

Overview of AI

Dr. Thomas R. Ioerger
CSCE 420 – Spring 2022
read: Ch. 1

What is AI?

- popular view
 - based on movies, sci-fi...
 - Hal from 2001; Data from Star Trek NG; Spielberg's AI; Ex Machina...
- landmark AI systems
 - Deep Blue (chess), AlphaGo
 - IBM Watson (general question-answering)
- real industrial AI applications
 - robotics, autonomous vehicles
 - mail sorting, OCR; face recognition; speech2text, translation...
 - medical diagnosis, chest radiography, fraud detection, legal precedent finding, drug design...
 - stock trading/portfolio risk mgmt/economic modeling/quant funds...

Intelligence

- What is Intelligence?
 - difficult to define
 - the thing that makes humans unique. what about animals, dogs, dolphins...?
- some properties of intelligent behavior
 - doing more than just reacting to the environment
 - acting without being programmed (autonomous decision-making)
 - having volition/motivation/intent
 - does this require free-will? consciousness? self-awareness?
 - learning from experience, adaptiveness
- What do “IQ” tests measure?
 - there are many dimensions to intelligence, not just factual knowledge or reasoning ability

Aspects of Intelligence

- rationality – doing the “right” or “optimal” thing
 - instead of just “reacting to the environment”
 - implies there are goals/objectives (*is there an objective framework?*)
- *thinking vs. doing* distinction
 - are outcomes all that matter, or is how we “think” about it relevant?

Thinking ideally (rationally) <ul style="list-style-type: none">• did we weigh all the pros and cons?• did we make the right inference?• did we do it for the right reasons?	Acting ideally (rationally) <ul style="list-style-type: none">• if we can quantify goals and costs,• we can define optimal actions• usually NP-hard
Thinking humanly <ul style="list-style-type: none">• can we emulate decisions a human would make?• practical choices in face of incomplete info.• humans make mistakes, are biased by emotions	Acting humanly <ul style="list-style-type: none">• play chess like a human (grandmaster)• think about “good” driving• are humans always the best measure?

1. Philosophical Perspective on AI

- attempts to define “correct thought” have a long history
 - Aristotle, Plato...Descartes, Kant, Carnap...Russell, Wittgenstein, Frege...
 - innate vs. empirical knowledge
- syllogisms – laws of thought, or patterns of correct thought
 - Categorical syllogisms: All A are B; x is A; therefore x is B.
 - Socrates is a man; men are mortal; therefore, Socrates is mortal.
 - Conditional syllogisms:
 - Only people with good grades get into college; Kate is in college; Kate must have gotten good grades.
 - Disjunctive syllogisms:
 - All living things are plants or animals; Snoopy is not a plant; therefore, Snoopy is an animal.
 - example of a fallacy:
 - * Insects crawl; Charlotte crawls; therefore, Charlotte is an insect.
 - [incorrect - she is a spider, which is an arachnid]

1. Philosophical Perspective on AI

- grounded understanding – mind/body problem (Descarte)
 - can you have an “isolated” intelligence, or must be connected to sensors?
- consciousness
 - is consciousness necessary?
 - what is it? introspection? self-awareness?

1. Philosophical Perspective on AI

- Turing Test (proposed by computer scientist Alan Turing)
 - how it works
 - 15 min conversation via terminal (blinded)
 - restricted to a topic/subject, like pets, or gardening (but not current events)
 - human assessors judge whether human or computer
 - statistical criterion – pass if can fool >50% of judges – indistinguishable from humans
 - advantages:
 - objective/empirical evaluation
 - integrates reasoning and understanding and communication
 - doesn't matter if it is conscious, or how it is “implemented” internally
 - disadvantages:
 - non-conversational aspects of intelligence – recognizing objects visually, tying shoes...
 - does computer have delays answers and make mistakes to emulate humans? does it have to appreciate humor? analogies?

1. Philosophical Perspective on AI

- Turing Test (proposed by computer scientist Alan Turing)
 - Loebner competition – annual competition, \$100,000 prize
 - no true winner as of 2019
 - chatbots
 - Eliza – emulates a psychiatrist; pattern-based: “Tell me more about X...” “How did that make you feel?”
 - Eugene Goostman – emulates a 13-yr old Ukrainian boy (which induces people to forgive his conversational mistakes)

Eliza transcript

- ELIZA: Please tell me what's been bothering you.
- USER: I feel happy to be writing right now, to be typing this as we speak. Lately, I've had writer's block.
- ELIZA: Can you elaborate on that ?
- USER: I only really feel happy when I'm writing.
- ELIZA: How long have you been writing ?
- USER: Always, but I've been writing seriously since I was eighteen.
- ELIZA: What incident are you thinking of ?
- USER: Um.
- ELIZA: What does that suggest to you ?
- USER: I don't think it's enough, to only be happy when you write.
- ELIZA: Does that trouble you ?
- USER: ...Um, because you can't write all the time. And then there's the real world, and you still have to deal with real people.
- ELIZA: How does that make you feel ?

Eugene Goostman transcript

- <https://www.coventry.ac.uk/primary-news/turing-test-transcripts-reveal-how-chatbot-eugene-duped-the-judges/>

JUDGE: Hello

EUGENE: Hello, I'm really glad to have the chance to chat with you! My guinea pig Bill sends his regards too!

JUDGE: Is Bill a male or a female?

EUGENE: Ask Bill personally, please.

JUDGE: Well I'd rather talk to you. What is your name?

EUGENE: Call me Eugene. I am glad to talk to you!

JUDGE: My name is Jane and I am female. How about you? What's your gender?

EUGENE: I'm a male. A "guy" I'd say.

JUDGE: Pleased to meet you Eugene. What's the weather like where you are?

EUGENE: Let's get on with our conversation!

JUDGE: Don't you like talking about the weather?

EUGENE: All these talks about weather is a waste of time.

JUDGE: What would you like to discuss?

EUGENE: I don't know... Better tell me more about yourself!

JUDGE: what is your gender

EUGENE: I'm a young boy, if you care to know. And not very ugly, by the way!

JUDGE: what are you aware of at this moment

EUGENE: I would rather not talk about it if you don't mind. Could you tell me what are you? I mean your profession.

JUDGE: how old are you

EUGENE: I'm a little boy 13 years old.

JUDGE: what are your qualifications

EUGENE: My qualifications? I'm not gonna tell you. Oooh. Anything else?

1. Philosophical Perspective on AI

- Can “intelligence” be a *computable* function? (posed by Bertrand Russell)
 - think of mapping from “situations” to “actions”
 - (remember – computable functions are expressible as finite combinations of simpler functions using operators like addition, comparison, composition, recursion...)
 - the answer is not clear, but it brings out issues of role of “free will”, etc
- Does intelligence rely on quantum effects? (Roger Penrose)
 - some things are physically unknowable, like joint position and momentum of an electron
 - could the (human) brain be exploiting uncertainty in states of water molecules?

2. Psychological Perspective on AI

- How do humans think?
 - Humans are our best examples of intelligent beings
 - Can we write programs that think/act like humans?
- Example: linguistics
 - How do humans parse and understand sentences?
 - How do we turn sentences into meaning?
 - How can we correct for mistakes? (grammatical, spelling...)
 - How do we resolve ambiguities? (I saw the man on the hill with the hat...)
 - How do we fill-in missing/foreign words, or recognize a pun?
 - How do children learn to speak (language acquisition)?
 - When/why do we commit a slip-of-the-tongue, or a faux pas?

2. Psychological Perspective on AI

- Humans are great at ...
 - making complex decisions in the face of uncertainty (situation assessment; multi-objectives; deadlines, costs, tradeoffs; ethical conflicts...)
 - interpreting perceptual input (recognizing a tiger, or the sound of a tiger), even if there is interference/noise, occlusions, shadows...
 - using context to resolve ambiguities
 - reasoning about other humans' intentions
 - discovering unexpected solutions (“Eureka” effect) (creativity – Douglas Hofstadter)
- But humans also have limitations:
 - humans make mistakes, and are subject to fatigue
 - limits on short-term memory; calculation speed; attention (distraction)
 - decisions can be influenced by emotions
 - biases: recency bias, expectation/confirmation bias
 - long-term memory is malleable, subjective, incomplete
 - decisions are often sub-optimal (satisficing=“good enough”)

2. Psychological Perspective on AI

- paradigms: Behaviorism vs. Information Processing
 - behaviorism: stimulus-response
 - Info Proc: can form internal representations of the world based on sensory input, reason about it, make predictions and decisions
- Cognitive Science
 - concept representations; inter-connections (links) among concepts
 - memory (models of short-term and long-term memory); activation and retrieval
 - focus of attention
 - mental imagery

2. Psychological Perspective on AI

- consciousness
 - debate among modern philosophers John Searles, Daniel Dennett
 - can it be explained as an “illusion” of perceiving stream of perceptual states?

2. Psychological Perspective on AI

- (Major paradigms for intelligence and problem-solving)
- **Symbolic AI** (John McCarthy; Herb Simon; Alan Newell)
 - Symbol Systems Hypothesis – thoughts are like discrete abstract tokens that relate to and generate other tokens (e.g. by reminding or classification or inference)
 - GPS (General Problem Solver) – algebra problems, checkers, question-answering...
 - ACT (Jon Anderson) – psychology-based simulation of memory and rules (production system) that simulates human subject task performance
 - SOAR (Alan Newell) – Universal Subgoaling (when stuck, humans create a new search space to figure out how to get unstuck)
 - Is this the (only) path to truly intelligent behavior?
 - strengths: tractable for computing (world idealized as discrete states, production rules)
 - weaknesses: 1) uncertainty/partial matching; 2) need to program a huge network of concepts and rules to do anything productive (“knowledge engineers”); 3) brittleness

2. Psychological Perspective on AI

- (Major paradigms for intelligence and problem-solving)
- **Connectionism** (Marvin Minsky)
 - perceptrons, neural networks, Deep Networks!
 - Is this the (only) path to truly intelligent behavior?
 - strength: ideal for perceptual and motor-control tasks; all you need is training examples
 - weakness: can't "explain" why decisions are made; concepts are distributed as weights throughout network, tuned through prior experience; output nodes are activated collectively from connections

3. Engineering Perspective on AI

- what do we need to built programs that solve complex problems?
- need more sophisticated *algorithms* in our toolbox
- performance is everything; it doesn't matter how it is achieved (thinking ideally or like a human is irrelevant; consciousness is not required)
- task-focused intelligence; “wholistic” (integrated) AI is not necessary

3. Engineering Perspective on AI

- Core Concepts
 - Search Algorithms
 - Knowledge Representation and Inference
 - Machine Learning

3. Engineering Perspective on AI

- Search
 - almost everything in AI reduces to a Search problem (in some space)
 - planning, learning, parsing, automated deduction...
 - we will study various Search Algorithms
 - especially Heuristic Search (A*)

3. Engineering Perspective on AI

- Edward Feigenbaum: “In the knowledge is the power.”
- Knowledge-based programming
 - rule bases
 - declarative programming vs. procedural programming: say “what” not “how”
- Knowledge Representation
 - formal logic(s) (propositional/Boolean, predicate/First-Order logic, epistemic logics, temporal logics, default logics, fuzzy logic...)
 - ontologies for qualitative reasoning about events, time, spatial relationships, processes...
 - challenge: how to express uncertainty? probability, confidence factors...

3. Engineering Perspective on AI

- Inference Algorithms
 - how can you generate deductions from a knowledge base?
- Is an action or decision a logical consequence of what I know?
- How to generate proofs?
- Is a formula satisfiable? Is there a faster way than try all combinations of variables?

- (Has applications to reasoning about mathematics, proving theorems in number/set/group theory, and computer program verification.)

3. Engineering Perspective on AI

- *Expertise* in high-performing humans
 - chess grandmasters; military generals; football coaches; air-traffic controllers; 5-star chefs...
 - typically accumulated from experience (hence, think: knowledge base)
- Expert Systems
 - example: MYCIN – asks questions to diagnose blood diseases and infections
 - commonly used ES shells: PROLOG, CLIPS, JESS
 - can often generate explanations of decisions
- Major limitation of ES: *brittleness*
 - ES often draw inaccurate/silly conclusions outside scope of what they were designed for

3. Engineering Perspective on AI

- Machine Learning
 - who says machines can't be adaptive and learn from experience?!
 - optimizing performance for an objective function
 - many learning methods:
 - statistics (regression)
 - concept generalization, features, classification trees
 - rule revision (e.g. adding antecedents, special conditions)
 - clustering, pattern recognition
 - neural networks
 - reinforcement learning: sequences of actions to achieve goals in dynamic environments
 - to develop/build/program/implement a system, you need to *train* it on *examples*

3. Engineering Perspective on AI

- Complexity
 - we will see that most “AI” problems are computationally difficult
 - often NP-complete or worse
 - huge search spaces (exponential or infinite)
 - much of AI focuses on using knowledge and heuristics to reduce complexity and find an acceptable solution (often sacrificing completeness/optimalty)
 - this connects AI to *optimization*
 - greedy algorithms; approximation algorithms
 - control theory; operations research