Probabilistic Graphical Models & Probabilistic AI

Ben Lengerich

Lecture 1: Introduction to PGMs

January 21, 2025

Reading: See course homepage

Class webpage

lengerichlab.github.io/pgm-spring-2025

WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON

Probabilistic Graphical Models

STAT 479 • Spring 2025 • UW-Madison

Many problems in fields such as artificial intelligence, statistics, computer vision, natural language processing, and computational biology can be viewed as the search for a coherent global conclusion from local information. The probabilistic graphical models (PGMs) framework provides a unified approach for solving this wide range of problems, enabling efficient inference, decision-making, and learning in systems with a large number of attributes and huge datasets.

This course will provide a strong foundation for applying PGMs to complex real-world problems, as well as addressing core research topics in graphical models. Students will learn to construct both Bayesian and Markov networks, perform inference, and apply learning algorithms. Toward the end of the course, we will also explore modern probabilistic AI techniques, including generative models and large language models (LLMs), emphasizing how PGMs underpin many state-of-the-art AI systems.

- **Time:** Tuesday/Thursday 11:00am-12:15pm
- **Location:** Microbial Sciences Building 1520
- **Discussion:** Canvas
- **HW submission:** Canvas
- **Contact:** Students should ask all course-related questions on Canvas, where you will also find announcements. Individual enquires can be directed to TA/instructor emails.

Instructor Benjamin Lengerich
Email: lengerich@wisc.edu
Office hours: TBD

TA Chenyang Jiang
Email:
Office hours:

Logistics

- Textbooks:
  - Daphne Koller and Nir Friedman, Probabilistic Graphical Models
  - M. I. Jordan, An Introduction to Probabilistic Graphical Models
- Class Announcements: **Canvas**
- Assignment Submissions: **Canvas**
- Instructor: **Ben Lengerich**
  - Office Hours: Thursday 2:30-3:30pm, 7278 Medical Sciences Center
  - Email: lengerich@wisc.edu
- TA: **Chenyang Jiang**
  - Office Hours: Monday 11am-12pm, 1219 Medical Sciences Center
  - Email: cjiang77@wisc.edu

Course Schedule / Calendar

Weeks	Lecture Dates	Topic	Assignments
Module 1: Foundations of PGMs, Exact Inference			
1-4	Jan 21- Feb 13	Course Introduction, Foundations of PGMs, Exact Inference	HWs 1, 2
4	Feb 13	**Quiz**	
Module 2: Learning			
5-9	Feb 18 - Mar 18	Parameter Learning, Structure Learning, Approximate Inference	HWs 3,4,5
9	Mar 20	**Midterm Exam**	
10	Mar 21 - Mar 30	Spring Recess	
Module 3: Modern Probabilistic AI			
11-14	Apr 1 - Apr 24	Deep Learning, LLMs from a GM perspective	Project Midway Report
15	Apr 29 - May 1	Project Presentations	Project Final Report

Grading

- **HW Assignments** (40% of grade)
  - Expecting 5 HW Assignments
  - All in Modules 1 and 2
- **Quiz** (10% of grade)
  - End of Module 1 (Feb 13)
  - Format: In-class, Closed-book.
- **Midterm Exam** (20% of grade)
  - End of Module 2 (Mar 20)
  - Format: In-class, Closed-book.
- **Final Project** (30% of grade)
  - Milestones will be due along the way semester
- **Lecture Notes** (up to 3% extra credit)

Grading Scale

- **No curving:**
  - A: 93–100%
  - AB: 88–92%
  - B: 83–87%
  - BC: 78–82%
  - C: 70–77%
  - D: 60–69%
  - F: Below 60%

Homework Policies

- Expecting 5 HW Assignments
- All in Modules 1 and 2
- Assignments must be submitted via Canvas by the deadline (typically Fridays at 11:59 PM).
- **Late submissions** will incur a penalty of 10% per day, up to three days, at which they will not be accepted.
- **Collaboration:** Students may collaborate on understanding the problems, but all submitted work must be individual. Use of AI tools is allowed, but answers must reflect your own understanding. You are responsible for ensuring no plagiarism.

Project

- **Proposal** (5%): Due Feb 21, 2025. Must outline the group members, problem statement, motivation, dataset(s).
- **Midway Report** (5%): Due Apr 11, 2025. Build on the proposal to have some preliminary dataset exploration and revised methods.
- **Presentation** (5%): During the last week of class, students will present their findings in a concise and engaging format.
- **Report** (15%): Due May 5, 2025. The report should detail the problem, methods, results, and conclusions.
- **Collaboration:** Teams of up to three students are allowed. Contributions of each individual must be clearly delineated in the report.

Extra Credit: Lecture Notes

- **Opportunity:** Students can sign up to write lecture notes for each class session. If the notes meet quality standards, students will receive up to 3% extra credit added to their final grade.
  - **Sign-up sheet**
  - You are expected to work together with the other scribe to generate 1 note document for the lecture.
  - Accepted notes will be hosted on our course webpage (optionally with your name attached, your choice).
- **Submission:** Lecture notes must be submitted within one week of the session.
  - Submit on the GitHub for the course webpage.
- **Template** is available on the webpage.

Attendance / Participation

- Attendance and participation are expected, though they are not part of the formal grade calculation.
- Contact the instructor in advance of extended absences.

Academic Integrity

- By virtue of enrollment, each student agrees to uphold the high academic standards of the University of Wisconsin–Madison; academic misconduct is behavior that negatively impacts the integrity of the institution. Cheating, fabrication, plagiarism, unauthorized collaboration and helping others commit these previously listed acts are examples of misconduct which may result in disciplinary action. Examples of disciplinary sanctions include, but are not limited to, failure on the assignment/course, written reprimand, disciplinary probation, suspension or expulsion.

Mental Health & Wellbeing

- Students often experience stressors that can impact both their academic experience and personal well-being. These may include mental health concerns, substance misuse, sexual or relationship violence, family circumstances, campus climate, financial matters, among others.

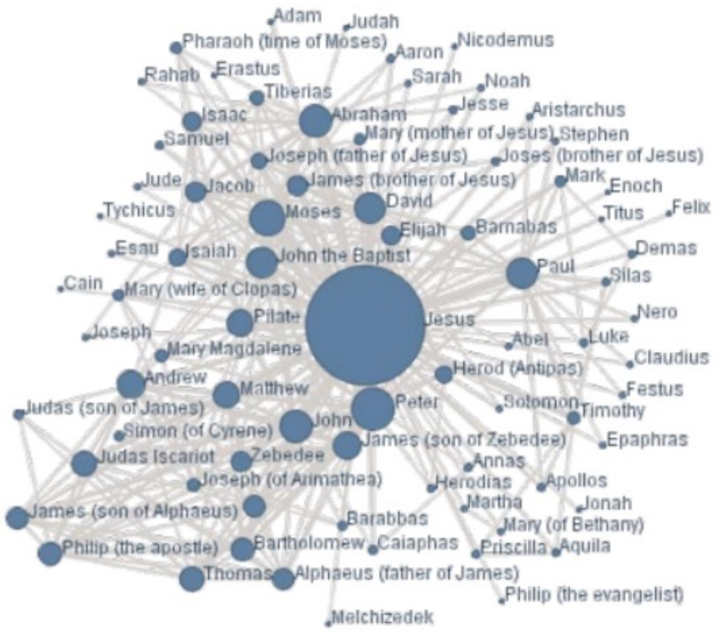
- UW–Madison students are encouraged to learn about and utilize the university's mental health services and/or other resources as needed. Students can visit **uhs.wisc.edu** or call **University Health Services** at **(608) 265-5600** to learn more.

Questions about Course Logistics?

Introduction to Probabilistic Graphical Models (PGMs)

What is a Graphical Model?

Graph



Model

$$\mathrm{M}_{\mathrm{G}}$$

Data

$$\mathrm{D} = \{X_1^{(i)}, X_2^{(i)}, \dots, X_m^{(i)}\}_{i=1}^{N}$$

The Fundamental Questions

- Representation
  - How to encode our domain knowledge/assumptions/constraints?
  - How to capture/model uncertainties in possible worlds?
- Inference
  - How do I answer questions/queries according to my model and/or based on observed data?

$$\text{e.g. } P(X_i|D)$$

- Learning
  - What model is "right" for my data?

$$\text{e.g. } M = argmax_{M \in \mathcal{H}} F(D; M)$$

The Fundamental Questions in Probabilistic Form

- Representation
  - What is the joint probability distribution of multiple variables?
    $$P(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$$
  - Assume Boolean X_i. How many state configurations of the joint probability?
    - 2^8
  - Can we disallow certain state configurations, and instead focus on a subset?

The Fundamental Questions in Probabilistic Form

- Inference
  - How do I answer questions/queries according to my model and/or based on observed data?

e.g. $P(X_i|D)$

  - Let's try $P(X_8|X_1)$:

$$P(X_8|X_1) = \frac{P(X_8, X_1)}{P(X_1)}$$
$$= \frac{\sum_{X_2}\sum_{X_3}\cdots\sum_{X_7} P(X_1, \ldots, X_8)}{\sum_{X_2}\sum_{X_3}\cdots\sum_{X_8} P(X_1, \ldots, X_8)}$$

  - Require summing over 2^6 configurations of unobserved variables
  - On the other hand, if X_i all independent: $P(X_8|X_1) = P(X_8)$
  - Graphical form allows us to work in between.

The Fundamental Questions in Probabilistic Form

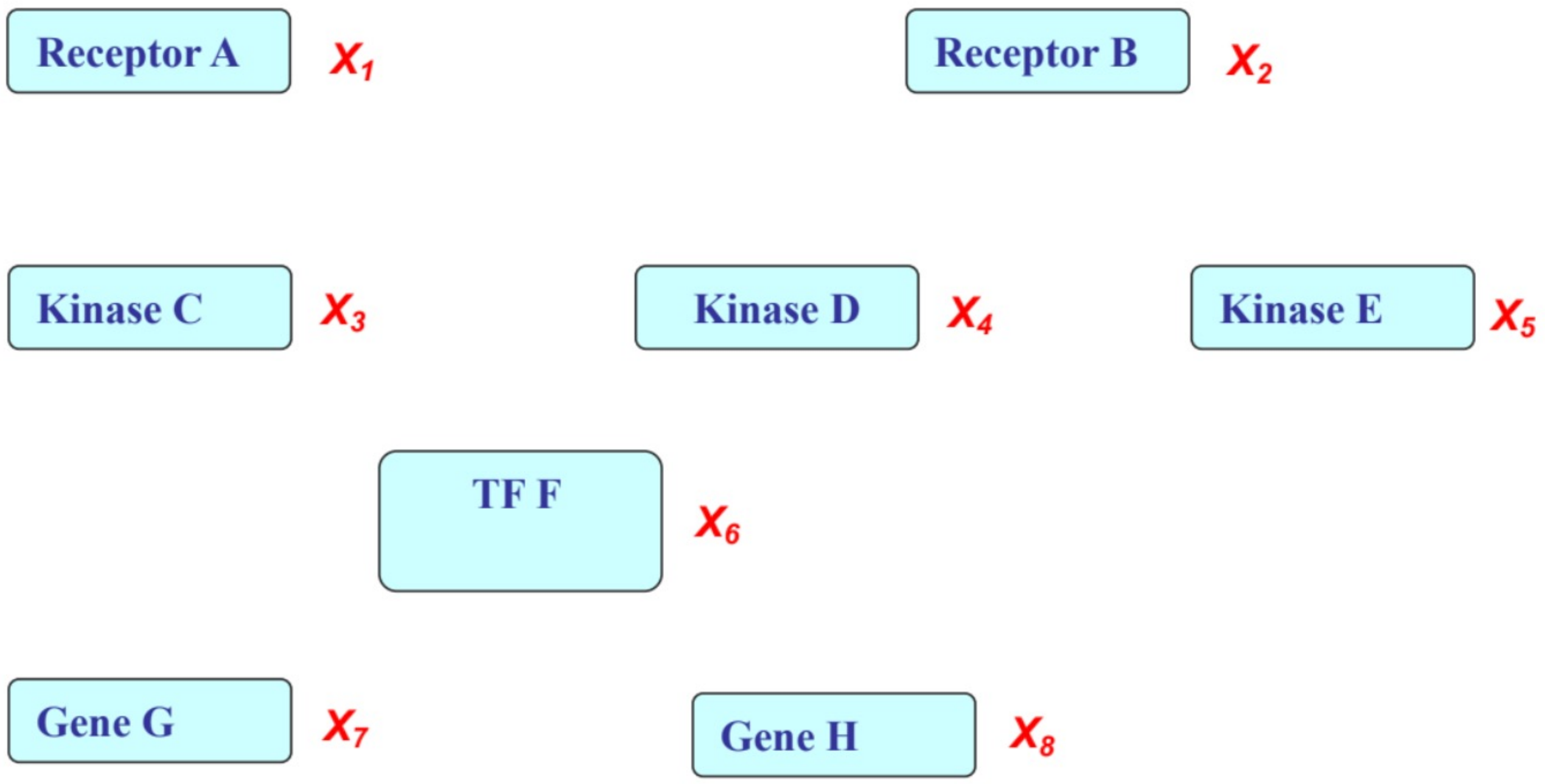
- Learning
  - What model is "right" for my data?

$$\text{e.g. } M = argmax_{M \in \mathcal{H}} F(D; M)$$

  - How can we constrain the hypothesis space $\mathcal{H}$ so that the search for the argmax is efficient?

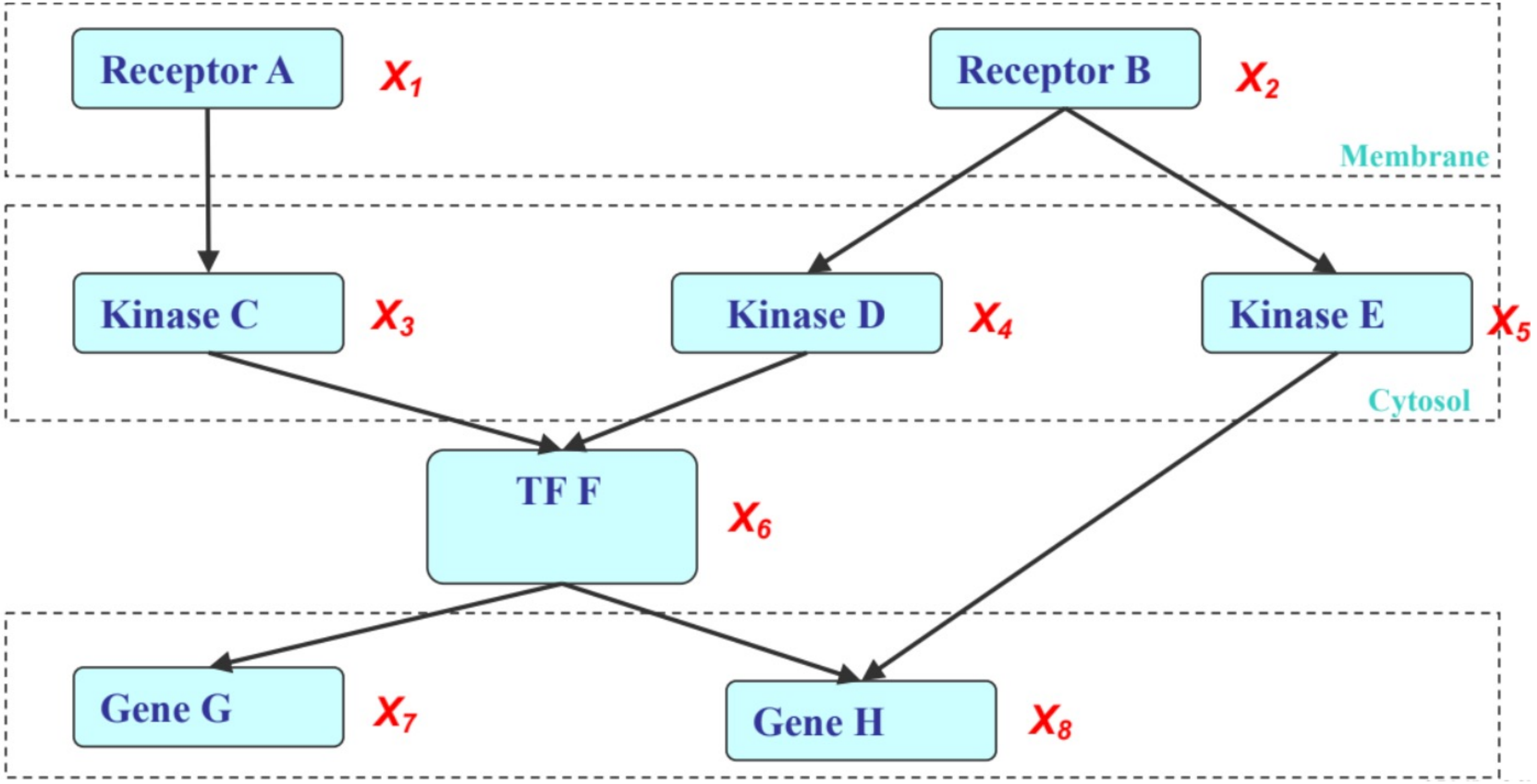
An example

- A possible world for cellular signal transduction:



An example

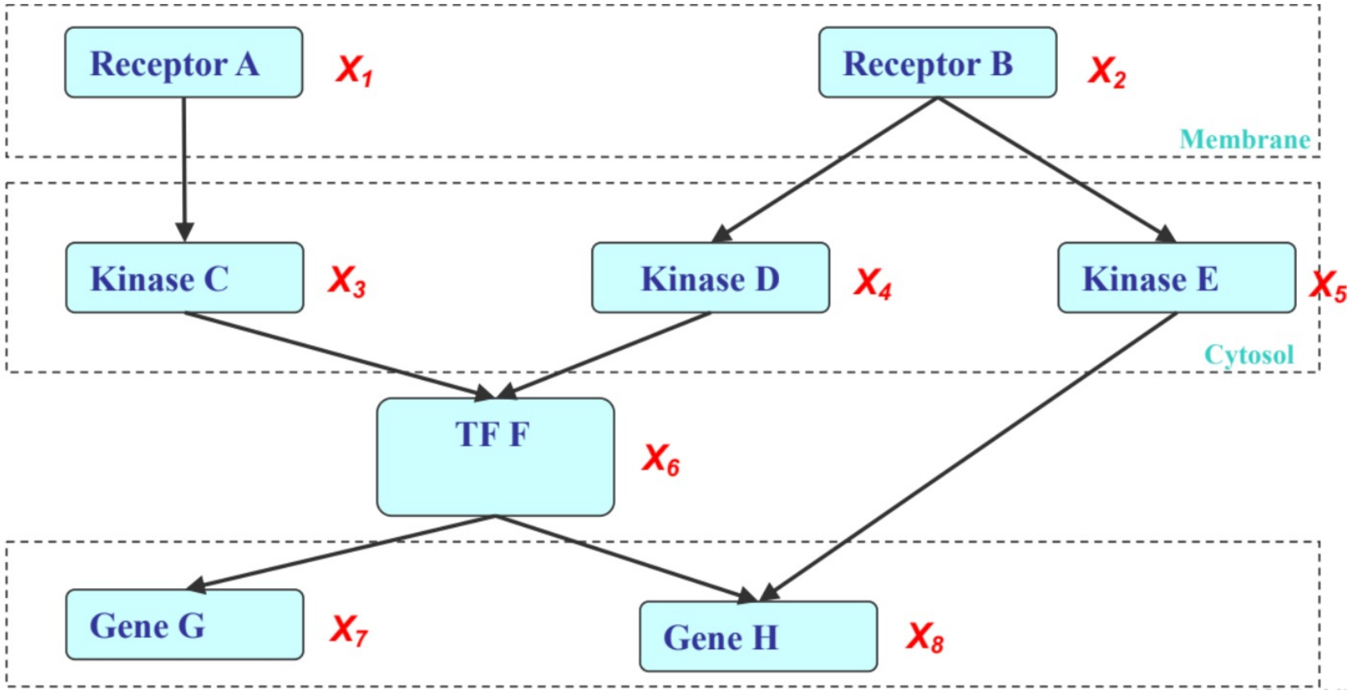
- Imposing structure of dependencies simplifies representation



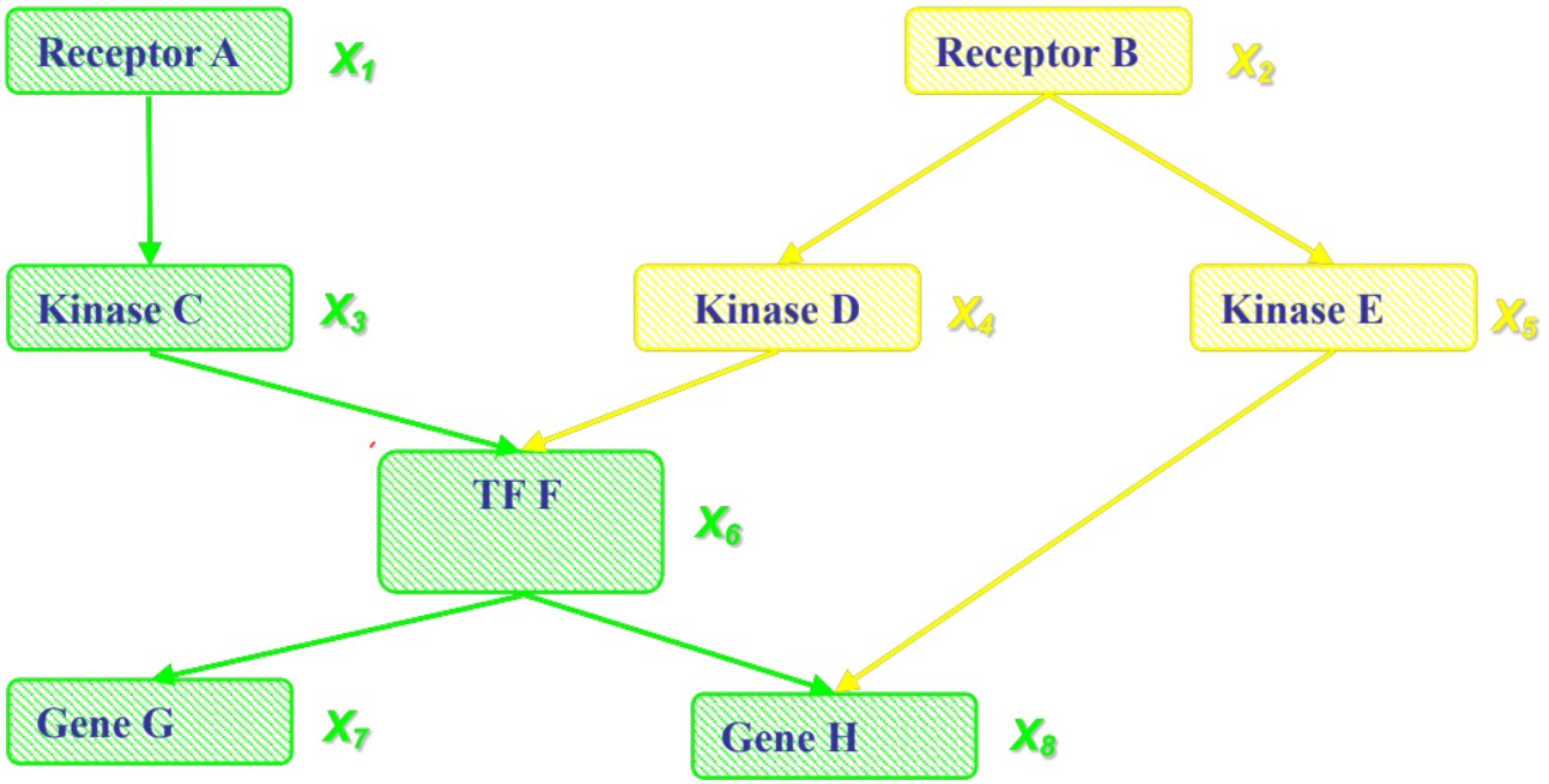
Probabilistic Graphical Models

- If the X_is are conditional independence (as encoded by a PGM), then the joint can be factorized into a product of simpler terms:

$$P(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) =$$
$$P(X_1)P(X_2)P(X_3|X_1)P(X_4|X_2)P(X_5|X_2)P(X_6|X_3, X_4)P(X_7|X_6)P(X_8|X_6, X_5)$$



Data Integration



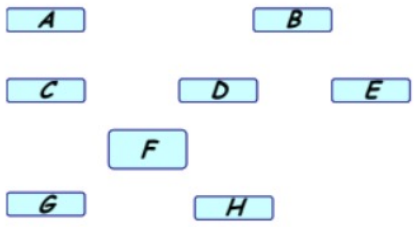
- Others?

Why PGMs?

- Incorporation of domain knowledge and causal (logical) structures
- Learning domain knowledge and causal structures from data
- Modular combination of heterogeneous parts – data fusion
- Bayesian Learning – can easily incorporate uncertainty, and can be specific to a particular part of the model

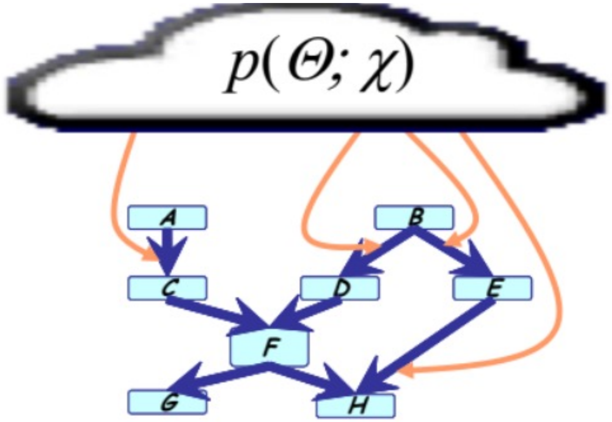
GMs and Bayesian Learning

- Let GM parameters Θ be a RV. We can use Bayesian reasoning to update our beliefs based on data:



$$\mathbf{A} = \begin{array}{l} (A,B,C,D,E,\ldots)=(T,F,F,T,F,\ldots) \\ (A,B,C,D,E,\ldots)=(T,F,T,T,F,\ldots) \\ \ldots\ldots\ldots \\ (A,B,C,D,E,\ldots)=(F,T,T,T,F,\ldots) \end{array}$$

$$\boxed{\Theta_{Bayes} = \int \Theta \, p(\Theta \mid \mathbf{A}, \chi) \, d\Theta}$$



C	D	P(F | C,D)	
c	d	0.9	0.1
c	$\bar{d}$	0.2	0.8
$\bar{c}$	d	0.9	0.1
$\bar{c}$	$\bar{d}$	0.01	0.99

$$\underbrace{p(\Theta \mid \mathbf{A}; \chi)}_{\text{posterior}} \propto \underbrace{p(\mathbf{A} \mid \Theta)}_{\text{likelihood}} \underbrace{p(\Theta; \chi)}_{\text{prior}}$$

So what is a PGM?

- PGM = Multivariate Statistics + Structure
- GM = Multivariate Objective Function + Structure

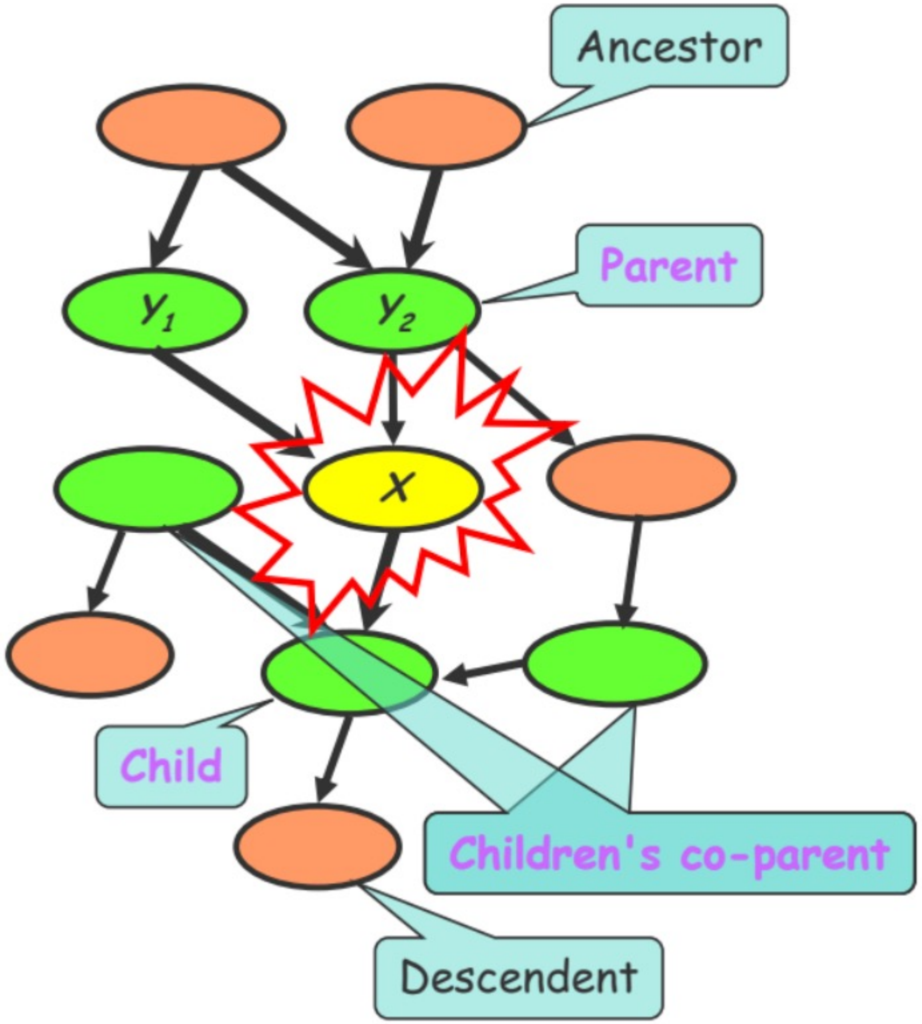
- Formally: A PGM is a **family of distributions** on a set of random variables that are compatible with all the probabilistic independence propositions encoded by a **graph** that connects these variables.

Two Types of GMs

- **Directed** edges give causality relationships (Bayesian Network or Directed Graphical Model)
- **Undirected** edges simply give correlations between variables (Markov Random Field or Undirected Graphical model):

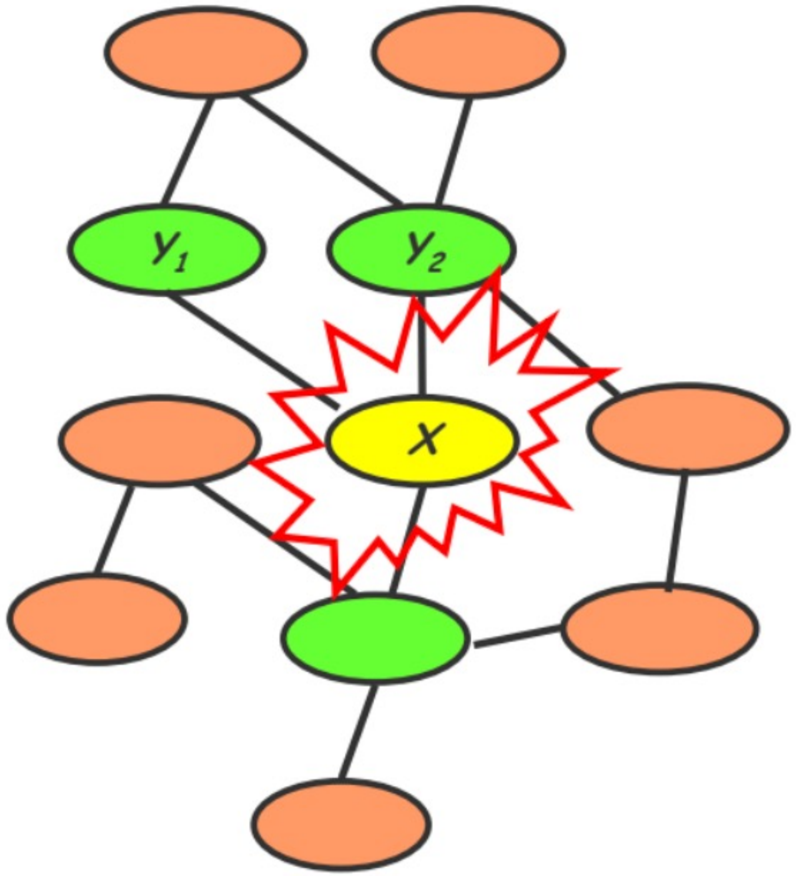
Bayesian Networks

- Structure: Directed Acyclic Graph
- A node is conditionally independent of all other nodes outside its Markov Blanket
- Local conditional distributions (CPD) and the DAG completely determine the joint distributions.
- Represent causality relationships and facilitate a generative process.
- Why acyclic?



Markov Random Fields

- Structure: Undirected Graph
- A node is conditionally independent of all other nodes given its Direct Neighbors
- Local contingency functions (potentials) and the cliques in the graph completely determine the joint distribution.
- Give correlations between variables, but no explicit way to generate samples.
- Can MRFs have cycles?

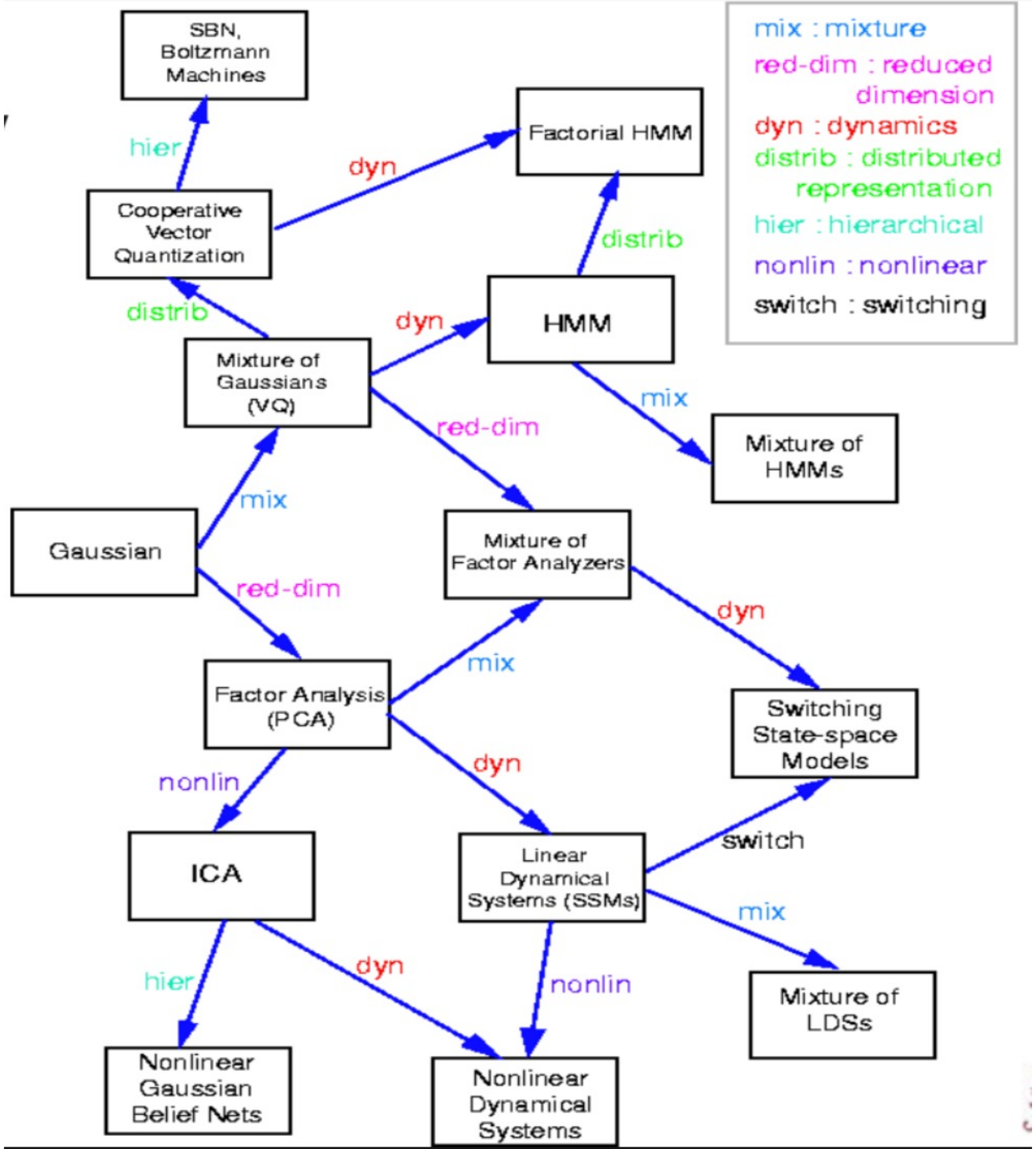


Towards structural specification of probability distributions

- Separation properties in the graph imply independence properties about the associated variables.
- For the graph to be useful, any conditional independence properties we can derive from the graph should hold for the probability distribution that the graph represents.
- **The Equivalence Theorem:**
  - For a graph G,
    - Let D1 denote the family of all distributions that satisfy I(G),
    - Let D2 denote the family of all distributions that factor according to G,
  - Then D1≡D2.

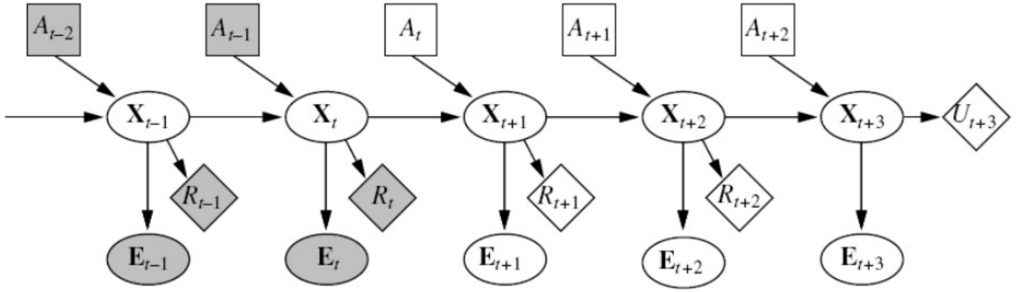
An (incomplete) genealogy of GMs

- Figure from Zoubin Ghahramani & Sam Roweis

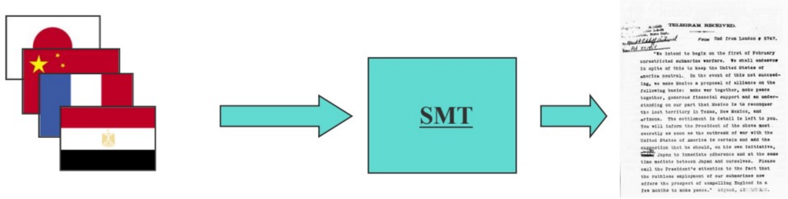


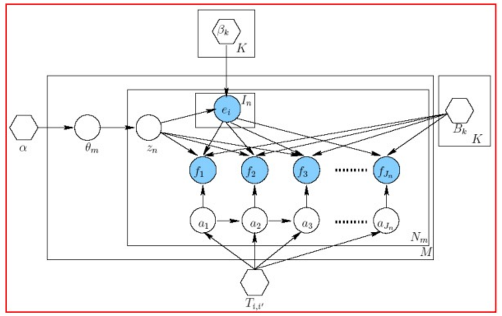
Fancier GMs

- **Reinforcement Learning:**
- Partially Observed Markov Decision Process (POMDP)



- **Machine Translation:**

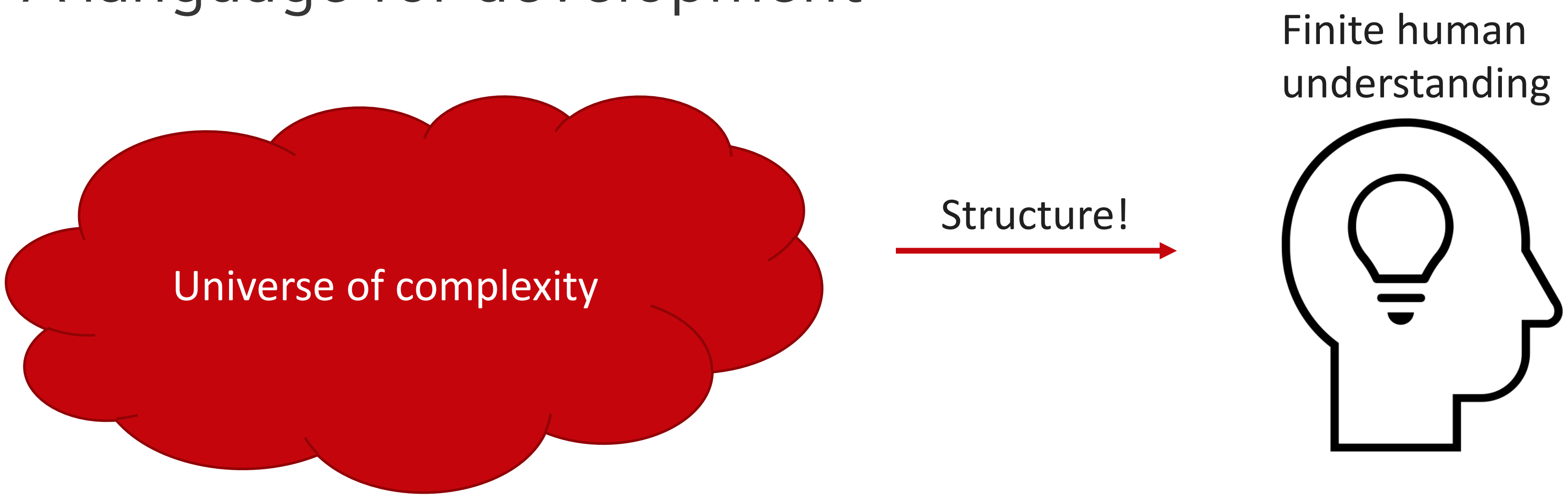




The HM-BiTAM model (B. Zhao and E.P Xing, ACL 2006)

Why GMs

- **What's the point of GMs in the AI era?**
- A language for communication
- A language for computation
- A language for development



Course Plan

Weeks	Lecture Dates	Topic	Assignments
Module 1: Foundations of PGMs, Exact Inference			
1-4	Jan 21- Feb 13	Course Introduction, Foundations of PGMs, Exact Inference	HWs 1, 2
4	Feb 13	**Quiz**	
Module 2: Learning			
5-9	Feb 18 - Mar 18	Parameter Learning, Structure Learning, Approximate Inference	HWs 3,4,5
9	Mar 20	**Midterm Exam**	
10	Mar 21 - Mar 30	Spring Recess	
Module 3: Modern Probabilistic AI			
11-14	Apr 1 - Apr 24	Deep Learning, LLMs from a GM perspective	Project Midway Report
15	Apr 29 - May 1	Project Presentations	Project Final Report

Questions?

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON