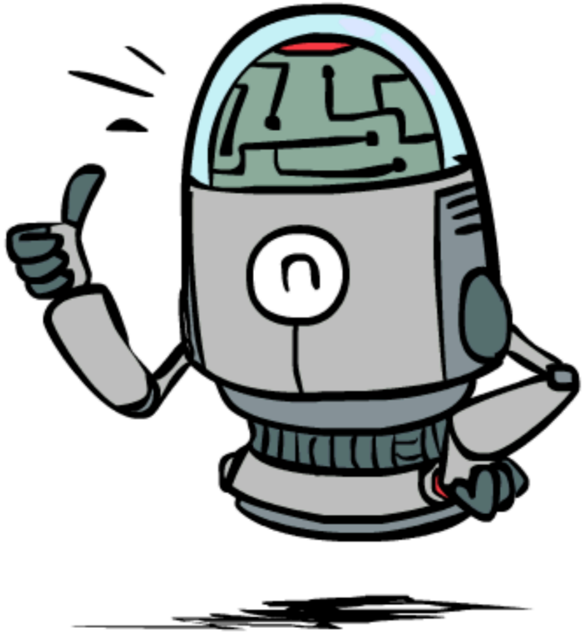
CS 5522: Artificial Intelligence II

Final Review



Instructor: Wei Xu

Ohio State University

[Cartoons were adapted from CS188 Intro to AI at UC Berkeley.]

Final Exam

- The final exam will be closed notes, books, laptops, and people. (You will be given an instructor-provided “cheat sheet”)

- 105 minutes. 12/12 Wednesday 12:00-1:45pm. This room.

- Preparation:
  - Lecture Slides
  - Written Homework
  - Example Exam
  - Project 3: Ghostbusters

Final Exam

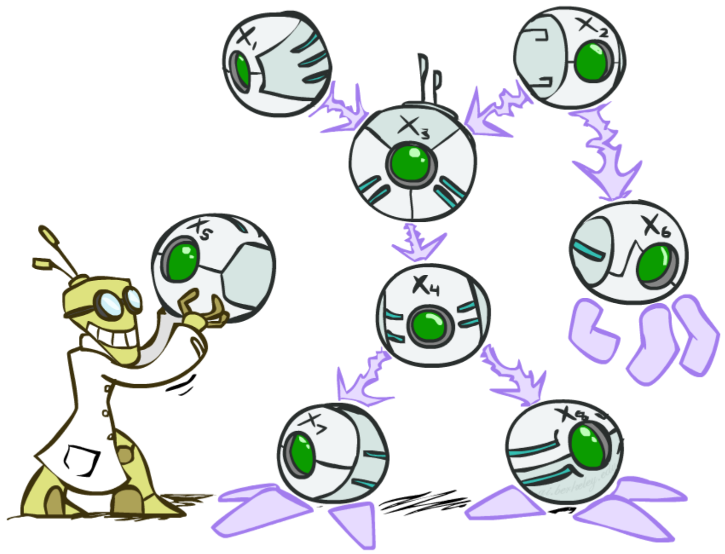
- Make sure you **understand** the fundamentals in addition to being able to procedurally execute algorithms.

- The exam will not test your knowledge of Python, however questions may assume **familiarity with the projects** and test ability of **writing pseudocode**.

- See **written homework** and **example exam** for examples

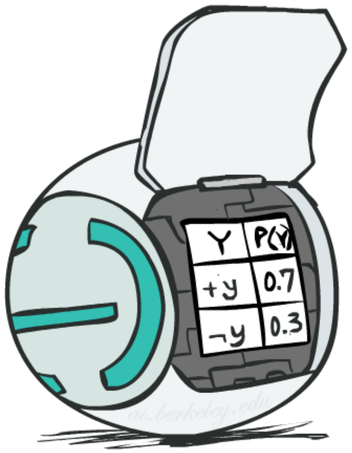
Possible Final Exam Topics

- Markov Models and HMMs:
  - Markov Model and HMM definition
  - Observation and Time Elapse
  - Exact Inference
  - Approximate Inference
  - Particle Filtering
  - Most Likely Explanation



- Bayes Nets:
  - Probabilistic Representations
  - D-Seperation
  - Inference
  - Variable Elimination

- Naive Bayes:
  - Classification
  - Parameter Estimation
  - Prediction



Possible Final Exam Topics

- Will **not** cover the follows:
  - Search
  - Reinforcement Learning
  - Neural Networks
  - Speech Recognition
  - Computer Vision
  - Natural Language Processing

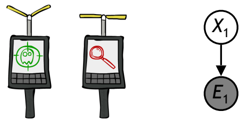
Office Hours

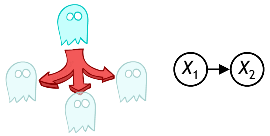
- Office Hour - Wednesday, 12/5 4-5pm (DL 495)
- Office Hour - Friday, 12/7 4-5pm (DL 495)

- Will also answer questions on Piazza

Instructor provided “cheat sheet”

Inference: Base Cases





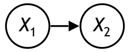
$$P(X_1|e_1)$$

$$P(x_1|e_1) = P(x_1, e_1)/P(e_1)$$
$$\propto_{X_1} P(x_1, e_1)$$
$$= P(x_1)P(e_1|x_1)$$

$$P(X_2)$$

$$P(x_2) = \sum_{x_1} P(x_1, x_2)$$
$$= \sum_{x_1} P(x_1)P(x_2|x_1)$$

Passage of Time



- Assume we have current belief P(X | evidence to date)

$$B(X_t) = P(X_t|e_{1:t})$$

- Then, after one time step passes:

$$P(X_{t+1}|e_{1:t}) = \sum_{x_t} P(X_{t+1}, x_t|e_{1:t})$$
$$= \sum_{x_t} P(X_{t+1}|x_t, e_{1:t})P(x_t|e_{1:t})$$
$$= \sum_{x_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})$$

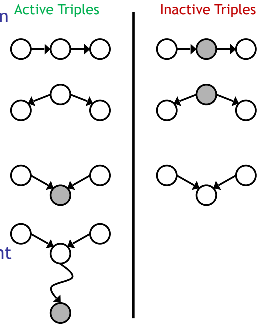
- Or compactly:

$$B'(X_{t+1}) = \sum_{x_t} P(X'|x_t)B(x_t)$$

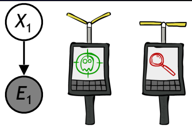
- Basic idea: beliefs get “pushed” through the transitions
  - With the “B” notation, we have to be careful about what time step t the belief is about, and what evidence it includes

Active / Inactive Paths

- Question: Are X and Y conditionally independent given evidence variables {Z}?
  - Yes, if X and Y “d-separated” by Z
  - Consider all (undirected) paths from X to Y
  - No active paths = independence!
- A path is active if each triple is active:
  - Causal chain A → B → C where B is unobserved (either direction)
  - Common cause A ← B → C where B is unobserved
  - Common effect (aka v-structure) A → B ← C where B *or one of its descendents* is observed
- All it takes to block a path is a single inactive segment



Observation



- Assume we have current belief P(X | previous evidence):

$$B'(X_{t+1}) = P(X_{t+1}|e_{1:t})$$

- Then, after evidence comes in:

$$P(X_{t+1}|e_{1:t+1}) = P(X_{t+1}, e_{t+1}|e_{1:t})/P(e_{t+1}|e_{1:t})$$
$$\propto_{X_{t+1}} P(X_{t+1}, e_{t+1}|e_{1:t})$$
$$= P(e_{t+1}|e_{1:t}, X_{t+1})P(X_{t+1}|e_{1:t})$$
$$= P(e_{t+1}|X_{t+1})P(X_{t+1}|e_{1:t})$$

- Or, compactly:

$$B(X_{t+1}) \propto_{X_{t+1}} P(e_{t+1}|X_{t+1})B'(X_{t+1})$$

- Basic idea: beliefs “reweighted” by likelihood of evidence
- Unlike passage of time, we have to renormalize

The Forward Algorithm

- We are given evidence at each time and want to know

$$B_t(X) = P(X_t|e_{1:t})$$

- We can derive the following updates

$$P(x_t|e_{1:t}) \propto_X P(x_t, e_{1:t})$$
$$= \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t})$$
$$= \sum_{x_{t-1}} P(x_{t-1}, e_{1:t-1})P(x_t|x_{t-1})P(e_t|x_t)$$
$$= P(e_t|x_t) \sum_{x_{t-1}} P(x_t|x_{t-1})P(x_{t-1}, e_{1:t-1})$$

We can normalize as we go if we want to have P(x|e) at each time step, or just once at the end…

- Forward Algorithm
- D-Seperation