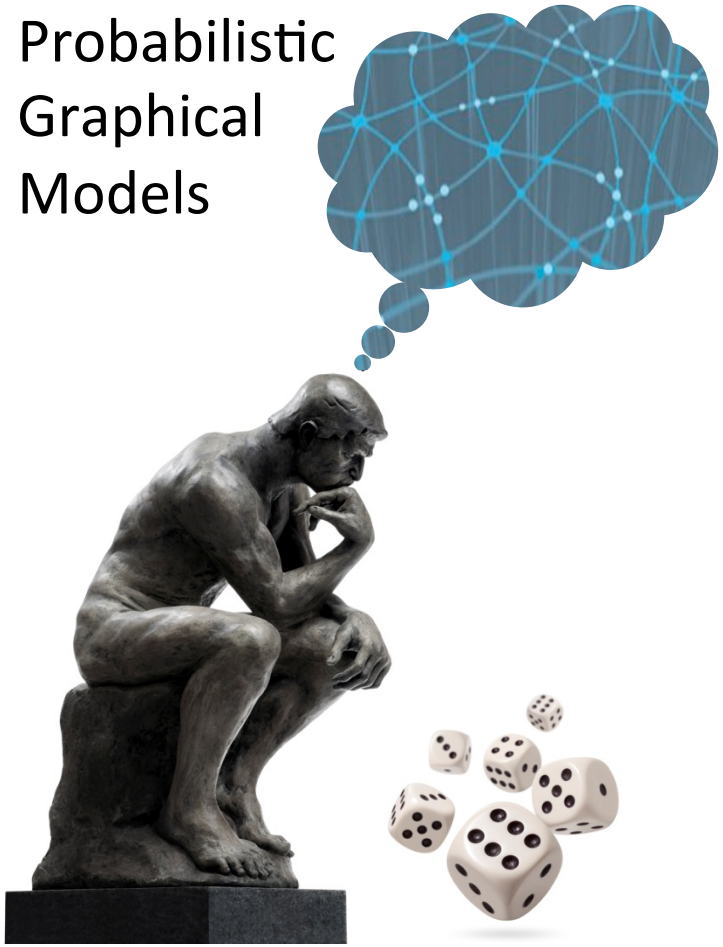
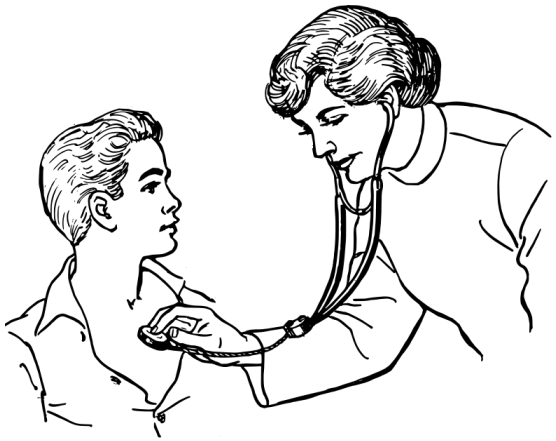
Probabilistic Graphical Models

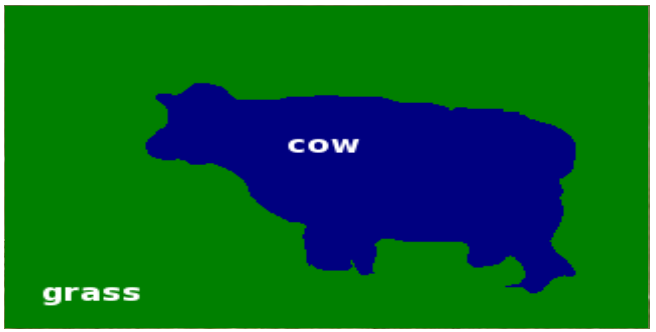


Introduction

Motivation and Overview



predisposing factors
symptoms
test results
diseases
treatment outcomes



millions of pixels or
thousands of superpixels

each needs to be labeled
{grass, sky, water, cow, horse, …}

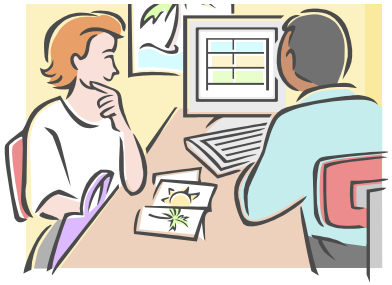
Daphne Koller

Probabilistic Graphical Models

Daphne Koller

Models

domain expert



Declarative representation

Data

elicitation

Model

Learning

Algorithm

Algorithm

Algorithm

Daphne Koller

Uncertainty

- Partial knowledge of state of the world
- Noisy observations
- Phenomena not covered by our model
- Inherent stochasticity

Probability Theory

- Declarative representation with clear semantics
- Powerful reasoning patterns
- Established learning methods

conditioning
decision making

Complex Systems

predisposing factors
symptoms
test results
diseases
treatment outcomes

class labels for
thousands of superpixels

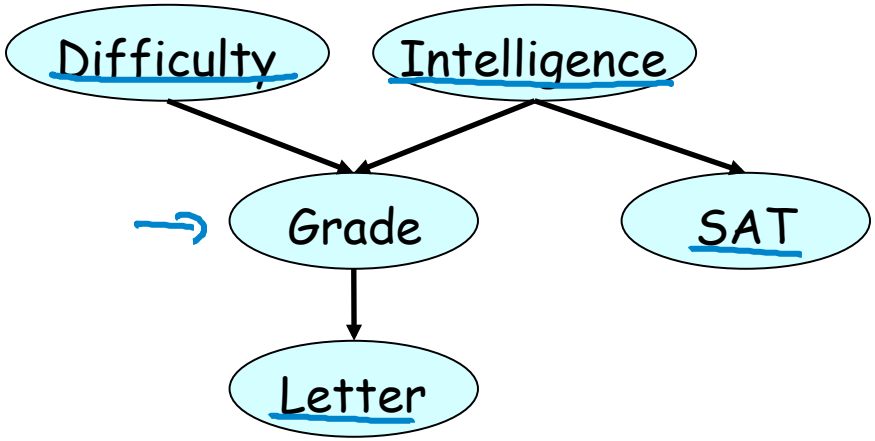
Random variables $X_1, \ldots, X_n$

Joint distribution $P(X_1, \ldots, X_n)$

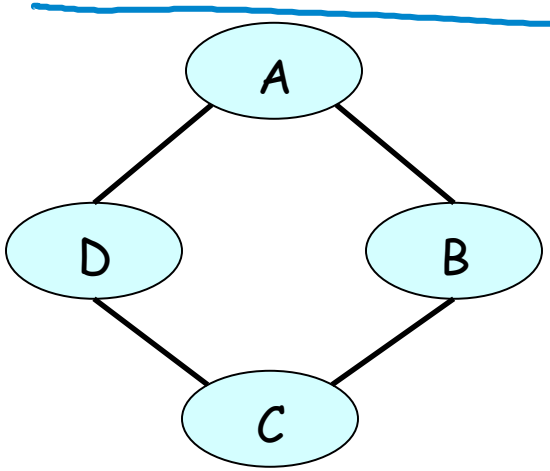
~ binary valued
distribution
over 2^n
possible states

Graphical Models

Bayesian networks

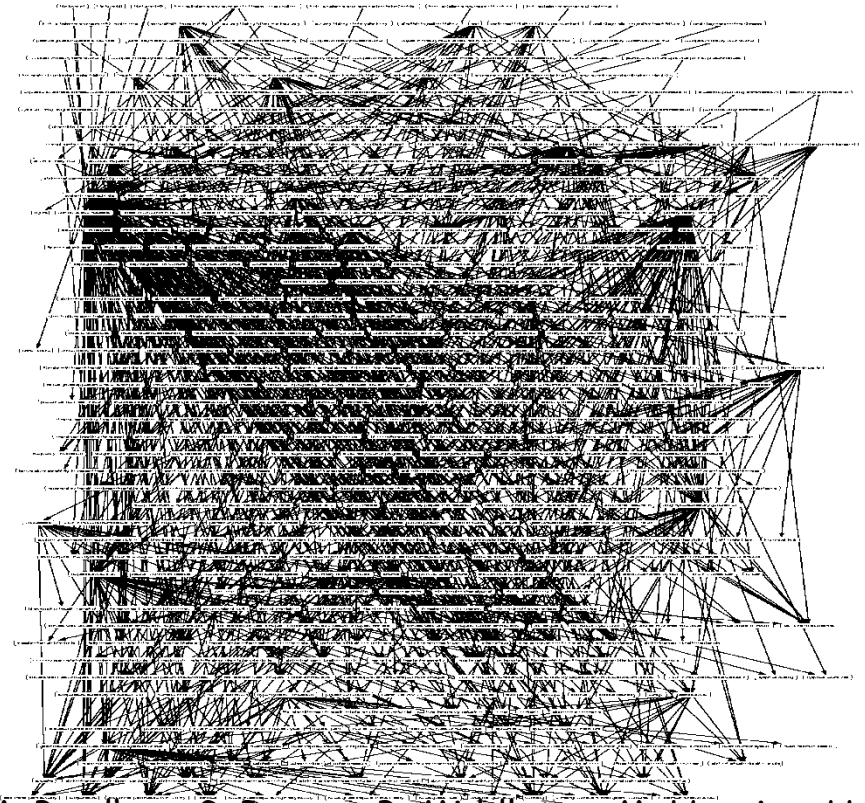


Markov networks

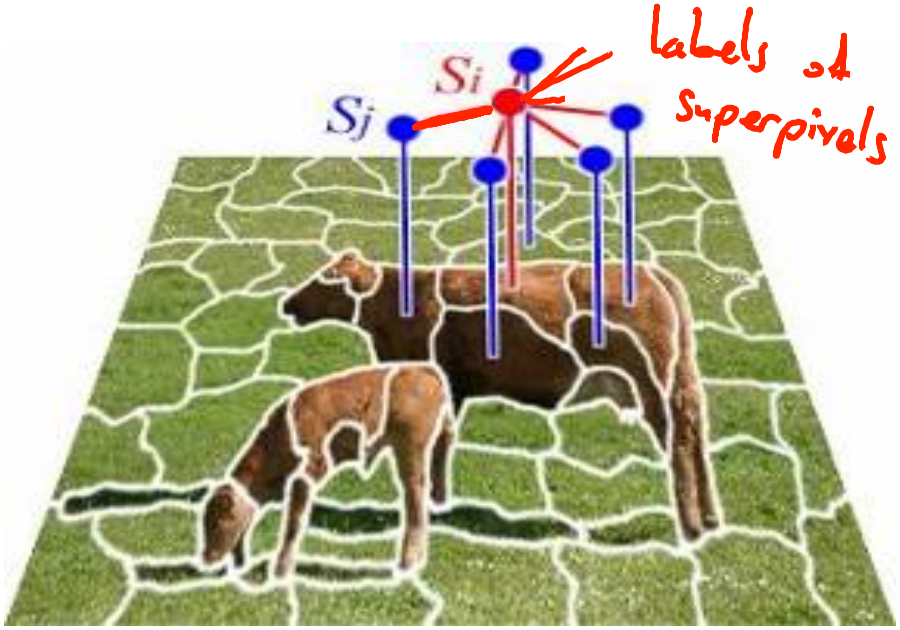


Graphical Models

diagnosis CPCS



M. Pradhan, G. Provan, B. Middleton, M. Henrion, UAI 94



Graphical Representation

- Intuitive & compact data structure
- Efficient reasoning using general-purpose algorithms
- Sparse parameterization
  - feasible elicitation ← by hand
  - learning from data ← automatically

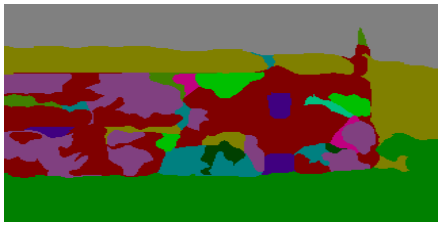
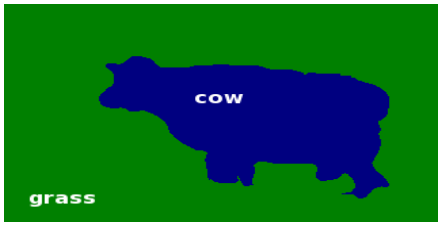
Many Applications

- Medical diagnosis
- Fault diagnosis
- Natural language processing
- Traffic analysis
- Social network models
- Message decoding
- Computer vision
  - Image segmentation
  - 3D reconstruction
  - Holistic scene analysis
- Speech recognition
- Robot localization & mapping

Image Segmentation





superpixels

machine learning
to separate superpixels

Daphne Koller

Thanks to: Eric Horvitz, Microsoft Research

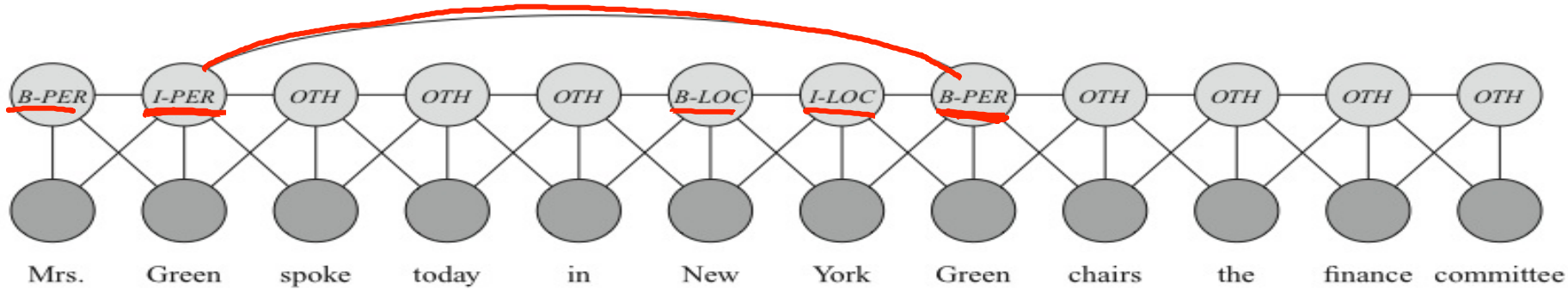
Medical Diagnosis

Daphne Koller

Textual Information Extraction

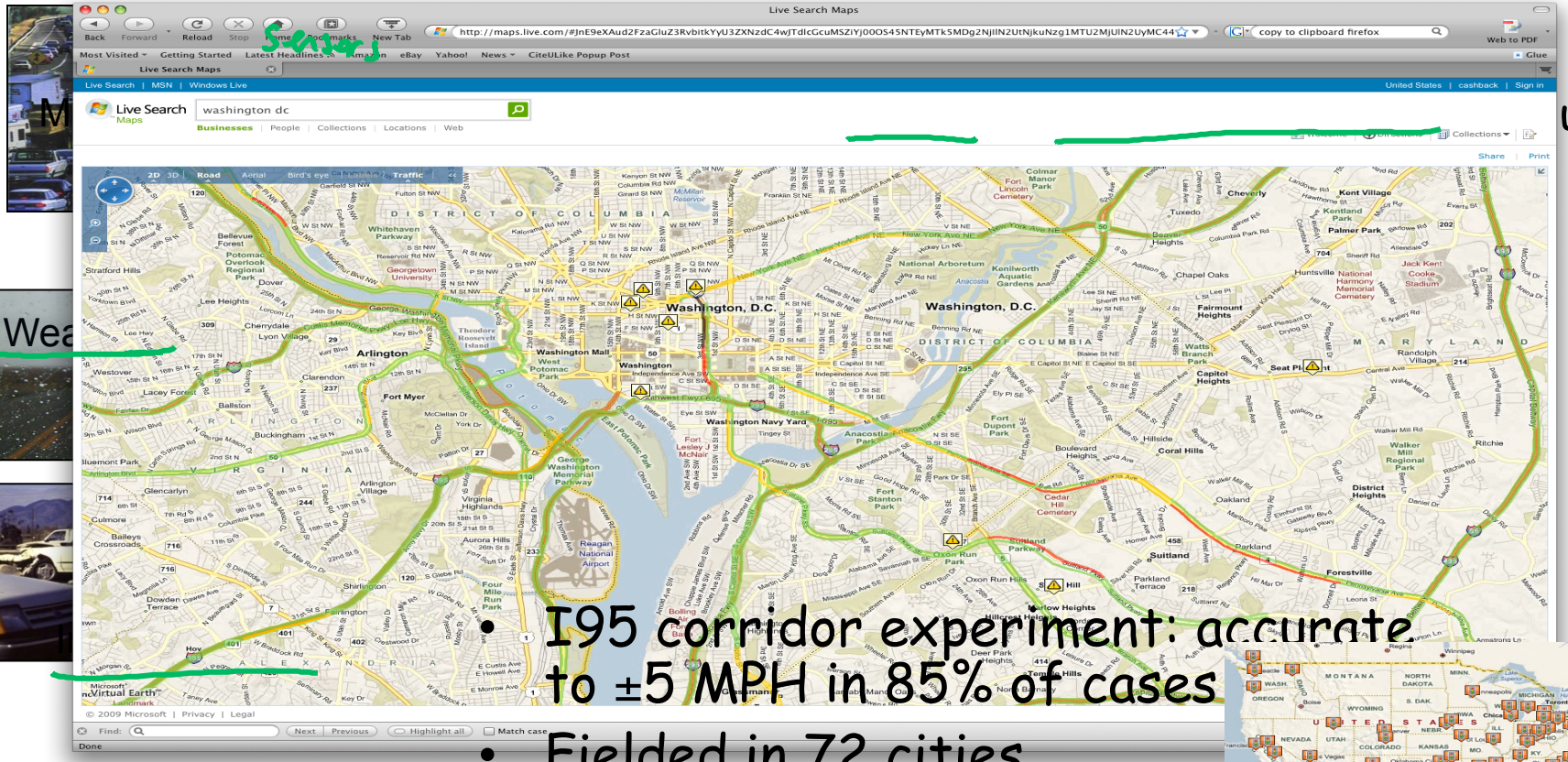
Mrs. Green spoke today in New York. Green chairs the finance committee.

Person Location Person Organization



Daphne Koller

Multi-Sensor Integration: Traffic

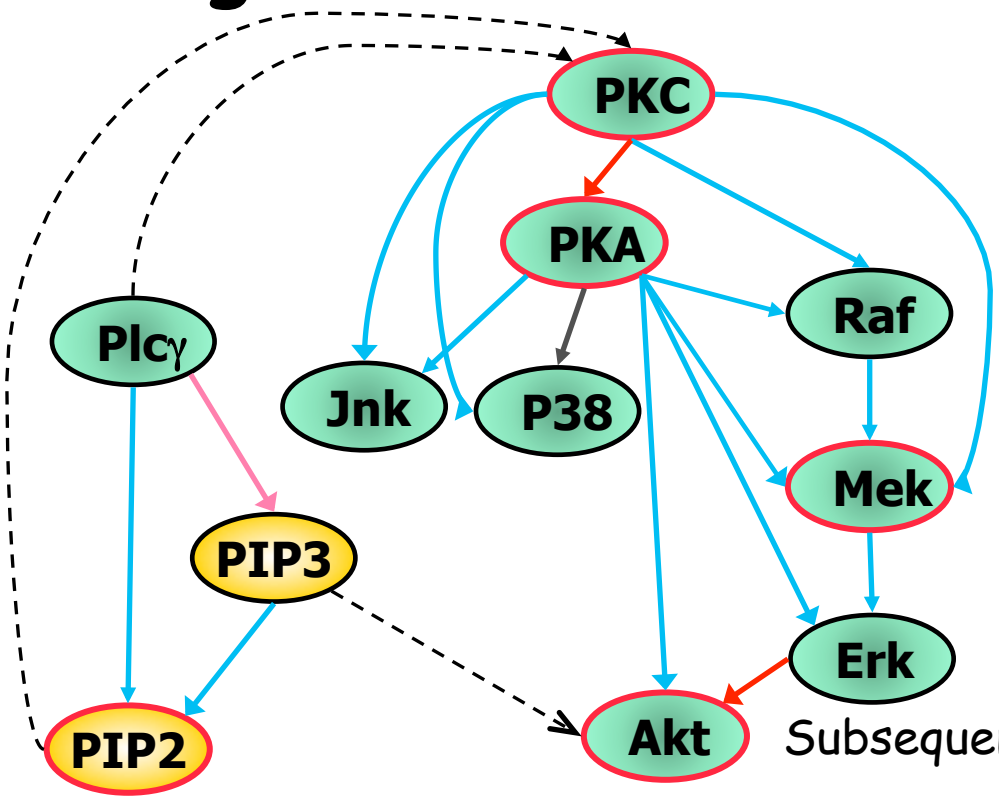


- I95 corridor experiment: accurate to ±5 MPH in 85% of cases
- Fielded in 72 cities

Thanks to: Eric Horvitz, Microsoft Research

This figure may be used for non-commercial and classroom purposes only.
Any other uses require the prior written permission from AAAS

Biological Network Reconstruction



Phospho-Proteins
Phospho-Lipids
Perturbed in data

Known	———	15/17
Supported	———	2/17
Reversed	———	1
Missed	- - -	3

Subsequently validated in wetlab

Causal protein-signaling networks derived from multiparameter single-cell data
Sachs et al., *Science* 2005

Daphne Koller

Overview

- Representation
  - Directed and undirected
  - Temporal and plate models
- Inference reasoning
  - Exact and approximate
  - Decision making
- Learning
  - Parameters and structure
  - With and without complete data