

Formulario de Aprobación Curso de Posgrado 2009

Asignatura:

Gramáticas formales para el lenguaje natural

Profesor de la asignatura ¹:

Doctora Dina Wonsever, Profesora Agregada, Instituto de Computación

Profesor Responsable Local ¹:

(título, nombre, grado, Instituto)

Otros docentes de la Facultad:

Ingeniera en Computación Aiala Rosá, Asistente, Instituto de Computación

Docentes fuera de Facultad:

Licenciada Marisa Malcuori, Profesora Adjunta, Instituto de Lingüística, Facultad de Humanidades y Ciencias de la Educación

Instituto ó Unidad:

Instituto de Computación

Departamento ó Area:

Departamento de programación, grupo PLN (Procesamiento de Lenguaje Natural)

¹ Agregar CV si el curso se dicta por primera vez.

(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

Fecha de inicio y finalización:

4 de setiembre e 2009 -30 de noviembre de 2009

Horario y Salón:

lunes, miércoles y viernes de 8:00 a 10:00, salón de posgrado del InCo

Horas Presenciales:

(sumar horas directas de clase – teóricas, prácticas y laboratorio – horas de estudio asistido y de evaluación)

Se deberán discriminar las mismas en el ítem Metodología de enseñanza.

82 horas

Nº de Créditos:

13 créditos

Público objetivo y Cupos:

(si corresponde, se indicará el número de plazas, mínimo y máximo y los criterios de selección. Si no existe indicación particular para el cupo máximo, el criterio general será el orden de inscripción en el Depto. de Posgrado, hasta completar el cupo asignado)

El curso está dirigido a estudiantes de posgrado en Informática y estudiantes avanzados de Ingeniería en Computación y de Lingüística. También se aceptarán estudiantes de posgrado de especialidades afines.

Se establece un cupo de 30 estudiantes debido a que se trata de un curso intensivo en taller y trabajos prácticos, de modalidad participativa y con discusión en clase de los trabajos realizados.

La selección de estudiantes se hará mediante un sorteo entre los inscriptos.

Objetivos:

Introducción a los principales temas y aplicaciones en Procesamiento de Lenguaje Natural y al rol de las gramáticas formales en este contexto.

Conocimiento de distintos enfoques y formalismos para la descripción sintáctico-semántica del lenguaje natural.

Experimentación con herramientas y gramáticas existentes. Desarrollo de nuevas reglas o extensiones por parte del estudiante. Conocimiento de medidas de evaluación, trabajo con conjuntos de prueba.

Conocimientos previos exigidos:

Lógica de predicados.
Probabilidad y estadística.
Gramáticas formales.
Programación.

Conocimientos previos recomendados:

Metodología de enseñanza:

(comprende una descripción de las horas de clase asignadas y su distribución en horas de práctico, horas de teórico, horas de laboratorio, etc. si corresponde)

El curso es participativo y está pensado para un grupo de no más de 30 personas. Se propone como metodología: presentación de temas por parte de los docentes, discusión de ejercicios prácticos en clase, experimentación con herramientas durante el curso, desarrollo de un proyecto de laboratorio.

Se estima un total de aproximadamente 207 horas de trabajo del estudiante, desglosadas en:

- 42 horas de clase,
- 35 horas de estudio personal,
- 70 horas de trabajo de laboratorio, de las cuales 25 corresponden a consultas con los docentes.
- 60 horas de elaboración de artículo, de las cuales 15 corresponden a consultas con los docentes.

El total de horas presenciales es de 82 horas.

Forma de evaluación:

El curso se evaluará a partir de:

- un proyecto: 30%
- una prueba individual escrita: 40%
- estudio en profundidad de algún tema del curso y elaboración de un artículo: 30%

Las tres instancias son obligatorias y se requerirá como mínimo el 60% del puntaje de la prueba individual escrita. El curso se aprueba obteniendo más del 60% de los puntos totales.

Temario:

- Introducción al área de Procesamiento de Lenguaje Natural. Motivación. Aplicaciones más comunes. Enfoques simbólicos, estadísticos, híbridos.
- Conceptos básicos de gramática. Nociones básicas de gramática del español.
- Distintos modelos para gramáticas formales:
 - Gramáticas libres de contexto
 - Gramáticas de dependencias
 - Gramáticas categoriales
 - Teoría de Gobierno y Ligadura
 - Gramáticas de restricciones, HPSG
 - Gramáticas probabilistas
 - Reglas contextuales
- Semántica del Lenguaje Natural
- Análisis sintáctico. Métodos de parsing. Generación de medidas de evaluación, conjuntos de prueba.
- Aplicaciones

Bibliografía:

(título del libro-nombre del autor-editorial-ISBN-fecha de edición)

Bibliografía básica:

- *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, (2nd edition). D. Jurafsky, James H. Martin. Prentice Hall, Upper Saddle River, N.J. ISBN-10: 0131873210. 2008.



Facultad de Ingeniería Comisión Académica de Posgrado

- *Introducción a la Teoría del Gobierno y la Ligadura de Noam Chomsky*. S. Costa y M. Malcuori. Serie "Avances de Investigación", Fac. de Humanidades y Ciencias. Montevideo. 1990.
- *Implementing Typed Feature Structure Grammars*. Ann Copestake (University of Cambridge) Stanford, CA: CSLI Publications (CSLI lecture notes, number 110). ISBN 1-57586-261-1. 2002.

Bibliografía complementaria:

- *Handbook of natural Language Processing*, R. Dale, H. Moisl, H. Somers, ed., Marcel Dekker, New York. ISBN 0824790006. 2000.
- *The Logic of Typed Feature Structures*, B. Carpenter, Cambridge University Press, USA. ISBN 0521419328. 1992.
- *Natural Language Understanding*, J. Allen, Benjamin/Cummings Publishing Company. ISBN 0805303340. 1995.

Durante el curso se indicará bibliografía adicional.
