

Knowledge Mining Biological Network Models

S.H. Muggleton

Department of Computing

Imperial College London

Abstract: In this talk we survey work being conducted at the Centre for Integrative Systems Biology at Imperial College on the use of machine learning to build models of biochemical pathways. Within the area of Systems Biology these models provide graph-based descriptions of bio-molecular interactions which describe cellular activities such as gene regulation, metabolism and transcription. One of the key advantages of the approach taken, Inductive Logic Programming, is the availability of background knowledge on existing known biochemical networks from publicly available resources such as KEGG and Biocyc. The topic has clear societal impact owing to its application in Biology and Medicine. Moreover, object descriptions in this domain have an inherently relational structure in the form of spatial and temporal interactions of the molecules involved. The relationships include biochemical reactions in which one set of metabolites is transformed to another mediated by the involvement of an enzyme. Existing genomic information is very incomplete concerning the functions and even the existence of genes and metabolites, leading to the necessity of techniques such as logical abduction to introduce novel functions and invent new objects. Moreover, the development of active learning algorithms has allowed automatic suggestion of new experiments to test novel hypotheses. The approach thus provides support for the overall scientific cycle of hypothesis generation and experimental testing.

Bio-Sketch: Professor Stephen Muggleton FEng holds a Royal Academy of Engineering and Microsoft Research Chair (2007-) and is Director of the Imperial College Computational Bioinformatics Centre (2001-) (www.doc.ic.ac.uk/bioinformatics) and Director of Modelling at the BBSRC Centre for Integrative Modelling at Imperial College. Prof. Muggleton's career has concentrated on the development of theory, implementations and applications of Machine

Learning, particularly in the field of Inductive Logic Programming. Over the last decade he has collaborated increasingly with biological colleagues, in particular Prof Mike Sternberg, on applications of Machine Learning to Biological prediction tasks. These tasks have included the determination of protein structure, the activity of drugs and toxins and the assignment of gene function. Previous posts were as Professor of Machine Learning at the Computer Science Department, University of York (1997-2001) ; Reader in Machine Learning and Research Fellow at Wolfson College Oxford (1993-1997); EPSRC Advanced Research Fellow (1993-1997); Visiting Associate Professor (Fujitsu Chair) at the University of Tokyo. EPSRC Post-doctoral Fellow and Turing Institute Fellow (1987-1992); PhD in Artificial Intelligence Edinburgh University (1986); BSc in Computer Science Edinburgh University (1983). Professional positions: Fellow of the American Association for Artificial Intelligence (2002-), Editor-in-Chief of the Machine Intelligence series; panel member for the DTI Functional Genomics initiative (2002-2005) and the BBSRC EBI Committee (2004-2006).