

Harnessing Artificial Intelligence

CS4000

WINTER 2024

Peter Denning
Producer

Our Purpose

- EXPLORE

- What is AI?
- What is ML?
- What are their benefits and risks?
- What is their military significance?

AND ALSO

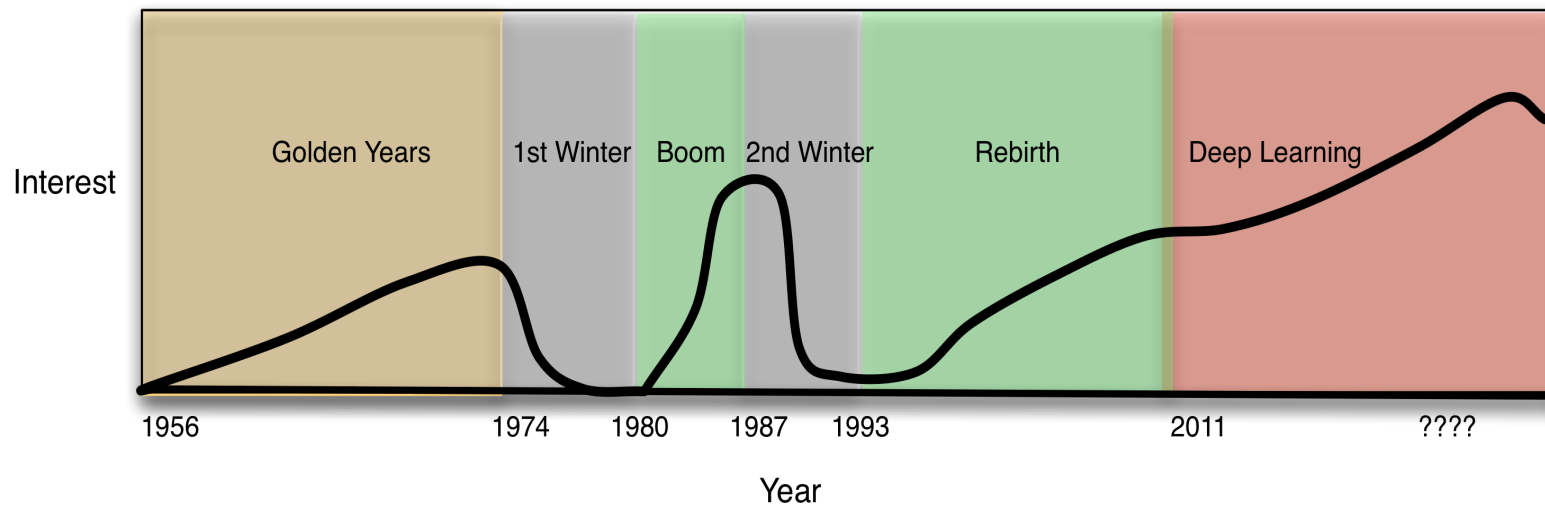
Establish a common language with precise definitions for discussing AI/ML.

Cut through the hype.

“Don’t mistake utility and speed for intelligence.”

AI field

- Founded 1956 to study how to build computing machines that perform intelligent functions
- Buoyed by enthusiasm
- Plagued by hype and over-promising



speech recognition
 board games (chess)
 language translation
 simple robots
 problem solving systems
 neural networks
 perceptrons

expert systems
 knowledge revolution
 multilevel neural networks
 statistical inference
 automatic classification
 sparse distributed memory
 Science Grand Challenges
 Strategic Computing Initiative
 Fifth Generation Project

data analytics
 deep learning
 image recognition
 voice transcription
 voice commands (Siri, Alexis)
 grandmaster game players (Chess, Go)
 human machine symbiosis
 driverless vehicles
 automatic fire control
 drone surveillance
 swarms
 cyber defense and offense



What is AI?

Popular notions:

- Self-driving cars and auto-pilots
- Thinking machines
- Conversation machines
- Puzzle-solvers
- Games
- Speech recognizers
- Vision recognizers
- Expert systems

Hollywood notions:

- Terminator
- Skynet
- I, Robot
- Battlestar Gallactica
- Enders Game
- Commander Data
- R2D2

Things that worry us about AI

- Automated weapon systems
- Drone swarms becoming WMD
- Mastering large battlespaces
- Unexplainable recommendations
- Fragile neural networks
- Mass surveillance and population control
- Mass unemployment
- Untrustworthy Generative AI
- Deepfakes, rampant disinformation
- Internet slime and data pollution

What is AI?

- It's

Computers doing human intelligent tasks

- But wait ...

What is intelligence?

What is intelligence?

- Problem: intelligence is ill-defined
- Some kinds of intelligence are not human – ants, bees, dogs, cats, chimpanzees, and others
- Elusive goalposts: “When a machine does it, we no longer think of it as intelligent”

- Intelligence is NOT speed
 - Machines do 1 Billion calculations per second
 - Humans do 1 calculation per second
- Your laptop and nearest supercomputer are not intelligent

Scientific Definition of Intelligence?

- Many fields have an orienting concern that cannot be defined precisely. Examples:

Artificial Intelligence

Biology

Neuroscience

Computer science

- Artificial Intelligence: when is a machine intelligent?
 - Turing tests? Logical reasoning? Associative memory?
Natural language? Achieving self-determined goals?
- Biology: when is an organism alive?
 - Checklist: Nutrition, respiration, movement, excretion,
growth, reproduction, sensitivity
- Neuroscience: when is someone conscious?
 - Correlates: eye movement, breath, brainwave measures
- Computing: what is information?
 - Symbolic? Tacit? Assessment? Substance? Force of nature?

Literal interpretations of AI and ML

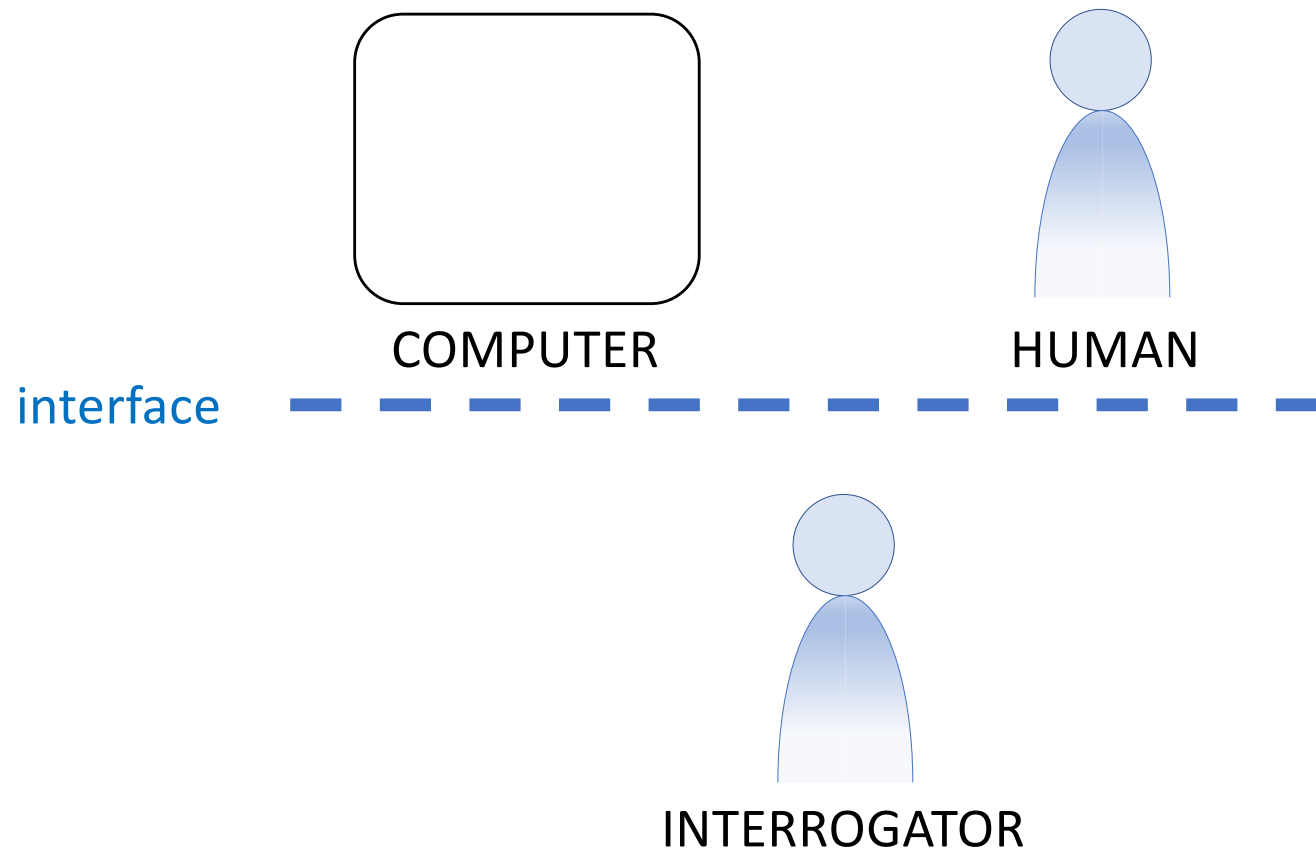
- AI = artificial intelligence
 - Designing intelligent machines
 - Thinking, intuition, making conjectures, imagination, intention, making and keeping commitments, consciousness, cognition, self-awareness
- ML = machine learning
 - Designing machines that learn new functions
 - Speech recognition, image recognition, games, planning, deep learning, conversations, driverless vehicles, automatic fire control, swarms

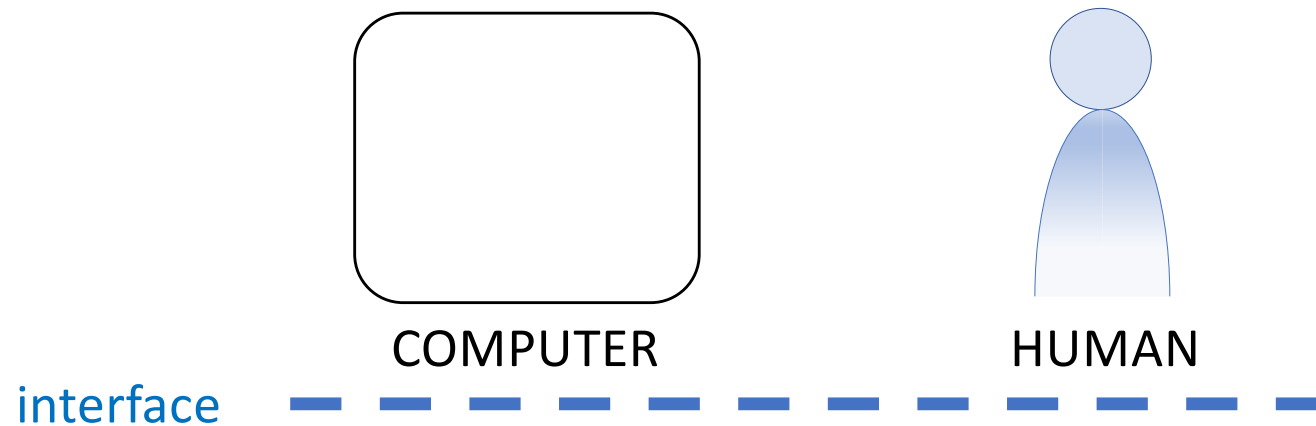
By these interpretations

- AI = artificial intelligence
 - No one has any idea how to do this
 - Little progress since 1950s
- ML = machine learning
 - Almost all recent AI progress is actually ML
 - Most military applications are ML

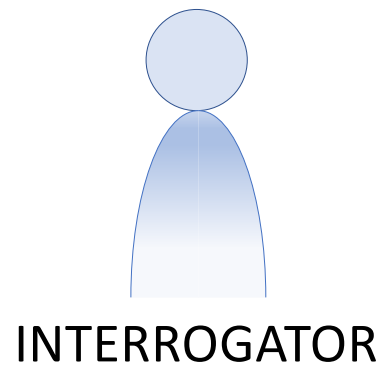
The Turing Test (1950)

- Alan Turing sidestepped defining intelligence by specifying an "imitation game"
- Can a conversation machine fool a human interrogator?





How long does it take
interrogator to distinguish
computer from human?



Even if indistinguishable,
is a simulation of
intelligence intelligent?

Joseph Weizenbaum's ELIZA program (1966)

- Very short program using keyword substitution to simulate Rogerian psychotherapy
- Many thought it was intelligent
- He intended it to discredit Turing Test

Sample on next page

YOU: my relationship is in trouble

ELIZA: Is it important to you that your relationship is in trouble ?

YOU: yes

ELIZA: You are sure.

YOU: yes

ELIZA: I understand.

YOU: can we discuss my spouse

ELIZA: Why do you say your spouse ?

YOU: that's why my relationship is in trouble

ELIZA: Is it important to you that your relationship is in trouble ?

YOU: yes, i want to fix it

ELIZA: You seem to be quite positive.

YOU: that's what i want

ELIZA: What would getting mean to you ?

YOU: acquiring

ELIZA: Does that have anything to do with the fact that your relationship is in trouble ?

YOU: yes, my spouse nags me

ELIZA: You are sure.

YOU: all the time

ELIZA: But your spouse.

YOU: no one else

ELIZA: Are you thinking of a very special person ?

YOU: my spouse

ELIZA: Please tell me what's been bothering you.

YOU: what's been bothering me

ELIZA: Why do you ask ?

YOU: what's been botherin me

ELIZA: Does that question interest you ?

YOU: what's been bothering me

ELIZA: What do you think ?

YOU: what's been bothering me

ELIZA: Does that question interest you ?

YOU: what's been bothering me

ELIZA: What comes to mind when you ask that ?

YOU: what's been bothering me

ELIZA: Are such questions much on your mind ?

YOU: what's been bothering me

ELIZA: What answer would please you most ?

YOU: what's been bothering me

ELIZA: What comes to mind when you ask that ?

YOU: what's been bothering me

ELIZA: Have you asked such questions before ?

YOU: what's been bothering me

ELIZA: What comes to mind when you ask that ?

YOU: what's been bothering me

ELIZA: Are such questions much on your mind ?

YOU: █

Turing test changed the (subjective) intelligence question to (objective) machine output question

But that did not help

We still don't know how to build a machine that passes the test

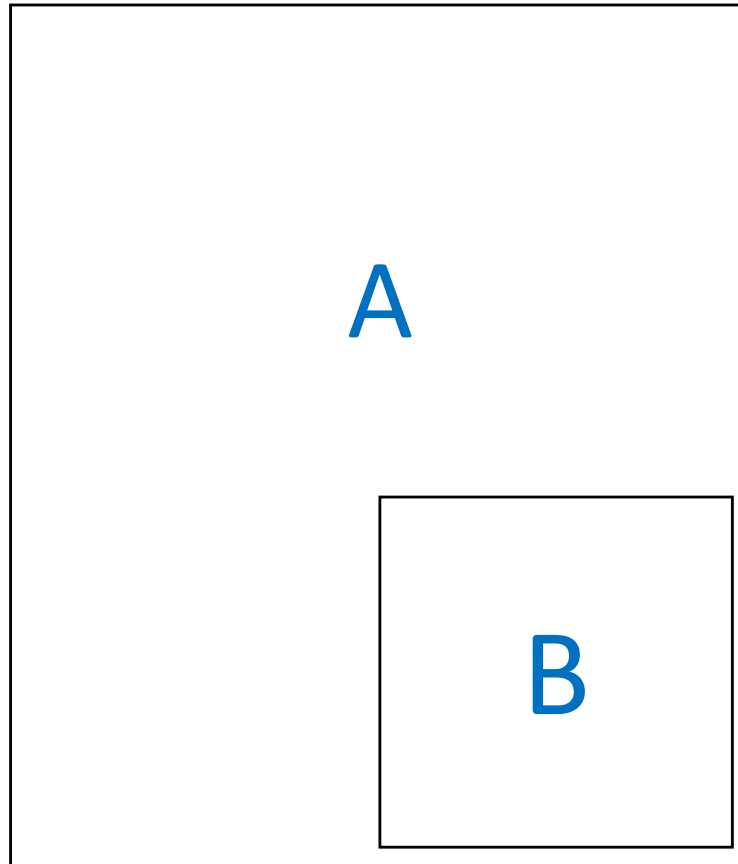
Generative AI (e.g., ChatGPT) a step closer?

What is a machine?

- Apparatus of components that automates a process humans want done
- Grants an advantage such as amplified force or speed that humans do not naturally have
- Human partially or fully out of the loop
- Computer is machine with hardware components controlled by software – implements input-output functions described by algorithms

Hierarchy of Learning Machines


- Classify existing machines by learning power
- Learn: acquire a new capacity for action
- Compare machine abilities, limitations, risks
- Precise language for evaluating learning machines



Machine A is more powerful than B:
A can learn functions B cannot

| | |
|---|---|
| 0 | Basic Automation |
| 1 | Rule-based systems |
| 2 | Supervised learning |
| 3 | Unsupervised and reinforcement learning |
| 4 | Generative AI |
| 5 | Human-machine interaction AI |
| 6 | Aspirational AI |

| | |
|---|---|
| 0 | Basic Automation |
| 1 | Rule-based systems |
| 2 | Supervised learning |
| 3 | Unsupervised and reinforcement learning |
| 4 | Generative AI |
| 5 | Human-machine interaction AI |
| 6 | Aspirational AI |



Different
degrees of
automation

Level 0 – Basic Automation

- Baseline – classical automation of unintelligent tasks
- Get a machine to perform a process
- Human (mostly) out of the loop
- May have feedback to ensure stable operation, but feedback does not change the function (no learning)
- ML may be used in some components
- Examples: autopilots, container shipping, robotic warehouse, voice response robots

Level 1 – Rule-based systems

- Programmed systems that make logical deductions from an input with help from a database of facts and rules of inference
- Early examples: checker and chess
- Expert systems – aim to perform like an expert in a domain
- Examples: medical diagnosis, image interpretation, equipment repair, logistics

Level 2 – Supervised learning

- Artificial neural networks
- Train the network to learn a particular input-output function given a large set of examples
 - When outputs are “labels” for inputs, network classifies
- Trained network approximates the desired function and may give inexact answers to some trained inputs
- What happens when network shown input not in the training set?
- Reliable training data a real problem
- Examples: success of image and speech recognition, language translation; weird ways the network can go wrong

Level 3 – Unsupervised and reinforcement learning

- Learn a function from given data without being shown examples or guidance from a trainer
- Avoid the “sufficient data” problem
- Classification of unknown data into similarity clusters
- Still need human interpretation of meanings of clusters
- Early example: AUTOCLASS (NASA astronomy)

- Board games such as Chess and Go
- Use reinforcement learning, machine v. machine
- Reinforcement: feedback that rewards moves associated with wins and penalizes moves associated with losses
- Example: AlphaZero (Chess, Go)
 - Learn grandmaster Chess in 4 hours
 - Learn grandmaster Go in 13 days

Level 4 – Generative AI

- Generate new texts, poems, pictures, images, music, art, and code that are taken by humans as “creative”
- Also known as GPT (as in ChatGPT) and Large Language Models (LLM)
- Powered by monster artificial neural network
 - GPT-4: 1 trillion parameters, trained from text corpus of hundreds of billions of words, 90-100 days of supercomputer time and \$100 millions worth of electricity to train
- Given an input text (“prompt”), generates output text most likely to have been said relative to the training data
- First neural network machines capable of human-like conversations

- Most promising uses: customer service chatbots, initial drafts of text and code, provocations, jumpstarting, amusements

Huge investments in the technology

- Untrustworthy: Hallucinations and fabrications common

Avoid using in critical and sensitive applications

Larger size does not seem to improve this

- IP issues due to scraped data

- Fears of data pollution (“AI slime”), deepfakes, misinformation, job loss, threat to humanity

Level 5 – Human-Machine Interaction AI

- A system designed with an interface that allows humans and machines to do what each is best at, producing a result that is better than any human or machine can do alone
- Aim to amplify or augment human intelligence
- Examples: freestyle Chess, navigation-assisted cars

Level 6 – Aspirational AI

- Nowadays called AGI – Artificial General Intelligence
- These are all the intelligent machines AI researchers have dreamt about building ... the holy grails ... but to date none exist
(and no one knows how to do them)
- Endless debates on whether such machines are possible
- No military AI depends on this level

The Context Problem

- To bring learning machines closer to intelligent machines we will need a solution to the Context Problem
- Context: human ability to sense issues unstated in the background of our experience, bringing them forth to give meaning to what is around us
- The Cyc Project: database of 30 million “common sense facts” did not help expert systems become experts

- Machines do not sense context
 - Their parts work “locally”: generate outputs (signals and symbols) and in response to inputs (signals and symbols)
 - Do not understand the meaning of any input or output
 - All components are “context free”
 - Source of their great speed
- Until we can formalize “bringing forth from the context” ... no way to get a machine to do it

CAUTIONS

1. Hype
2. Anthropomorphizing
3. Overbroad definitions
4. Master Algorithm

1. – Hype

Hype: a common bad habit of claiming more than machines can deliver

- Hype bandwagon: drop the term “AI” liberally into all presentations and proposals so as not to be left out of the AI game, even when AI has no known technology to do what you’re discussing

2. – Anthropomorphizing

Anthropomorphizing: attributing human attributes of intelligence to machines

- We love to project our ideas of intelligence into the machine:
 - “The computer is an **electronic brain**”
 - “The computer **knew** where I was”
 - “The computer **thought** the input was wrong”
 - “The computer **understood** my problem”
- But it’s not a brain
- It does not know, think, care, or understand

Humans

--

Machines

- Social communities
- Care, Empathy, Compassion
- Commitments
- Judgments
- Invention
- Imagination
- Sensitive to context

- Calculations
- Logic
- Search
- Retrieval
- Comparisons
- Never bored
- Context free

3. – Overbroad definitions

- Example: “AI = machines with ability to perform functions normally considered intelligent”
 - No agreement on what functions make intelligence (if any)
 - Some people maintain that all human functions are intelligent and therefore all human-developed machines are AI, back to time immemorial
 - Some maintain that machine functions cannot be intelligent
 - Some claim that all software is AI (most software is not)

4. – Master Algorithm

- Claim that deep learning is the “master algorithm” for all AI
 - Alan Turing: there can be no such thing as a master algorithm
- Deep learning is level 2 in hierarchy, many more powerful kinds of machines above it
- Related claim: automated systems prior to arrival of deep learning were “handcrafted” – i.e., built by ad hoc methods
 - Pre-AI systems are well engineered
 - Today’s deep learning systems are handcrafted

The Best Way Forward

- Make careful assessments whether AI can help a military system by looking at the requirements and evaluating the machines, algorithms, and data used to realize the requirements
- What machines of the hierarchy, and what other machines, must be combined to do the military job?