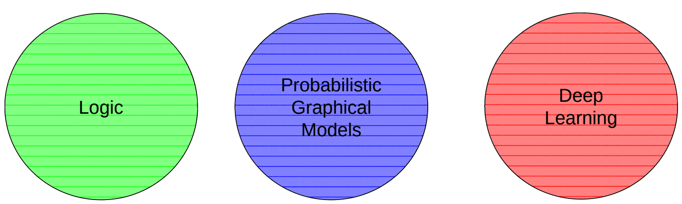
CS 536 Course Overview

Introduction

- Probabilistic graphical models encode joint probability distributions as a graph structure
- Why study Bayesian networks?
  - One of the major machine learning approaches
  - Nice framework for building probabilistic models
  - Lots of interesting connections with other areas of AI (e.g. deep learning and logic)

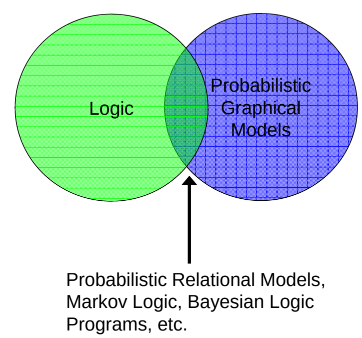
The Really Big Picture

Three major approaches influencing artificial intelligence over the years:



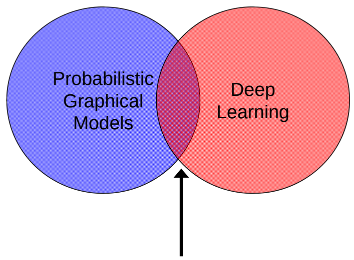
The Really Big Picture

Probabilistic graphical models overlap with Logic



The Really Big Picture

And also with deep learning…



Logistic Regression, Restricted Boltzman Machines, Variational Autoencoders, etc.

Graphical Models and Deep Learning

Graphical Models

- Compact representation of a joint probability distribution
- Can incorporate domain knowledge via priors
- The hard part: developing efficient inference algorithms
- May not need as much data as deep learning

Deep Learning

- Function approximator: maps inputs to outputs
- Doesn't always represent a probability distribution
- Great for learning representations
- Can be awkward to incorporate domain knowledge
- The hard part: architecture building, hyperparameter tuning
- Needs lots of data

Introduction

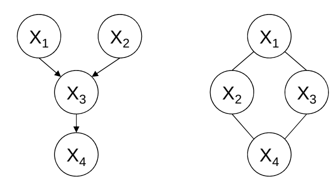
This course is about probabilistic graphical models. Specifically, it covers:

1. Representation
2. Inference
3. Learning
4. Connections with Deep Learning

Representation

Joint Probability Distribution
$P(X_1, \ldots, X_m)$

Graph Structure

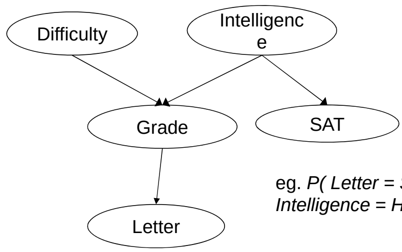


Directed Graphical Model eg. Bayesian network

Undirected Graphical Model eg. Markov Network

Inference

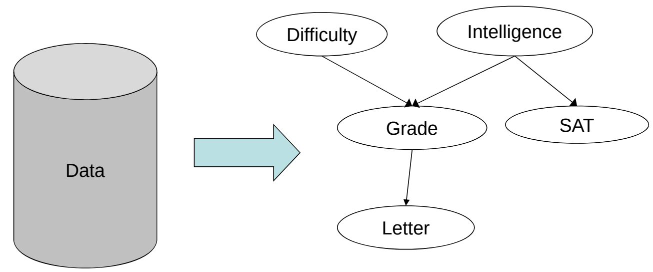
Given a graphical model, we would like to compute probabilistic queries of the form *P(**X** | **E**)*



eg. *P(Letter = Strong | Difficulty = High, Intelligence = High)*

9

Learning



How do I learn the structure and/or parameters of a Bayesian network from data?

10

Connections with Deep Learning

Topics include:

- Restricted Boltzmann machines, Deep Boltzmann machines
- Variational inference and Variational Autoencoders
- (If we have time) Sum-Product networks

11

Uses of Bayesian Networks at OSU

- Species Distribution Modelling
- Insider threat detection
- Activity Monitoring
- Intelligent Desktop Assistants
- Real-time strategy games
- Image recognition and segmentation
- Sensor data cleaning
- More on its way!

12

Quick Stats Review

Suppose two random variables X and Y are (marginally) independent of each other. Then:

1. $P(X, Y) = P(X)P(Y)$
2. $P(X|Y) = P(X)$
3. $P(Y|X) = P(Y)$

Suppose two random variables X and Y are conditionally independent of each other given Z. Then:

1. $P(X, Y|Z) = P(X|Z)P(Y|Z)$
2. $P(X|Y, Z) = P(X|Z)$
3. $P(Y|X, Z) = P(Y|Z)$

13

Quick Stats Review

Chain rule of probability:

$$P(X, Y) = P(X|Y)P(Y)$$

Bayes rule:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

14