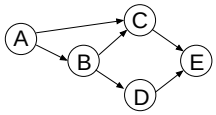
CSC412: Uncertainty & Learning in AI

- Instructor: Sam Roweis
- Tutor: Y.W. Teh
- Lectures: MW10, Pratt 266
- Tutorial: F10, Pratt 266
- 4 Assignments (50%), late work not accepted
- Midterm Test (25%), Final Test (25%), NO EXAM
- some programming in MATLAB on CDF
- text: draft of Jordan: *Introduction to Graphical Models*
- http://www.cs.toronto.edu/~roweis/csc412/
- csc412@cs.toronto.edu
- Office Hour: ?

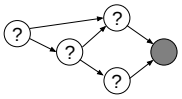
Uncertainty in AI

- Probabilistic methods can be used to:
  - make decisions given partial knowledge of the world
  - account for noisy sensors or actuators
  - explain phenomena not part of our models
  - describe inherently stochastic behaviour in the world



The Power of Learning

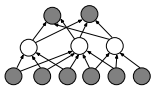
- Databases
  - traditional DB technology cannot answer queries about items that were never in the dataset
  - graphical models are like probabilistic databases



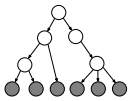
- Expert Systems
  - old expert systems needed hand coding of knowledge and of output semantics
  - probabilistic models automatically construct rules and support all types of queries

Three Main Problems

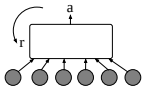
- Supervised Learning:
  - classification
  - regression/prediction



- Unsupervised Learning
  - density modeling
  - clustering/grouping
  - dimensionality reduction



- Reinforcement Learning



Machine Learning Applications

- Expert Systems
  - Medical diagnosis and image analysis
  - Troubleshooting and fault correction
  - Plant and vehicle monitoring (e.g. space shuttle Vista)
  - Freeway traffic
- Pattern Recognition and Prediction
  - Speech and handwriting recognition
  - Credit card fraud, insurance premiums
  - Biological sequence classification
- Information Retrieval
  - search (e.g. Google)
  - collaborative filterting (e.g. Amazon)

Other Related Applications

- Stochastic Signal Processing
  - denoising, source separation, scene analysis, morphing
- Data compression & coding
  - state of the art models use learning methods
- Decision Making, Planning
  - use both utility and probability, e.g. influence diagrams
- Adaptive Software Agents
  - action choice under limited resources and reward signals

Representation

- Key issue: how do we represent info about the world?
  - E.g. for an image, do we list all the pixel values? If so, in what order? Does it matter?

→ 127, 254, 3, 18, ...

- Must **exploit regularities** in the data
  - we will focus on "conditional independence" properties

Model Choice & Prior Assumptions

- You can't learn anything about the world without making some assumptions.
  (Although you *can* memorize what you see.)
- Representation is one kind of assumption.
- The model we chose is equally important, and closely related to the choice of representation.

In this course

- Using probabilities to represent beliefs
- Graphical models as structured representations of large joint probability distributions.
- Statistical parameter estimation for simple classification, regression and density models.
- Inference of hidden variables.
- General parameter learning for latent variable models.
- EM Algorithm
- Junction Tree Algorithm
- Mixture models, factor analysis, PCA, HMMs