

# Abductive and Inductive Inference for Integrative Systems Biology

A. Tamaddoni-Nezhad<sup>1,2,\*</sup> P. Hitchen<sup>2,3</sup> E. Kay<sup>2,4</sup> V. Lesk<sup>2,3</sup> F. Turner<sup>2,3</sup>  
A. Dell<sup>2,3</sup> C. Rawlings<sup>1,5</sup> M. Sternberg<sup>2,3</sup> B. Wren<sup>2,4</sup> S. Muggleton<sup>1,2</sup>

(1) Dept. of Computing, Imperial College London (2) Centre for Integrative Systems Biology, Imperial College London  
(3) Division of Molecular Biosciences, Imperial College London (4) Dept. of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine (5) Biomathematics and Bioinformatics Dept., Rothamsted Research Harpenden

## Abstract:

A central challenge in computational systems biology is to build models that will allow us to “read across” between different ‘omics datasets and to use diverse biological prior knowledge. In this talk we present a framework for integrative systems biology which uses inductive/abductive logic programming [1,2]. In this talk we first give an overview of abduction and induction and discuss their similarities and differences within a logic-based framework. We will review two previous projects where abduction and induction have been used for learning metabolic network models from biological data and prior knowledge [3,4,5]. We will also report on our current research in which we use inductive/abductive logic programming to build predictive models of genotype-phenotype relations in *Campylobacter jejuni*. These models will be used to test the hypothesis that mutations in *Campylobacter jejuni* result in changes in the surface glycans and therefore changes the host responses to the infection. This modeling involves two main parts: (i) a machine learning technique which can infer hypotheses about the genome of the mutants from the observed changes in the surface glycome and (ii) a cycle of model generation/revision and experimentation which allows the predictions of the model to be experimentally tested and the model to be revised accordingly. Inductive/abductive logic programming will be used to infer hypotheses from observed changes in metabolites, glycans and gene expressions profiles (empirical data) together with background knowledge which includes known metabolic networks, glycan structures and enzyme/gene functions extracted from biological databases such as KEGG [6]. We present the results from an initial inductive/abductive model for genotype-phenotype relation in *Campylobacter jejuni* and discuss the future plans.

## References:

1. S.H. Muggleton. Inductive Logic Programming. *New Generation Computing*, 8(4):295-318, 1991.
2. A.C. Kakas, R.A. Kowalski, and F. Toni. Abductive Logic Programming. *Journal of Logic and Computation*, 2(6), 719-770, 1992.
3. C.H. Bryant, S.H. Muggleton, S.G. Oliver, D.B. Kell, P. Reiser, and R.D. King. Combining inductive logic programming, active learning and robotics to discover the function of genes. *Electronic Transactions in Artificial Intelligence*, 5-B1(012):1-36, November 2001
4. R.D. King, K.E. Whelan, F.M. Jones, P.K.G. Reiser, C.H. Bryant, S.H. Muggleton, D.B. Kell, and S.G. Oliver. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*, 427:247-252, 2004.
5. A. Tamaddoni-Nezhad, R. Chaleil, A. Kakas, and S.H. Muggleton. Application of abductive ILP to learning metabolic network inhibition from temporal data. *Machine Learning*, 64:209-230, 2006.
6. M. Kanehisa, S. Goto. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28:27-30, 2000.

\* Email address for correspondence: atn@doc.ic.ac.uk