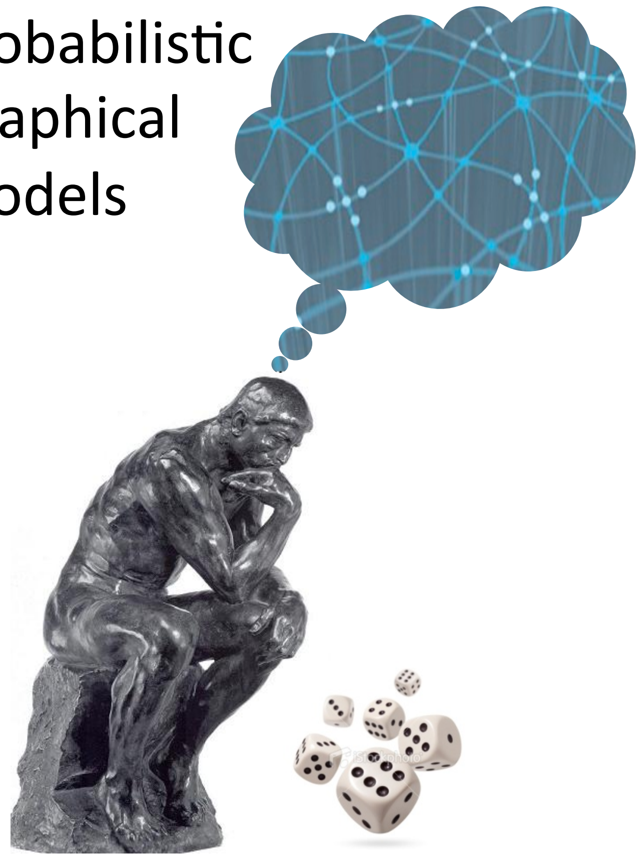
Probabilistic
Graphical
Models

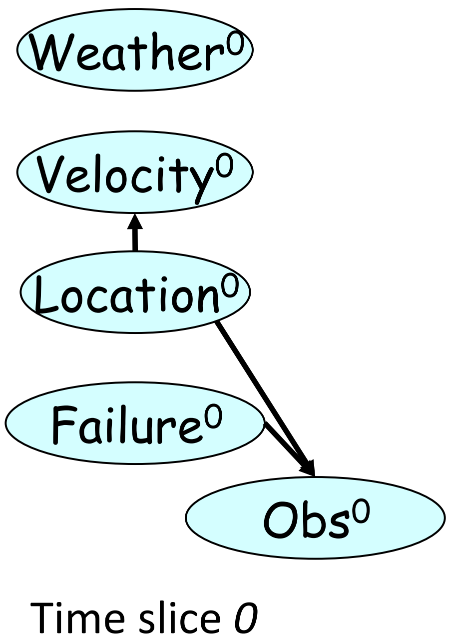


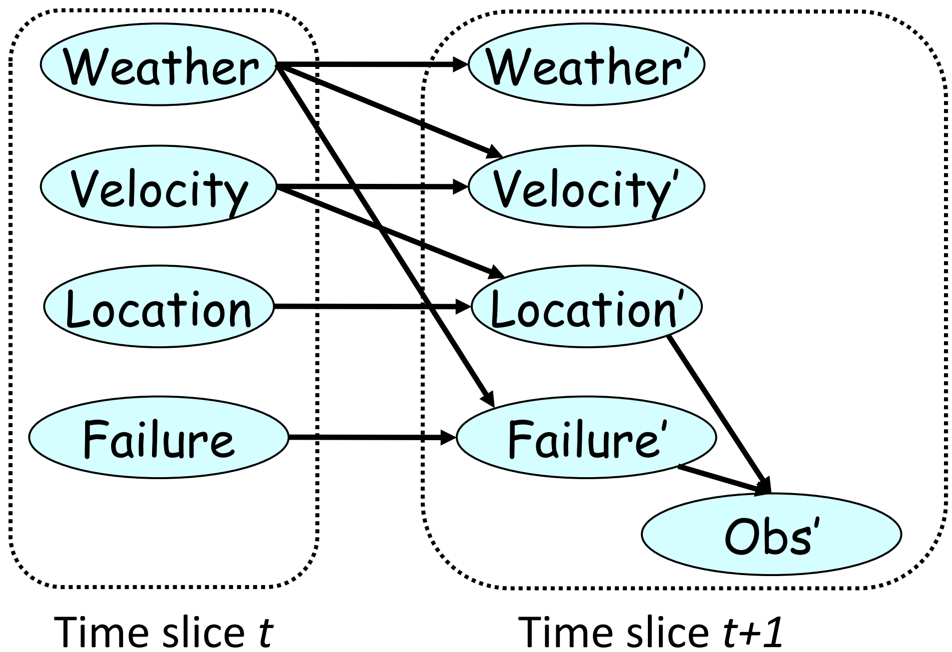
Inference

Template Models

# Inference in Template Models

# DBN Template Specification

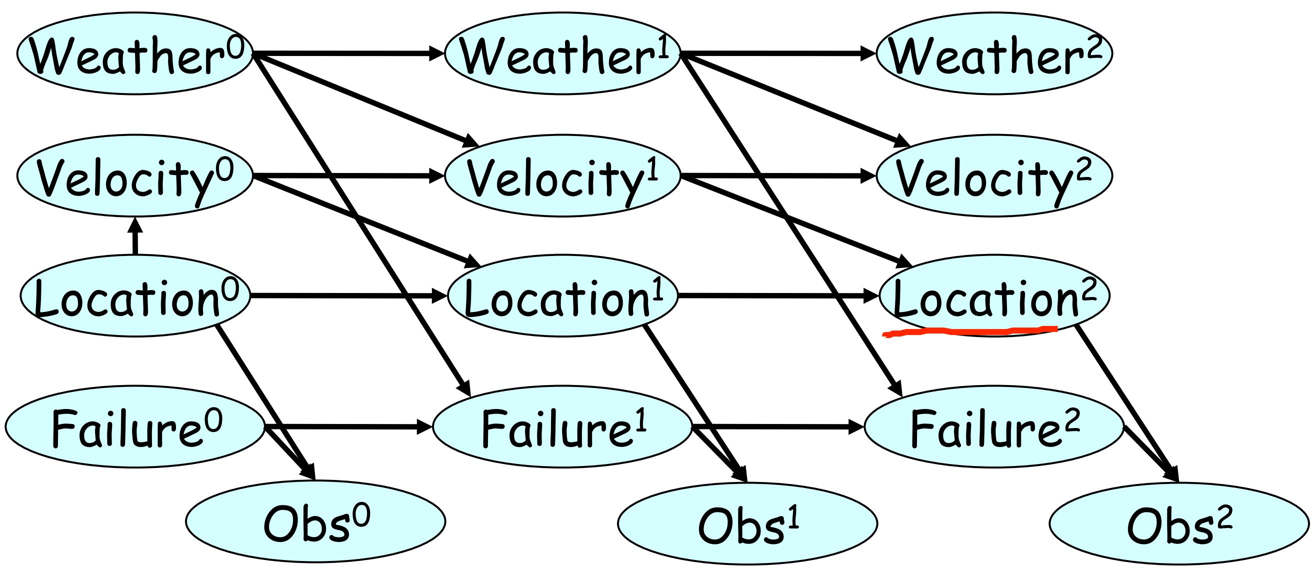



2TBN

Daphne Koller

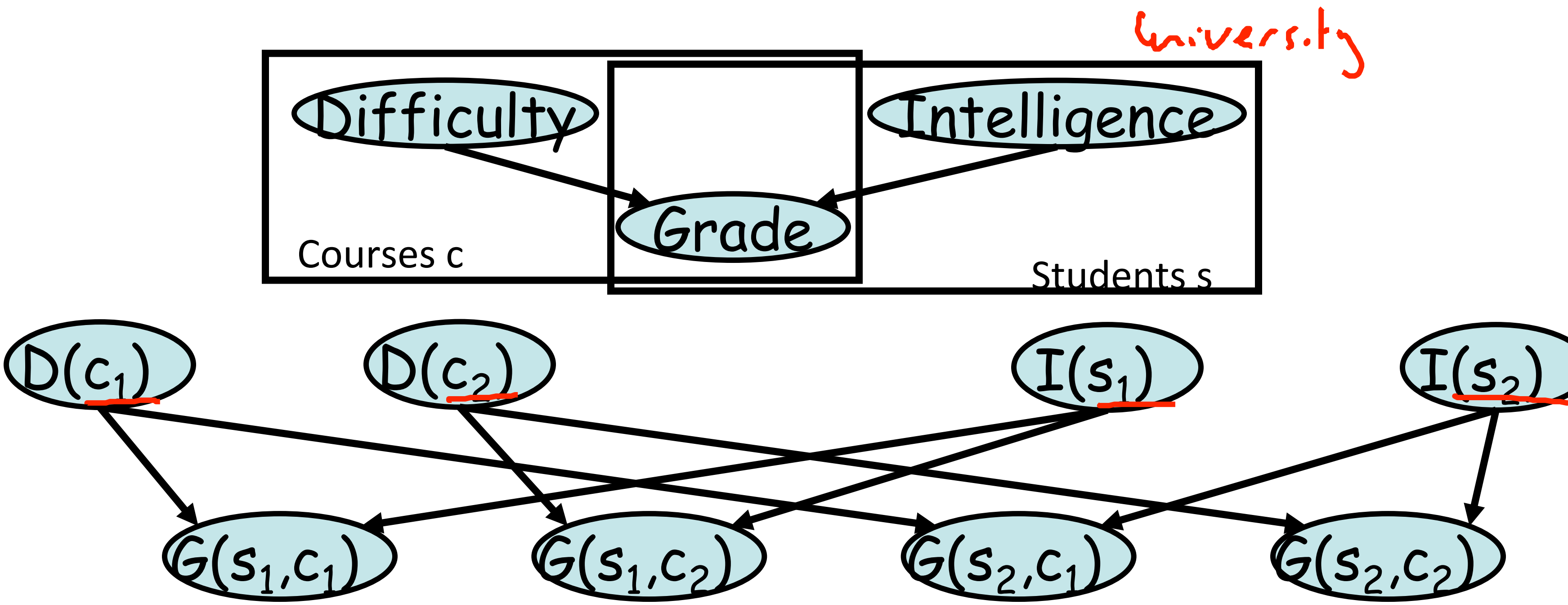
# Ground Bayesian Network




Can unroll DBN for given trajectory and run inference over ground network

Daphne Koller

# Plate Model



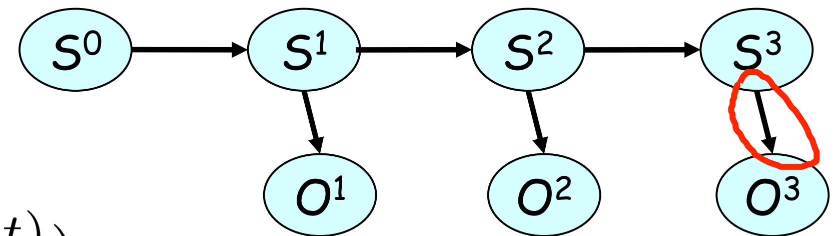

Can unroll plate model for given set of objects and run inference over ground network

Daphne Koller

# Belief State Tracking

$o^1, o^2, \ldots, o^t$

$$\sigma^{(t)}(S^{(t)}) = P(S^{(t)} \mid \boldsymbol{o}^{(1:t)})$$




t

$$\sigma^{(\cdot t+1)}(S^{(t+1)}) \triangleq P(S^{(t+1)} \mid o^{(1:t)})$$

$$= \sum_{S^{(t)}} P(S^{(t+1)} \mid S^{(t)}, o^{(1:t)}) P(S^{(t)} \mid o^{(1:t)})$$

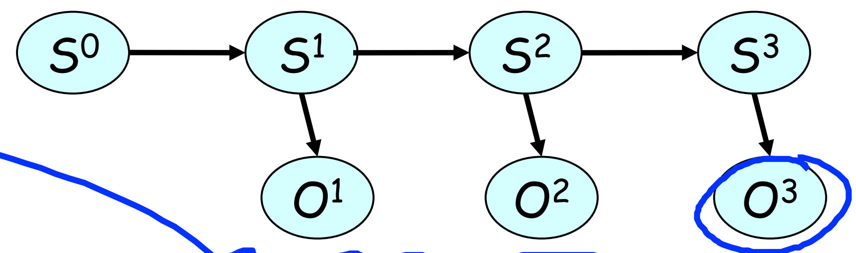
$$= \sum_{S^{(t)}} P(S^{(t+1)} \mid S^{(t)}) \sigma^{(t)}(S^{(t)})$$

transition model

# Belief State Tracking

$$\sigma^{(t)}(S^{(t)}) = P(S^{(t)} \mid \boldsymbol{o}^{(1:t)})$$

$$\sigma^{(\cdot t+1)}(S^{(t+1)}) \triangleq P(S^{(t+1)} \mid o^{(1:t)})$$




$$\sigma^{(t+1)}(S^{(t+1)}) = P(S^{(t+1)} \mid o^{(1:t)}, o^{(t+1)})$$

$$= \frac{P(o^{(t+1)} \mid S^{(t+1)}, \cancel{o^{(1:t)}})P(S^{(t+1)} \mid o^{(1:t)})}{P(o^{(t+1)} \mid o^{(1:t)})}$$

observation

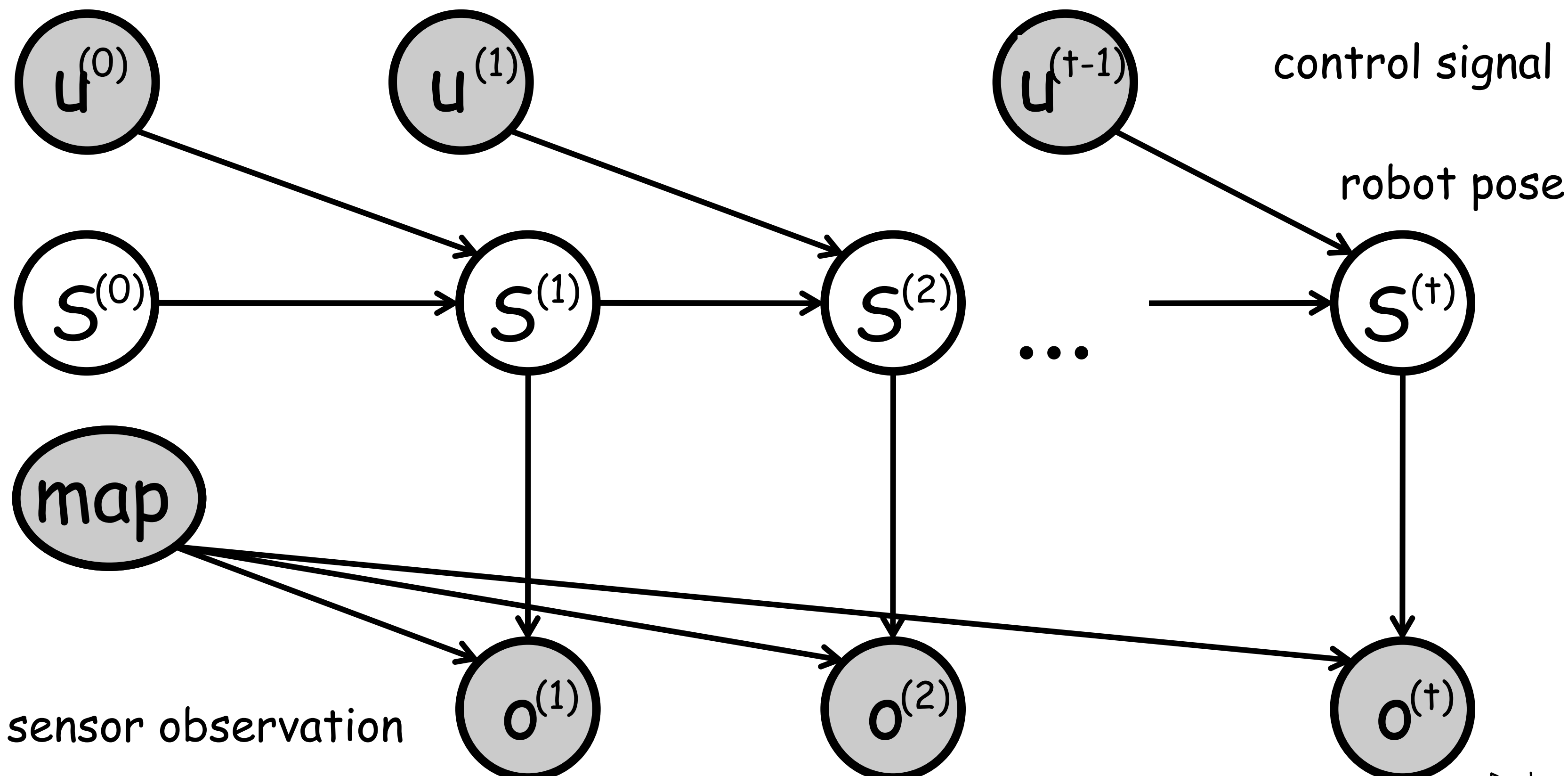
$$= \frac{P(o^{(t+1)} \mid S^{(t+1)})\sigma^{(\cdot t+1)}(S^{(t+1)})}{P(o^{(t+1)} \mid o^{(1:t)})}$$

numerator + normalizing

normalizing constant

Daphne Koller

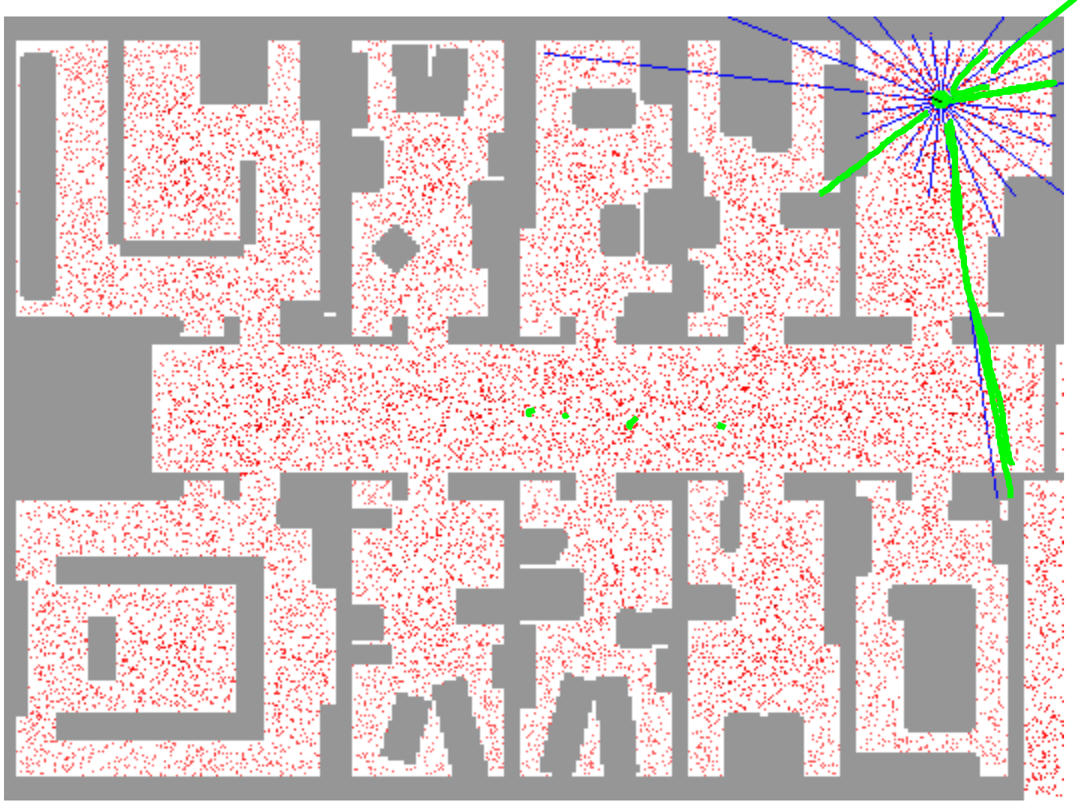
# Robot Localization




Daphne Koller

# Robot Localization

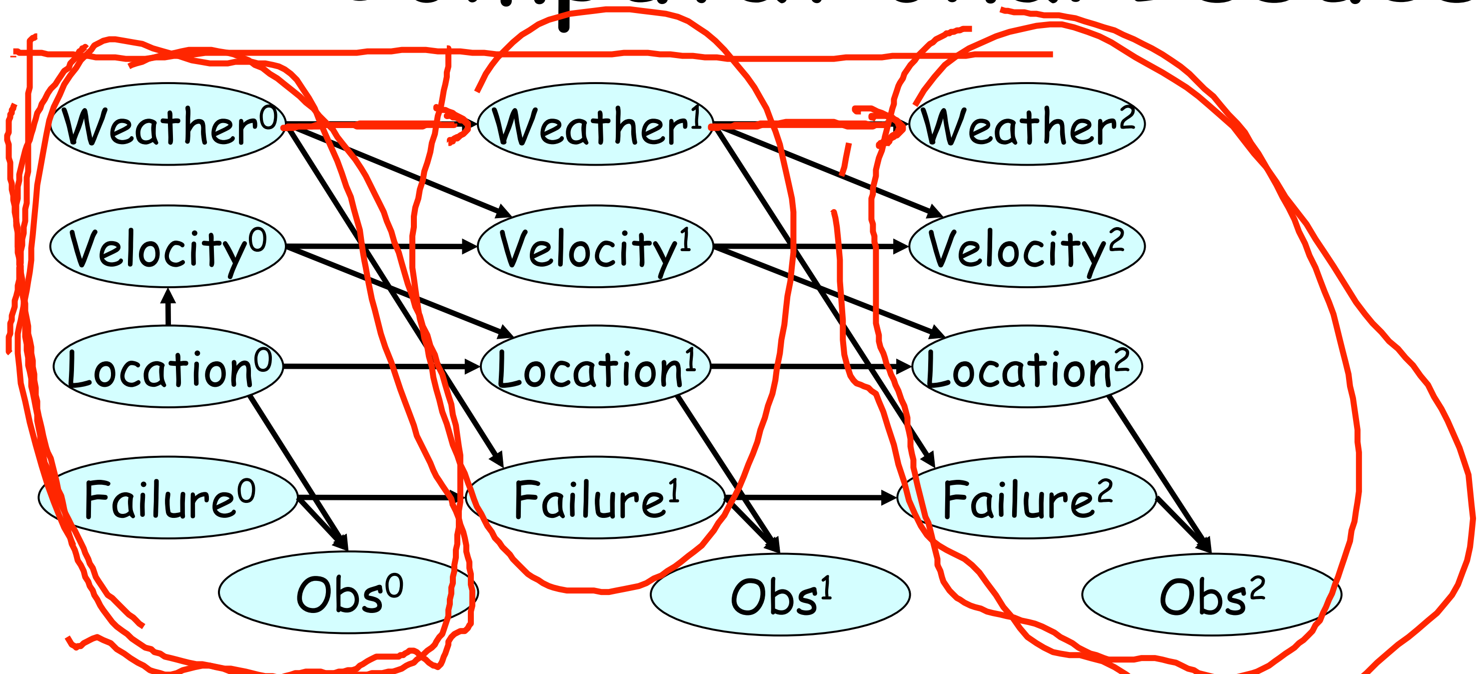
robot true position



Fox, Burgard, Thrun

Daphne Koller

# Computational Issues

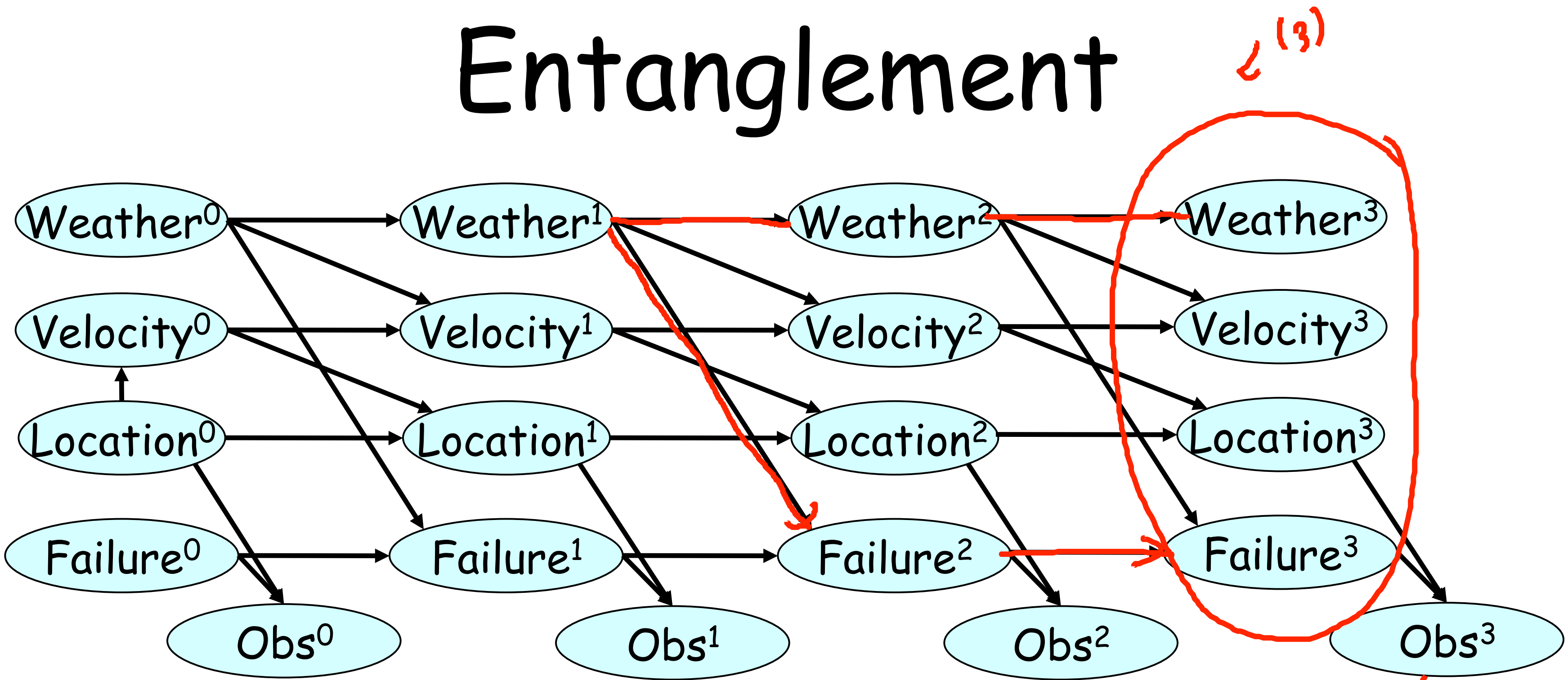



Minimal sepset must separate future from past

⇨ must involve at least all of the persistent variables

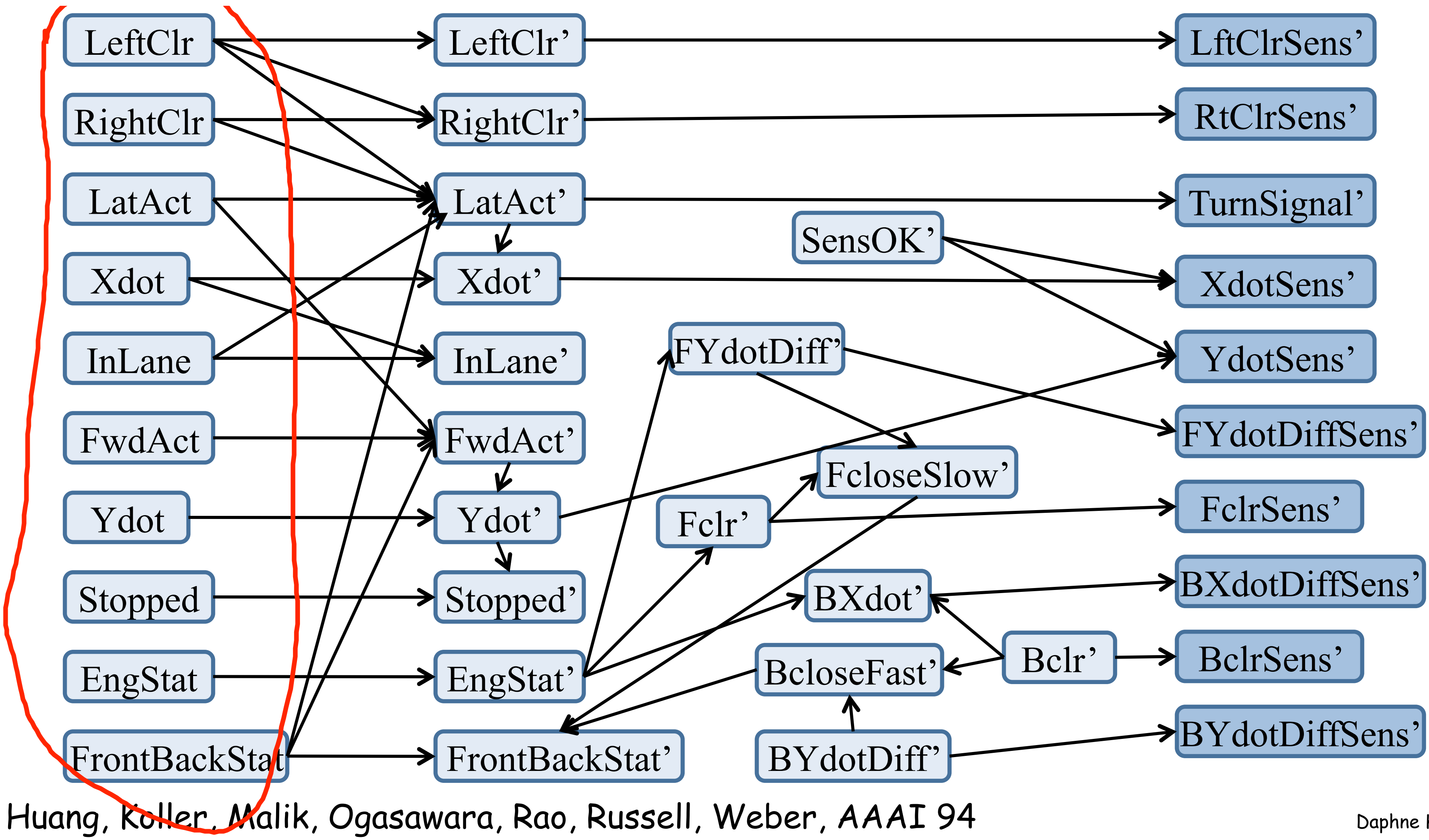
Daphne Koller

# Entanglement



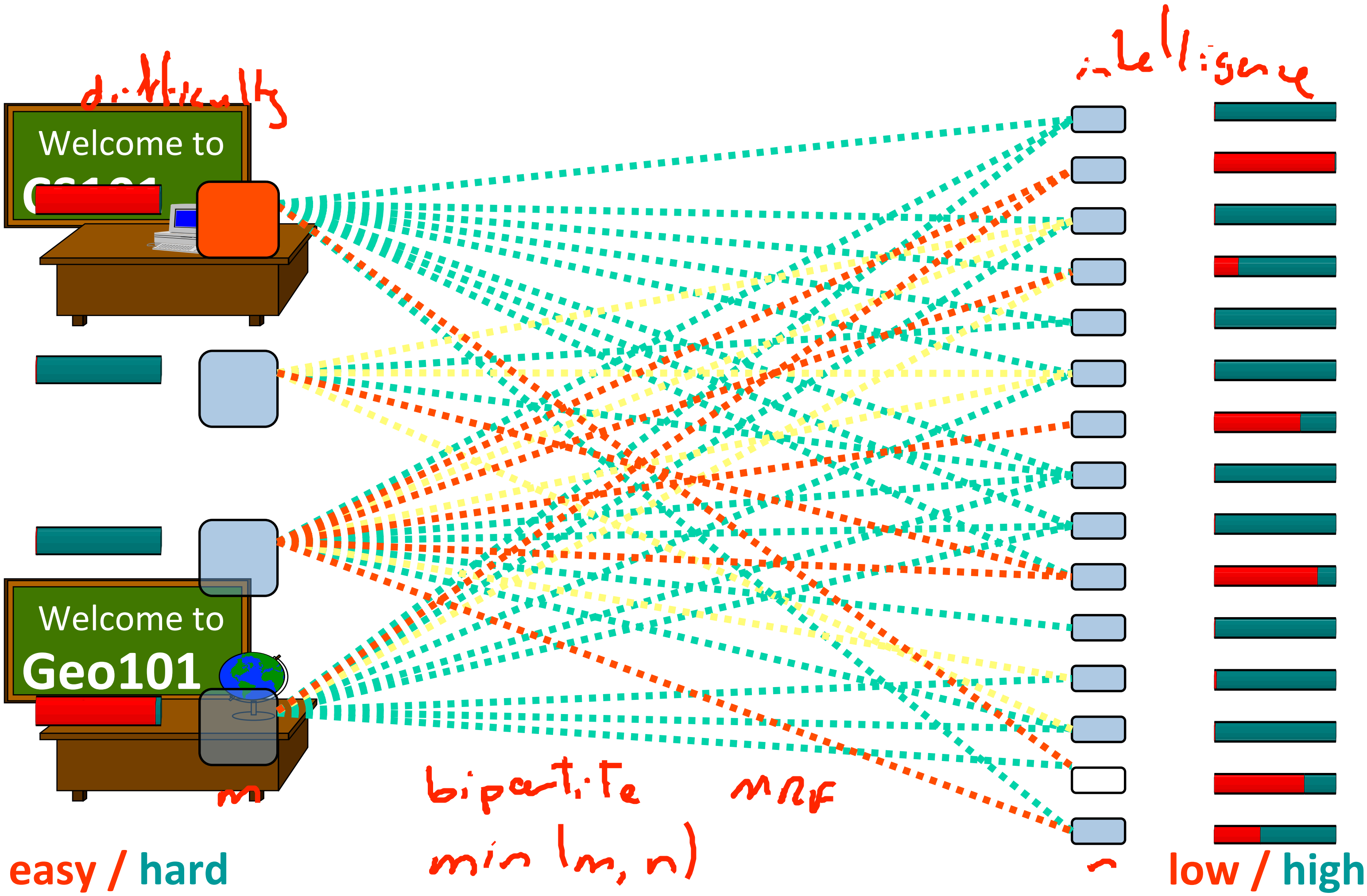

exact belief state is fully correlated in most cases

Daphne Koller




Huang, Koller, Malik, Ogasawara, Rao, Russell, Weber, AAAI 94

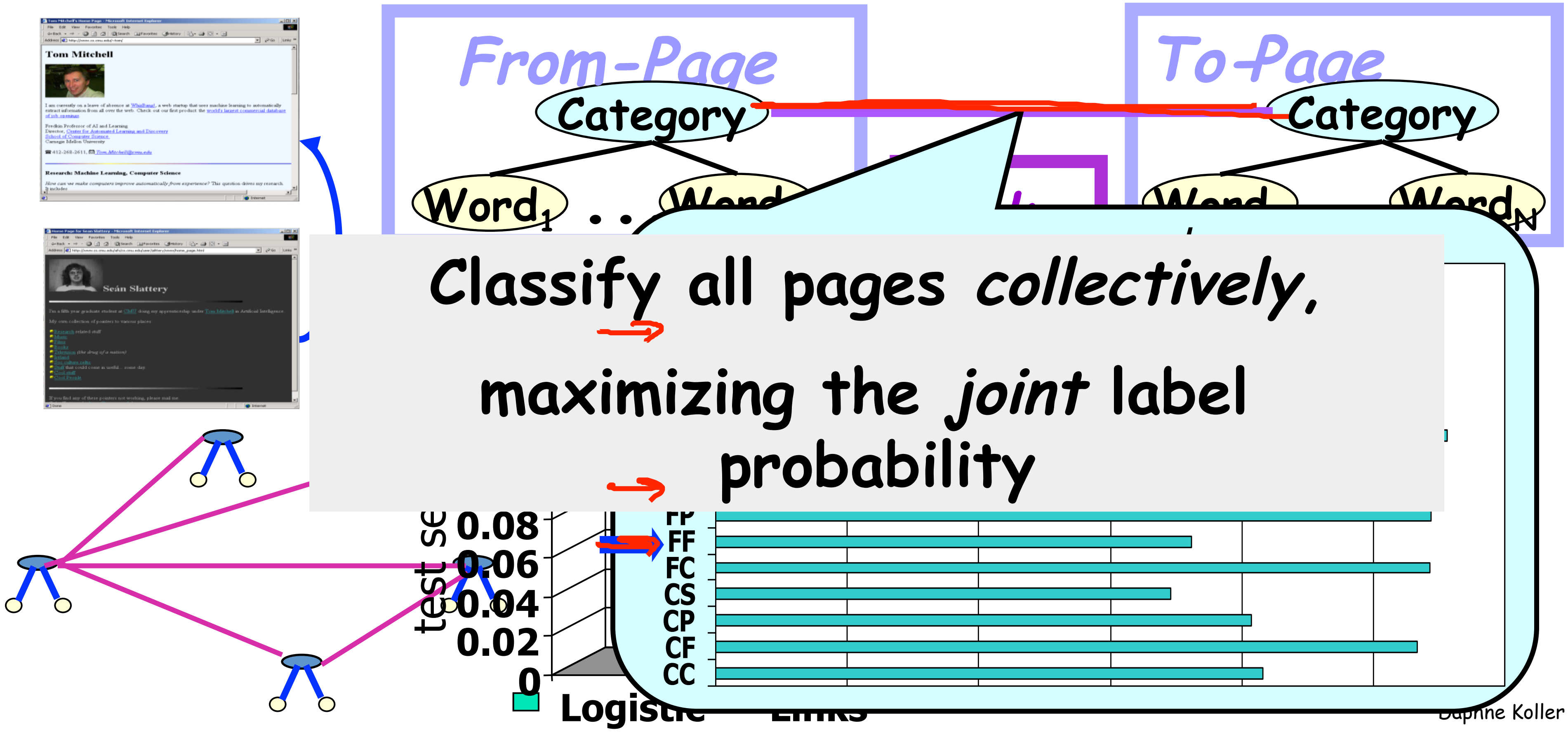
Daphne Koller




Daphne Koller

webkb [mitchell (Craven et al, Proc AAAI98; Tasker et al, UAI2002)

# Collective Webpage Classification

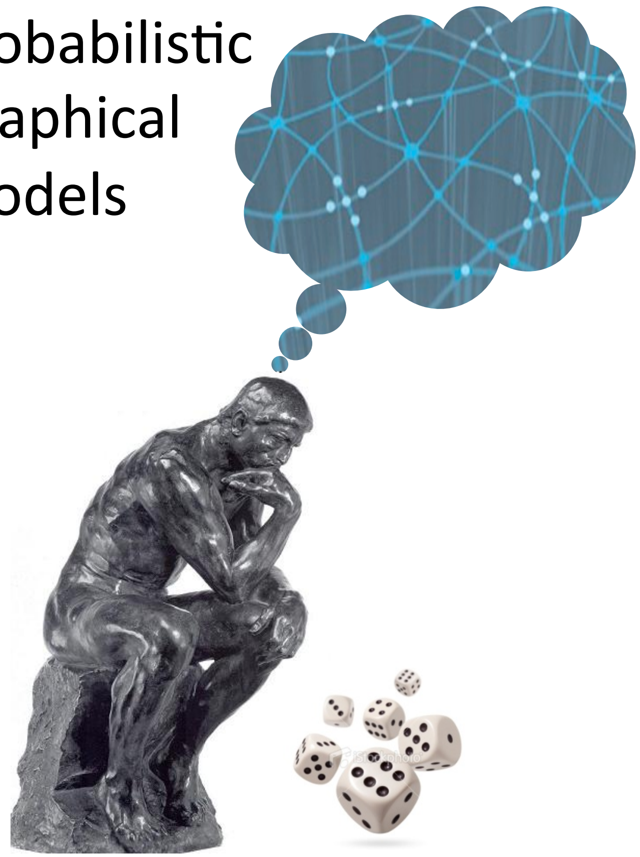



**Classify all pages *collectively*, maximizing the *joint* label probability**

# Summary

- Inference in template and temporal models can be done by unrolling the ground network and using standard methods
- Temporal models also raise new inference tasks, such as real-time tracking, which require that we adapt our methods
- Moreover, ground network is often large and densely connected, requiring careful algorithm design and use of approximate methods

Probabilistic
Graphical
Models



Inference

Summary

# Inference Methods and Evaluation

# MAP vs Marginals

**Marginals**

- Less fragile
- Confidence in answers
- Supports decision making

**MAP**

- Coherent joint assignment
- More tractable model classes
- Some theoretical guarantees

---

Approximate inference

- Errors are often attenuated
- Ability to gauge whether algorithm is working

Daphne Koller

# Algorithms for Marginals

- Exact inference

  fits in memory ⇒ exact inference

- Loopy message passing
- Sampling methods

# Algorithms for MAP

- Exact inference

  low treewidth
  associative models
  ⋮

- Optimization methods:
  - exact or approximate

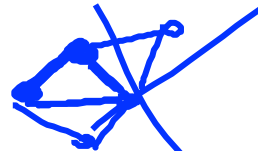
  (dual decomposition)

- Search-based methods (including sampling)

  hill-climbing
  MCMC

# Factors in Approximate Inference

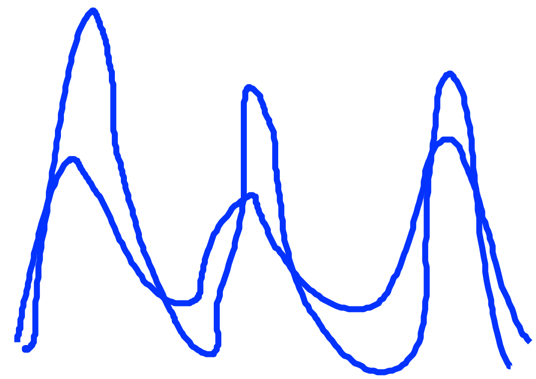
- Connectivity structure
- Strength of influence
- Opposing influences
- Multiple peaks in likelihood



message passing

convergence accuracy

coupling between variables



Daphne Koller

# So, now what?

- Identify "problem regions" in network
- Try to make inference in these regions more exact
  - Larger clusters in cluster graph
  - Proposal moves over multiple variables
  - Larger "slave" in dual decomposition