Introduction to Probabilistic Graphical Models

The actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have to reason on. Therefore the *true logic* for this world is the calculus of *Probabilities*, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable mans mind. – James Clerk Maxwell

Study group web site: `http://research.ee.sun.ac.za/pgms/`

Note: Since I'll be updating the instructions as the course proceeds, the updated instructions might only become available directly before/at the time of the meeting.

Organised by: JA du Preez, Phone: x4342, Email: dupreez at sun

Language: Afrikaans or English, depending on participants.

Format: This is intended as a guided self-study, with group support, to get you started in this field. It is NOT a formal offering by US, there is NO formal evaluation and you will NOT be able to get formal credits for this. There will be weekly reading material and exercises (that will involve coding) to do, followed by a discussion at the next session. The work load should be around 5-10 hours per week.

Prerequisites: You will need to be familiar with basic probability theory and should be able to write and use software – the suggested PGM toolbox is written in python.

Content: We will cover the following on an introductory level:

- **Representation:** Reasoning patterns, Bayes Nets, Markov Random Fields, Log-linear models, Chain graphs, Templates and Temporal models, Conditional Random Fields
- **Inference:** Variable Elimination, Sum Product, Max Product, Max Sum, Junction Tree, Variational techniques
- **Learning:** Maximum Likelihood, Maximum Posterior, Bayesian Learning. If time allows, graphical structure learning.

Primary resources:

- **Book:** Bayesian Reasoning and Machine Learning, David Barber, Cambridge University Press, 2012. Free pdf download at `http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Brml.Online`
- **Video lectures:** Koller's PGM lecture series is a very useful resource. You can access them either through the Coursera course, or retrieve them from YouTube `https://youtu.be/6ODl1rxoT14?list=PL50E6E80E8525B59C`
- **Code:**
  - **python:** We will make use of python-based `pgmpy` toolbox. You can find it on github at `https://github.com/pgmpy`, documentation at `http://pgmpy.org/`. For this you will also need a good up to date jupyter/ipython3 installation. Get it at `https://www.continuum.io/downloads`.

    Note: I am currently unsure as to how well `pgmpy` will stack up to more serious computational tasks. Hopefully we will establish this as we proceed with the course. However, it is a splendid vehicle for the small educational problems we use in this course.
  - **c++:** If you have an appetite for more serious computational tasks, you might want to consider the Stellenbosch university's `emdw` toolbox. For this you will need Linux and a fairly new g++ compiler. Contact me (Johan du Preez) to get access to the toolbox.
  - **octave:** We have used Barber's `BRML` toolbox in previous runs of this course. If you do want to go this way you can download it from `http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Brml.Software`.

Other resources:

- **Online course:** Duration 10 weeks, you will need to invest about 20 hours per week. See `https://www.coursera.org/` for the next round of it.
- **Supplementary book 1:** Probabilistic Graphical Models: Principles and Techniques, Daphne Koller and Nir Friedman, MIT Press, 2009.
- **Supplementary book 2:** Information Theory, Inference and Learning Algorithms, David JC MacKay, Cambridge University Press, 2003. Free pdf download at `http://www.inference.phy.cam.ac.uk/mackay/itila/book.html`