

# Programa de Asignatura

2  
as

**Nombre de la asignatura**

Análisis Semántico de Lenguaje Natural

**Créditos**

8 créditos

**Objetivos de la asignatura**

La temática del curso refiere a métodos que permiten la inserción de comunicaciones en lenguaje natural en diversos tipos de aplicaciones. Se trata de un área muy demandada y activa actualmente en el mundo, debido a la posibilidad de acceso a muy grandes volúmenes de información y a la penetración cada vez mayor de las TICs en la vida corriente de la gente.

Se pretende que el estudiante realice adquisición de conocimientos y experimentación práctica en distintos tipos de representaciones semánticas para el lenguaje natural y en métodos y procesos para su obtención a partir de enunciados lingüísticos.

Deberá conocer además cuales son los recursos existentes, las metodologías de construcción de nuevos recursos y los métodos de transferencia desde otras lenguas.

**Metodología de enseñanza**

En el curso se realizarán tres tipos de actividades: clases teórico-práctica impartidas por el docente, clases de presentación y discusión de artículos por parte de los estudiantes y realización de un proyecto con escritura de un artículo científico sobre este.

- Horas clase (teórico): 28
- Horas consulta: 4
- Horas evaluación: 4
  - Subtotal horas presenciales: 36
- Horas estudio: 30
- Horas proyecto final/monografía: 60
  - Total de horas de dedicación del estudiante: 126

**Temario**

- Revisión de nociones básicas de Procesamiento de Lenguaje Natural
- Representaciones semánticas en PLN: semántica lógica/composicional, semántica léxica, etiquetas semánticas en dependencias, semántica específica orientada a acciones un dominio.
- Anotación semántica, esquema AMR, esquemas particulares para algunos fenómenos: temporalidad, relaciones espaciales, afectividad, otros. Esquemas transversales y esquemas orientados a dominio.
- Recursos para el análisis semántico:
  - Repositorios léxicos, relaciones entre unidades léxicas: diccionarios, tesauros, bases de datos léxicas (WordNet). Repositorios para el español y el inglés
  - Corpora: texto puro, texto anotado, representaciones vectoriales.
- Construcción de representaciones semánticas para expresiones lingüísticas (*parsing semántico*):
  - Aprendizaje supervisado en CCG
  - Aprendizaje semi-supervisado
  - Algoritmos basados en grafos

## Bibliografía y referencias

Consulta (disponible en la biblioteca del InCo)

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

### Artículos

Liang, P., Jordan, M. I., and Klein, D. (2011). Learning dependency-based compositional semantics. In Proceedings of ACL-11 pages 590-599. (Disponible web, [http://www.mitpressjournals.org/doi/pdf/10.1162/COLI\\_a\\_00127](http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_a_00127))

Liang, P. and Potts, C. (2015). Bringing machine learning and compositional semantics together. *Annual Reviews of Linguistics*, 1(1):355-376. (disponible web en versión casi final, <http://web.stanford.edu/~cgpotts/manuscripts/liang-potts-semantics.pdf>)

Graph-based Algorithms for Semantic Parsing Jeffrey Flanigan, Samuel Thomson, David Bamman, Jesse Dodge, Manaal Faruqi, Brendan O'Connor, Nathan Schneider, Swabha Swayamdipta, Chris Dyer and Noah A. Smith ACL 2014 (disponible web)

Wong, Y. W. and Mooney, R. J. (2006). Learning for semantic parsing with statistical machine translation. In Proceedings of HLT-NAACL-2006, pages 439-446. (disponible web : <http://dl.acm.org/citation.cfm?id=1220891>)

*(Se verán artículos adicionales durante el curso)*

### Conocimientos previos exigidos y recomendados

#### Conocimientos previos exigidos:

Lógica de predicados  
Probabilidad y estadística  
Programación  
Procesamiento de Lenguaje Natural

#### Conocimientos previos recomendados:

Además de los conocimientos exigidos es recomendable tener conocimientos y experiencia en Aprendizaje Automático

## Anexos

### 1) Cronograma tentativo

- Mes 1 : 20 hs de clases teóricas dictadas por el docente. Asignación de artículos y proyectos a grupos
- Mes 2 : Realización del proyecto, preparación de exposición de artículo, reuniones de consulta con el docente
- Mes 3 : Exposición de artículos, exposición de proyectos, prueba escrita

### 2) Modalidad del curso y procedimiento de evaluación

El curso es presencial. Los estudiantes trabajarán en equipo.

Las actividades son:

- clases teóricas de exposición y discusión de temas, impartidas tanto por el docente como por los estudiantes (en las presentaciones de artículos)
- realización de un proyecto y artículo vinculado
- instancia individual de evaluación escrita.

El curso se evalúa en base al proyecto / artículo científico final (contribuye con 30% al puntaje total), la presentación de un artículo de la bibliografía indicada en clase (contribuye con 30% al puntaje total) y una prueba individual escrita (contribuye con 40% al puntaje total). Cada una de estas partes tiene un mínimo de suficiencia de un 60%; es imprescindible llegar a estos mínimos para la aprobación del curso.

I

### 3) Materia

#### **Ingeniería en Computación(plan 97)**

Inteligencia Artificial y Robótica

#### **Licenciatura en Computación**

Inteligencia Artificial y Robótica

### 4) Previaturas

#### **Ingeniería en Computación (plan 97)**

Una al menos de:

- Introducción al Procesamiento de Lenguaje Natural
- Gramáticas Formales para el Lenguaje Natural

#### **Licenciatura en Computación**

Una al menos de:

- Introducción al Procesamiento de Lenguaje Natural
- Gramáticas Formales para el Lenguaje Natural

Esta asignatura no adhiere a resolución del consejo sobre condición de libre

APROB. RES. CONSEJO DE FAC. ING.

de fecha 26.7.16 Exp. 060120-001525-16