

Escuela de Ciencias Informáticas 2016

Curso propuesto: Problemas de integración en bases de conocimiento masivas para la Web Semántica

Idioma: Castellano

Profesores:

Dra. Maria Vanina Martinez y Dr. Gerardo I. Simari

Instituto de Ciencias e Ingeniería de la Computación (CONICET–Universidad Nacional del Sur), Bahía Blanca, Argentina

Resumen del curso:

Se introducen los fundamentos lógicos de formalismos utilizados en la representación del conocimiento y razonamiento en bases de conocimiento, con un enfoque en ontologías para la Web Semántica. Los temas principales son la integración del conocimiento – que da lugar a incertidumbre e inconsistencia – y la complejidad computacional asociada a los problemas de razonamiento y respuesta a consultas en este entorno. Dada la proliferación de lenguajes y formalismos (sobre todo aquellos relacionados con la Web), es necesario dotar a los futuros profesionales de un buen entendimiento de las herramientas disponibles para el manejo de información que no puede ser explotada adecuadamente en bases de datos tradicionales. Además, el enfoque en el análisis del costo computacional del uso de estas herramientas apunta a lograr una comprensión de los compromisos entre su aplicabilidad en la práctica (para poder manejar grandes cantidad de datos) y el poder expresivo de los lenguajes de representación subyacentes.

Se asumirán conocimientos básicos de bases de datos (modelo de datos relacional, álgebra relacional, SQL), y teoría de complejidad (clases básicas como P y NP, y reducciones).

Índice breve del programa del curso:

Día 1: Introducción a las bases de datos, teoría de modelos finitos, y complejidad descriptiva (complejidad de datos, complejidad de consulta y complejidad combinada). Consultas conjuntivas: complejidad y optimización. SQL y consultas de lógica de primer orden: complejidad y límites del poder expresivo.

Día 2: Lenguajes ontológicos: Datalog, la familia Datalog+/- y lógicas de descripción. OBDA (acceso a datos via ontologías). Complejidad y límites del poder expresivo de cada lenguaje.

Día 3: Problemas de integración e intercambio de datos (*data exchange*).

Día 4: Manejo de inconsistencia en lenguajes ontológicos: diferentes semánticas para contestar consultas ante la presencia de inconsistencia. Ventajas y desventajas desde el punto de vista semántico y computacional. Tratabilidad computacional como principal obstáculo: aproximaciones y algoritmos de reescritura. Aproximaciones novedosas al problema.

Día 5: Manejo de incertidumbre en lenguajes ontológicos. Modelos probabilísticos: Lógica de Markov, Redes Bayesianas, y modelos probabilísticos computacionalmente tratables. Extensiones probabilísticas a lenguajes ontológicos (lógicas de descripción y Datalog+/-) para el modelamiento de la incertidumbre en la Web. Algoritmos para contestar consultas de ordenamiento (*ranking*), top-k, y consultas conjuntivas probabilísticas.

Bibliografía:

1. Abiteboul, Serge, Richard Hull, and Victor Vianu. *Foundations of databases*. Vol. 8. Addison-Wesley, 1995.
2. Kolaitis, Phokion. "On the expressive power of logics on finite models." *Finite Model Theory and its Applications* (2007): 27-123.
3. Johnson, David S. "A catalog of complexity classes." *Handbook of theoretical computer science 1* (1990): 67-161.
4. Lembo, Domenico, Lenzerini, Maurizio, Rosati, Riccardo, Ruzzi, Marco, and Savo, Domenico Fabio: Inconsistency-Tolerant Semantics for Description Logics. Proc. of RR 2010: 103-117
5. Rosati, Riccardo: On the Complexity of Dealing with Inconsistency in Description Logic Ontologies. Proc. of IJCAI 2011: 1057-1062
6. Lukasiewicz, Thomas, Martinez, Maria Vanina, and Simari, Gerardo I.: Inconsistency Handling in Datalog+/- Ontologies. Proc. of ECAI 2012: 558-563
7. Biennu, Meghyn: On the Complexity of Consistent Query Answering in the Presence of Simple Ontologies. Proc. of AAAI 2012
8. Gottlob, Georg, Lukasiewicz, Thomas, Martinez, Maria Vanina, and Simari, Gerardo I.: Query Answering Under Uncertainty in Datalog+/- Ontologies, *Annals of Mathematics and Artificial Intelligence*, In Press.
9. Gutiérrez-Basulto, Víctor, Jung, Jean Christoph, Lutz, Carsten, and Schröder, Lutz: A Closer Look at the Probabilistic Description Logic Prob-EL. Proc. of AAAI 2011
10. Richardson, Matthew and Domingos, Pedro: Markov logic networks. *Machine Learning* 62(1-2): 107-136 (2006)
11. Domingos, Pedro , Webb, William Austin: A Tractable First-Order Probabilistic Logic. Proc. of AAAI 2012
12. Pearl, Judea . *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Representation and Reasoning Series (1988).
13. R. Fagin, P. G. Kolaitis, R. J. Miller, L. Popa, Data exchange: Semantics and query answering, *Theor. Comput. Sci.* 336 (1) (2005) 89–124.
14. R. Fagin, B. Kimelfeld, P. G. Kolaitis, Probabilistic data exchange, *J. ACM* 58 (4) (2011) 15:1–15:55.