Avd. Matematisk statistik

KTH Teknikvetenskap

A GRADUATE COURSE IN BAYESIAN NETWORKS and STATISTICAL GENETICS (7.5 p.):

lp 4 2010

FIRST LECTURE

Day: TUESDAY 30TH OF MARCH, 2010
Time: 10.15 A.M.
Room: ROOM 3733. Institutionen för matematik, KTH,

COURSE DESCRIPTION

Bayesian networks provide compact and powerful representations for expressing joint probability distributions and for doing probabilistic inference. They have become important in the biological sciences, e.g., for inferring gene regulatory networks and in analysis of genetic data like gene expression data.

The course is of interest for geneticists, computer scientists and others who work with, e.g., modelling and analysis of multilocus local dependencies, or modelling of highly complex systems. The course presents the algorithms for updating of probabilistic uncertainty in response to evidence, and statistical learning of model parameters and structures.

- One brief general description of the topics is in C.J. Needham, J.R. Bradford, A.J. Bulpitt, D.R. Westhead: A Primer on Learning in Bayesian Networks for Computational Biology. *PLoS Computational Biology*,
  `http://www.cs.rice.edu/~nakhleh/COMP572/Papers/Lecture5-2.pdf`

- A field of application is described in
  C.J. Verzill, N. Stallard, and J.C. Whittaker: Bayesian Graphical Models for Genomewide Association Studies. *The American Journal of Human Genetics*, 2006, Vol. 79, pp. 100−112.

The course is preliminarily recommended to award 7.5 p..

COURSE SYLLABUS: The course will consist of three parts: (I) an introduction to probabilistic methods and concepts and learning theory associated with Bayesian networks, and (II) basics of Bayesian networks and (III) further properties of Bayesian networks.

(I)
- Probability basics:conditional probability, probabilistic independence, Bayes formula
- Statistics basics: maximum likelihood, Markov chain Monte Carlo

(II)
- definition, directed acyclic graphs, and basic representations
- causality
- inference in Bayesian networks

(III)
- and d-separation, conditional independence
- Markov properties for directed acyclic graphs and faithfulness.
- Learning about probabilities
- Structural learning; MDL, predictive inference
- Causality and intervention calculus

(IV) Possible additional topics

PREREQUISITES: basic probability and statistics (will be reviewed)

LITERATURE: The course textbook is

- R.E. Neapolitan: *Probabilistic Methods for Bioinformatics with an Introduction to Bayesian Networks* . Morgan and Kaufman, 2009. ISBN-13: 978-0123704764

Additional material will be distributed during the lectures or made available on the course home page. Further material is found in

- T. Koski & J. Noble: *Bayesian Networks. An Introduction* . John Wiley & Sons, 2009. ISBN-13: 978-0470743041

HOME PAGE: The home page of the course is

- `http://www.math.kth.se/~tjtkoski/banetki2010`

EXAMINATION: Home assignments, presentations.

COURSE LEADER Prof. Timo Koski, Inst. för matematik, KTH, tjtkoski@kth.se, tfn: