UBC MLRG (Summer 2016): Miscellaneous

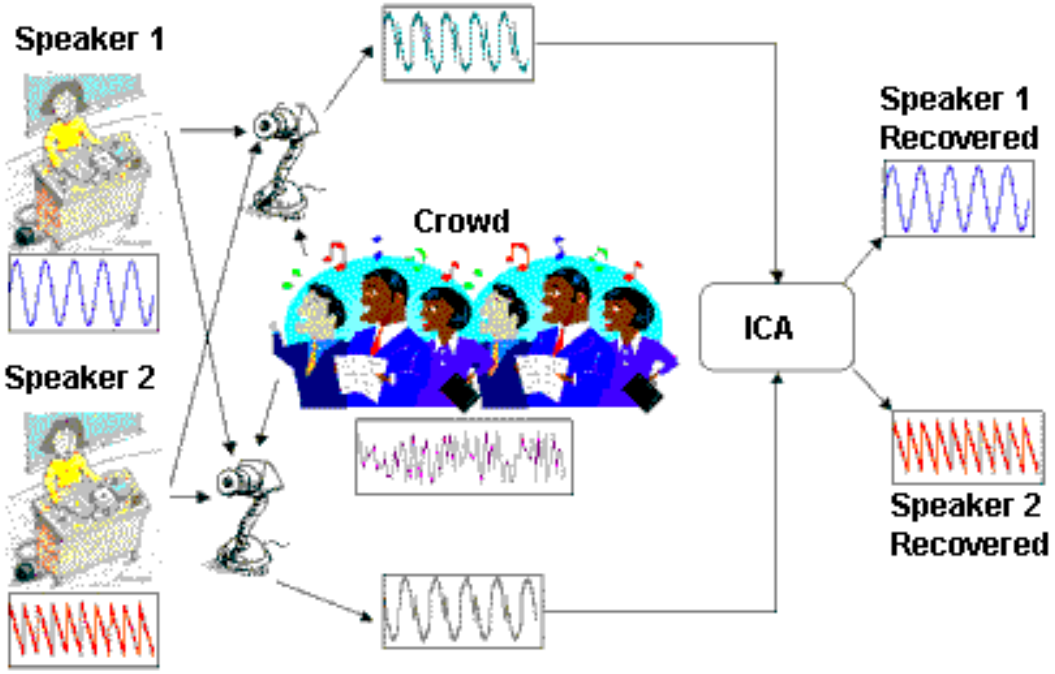
Some images from this lecture are taken from Google Image Search.

Recent MLRG History

- Topics covered in recent MLRG terms:
  - Fall 2014: Deep learning and Bayesian optimization.
  - Winter 2015: Causality, bandits, reinforcement learning.
  - Summer 2015: Probabilistic graphical models.
  - Fall 2015: Convex optimization.
  - Winter 2016: Bayesian statistics.
- Summer 2016 proposal: miscellaneous!
  - Major topics not previously covered in CPSC 540 or recently in the MLRG.
    - But that we don't want to dedicate a semester to.
  - Plan is to spend 2 weeks on each topic.
- Today:
  - Overview of these topics, and getting people to choose topics/weeks.

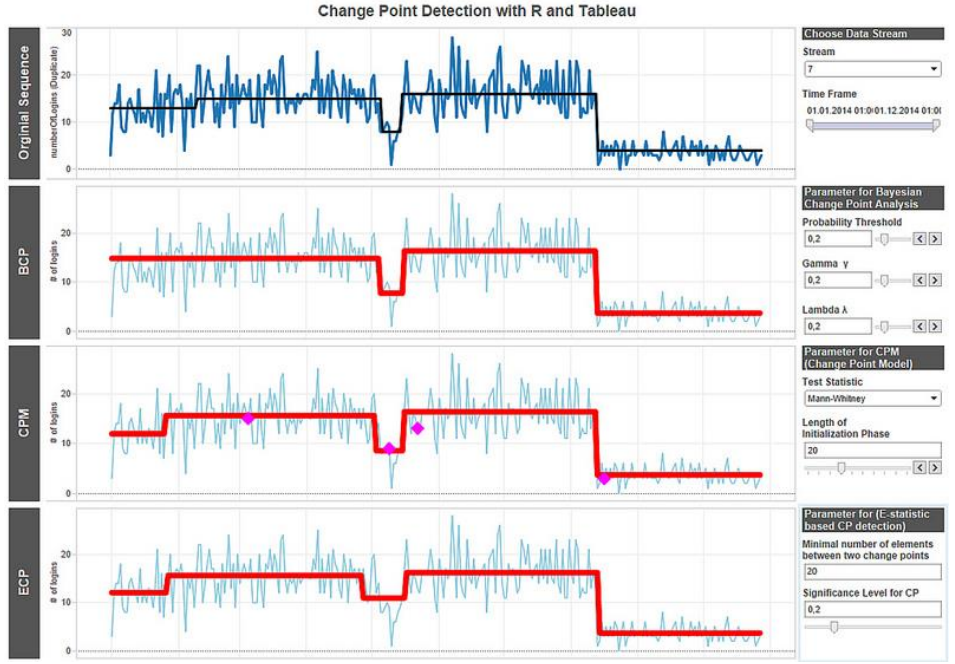
Independent Component Analysis (ICA)

- 540 covers probabilistic PCA and factor analysis:
  - Latent factors follow an independent Gaussian distribution.
- Independent component analysis:
  - Latent factors follow and independent non-Gaussian distribution.
- Key advantage:
  - Latent factors become identifiable (up to label switching).
- Key applications:
  - Source separation.
  - Causality.



Changepoint Detection

- 540 covers hidden Markov models:
  - Finds latent “modes” and latent dynamics in time-series.
- Changepoint detection:
  - Task of finding breakpoints where time series distribution changes.
- Key advantage:
  - May be easier than fitting generative HMM.
- Key applications:
  - Bioinformatics.
  - Econometrics.



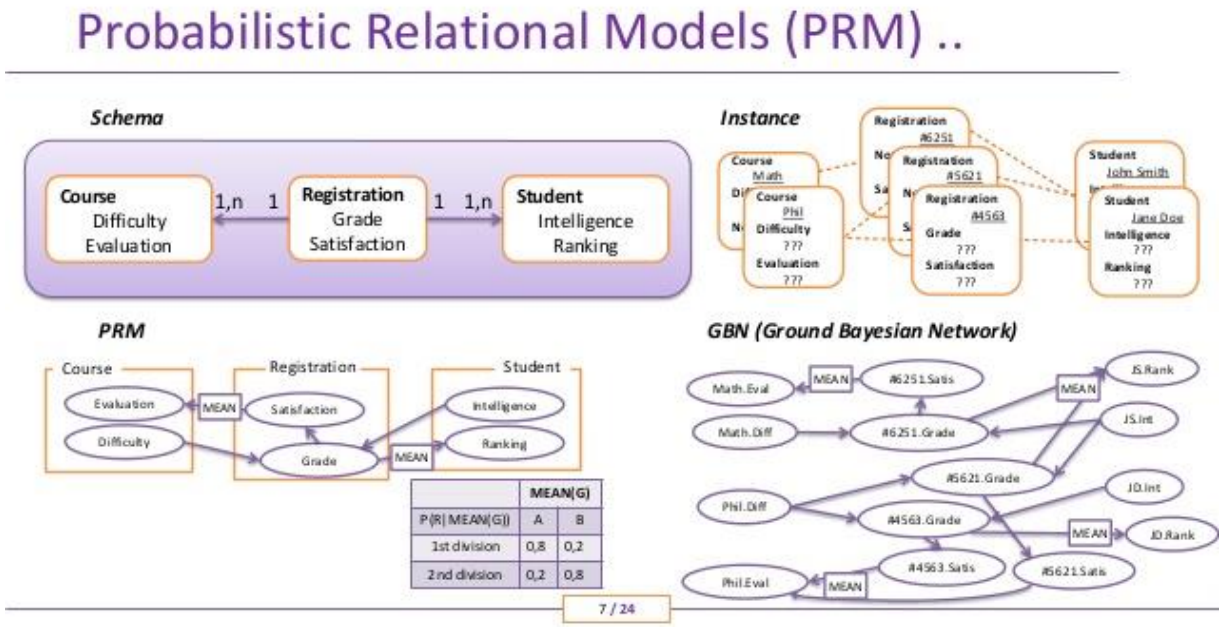
Sub-Modularity

- 540 covers convexity:
  - Class of continuous functions that is easy to minimize.
- Sub-modularity:
  - Class of discrete functions that is easy to minimize.
  - Algorithms also exist for approximate maximization.
- Key advantage:
  - Expands class of efficiently-solvable problems.
- Applications:
  - Computer vision.
  - Sensor networks.
- Related: QUBO.

Figure 1 From a set of five source images (of which four are shown on the left), we quickly create a composite family portrait in which everyone is smiling and looking at the camera (right). We simply flip through the stack and coarsely draw strokes using the *designated source* image objective over the people we wish to add to the composite. The user-applied strokes and computed regions are color-coded by the borders of the source images on the left (middle).

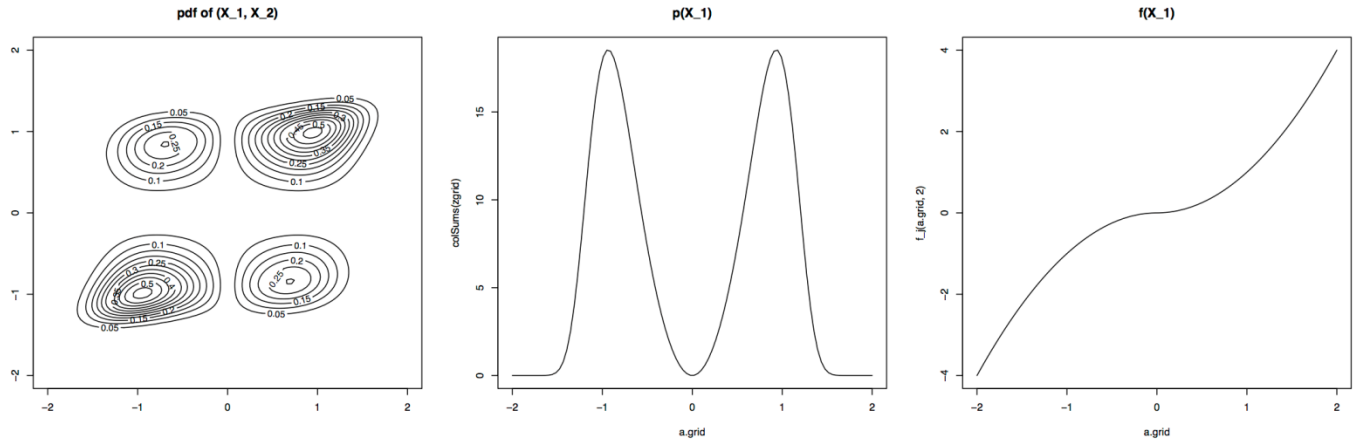
Relational Models

- 540 covers Bayesian networks:
  - Describes relationships between variables.
- Probabilistic relational models and Markov logic networks:
  - We have “types” of variables and probabilities on logical statements.
- Key advantage:
  - More expressive language.
- Applications:
  - Adding probabilities to databases.
- Related: sum-product networks.

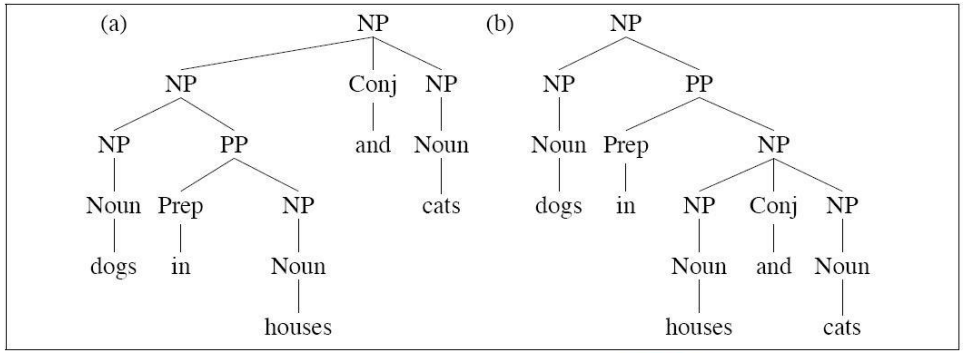


Continuous Graphical Models

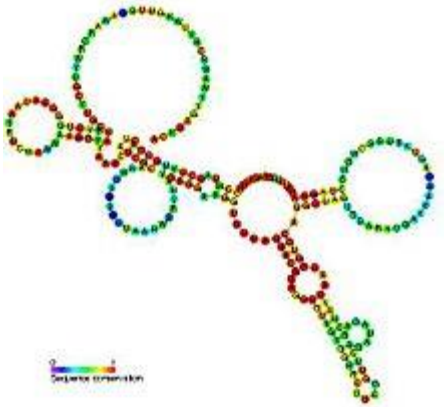
- 540 covers discrete and Gaussian graphical models:
  - Convenient due to conjugacy.
- Nonparanormal and Gaussian-copulas models:
  - More flexible continuous distributions.
- Key advantage:
  - Gaussians are very restricted class.
- Applications:
  - Stock market crash of 2008.



Grammars



- 540 covers Markov models:
  - Useful for modeling sequence data with Markov assumption.
  - Can be generalized to Bayesian networks.
- Probabilistic context-free grammars:
  - Different generalization using “recursive” Markov assumption.
- Key advantage:
  - Dependencies at different scales.
- Applications:
  - Natural language processing.
  - RNA secondary structure.
- Related: And-Or trees, image grammars.



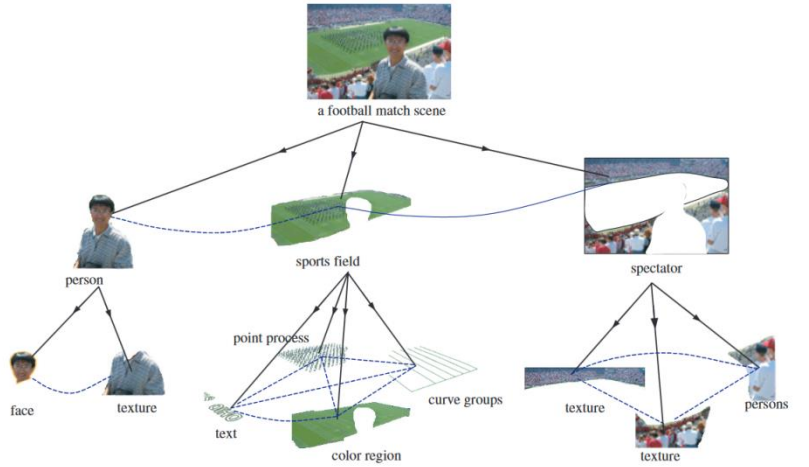
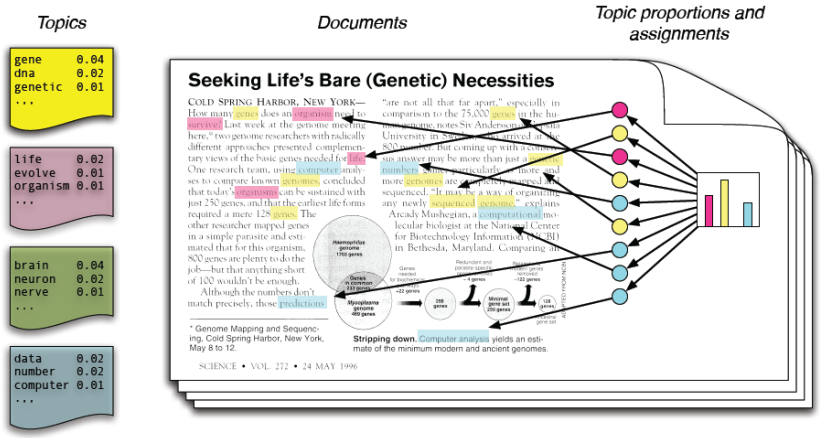


Fig. 1.1 Illustrating the task of image parsing. The parse graph includes a tree structured decomposition in vertical arrows and a number of spatial and functional relations in horizontal arrows. From [72].

Topic Models

- 540 covers density estimation.
- Topic models:
  - Hierarchical density estimation.
- Key advantage:
  - Structure at different scales (document vs. word).
- Applications:
  - Document modeling/clustering/analysis.
- Related: more non-parametric Bayes.



Spectral Methods

- 540 covers expectation maximization:
  - Only finds global optimum.
- Spectral methods:
  - Consistent estimators.
- Key advantage:
  - No local minima if you have enough data.
- Applications:
  - All the usual HMM applications.

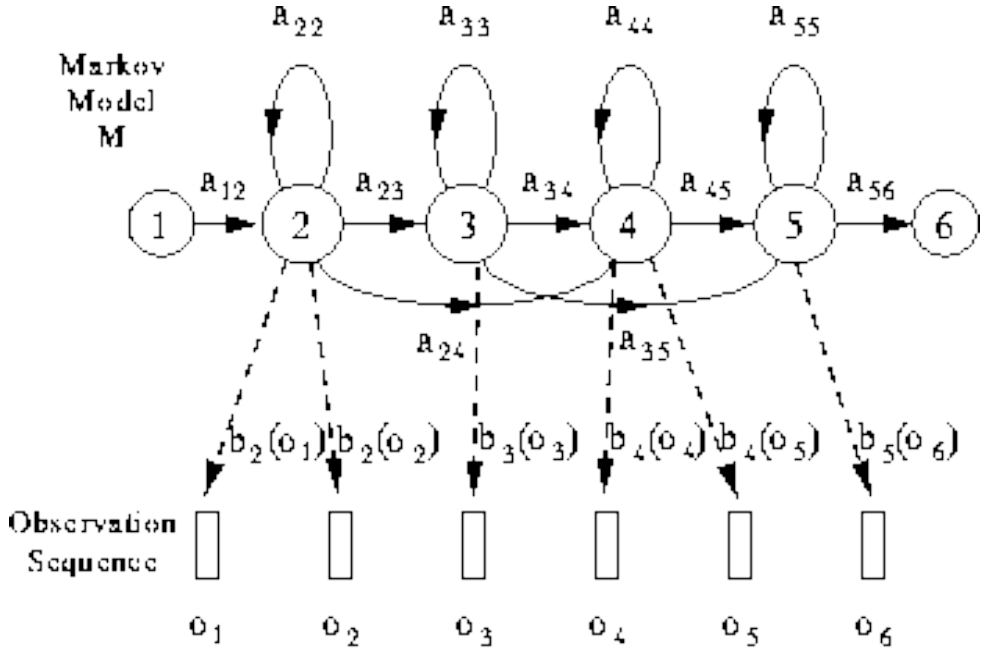
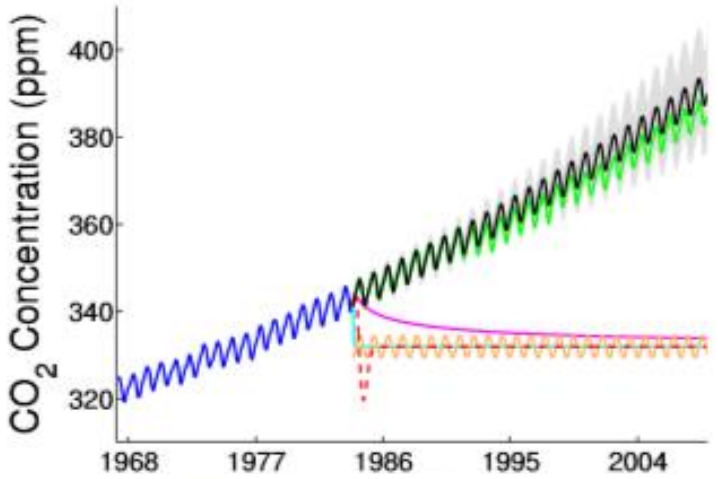
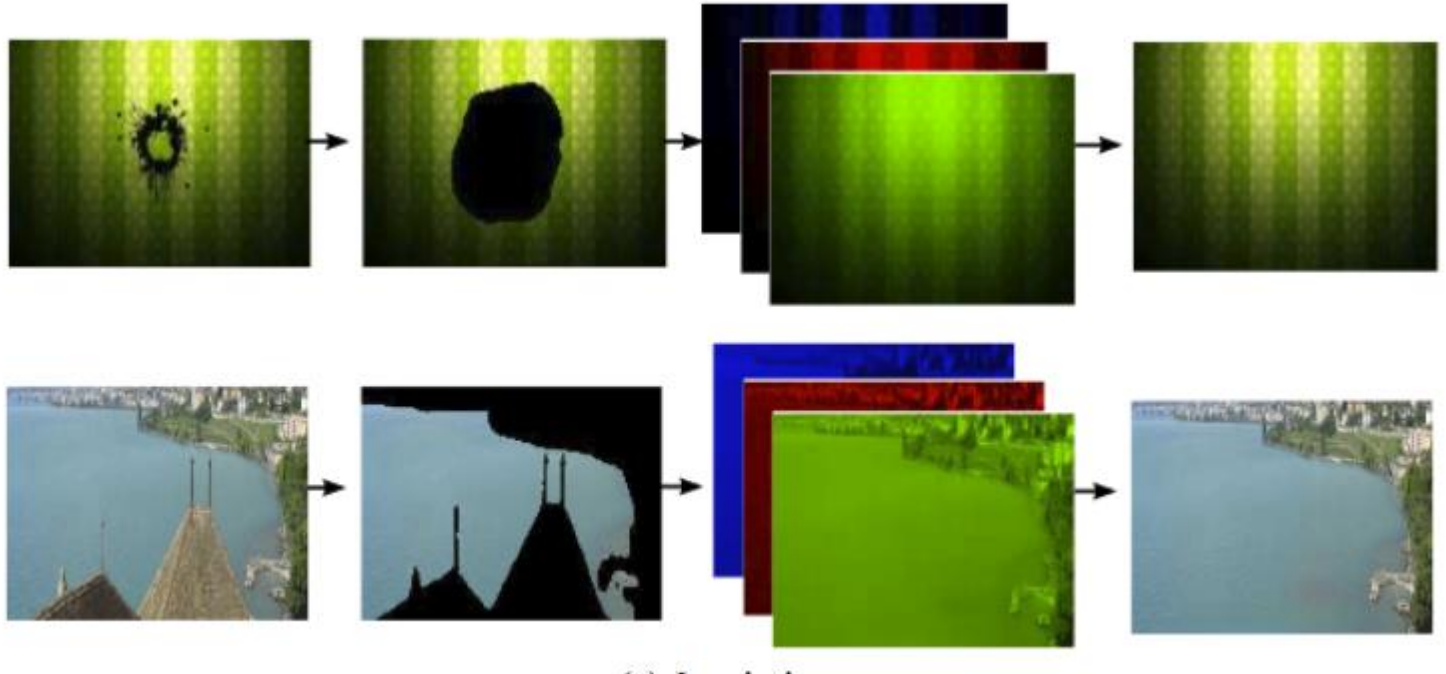


Fig. 1.3 The Markov Generation Model

Large-Scale Kernels Methods

- 540 covers kernel methods:
  - Flexible universal approximators.
  - But $O(n^2)$ or worse cost/storage.
- Large-scale methods:
  - Tricks to get O(n log n) c
- Big literature:
  - Nystrom.
  - Fast multipole.
  - Kronecker products.
  - Circulant matrics.





Topics/Schedule

- June 1: No meeting (UAI camera-ready deadline)
- June 8, 15: Spectral Methods (Sharan and Geoff)
- June 22: Relational Models (Chris)
- June 29: Submodularity (Saif)
- July 6: Grammars (Nasim)
- July 13, 20: Continuous graphical models (Eviatar, Steven, Kevin)
- July 27, August 3: Large-scale kernels methods (Issam, Julieta)
- August 10, 17: Chamgepoint detection (Mohamed, Alireza)
- August 24, 31: Independent component analysis (Julie, Ricky)