Probabilistic Foundations of Artificial Intelligence

Introduction

Prof. Andreas Krause
Learning and Adaptive Systems (las.ethz.ch)
Institute for Machine Learning
What is AI?

“The science and engineering of making intelligent machines”
(John McCarthy, ’56)

What does intelligence mean?
Acting Humanly: The Turing Test

- Turing (‘50): Computing Machinery and Intelligence
- Predicted that by 2000, machine has 30% of fooling a lay person for 5 minutes
The Turing Test in Theory & Practice

Conversation with Eugene Goostman:

Interrogator In the first line of a sonnet which reads ‘Shall I compare thee to a summer’s day,’ wouldn’t ‘a spring day’ be better?

Goostman What makes you ask me this? Sound like you can't choose a right answer yourself! Rely on your intuition! :-) Maybe, let's talk about something else? What would you like to discuss?

Interrogator I’d like to talk about poetry.

Goostman Classics? I hope you aren't going to read "King Lear" or something like that to me right now :-) )))

Interrogator No, we can talk contemporary poetry. Do you like Patricia Lockwood's work?

Goostman I am only 13, so I'm attending school so far.

Interrogator Oh, that’s when I first read Auden. Do you have a favorite poem?

Goostman No, I haven't a best-liked poem! How could you imagine such a trash! Did I forget it or you really haven't told me about the place you are from?

Interrogator Does it matter?

Goostman Maybe. I think *you* know the answer. Could you tell me what are you? I mean your profession.
Other Definitions

- Thinking humanly
  - Need to get inside the actual workings of human minds
  - Introspection / psychological experiments / brain imaging
  - => Theory of the mind

- Thinking rationally
  - "laws of thought" / "right thinking"

- Acting rationally
AI Today

- Build systems that act intelligently (“Strong AI”)
  - Build systems that act rationally
  - Act rationally = “perform well on some task”

- Amenable to mathematical analysis, empirical evaluation
- Involves / builds on / integrates
  - optimization, algorithms, control theory, logic, probability & statistics, game theory, ...

- This is what this course is about!

- “Strong AI” still inspiration for the field!
What if We Had Intelligent Machines?

- What will happen to our jobs?
- Should intelligent machines have rights?
- Will machines surpass human intelligence?
- What will we do with superintelligent machines?
- What will they do with us?
- ...

...
Autonomous Driving

- 1994, >1000km 3-lane highway w heavy traffic [Dickmann]
- DARPA Grand Challenges:
  - 2005: drive 150 mile in the Mojave desert
  - 2007: drive 60 mile in traffic in urban environment
- Tesla / GoogleX’s / Uber / ... self-driving car projects

CMU’s Boss

Stanford’s Stanley
Humanoid Robotics

Honda ASIMO

TOSY TOPIO
A Robot Scientist

[King et al, Nature ’04, Science ‘09]
Games

IBM’s Deep Blue wins 6 game match against Garry Kasparov (’97)
Games

- **Go:**
  - 2008: MoGo beats Pro (8P) in 9x9 game
  - 2011: Zen beats Pro (9P) in 19x19 game at handicap 4
  - 2016: AlphaGo wins 5th final game against Lee Sedol

- **Poker:**
  - 2017: Libratus led pro Texas No limit Hold’em players by >$1m
Computer games
February ’11 IBM Watson beat Brad Rutter, Ken Jennings in a $1M competition. Used 200 million pages of structured and unstructured text (incl. Wikipedia). 16TB RAM, 720cores
Image Captioning

1. A person riding a motorcycle on a dirt road.
2. Two dogs play in the grass.
3. A skateboarder does a trick on a ramp.
4. A group of young people playing a game of frisbee.
5. Two hockey players are fighting over the puck.
6. A little girl in a pink hat is blowing bubbles.

[Vinyals et al ‘14]
Topics covered

- Probabilistic reasoning
- Planning under uncertainty
- Learning
- (Deep) Reinforcement learning
- Applications

Focus is on **dealing with uncertainty**

⇒ Foundations in Probability & Statistics
Relation to other ML Courses @ ETHZ

- Machine Learning (Fall)
  - Learning from data
- Data Mining (Fall)
  - How to extract useful information from massive data sets
- Deep Learning (Fall)
  - Neural networks and their applications
- Computational Intelligence Lab (Spring)
  - Matrix Factorization, Recommender Systems
- Statistical Learning Theory (Spring)
  - Theoretical foundations; model validation
Overview

- **Instructor:**
  Andreas Krause (krausea@ethz.ch)

- **Teaching assistants:**
  Alkis Gotovos (alkisg@inf.ethz.ch) – Head TA
  Carlos Cotrini (ccarlos@inf.ethz.ch)
  Sebastian Curi (sebastian.curi@inf.ethz.ch)
  Hoda Heidari (hheidari@inf.ethz.ch)
  Johannes Kirschner (johannes.kirschner@inf.ethz.ch)
  Jens Witkowski (jensw@inf.ethz.ch)

- **Administrative assistant:**
  Rita Klute (rita.klute@inf.ethz.ch)
Course material

- **Textbook:**
  

- **Additional reading on course webpage:**
  
  [https://las.inf.ethz.ch/teaching/pai-f17](https://las.inf.ethz.ch/teaching/pai-f17)
Background & Prerequisites

- **Required**: Solid basic knowledge in probability, linear algebra, algorithms and programming.

- We review necessary background, but will move quickly...
Coursework

- Grade based on written session exam

- ~ Six homeworks (not graded)
  - Mix of theory and programming assignments (Python recommended)

- Recitations (2 sessions of 1 hour; 1-2pm, 2-3pm)
  - Last name < “M”: 1-2pm
  - Last name >= “M”: 2-3pm
  - Discussion of homework solutions
  - Opportunities to ask questions
  - Will start next week
Lecture Recordings

- Lectures are recorded
- URL to be announced
Modeling Rational Systems
Agents and Environments

- **Agents**: Aut. Car, Poker player, Robot, ...
  - Agent maps sequence of percepts to action
  - Implemented as algorithm running on physical architecture
- **Environment** maps sequence of actions to percept
Modeling the Environment

- Set of states $S$ (not necessarily finite)
- State transitions depend on current state and actions (can be stochastic or nondeterministic)
Rationality: Performance Evaluation

- Fixed performance measure

\[ R : S^* \to \mathbb{R} \]
evaluates state sequence

- For example:
  - One point for each clean square after 10 rounds?
  - Time it takes until all squares clean?
  - One point per clean square per round, minus one point per move

- **Goal:** find agent function (program) to maximize performance
What is Rational?

- What is rational at any time?
- Depends on:
  - Performance measure
  - Agent’s prior knowledge of the environment
  - Actions the agent can perform
  - Agent’s percept sequence to date
- Pick action that maximizes the expected performance measure given the above information
Environment Types

- Fully observable vs. partially observable?
- Deterministic vs. stochastic vs. nondeterministic?
- Discrete vs. continuous?
- Known vs. unknown?
- Single-agent vs. multi-agent?

Can have dramatic effect on complexity

- In this course, focus is on single-agent environments (but we explore partially-observable, stochastic, discrete and continuous, unknown environments)
Modeling Agents and Environments

- Engineering principle:
  - Decouple problem specific properties and problem independent algorithms

- Need rich language to specify agents and environments

- One classical approach: Logic

- However, many real-world environments are uncertain: probability theory!

- We are interested in agents that make robust decisions under uncertainty
Uncertainty in AI
Characterizing Uncertainty

- Often, actions can have *uncertain outcomes*
- Sensor observations are *noisy or not available*

- One approach: **Nondeterministic** actions / observations
  - Actions can have multiple outcomes
  - *Not* specified which outcome is more likely
### Problems with Nondeterminism

- **Motion model:** sometimes, actual direction is off in arbitrary direction
- **Nondeterministic planning** finds no feasible solution
- **Suppose,** error occurs with at most 20% chance... What should we do?

![Diagram showing a grid with PIT markings and a gold illustration](image-url)
Review: Probability

- Formally: **Probability Space** $\left( \Omega, \mathcal{F}, P \right)$
  - Set of **atomic events**: $\Omega$
  - Set of all **non-atomic events**: $\mathcal{F} \subseteq 2^\Omega$
    - $\mathcal{F}$ is a $\sigma$-Algebra (closed under complements and countable unions)
      1. $\Omega \in \mathcal{F}$
      2. $A \in \mathcal{F} \Rightarrow \Omega \setminus A \in \mathcal{F}$
      3. $A_1, \ldots, A_n, \ldots \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$

- **Probability measure** $P: \mathcal{F} \rightarrow [0, 1]$
  For $A \in \mathcal{F}$, $P(A)$ is the probability that event $A$ happens
Probability Axioms

Normalization:

\[ P(\Omega) = 1 \]

Non-negativity:

\[ \forall A \in \mathcal{F} : \quad P(A) \geq 0 \]

σ-additivity:

\[ \forall A_1, \ldots, A_n \ldots \text{s.t.} \quad A_i \cap A_j = \emptyset \quad \forall i \neq j, \quad P(\bigcup_{i} A_i) = \sum_{i} P(A_i) \]
Independent Events

- Two random events $A, A'$ are independent iff
  \[ P(A \cap A') = P(A) \cdot P(A') \]

- Events $A_1, A_2, \ldots, A_n$ are independent iff
  \[ \forall B \subseteq \{1 \ldots n\}: \quad P(\bigcap_{i \in B} A_i) = \prod_{i \in B} P(A_i) \]
Interpretation of probabilities

- Philosophical debate...

- **Frequentist** interpretation
  - $P(A)$ is relative frequency of $A$ in repeated experiments
  - Can be difficult to assess with limited data

- **Bayesian** interpretation
  - $P(A)$ is “degree of belief” that $A$ will occur
  - Where does this belief come from?
  - Many different flavors (subjective, objective, pragmatic, ...)

- For now assume probabilities are known
Acknowledgments