Communication

- Email
- Office hours
- Course website (http://cs.allegheny.edu/sites/jjumadinova/teaching/310)
- Slack channel (https://cs310s2020.slack.com/)
  Slack Invitation Link
- GitHub Organization (https://github.com/allegheny-computer-science-310-s2020)
Conceptual topics of this course

- Ethics
Conceptual topics of this course

- Ethics
- Learning
Conceptual topics of this course

- Ethics
- Learning
- Problem Solving
Conceptual topics of this course

- Ethics
- Learning
- Problem Solving
- Uncertainty
Conceptual topics of this course

- Ethics
- Learning
- Problem Solving
- Uncertainty
- Communication
Conceptual topics of this course

- Ethics
- Learning
- Problem Solving
- Uncertainty
- Communication
- Reasoning
Applications of this course

- Game Development
- Computer Vision
- Natural Language Processing
- Virtual Reality
- and more!
Applications of this course

- Game Development
- Computer Vision
- Natural Language Processing
- Virtual Reality
- and more!
Technologies of this course

This will be updated based on the results of the survey
Technologies of this course

This will be updated based on the results of the survey

- Open CV
- TensorFlow
- NLTK
- Google VR
- ...

Please fill out the Background Survey
Technologies of this course

This will be updated based on the results of the survey

- Open CV
- TensorFlow
- NLTK
- Google VR
- ...

Please fill out the: Background Survey
What is AI?

AI is the ability of a computer system to solve problems and to perform tasks that would otherwise require human intelligence.
What is AI?

AI is the ability of a computer system to solve problems and to perform tasks that would otherwise require human intelligence.

“a branch of computer science that studies the properties of intelligence by synthesizing intelligence”
What is AI?

AI RESEARCH PRIORITIES
- Integrated Intelligence
  - Science of integrated intelligence
  - Contextualized AI
  - Open knowledge repositories
  - Understanding human intelligence
- Meaningful Interaction
  - Collaboration
  - Trust and responsibility
  - Diversity of interaction channels
  - Improving online interaction
- Self-Aware Learning
  - Robust and trustworthy learning
  - Deeper learning for challenging tasks
  - Integrating symbolic and numeric representations
  - Learning in integrated AI robotic systems

CROSS-CUTTING ISSUES
- Security & vulnerability, ethics, resources (data, hardware, software, storage, people...)

ASPIRATIONS
- Reduced healthcare cost
- Universal personalized education
- Evidence-driven social opportunity
- Accelerated scientific discovery
- Unprecedented innovation for businesses
- National defense and security

AI-driven capabilities
- Behavioral health coaches
- High payoff experiments
- Opportunistic education
- Resolve supply chain delays
- At-home robot caregivers/helpers
- Effective natural disaster response
- Novel business processes
- Address food and water insecurity
- Resilient cyber-physical systems

DATA-DRIVEN AI METHODS ARE HIGHLY EFFECTIVE BUT HAVE IMPORTANT FLAWS
- Industry focuses largely on practical, near-term solutions using massive proprietary resources
- Academia asks many of the fundamental long-term questions that lay the foundations for AI

ARTIFICIAL INTELLIGENCE (AI) LANDSCAPE

Janyl Jumadinova
Artificial Intelligence
January 13–15, 2020
Why AI?

THE ECONOMIC IMPACT OF ARTIFICIAL INTELLIGENCE

Projected Global Economic Effects of AI by 2030

- NORTH AMERICA: $3.7T
- NORTHERN EUROPE: $1.8T
- LATIN AMERICA: $0.5T
- SOUTHERN EUROPE: $0.7T
- REST OF WORLD: $1.2T
- CHINA: $7T
- DEVELOPED ASIA: $0.9T

Source: PwC
Executive Order on Maintaining American Leadership in Artificial Intelligence

Infrastructure & Technology | Issued on: February 11, 2019
Executive Order on Maintaining American Leadership in Artificial Intelligence

Tech-driven change a key priority for new EC president
Why AI?

- **Boost Health and Quality of Life:** Prevention of illness and elderly ailments, mental/behavioral health, reducing cost (25% feasible) while improving care, remote patient care.
- **Accelerate Scientific Discovery and Technological Innovation:** Biomedical, environmental, new materials, personalized services, robotics, self-driving cars, etc.
- **Lifelong Education and Training:** Personalized, scalable education support. Improve the AI knowledge and skills of people who will lose jobs. Training next generation of AI specialists, data scientists, and software engineers.
- **Evidence-Driven Social Opportunity:** Engaging and empowering disadvantaged communities, improving civic and political discourse.
- **Reinvent Business Innovation and Competitiveness:** Evidence-driven companies, which would increase productivity and value and open new sectors/products.
- **Transform National Defense and Security:** AI-driven systems can compensate for a relatively small cyber defense workforce, adversarial reasoning.
Google’s AI for mammograms doesn’t account for racial differences

By Amits Khaled • January 9, 2020
Side Effects of AI

Google’s AI for mammograms doesn’t account for racial differences
By Amrita Khalid - January 9, 2020

Facial-Recognition Software Suffers From Racial Bias, U.S. Study Finds
Report amplifies concerns about the rollout of algorithms that analyze images

Janyl Jumadinova
Artificial Intelligence
January 13–15, 2020
Side Effects of AI

Google’s AI for mammograms doesn’t account for racial differences

By Amrita Khalid • January 9, 2020

Facial-Recognition Software Suffers From Racial Bias, U.S. Study Finds

Report amplifies concerns about the rollout of algorithms that analyze images

Are robots sexist? UN report shows gender bias in talking digital tech
Side Effects of AI

Google’s AI for mammograms doesn’t account for racial differences

By Amrita Khalid • January 9, 2020

Facial-Recognition Software Suffers From Racial Bias, U.S. Study Finds

Report amplifies concerns about the rollout of algorithms that analyze images

Are robots sexist? UN report shows gender bias in talking digital tech

Artificial Intelligence / Machine Learning

Training a single AI model can emit as much carbon as five cars in their lifetimes

Janyl Jumadinova

Artificial Intelligence

January 13–15, 2020
Robots as Moral Agents

Is it possible to construct some kind of ‘artificial moral agents’? (implicit vs. explicit)
Robots as Moral Agents

Is it possible to construct some kind of ‘artificial moral agents’? (implicit vs. explicit)

If so,
Which moral code should they be programmed with?
Robots as Moral Agents

Moral Machine
What is AI?

January 15, 2020
What is AI?

- making computers that think?
What is AI?

- making computers that think?
- the automation of activities we associate with human thinking, like decision making, learning, ...?
What is AI?

- making computers that think?
- the automation of activities we associate with human thinking, like decision making, learning, ...?
- the art of creating machines that perform functions that require intelligence when performed by people?
What is AI?

- making computers that think?
- the automation of activities we associate with human thinking, like decision making, learning, ...?
- the art of creating machines that perform functions that require intelligence when performed by people?
- the study of mental faculties through the use of computational models?
What is AI?

- the study of computations that make it possible to perceive, reason and act?
- a field of study that seeks to explain and emulate intelligent behaviour in terms of computational processes?
- a branch of computer science that is concerned with the automation of intelligent behaviour?
- anything in Computing Science that we don’t yet know how to do properly? (!)
Weak AI vs. Strong AI

- **Strong (General) AI:** Computer software + hardware alone can emulate a human mind. There is no fundamental difference between man and machine.
Weak AI vs. Strong AI

- **Strong (General) AI:** Computer software + hardware alone can emulate a human mind. There is no fundamental difference between man and machine.

- **Weak (Narrow) AI:** Computer software + hardware alone can simulate every aspect of a human mind. Only people can think, machines cannot.
# What is AI?

<table>
<thead>
<tr>
<th>THOUGHT</th>
<th>Systems that think like humans</th>
<th>Systems that think rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEHAVIOUR</td>
<td>Systems that act like humans</td>
<td>Systems that act rationally</td>
</tr>
<tr>
<td>HUMAN</td>
<td>RATIONAL</td>
<td></td>
</tr>
</tbody>
</table>
Turing (1950) “Computing machinery and intelligence”:

- “Can machines think?” $\rightarrow$ “Can machines behave intelligently?”
- Operational test for intelligent behavior: the *Imitation Game*
Thinking humanly: Cognitive Science

Requires scientific theories of internal activities of the brain

- What level of abstraction? “Knowledge” or “circuits”?
- How to validate? Requires
  1. Predicting and testing behavior of human subjects (top-down) or
  2. Direct identification from neurological data (bottom-up)

Both approaches (roughly, *Cognitive Science* and *Cognitive Neuroscience*) are now distinct from AI
Thinking and Acting rationally

- *Normative* (or *prescriptive*) rather than *descriptive*
- Direct line through mathematics and philosophy to modern AI
Thinking and Acting rationally

- *Normative* (or *prescriptive*) rather than *descriptive*
- Direct line through mathematics and philosophy to modern AI

- **Rational** behavior: *doing the right thing*
- The right thing: that which is expected to maximize goal achievement, given the available information
A (Very Short) History of AI

1642 - First mechanical calculating machine built by French mathematician and inventor Blaise Pascal.

1837 - First design for a programmable machine, by Charles Babbage and Ada Lovelace.

1943 - Foundations of neural networks established by Warren McCulloch and Walter Pitts, drawing parallels between the brain and computing machines.

1950 - Alan Turing introduces a test—the Turing test—as a way of testing a machine's intelligence.

1955 - 'Artificial intelligence' is coined during a conference devoted to the topic.

1965 - ELIZA, a natural language program, is created. ELIZA handles dialogue on any topic; similar in concept to today's chatbots.

2009 - Google builds the first self-driving car to handle urban conditions.

2011 - IBM's Watson defeats champions of US game show Jeopardy!

2011-2014 - Personal assistants like Siri, Google Now, Cortana use speech recognition to answer questions and perform simple tasks.

2014 - Ian Goodfellow comes up with Generative Adversarial Networks (GAN).

2016 - AlphaGo beats professional Go player Lee Sedol 4-1.

1980s - Edward Feigenbaum creates expert systems which emulate decisions of human experts.

2018 - Most universities have courses in Artificial Intelligence.
<table>
<thead>
<tr>
<th>Philosophy</th>
<th>logic, methods of reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mind as physical system</td>
</tr>
<tr>
<td></td>
<td>foundations of learning, language, rationality</td>
</tr>
</tbody>
</table>
| **Philosophy** | logic, methods of reasoning  
mind as physical system  
foundations of learning, language, rationality |
| **Mathematics & probability** | formal representation and proof  
algorithms, computation, (un)decidability, (in)tractability |
<table>
<thead>
<tr>
<th>AI Connections</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Philosophy</td>
<td>logic, methods of reasoning</td>
</tr>
<tr>
<td></td>
<td>mind as physical system</td>
</tr>
<tr>
<td></td>
<td>foundations of learning, language, rationality</td>
</tr>
<tr>
<td>Mathematics &amp; probability</td>
<td>formal representation and proof</td>
</tr>
<tr>
<td></td>
<td>algorithms, computation, (un)decidability, (in)tractability</td>
</tr>
<tr>
<td>Psychology</td>
<td>adaptation</td>
</tr>
<tr>
<td></td>
<td>phenomena of perception and motor control</td>
</tr>
<tr>
<td></td>
<td>experimental techniques (psychophysics, etc.)</td>
</tr>
</tbody>
</table>
| **Philosophy** | logic, methods of reasoning  
|              | mind as physical system  
|              | foundations of learning, language, rationality |
| **Mathematics & probability** | formal representation and proof  
|                        | algorithms, computation, (un)decidability, (in)tractability |
| **Psychology** | adaptation  
|                | phenomena of perception and motor control  
<p>|                | experimental techniques (psychophysics, etc.) |
| <strong>Economics</strong> | formal theory of rational decisions |</p>
<table>
<thead>
<tr>
<th><strong>Philosophy</strong></th>
<th>logic, methods of reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mind as physical system</td>
</tr>
<tr>
<td></td>
<td>foundations of learning, language, rationality</td>
</tr>
<tr>
<td><strong>Mathematics &amp; probability</strong></td>
<td>formal representation and proof</td>
</tr>
<tr>
<td></td>
<td>algorithms, computation, (un)decidability, (in)tractability</td>
</tr>
<tr>
<td><strong>Psychology</strong></td>
<td>adaptation</td>
</tr>
<tr>
<td></td>
<td>phenomena of perception and motor control</td>
</tr>
<tr>
<td></td>
<td>experimental techniques (psychophysics, etc.)</td>
</tr>
<tr>
<td><strong>Economics</strong></td>
<td>formal theory of rational decisions</td>
</tr>
<tr>
<td><strong>Linguistics</strong></td>
<td>knowledge representation</td>
</tr>
<tr>
<td></td>
<td>grammar</td>
</tr>
</tbody>
</table>
## AI Connections

<table>
<thead>
<tr>
<th>Science</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Philosophy</strong></td>
<td>logic, methods of reasoning</td>
</tr>
<tr>
<td></td>
<td>mind as physical system</td>
</tr>
<tr>
<td></td>
<td>foundations of learning, language, rationality</td>
</tr>
<tr>
<td><strong>Mathematics</strong></td>
<td>formal representation and proof</td>
</tr>
<tr>
<td><strong>&amp; probability</strong></td>
<td>algorithms, computation, (un)decidability, (in)tractability</td>
</tr>
<tr>
<td><strong>Psychology</strong></td>
<td>adaptation</td>
</tr>
<tr>
<td></td>
<td>phenomena of perception and motor control</td>
</tr>
<tr>
<td></td>
<td>experimental techniques (psychophysics, etc.)</td>
</tr>
<tr>
<td><strong>Economics</strong></td>
<td>formal theory of rational decisions</td>
</tr>
<tr>
<td><strong>Linguistics</strong></td>
<td>knowledge representation</td>
</tr>
<tr>
<td></td>
<td>grammar</td>
</tr>
<tr>
<td><strong>Neuroscience</strong></td>
<td>plastic physical substrate for mental activity</td>
</tr>
<tr>
<td>Field</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| **Philosophy**   | logic, methods of reasoning  
                  mind as physical system  
                  foundations of learning, language, rationality |
| **Mathematics & probability** | formal representation and proof  
                  algorithms, computation, (un)decidability, (in)tractability |
| **Psychology**   | adaptation  
                  phenomena of perception and motor control  
                  experimental techniques (psychophysics, etc.) |
| **Economics**    | formal theory of rational decisions |
| **Linguistics**  | knowledge representation  
                  grammar |
| **Neuroscience** | plastic physical substrate for mental activity |
| **Control theory** | homeostatic systems, stability  
                  simple optimal agent designs |