

Procesamiento del Lenguaje Natural

Curso 2006-2007

Arantza Díaz de Ilarraza

jipdisaa@si.ehu.es

German Rigau

german.rigau@ehu.es

IXA group

Ingeniería en Informática

Índice

- Procesamiento del Lenguaje Natural (PLN)
 - Objetivos
 - Temario
 - Destrezas
 - Bibliografía
 - Nota, horarios
- Motivación al PLN

Objetivos

- Presentar la Ingeniería Lingüística (procesamiento de la lengua, lenguaje natural o lenguaje humano) en sus tres vertientes:
 - 1. métodos y herramientas
 - 2. recursos y formas de adquirirlos
 - 3. aplicaciones.
- Conocer las dos aproximaciones que se usan en PLN:
 - la aproximación basada en el conocimiento (básicamente lingüístico)
 - la aproximación basada en métodos empíricos (básicamente de tipo estadístico o de Aprendizaje Automático).

Temario

- Introducción al PLN
- Niveles básicos del Procesamiento Lingüístico
- Tratamiento sintáctico
- Tratamiento semántico
- Pragmática
- Generación
- Métodos descriptivos (empíricos)
- Aplicaciones

Conocimientos y destrezas que se adquieren

- Analizar un problema, identificar las posibilidades y ventajas de aplicar técnicas de PLN para abordarlo.
- Elección de las técnicas más adecuadas por cada tarea.
- Capacidad de integrar componentes ya existentes de PLN (tanto recursos lingüísticos como herramientas de tratamiento) en aplicaciones reales.
- Capacidad para diseñar y construir componentes software que resuelvan problemas básicos de PLN.
- Capacidad para diseñar y llevar a término experimentos necesarios por la aplicación de los métodos empíricos de PLN, y de analizar los resultados.

Bibliografía

- D. Jurafsky, James H. Martin Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Prentice Hall, Upper Saddle River, N.J. , 2000.
- C. Manning, H. Schütze Foundations of statistical Natural Language Processing, MIT Press Cambridge, Mass., 1999.
- R. Dale, H. Moisl, H.Somers, ed. Handbook of natural Language Processing, Marcel Dekker, New York, 2000.
- R. Mitkov (editor) The Oxford handbook of Computational Linguistics, Oxford University Press, 2004.

Nota y horarios

- $\text{Nota PLN} = 0.6 * E + 0.4 * L$
 - E: examen
 - L: Laboratorio

- Horario
 - Martes: 11:30h a 13:00h (laboratorio)
 - Jueves: 09:00h a 10:15h
 - Viernes: 11:30h a 12:45h

La motivación

- Google y el PLN