



ESCUELA SUPERIOR POLITÉCNICA ÁREA TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES

PROGRAMA DE LA ASIGNATURA

Asignatura:	PROCESAMIENTO DEL LENGUAJE NATURAL
Código:	1011002811
Titulación:	Ingeniero en Informática
Profesor/a:	José María Gómez Hidalgo
Departamento:	Sistemas Informáticos
Curso:	2006-07
Cuatrimestre/Semestre:	Primer cuatrimestre
Grupo/s:	M51
Créditos:	6
Plan:	08/11/2001

1. Course Description

Natural Language Processing, Computational Linguistics, Automatic Text Processing, Machine Translation, Information Extraction, Prolog, Logic Programming, Language Analysis, Morphology, Part-Of-Speech Tagging, Syntax, Parsing, Unification Grammars, Definite Clause Grammars, Semantics, Word Sense Disambiguation, Pragmatics, Ellipsis, Anaphora Resolution.

2. Objetivos

Una de las ramas más importantes de la Inteligencia Artificial es aquella orientada a facilitar la comunicación hombre-computadora por medio del lenguaje humano, o natural. El Procesamiento del Lenguaje Natural (PLN) es la disciplina encargada de producir sistemas informáticos que posibiliten dicha comunicación, por medio de la voz o del texto. Se trata de una disciplina tan antigua como el uso de las computadoras (años 50), de gran profundidad, y con aplicaciones tan importantes como la traducción automática o la búsqueda de información en Internet.

El objetivo principal de esta asignatura es dotar a los alumnos de conocimientos sobre las técnicas que se emplean para el desarrollo de este tipo de sistemas, abarcando tanto aspectos que conciernen a los métodos estadísticos de tratamiento del texto (resolución de la ambigüedad, etiquetado sintáctico, etc.) como los relativos a los métodos basados en conocimiento (análisis sintáctico y gramáticas de unificación, resolución de la anáfora, etc.).

3. Competencias

El tipo de cualificación exigida a un profesional para su éxito en la empresa no sólo depende de sus conocimientos académicos, sino también y de manera crítica, de sus habilidades a nivel de responsabilidad, iniciativa, liderazgo, comunicación, y otras.

El Plan de Competencias de la Universidad Europea de Madrid (UEM) ha sido diseñado con el fin de potenciar la formación de nuestros alumnos en este último aspecto, aumentando su empleabilidad y su probabilidad de éxito laboral. La UEM pone a disposición de sus alumnos una serie de cursos específicos orientados a este fin. Asimismo, los profesores asumimos el compromiso de potenciar aquellas competencias que sean más afines al tipo de asignatura, en cada una de ellas. En esta asignatura, pretendemos potenciar las siguientes competencias profesionales:

TRABAJO EN EQUIPO, por medio de la realización de prácticas en grupo.

INICIATIVA, por medio de la realización de trabajos voluntarios.

CONCIENCIA DE LOS VALORES ÉTICOS, por medio de la sensibilización de los asistentes sobre la influencia de sus actos en los demás, y de la utilización de software legal en el laboratorio.

4. Programa

1. Introducción al Procesamiento del Lenguaje Natural (PLN).
2. Léxico y morfología.
3. Técnicas de análisis sintáctico.
4. Representación del significado y semántica.
5. Interpretación contextual y pragmática.
6. Aplicaciones 1: Traducción automática.
7. Aplicaciones 2: Extracción de información.

5. Metodología

En esta asignatura se alternan las clases magistrales con ayuda de un proyector de transparencias, con las discusiones en el aula, las clases de problemas y las prácticas regladas en grupo en el laboratorio. Las metodologías más empleadas en esta asignatura son:

El APRENDIZAJE COOPERATIVO, tanto en grupos informales en las discusiones en el aula, como en grupos formales para la realización de prácticas en el laboratorio.

El APRENDIZAJE BASADO EN PROBLEMAS, planteados por el profesor para demostrar los conceptos y técnicas más relevantes en la asignatura.

6. Evaluación

La evaluación en CONVOCATORIA ORDINARIA admite dos posibilidades: evaluación continua y evaluación global.

La EVALUACIÓN CONTINUA se dirige a aquellos alumnos que asisten a más del 85% de las clases, y dispone como elementos de juicio ponderados: la participación en la clase, las calificaciones en los exámenes parciales, las calificaciones en las prácticas, y la realización de trabajos obligatorios y voluntarios. Es preciso obtener una calificación de cinco o superior en cada uno de estos apartados, para que la asignatura se considere superada. La parte teórica computa un 75% de la nota global, y la práctica un 25%. La evaluación continua prosigue hasta la CONVOCATORIA EXTRAORDINARIA, es decir, los alumnos sólo deberán examinarse de aquellos contenidos no superados en CONVOCATORIA ORDINARIA.

La EVALUACIÓN GLOBAL se dirige a aquellos alumnos que, por causa de sus faltas de asistencia, o bien por decisión propia, han perdido el derecho a la evaluación continua. Esta evaluación toma la forma de un único examen escrito sobre la totalidad de los contenidos de la asignatura, tanto teóricos como prácticos, incluyendo posiblemente los trabajos realizados por los propios compañeros. Es preciso obtener una calificación de cinco o superior en dicho examen para aprobar la asignatura.

La evaluación global en CONVOCATORIA EXTRAORDINARIA es idéntica a la evaluación global en convocatoria ordinaria.

7. Normativa específica

Además de las normas generales de la UEM (consultar la guía académica en <http://www.uem.es>), las siguientes normas son de aplicación en la presente asignatura:

Los alumnos deberán rellenar una FICHA ESPECIAL entregada por el profesor con el fin de comunicar cualquier solapamiento, provista de la firma del profesor de la asignatura con la que se solapa la presente.

Los alumnos usarán el CORREO ELECTRÓNICO proporcionado por la UEM para comunicarse con el profesor, en relación con la asignatura. Cualquier comunicación recibida desde una cuenta externa a la UEM será ignorada.

8. Bibliografía

BIBLIOGRAFÍA GENERAL

C. GAZDAR, C. MELLISH, Natural Language Processing in Prolog, Addison Wesley, 1989 (P98 .G39 Biblioteca UEM, versión en POP-11 / Disponible en <http://www.informatics.susx.ac.uk/research/nlp/gazdar/nlp-in-prolog/>).

C. MANNING, H. SCHÜTZE, Foundations of Statistical Natural Language Processing, The MIT Press, 1999 (P98.5 .S83 M36 Biblioteca UEM).

BIBLIOGRAFÍA COMPLEMENTARIA

J. ALLEN, Natural Language Understanding, Benjamin/Cummings, 1987 (QA76.7 .A45 Biblioteca UEM).

E. CHARNIAK, Statistical Language Learning, MIT Press, 1993 (P98.5 .S83 C43 Biblioteca UEM).

P. JACKSON, I. MOULINIER, Natural language processing for online applications: text retrieval, extraction and categorization, John Benjamins Pub. 2002 (QA76.9 .N38 J33 Biblioteca UEM).

B. KRENN, C. SAMUELSSON, The Linguist's Guide to Statistics, sin publicar, 1997 (Disponible en http://www.coli.uni-sb.de/~krenn/stat_nlp.ps.gz).

L. MORENO, M. PALOMAR, A. MOLINA, A. FERRÁNDEZ, Introducción al Procesamiento del Lenguaje Natural, Servicio de Publicaciones Universidad de Alicante, 1999 (P98 .I58 Biblioteca UEM).

G. SALTON, Introduction to Modern Information Retrieval, McGraw-Hill, 1983.

G. SALTON, Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer, Addison Wesley, 1989.

S. SHAPIRO, Encyclopedia of artificial intelligence, John Wiley & Sons, 1992 (Q334.2 .E53 Biblioteca UEM).

C.J. VAN RIJSBERGEN, Information Retrieval, Butterworths, 1979 (Disponible en <http://www.dcs.gla.ac.uk/Keith/Preface.html>).

I. WITTEN, E. FRANK, Data Mining: Practical Machine Learning Tools and Techniques with Java Applications, Morgan Kaufmann Publishers, 1999 (QA76.9 .D3 W58 Biblioteca UEM).

9. Enlaces

Disponibles a través de la página web del profesor (<http://www.esi.uem.es/~jmgomez>).

10. Contacto y tutorías

El horario reglado de tutorías es: Miércoles de 18:30 a 19:30.

Profesor	José María Gómez Hidalgo
Despacho	C36
Teléfono	91 211 5670
Correo electrónico	jmgomez@uem.es
Página Web	http://www.esp.uem.es/~jmgomez