CSCI E-82a
Probabilistic Programming and AI
Introduction

Steve Elston

HARVARD
Extension School

Why Probabilistic AI?

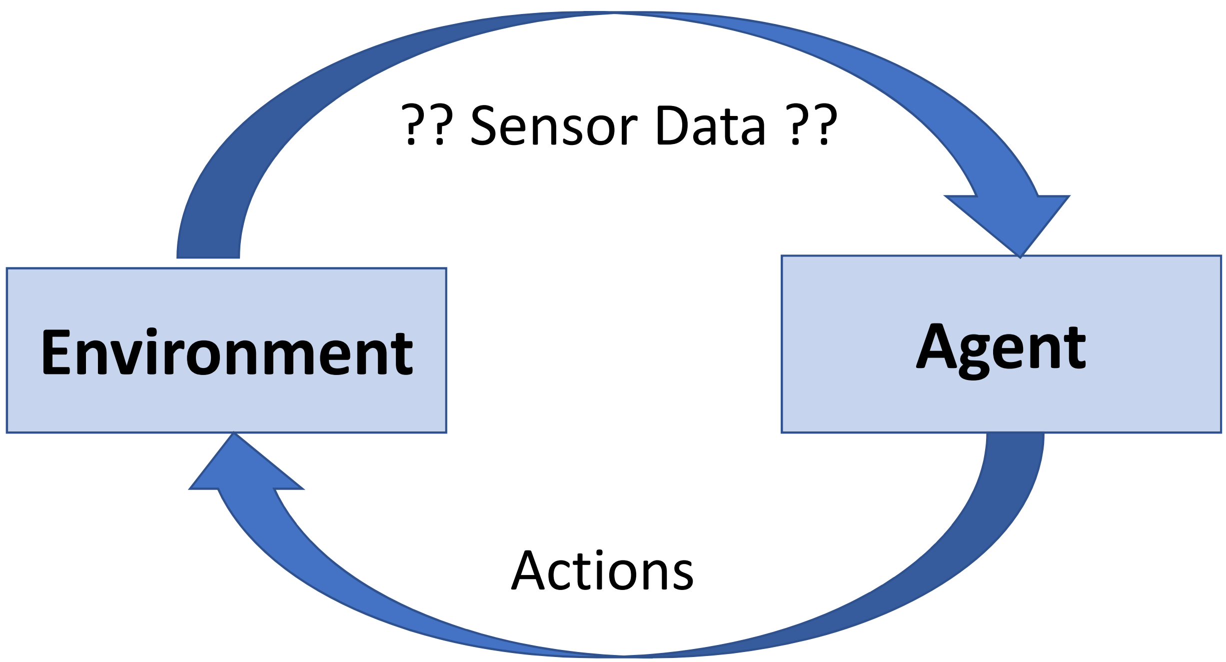
The common theme of this course is making **optimal decisions** in **complex and uncertain environments**

- Intelligent agents must interact with a complex world
- Complex environments lead to uncertainty
- Agents require algorithms that deal with uncertainty
- Probabilistic models, such as **Bayesian models** and **Markov decision processes (MDP)**, allow us to address these problems

Why Probabilistic AI?

Intelligent agent interacts with uncertain environment

- Information from the environment is incomplete and prone to errors
- Agent must take optimal actions given uncertain information



The Intelligent Agent

Fundamental functions of a probabilistic intelligent agent

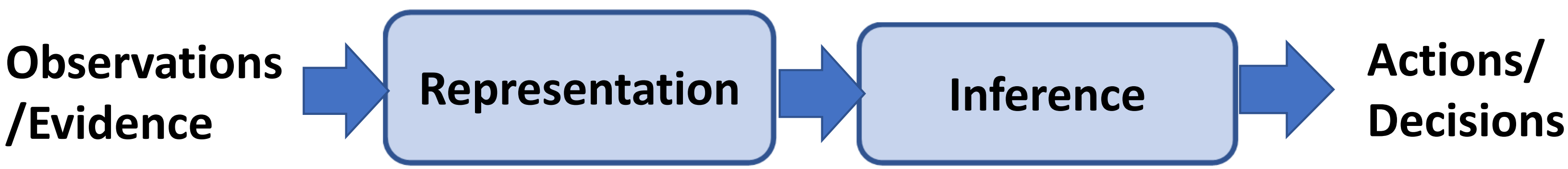
- **Representation**: A good representation is often the key to good machine intelligence. A good representation is a mapping of the model and the environment. Good representation is key to effective AI!
- Representations are often **approximate** given **high complexity** of real world

Representation

The Intelligent Agent

Fundamental functions of a probabilistic intelligent agent

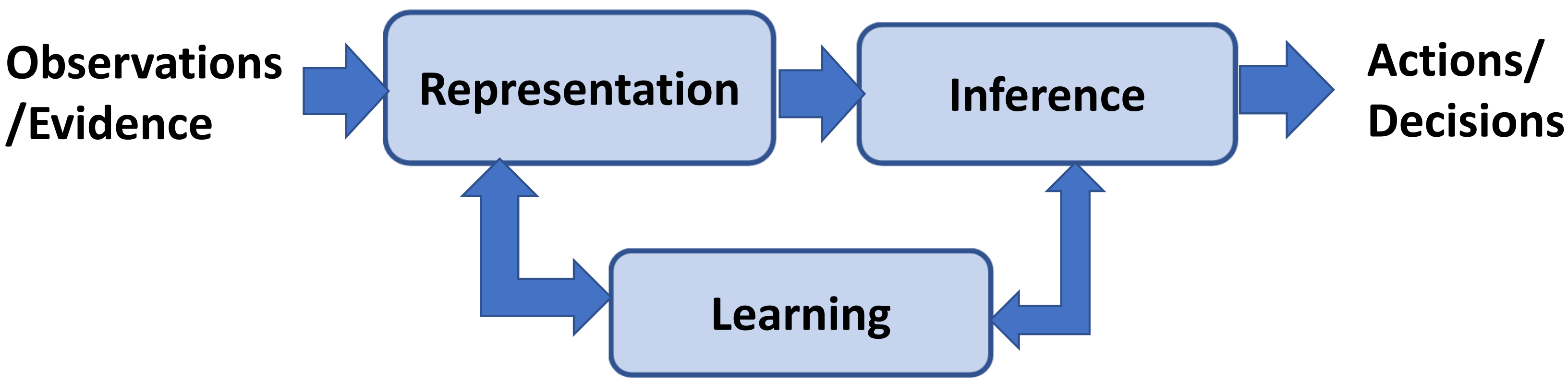
- **Inference or Reasoning:** The process of computing actions or decisions from **queries** of the model given the **evidence**. In the simplest form a query returns a mathematical result, such as the **marginal probability distribution** or the **maximum a posteriori** value.
- Reasoning computes a specific action which is applied to the environment.



The Intelligent Agent

Fundamental functions of a probabilistic intelligent agent

- **Learning**: The agent performs **learning** using data or **evidence** to update the model. The evidence is observed by **sensors** which provide information to the model on the **state of the environment**



Uncertainty in the Environment

Agent must navigate to destination

- Plans optimal route
- How much does the traffic volume change?
- Does the plan account for road repair?
- Does an accident block a route?
- In other words, which decisions are required to minimize travel time?
- Poor response to unexpected information is known a **brittleness** in a model

Uncertainty in the Environment

Integrate sensors for collision avoidance in self-driving car

- Sensors have different range and accuracy
- How are sensors affected by fog, rain or darkness?
- How accurate is traffic sign recognition?
- What is the response of each sensor to snow and ice covered roads?
- In other words, what is the posterior probability that a change in speed or direction is required?

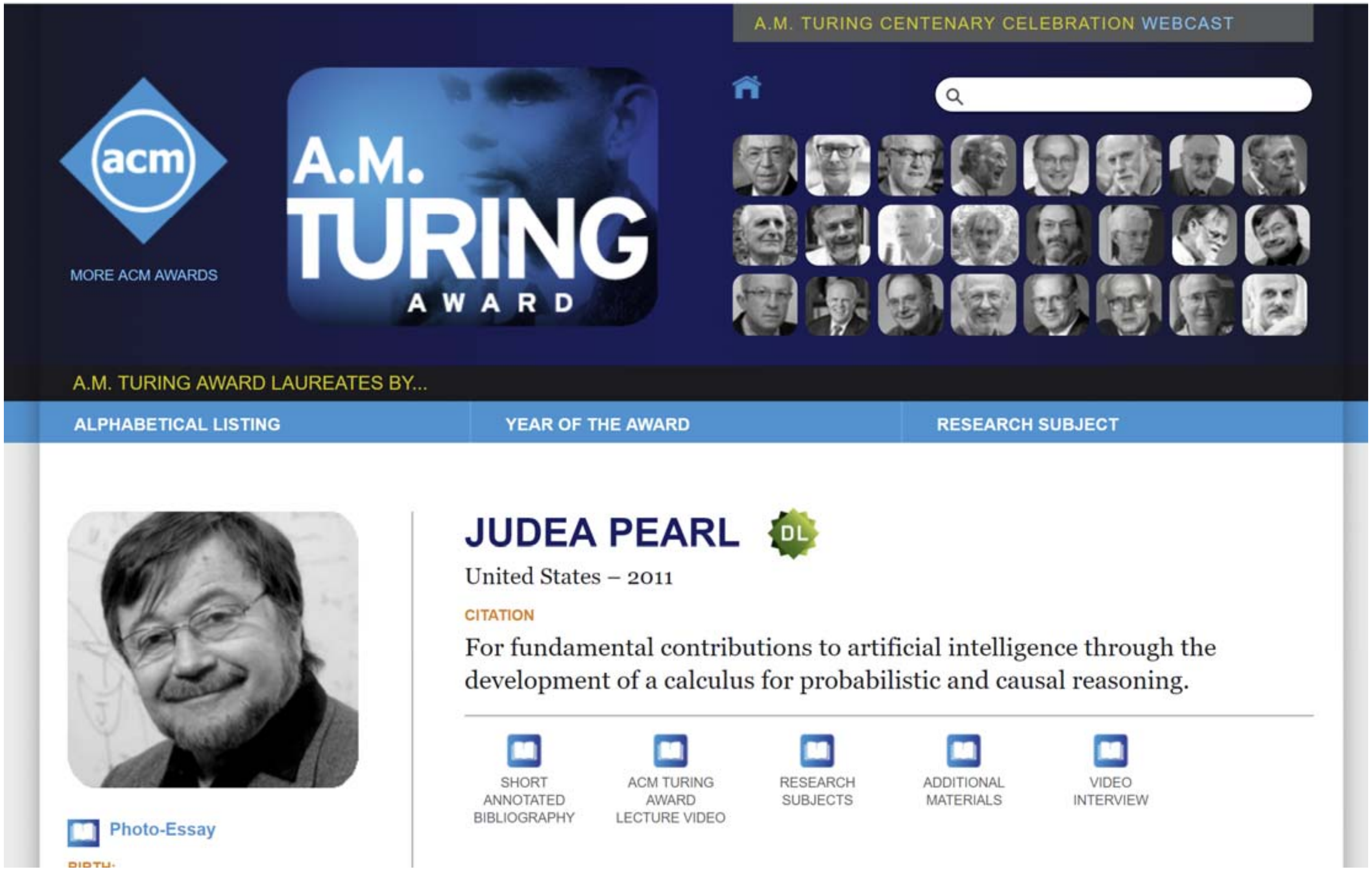
Uncertainty in the Environment

Unobservable information adds to uncertainty

- The intentions of other drivers
- The cards held by other players in a game of poker
- The spot price of wheat in the future
- Net result is **incomplete information**

Probabilistic Reasoning Recognized as Fundamental Method

Check out videos: https://amturing.acm.org/award_winners/pearl_2658896.cfm



About Your Instructor

- Principle Consultant at Quantia Analytics
- Instructor, Harvard Extension School, University of Washington
- MS and PhD in Geophysics from Princeton University
- Work in machine learning starting in 1980s
- Co-founded analytics businesses
- Worked in a number of areas:
  - Capital markets risk
  - Image analysis
  - Fraud detection
  - Forecasting
  - Failure prediction

About Your Teaching Fellow

- Sarah Asano – asano.sar@gmail.com
- Electro-Optical Engineer, Lockheed Martin, Sunnyvale, California
- MS Robotics, Carnegie Mellon University
- BS Mechanical Engineering, California Institute of Technology
- Experience in:
    - App development
    - Game development
    - Internet of things
    - Embedded systems
    - Robots

About This Course

Focus on two different classes of probabilistic algorithms

- Graphical models
  - Efficient method to compute posterior probabilities distributions
  - Sequential decision models
  - Explainable models
- Reinforcement learning algorithms
  - Agent learns by experience
  - Model free
  - Learn policy for complex and stochastic environment
- Models related through Markov Decision Processes (MDP)

About This Course

Grading is based on hands on work and class participation

- Homework assignments – 70%
  - Assignment most weeks
  - Focus on hands-on coding
  - **Read directions carefully and answer all questions;** don’t miss points**!**
- On campus weekend – 30%
  - 9am – 5 pm Dec 7-8. **You must attend the entire session** for course credit!
  - Meet at one Braddle Square, Cambridge
  - Team challenges
  - Book rooms, etc. early

About This Course

Course participation

- Your participation important to get maximum value from this course!
  - Students who attend lection and sections tend to do better.
- On-line lecture – Wednesdays 5:50 – 7:50 pm US Eastern Time
  - Lecture focused on theory
  - Lectures will be recorded
  - Please remind your instructor to record!!
- Section – TBD
  - Section focused on code, questions and homework
  - Perhaps, some background supplement for theory

About This Course

Text Books

- Readings are from two text books
- Both available at the Coop: https://tinyurl.com/300-F19-CSCI-E-82A-1
- Or free pdf downloads
  - Bayesian Reasoning and Machine Learning, Barber, 2012, Cambridge University Press: http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/091117.pdf
  - Reinforcement Learning, an introduction, Second edition, Sutton and Barto, 2018, MIT Press: https://mitpress.ublish.com/book/reinforcement-learning-an-introduction-2

About This Course

Other reference sources I draw material from:

- *Artificial Intelligence*, A Modern Approach, Stuart Russell and Peter Norvig, Prentice Hall, Third edition, 2010
- *Probabilistic Graphical Models*, Principles and Techniques, Daphne Koller and Nir Freedman, MIT Press, 2009
- *Decision Theory Under Uncertainty: Theory and Applications*, Kochenderfer, et. al., MIT Press, 2015.
- *Machine Learning: A Probabilistic Perspective*, Murphy, MIT Press, 2012.
- *Deep Learning*, Ian Goodfellow, Yushua Bengio, and Arron Courville, MIT Press, 2016

About This Course

Getting help with this course – essential component of class participation

1. Plan to attend the section
   - Bring your questions for class discussion
2. Use Piazza – https://piazza.com/class#fall2019/cscie82a
   - Access code: cscie82a
   - Ask questions
   - Answer questions
3. Email Steve – stephen.elston@quantia.com
   - Please only ask **questions of a private nature**; e.g. grading questions
   - Please direct general questions on course material and homework to the aforementioned venues – if you have a question, others likely will too!
4. Grading questions: email Sarah – asano.sar@gmail.com

Course Materials

- Obtain course materials from course Github repository
  - https://github.com/StephenElston/CSCI_E_82A_Probabalistic_Programming
  - Jupyter notebooks with review of theory and code
  - Slides
  - Course material will be updated regularly – **plan on doing a pull regularly**
- Homework assignments will be at: https://github.com/StephenElston/CSCI_E_82A_Probabalistic_Programming/Homework
- Submit completed homework and receive grades in Canvas

First Assignments

- Lesson 0– Self-paced
  - Review of probability concepts – In Github repository
  - Not graded
  - **Decide if this class is for you!**
- Homework 1 – Directed graphical models
  - Due September 18 at 24:00 (midnight) US Eastern Time

AI Is Still A Work In Progress!!

Views of 21st century AI

HAL, open the pod bay door!

As envisioned by Stanley Kubrick
c.a. 1967

Closer to reality?

I’m sorry, but I can’t do that Dave

Did you want the list of pod doors for sale on Ebay Dave?