LA MATERIA: Optativa
6. NUMERO DE CODIGO DE CARRERA: 18
7. NUMERO DE CODIGO DE MATERIA: C .....
8. PUNTAJE: 3
9. PLAN DE ESTUDIOS AÑO: 1993) -1904
10. DURACION DE LA MATERIA: cuatrimestral
11. HORAS DE CLASE SEMANAL:  
a)TEORICAS: 3 hs      c)PRACTICAS: 3 hs  
b)LABORATORIO: hs      d)SEMINARIOS: hs
12. CARGA HORARIA TOTAL SEMANAL: 6 hs, Lunes y Jueves de 19 a 22 hrs.
13. ASIGNATURAS CORRELATIVAS: Probabilidades y estadística
14. FORMA DE EVALUACION: 2 parciales, promocional.
15. PROGRAMA Y BIBLIOGRAFIA: Adjuntas a esta hoja.

Fecha: 27/05/2002

## Introducción a las Tecnologías del Habla – Programa y Bibliografía

### Objetivos:

- Proporcionar los fundamentos lingüísticos utilizados en las tecnologías del habla.
- Introducir en los conceptos fundamentales y métodos de procesamiento de la voz y del lenguaje utilizados en las tecnologías del habla.
- Aplicar los conocimientos adquiridos a desarrollos relativos a la síntesis y el reconocimiento automático de la voz.
- Brindar conocimientos sobre el estado del arte de las tecnologías de habla.

Docentes: Alvarez, José Angel y Renato, Alejandro (Fac. de Medicina, UBA)

Dedicación: 6 horas semanales

Correlatividades: Probabilidades y estadística

### Contenidos

#### Unidad 1

Motivaciones para las tecnologías del habla. Arquitecturas de los sistemas de procesamiento del lenguaje oral. Estructura del lenguaje oral. Acústica. Fonética y fonología. Silabas y palabras. Sintaxis y semántica.

#### Unidad 2

Elementos básicos de estadística. Probabilidades e inferencia estadística. Teoría de la información: entropía, entropía condicional, codificación e información mutua. Elementos de reconocimiento de patrones. Teoría bayesiana de decisión. Construcción de clasificadores. Métodos de estimación no supervisados. Árboles de clasificación y regresión.

#### Unidad 3

Elementos de procesamiento digital de señales. Sistemas y señales digitales. Transformadas de frecuencias continuas y no continuas. Filtros y ventanas digitales. Procesamiento digital de señales analógicas. Procesamiento de tasa múltiple. Bancos de filtros. Procesos estocásticos.

#### Unidad 4

Representaciones de la señal de habla. Análisis de Fourier de lapsos breves. Modelo acústico de la producción del habla. Codificación lineal predictiva (LPC). Procesamiento cepstral.

Representaciones motivadas perceptualmente. Formantes. Altura tonal. Codificación del habla: características y tipos de codificadores del habla.

#### Unidad 5

Modelos markovianos ocultos (HMM). Cadenas markovianas. Modelo oculto. Algoritmos Forward, Forward-Backward, Viterbi, Baum-Welch. HMMs continuos y semicontinuos. Cuestiones y limitaciones.

#### Unidad 6

Modelización acústica. Variabilidad en la señal de habla. Errores de reconocimiento. Extracción de rasgos. Modelización fonética: selección de unidades. Peso de rasgos acústicos. Técnicas adaptativas. Técnicas alternativas al HMM. Robustez ambiental.

#### Unidad 7

Modelización del lenguaje. Teoría de lenguajes formales y parsing. Modelos estocásticos del lenguaje (PCFG y n-gramas). Medida de complejidad de los modelos de lenguaje. Suavizamiento de modelos de n-gramas. Modelos adaptativos. Cuestiones prácticas.

  
 Dr. Guillermo Alvarez  
 Director Adjunto  
 Depto. de Computación  
 C.M.U.R.A.

**Unidad 8**

Algoritmos básicos de búsqueda para el reconocimiento del habla: Viterbi, Stack Decoding (A\*). Algoritmos de búsqueda con vocabularios grandes. Manipulación eficiente de léxico en árbol. Otras estrategias eficientes de búsqueda. Evaluación de algoritmos de búsqueda.

**Unidad 9**

Síntesis de habla. Sistemas TTS. Análisis textual y fonético. Léxico, estructura de un documento, normalización. Análisis lingüístico. Desambiguación de homógrafos. Análisis morfológico. Conversión letra a sonido (LTS). Evaluación de sistemas TTS.

**Unidad 10**

Síntesis de habla. Prosodia. Esquema de generación de prosodia. Estilo de habla. Prosodia simbólica. Asignación de duraciones. Generación de altura tonal (pitch). Lenguajes de marcación de prosodia. Evaluación de la prosodia. Síntesis de formantes del habla. Síntesis por concatenación. Modificación prosódica del habla. Modelos de fuente-filtro de la modificación prosódica.

**Unidad 11**

Comprepción del lenguaje hablado. Lenguajes hablados versus escritos. Estructura del diálogo. Interpretación de oraciones. Estructura del discurso. Gestión del diálogo. Generación de respuestas. Evaluación. Aplicaciones e interfaces de usuario. Aplicaciones típicas. Diseño de interfaces de habla. Internacionalización.

**Bibliografía**

- Allen, James (1995). *Natural Language Understanding*, Redwood City, Benjamin/Cummings, 2a. ed.
- Allen J., Hunnicut S. and Klatt, D. *Text-to-speech: the MITalk system*, Cambridge University Press, 1987.
- Breiman, L., Friedman, J. Olshen, R. and Stone, C. *Classification and regression trees*, Wadsworth and Brooks, Pacific Grove, CA. 1984.
- Charniak, Eugene (1993). *Statistical Language Learning*, Cambridge, MIT Press.
- Clark, John y Yallop, Colin (1995). *An Introduction to Phonetics and Phonology*, Blackwell.
- D'Introno y otros (1995). *Fonética y fonología española*, Madrid, Cátedra.
- Dutoit, T. *An introduction to Text-to-Speech Synthesis* Kluwer Academic Publishers, 1997.
- Fant, Gunnar (1960). *Theory of Acoustical Production*.
- Fosler-Lussier, Eric (1998). "Markov Models and Hidden Markov Models: A Brief Tutorial", International Computer Science Institute, TR-98-041.
- Huang, Xuedong; Acero, Alex y Hon, Hsiao-Wuen. *Spoken Language Processing: A Guide to Theory, Algorithm and System Development*. Prentice Hall, Upper Saddle River, New Jersey, 2001.
- Jelinek, Frederik. *Statistical Methods for Speech Recognition*. Cambridge, MIT Press, 1998.
- Jurafsky, Daniel y Martin, James H. *Speech and Language Processing - An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, Upper Saddle River, New Jersey, 2000.
- Lieberman, Philip y Blumstein, Sheila E. *Speech physiology, speech perception, and acoustic phonetics*. Cambridge University Press, Nueva York, 1988.
- Manning, Christopher D. y Schütze, Hinrich. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, 1999.
- Proakis, John G. y Manolakis, Dimitris G. *Tratamiento digital de señales - Principios, algoritmos y aplicaciones*. Prentice Hall, Madrid, 1998, 3a ed.
- Ritchie, Graeme D. y otros (1992). *Computational Morphology*, Cambridge, MIT Press
- Sells, Peter y otros (1991). *Foundational Issues in Natural Language Processing*, Cambridge, MIT Press.
- Young, Steve y otros (2000). *The HTK book. 1995-1999*, Microsoft Corporation.

Jr. Guillermo Jurafsky  
Director Adjunto  
Dept. de Computación  
F C E y N. I I R A

com 2002

14