A Hierarchy of AI Machines

Peter Denning

Objective

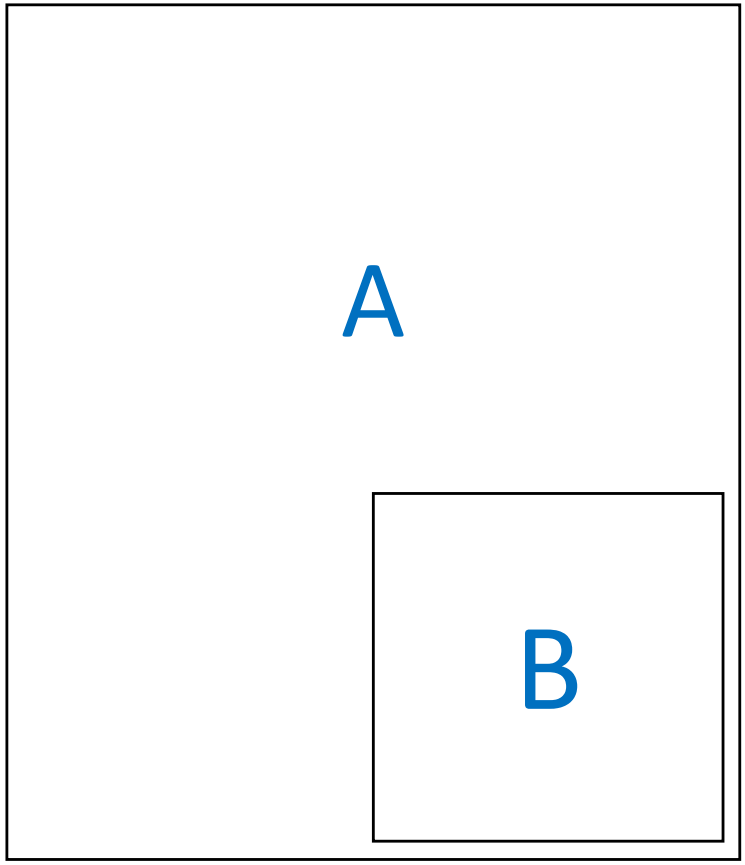
Classify existing AI machines by learning power

- Compare their abilities, limitations, and risks
- Give examples
- Precise, common language for discussing AI machines

Learn: acquire a new capacity for action

Machine A is more powerful than B:

A can learn functions B cannot


A
B

0	Basic Automation
1	Rule-based systems
2	Supervised learning
3	Unsupervised learning
4	Human-Machine Teaming
5	Aspirational AI

Why not classify by Intelligence?

- Intelligence ill defined
  - Disagreements about which animals are intelligent
  - The “moving goalposts” problem with machines
  - Near universal agreement that existing machines not intelligent
  - The Turing Test and Eliza problem

```
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:
```

- The best we can do is classify by correlates of intelligence -- abilities for
  - Learning
  - Problem-solving
  - Conversation
  - Thinking
- Learning is most easily quantified

Similar to “Life Problem”

- Biologists: which organisms are living?
- 7 criteria:
  - Nutrition, respiration, movement, excretion, growth, reproduction, sensitivity
- Use to assess whether life is present

Similar to “Consciousness Problem”

- Neuroscientists and Brain scientists
  - When is a person conscious?
  - What kinds of consciousness do animals have?
- What are “neurological correlates of consciousness” (NCC) – measurable neural or brain activities that reliably tell when someone is conscious?
- Many claim: Consciousness is widespread but not computable – machines can’t be conscious

0	Basic Automation
1	Rule-based systems
2	Supervised learning
3	Unsupervised learning
4	Human-Machine Teaming
5	Aspirational AI

Level 0 – Basic Automation

- Baseline – not intelligent
- Get a machine to perform a process, human out of the loop
- May have feedback to ensure stable operation, but feedback does change the function (no learning)
- AI may be used in some components
- Dr Joshua Kroll will discuss this with examples such as 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 deduction
- Early examples: checker and chess
- Expert systems – perform like an expert in a domain
- Dr Vinnie Monaco will discuss this with examples in medicine, equipment repair, logistics

Level 2 – Supervised learning

- Most commonly neural networks
- Train the network to learn a particular input-output function given a large set of examples
  - Outputs are “labels” for inputs, network classifies
- Trained network approximates the desired function and may give inexact answers to some inputs
- What happens when network shown an input not in the training set?
- Dr Marko Orescanin will discuss this with examples of successes in image recognition and weird ways the network can go wrong

Level 3 – Unsupervised learning

- Learn a function from given data without being shown examples or guidance from a trainer
- Classification of unknown data into clusters
- Board games such as Chess and Go
- Dr Chris Darken will discuss with examples that demonstrate advanced statistical inference and reinforcement learning

Level 4 – Human-Machine Teaming

- A system is 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
- Dr Rudy Darken will discuss with examples such as freestyle Chess and navigation-assisted cars

Level 5 – Aspirational AI

- These are all the machines AI researchers have dreamt about building … the holy grails … but to date do not exist

  (and no one knows how to do them)

- Dr Neil Rowe will discuss these aspirations including thinking machines, sentient machines, and conscious machines, and assess the likelihood we'll accomplish these dreams

Web Sites

https://nps.edu/web/ai-consortium

https://nps.edu/web/harnessing-ai-course