



MAESTRIA EN CIENCIAS DE LA COMPUTACIÓN

Área: Bases de Datos y Recuperación de Información

Programa de Asignatura: Tópicos Selectos BD-C
(Procesamiento del lenguaje natural basado en grafos)

Código: MCOM 22233

Tipo: Optativa

Créditos: 9

Fecha: Noviembre 2012



1. DATOS GENERALES

Nombre del Programa Educativo:	Maestría en Ciencias de la Computación
Modalidad Académica:	Escolarizada
Nombre de la Asignatura:	Tópicos Selectos BD-C Procesamiento del lenguaje natural basado en grafos
Ubicación:	Tercer semestre (Optativa)

2. REVISIONES Y ACTUALIZACIONES

Autores:	Dra. Darnes Vilariño Ayala Dr. David Eduardo Pinto Avendaño Dra. Josefa Somodevilla García
Fecha de diseño:	Noviembre 2012
Fecha de la última actualización:	Marzo 2017
Revisores:	Dra. Darnes Vilariño Ayala Dr. David Eduardo Pinto Avendaño Dra. Josefa Somodevilla García
Sinopsis de la revisión y/o actualización:	Se introdujeron nuevos contenidos y se actualizó la bibliografía



3. OBJETIVOS:

Educacional: Que el alumno conozca la forma en que se combinan los algoritmos basados en grafos para el tratamiento y procesamiento de la información. Que conozca como las aplicaciones de procesamiento de texto pueden modelarse mediante grafos. Que aprenda a construir grafos que le permitan la representación, sintáctica, léxica y semántica de los textos.

General: Conocer como la teoría de grafos puede ser aplicada directamente en la representación de los textos y manejar los diferentes algoritmos que le permitan descubrir sobre los textos las diferentes relaciones que se establecen

Específicos:

1. Revisar y repasar toda la terminología de la teoría de grafos.
2. Conocer y aplicar los algoritmos basados en grafos para el descubrimiento de las relaciones que se establecen en los textos.
3. Conocer y aprender a manejar las redes de información.
4. Conocer y aprender a procesar documentos escritos en lenguaje natural mediante grafos semánticamente.
5. Conocer y aprender a manejar textos, mediante su representación sintáctica sobre grafos.
6. Ser capaz de aplicar los conocimientos adquiridos en aplicaciones del procesamiento del Lenguaje Natural.



4. CONTENIDO

Unidad	Contenido Temático/Actividades de aprendizaje
1. Introducción a la Teoría de Grafos	1.1. Terminología y notaciones 1.2. Propiedades y Tipos de Grafos 1.3. Representación de Grafos 1.4. Valores y Vectores Propios
2. Bases de conocimiento basadas en grafos	2.1. Modelos de información basados en grafos 2.2. Alimentación de una base de conocimientos basada en grafos 2.3. Búsquedas de información escrita en lenguaje natural en un grafo 2.4. Visualización de la información 2.5. Inferencia basada en grafos
3. Redes	3.1. Redes aleatorias 3.2. Redes de lenguaje 3.2.1 Redes de co-ocurrencias 3.2.2 Redes de dependencia sintáctica 3.2.3 Redes semánticas 3.2.4 Redes de similitud
4. Procesamiento de Lenguaje Natural Basado en grafos	4.1. Semántica 4.1.1 Clases semánticas 4.1.2 Detección de sinónimos 4.1.3 Distancias semánticas 4.1.4 Implicación textual 4.1.5 Desambiguación del sentido de las palabras 4.1.6 Desambiguación de nombres 4.1.7 Sentimientos y subjetividad 4.2. Sintáctica 4.2.1 Etiquetamiento 4.2.2 Analizadores de dependencias 4.2.3 Frases preposicionales adjuntas 4.2.4 Resolución de co-referencias 4.3. Aplicaciones



Bibliografía	
Básica	Complementaria
<ul style="list-style-type: none"> • Rada, M. y D. Radev. Graph-based natural Language Processing and Information retrieval. Cambridge University Press. 2011 • Manning, D. C., and H. Schütze. Foundations of statistical natural language processing. MIT Press. 1999. • Christopher M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA. 2006. • Jurafsky, Daniel, and James H. Martin. 2nd edition. Pearson New International Edition. 2013. • Ian Robinson, Jim Webber, Emil Eifrem, Graph Databases, O'Reilly Media, 2013. 	<ul style="list-style-type: none"> • Alexander Clark, Chris Fox, Shalom Lappin. The Handbook of Computational Linguistics and Natural Language Processing. Blackwell Handbooks in Linguistics, 2012. • Thomas M. Mitchell. Machine Learning (1 ed.). McGraw-Hill, Inc., New York, NY, USA. 1997. • Richard O. Duda, Peter E. Hart, and David G. Stork. Pattern Classification (2nd Edition). Wiley-Interscience. 2000.

5. CRITERIOS DE EVALUACIÓN

Criterios	Porcentaje
• Exámenes	15%
• Participación en clase	10%
• Tareas	20%
• Exposiciones	20%
• Trabajo de investigación y/o de intervención	20%
• Proyecto final	15%
Total	100%