

CS7IS2: Artificial Intelligence

Lecture 0: Intro and Logistics

Ivana.Dusparic@scss.tcd.ie

Lectures

- › Thursdays 10-11 LB01
- › Fridays 2-3 LB08
- › Starting online, moving to hybrid, recorded
- › TurningPoint – please install app when we move in-person!

Assignments

- › Marks exam : coursework = 50 :50
- › Supplemental exam: 100% exam
- › Coursework
 - 20% individual programming assignment + demo – due week 6
 - 30% group project + presentation – due week 11
- › Deadline
 - No extensions (apart from medical cert or covid equivalent, note from tutor etc)
 - Late submissions: mark docked 33% per day
- › Plagiarism – all submissions will be run through plagiarism checkers and violations dealt with in line with college policies
 - <https://libguides.tcd.ie/friendly.php?s=plagiarism/levels-and-consequences>

Course Material

- › Lecture notes and assignments will be posted on Blackboard
- › Artificial Intelligence: A modern approach. Russel and Norvig, 4th edition. 2020
- › Artificial Intelligence: Foundations of Computational Agents, Poole and Mackworth. 2nd edition 2018
- › <http://norvig.com/> - link to pdf of the book etc
- › <https://people.eecs.berkeley.edu/~russell/>

Additional Optional Reading List

- › Artificial Intelligence: A Very Short Introduction – Margaret Boden
- › Rebooting AI: Building Artificial Intelligence We Can Trust – Gary Marcus and Ernest Davis
- › Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way – Virginia Dignum
- › The Book of Why: The New Science of Cause and Effect – Judea Pearl and Dana Mackenzie
- › Reinforcement Learning: An Introduction – Sutton and Barto
- › Human Compatible: AI and the Problem of Control – Stuart Russell
- › Possible Minds: 25 Ways of Looking at AI – John Brockman
- › Self Comes to Mind: Constructing the Conscious Brain - Antonio Damasio
- › Superintelligence: Paths, Dangers, Strategies – Nick Bostrom
- › Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy – Kathy O’Neill
- › Yoshua Bengio and Gary Marcus on the Best Way Forward for AI
 - <https://medium.com/@Montreal.AI/transcript-of-the-ai-debate-1e098eeb8465>
- › Coded Bias – movie available on Netflix

So what are we
actually going to
learn?



What to teach in an AI course?

- › A survey of current practice and teaching of AI.
 - Wollowski, Michael and Selkowitz, Robert and Brown, Laura E. and Goel, Ashok and Luger, George and Marshall, Jim and Neel, Andrew and Neller, Todd and Norvig, Peter.
 - AAAI'16 Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence
 - <https://dl.acm.org/citation.cfm?id=3016491>

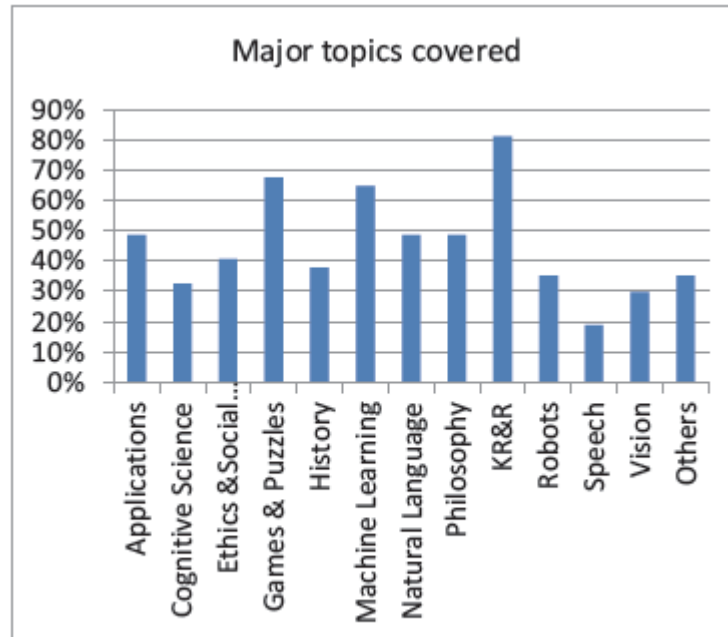


Figure 2: Major topics covered in AI courses

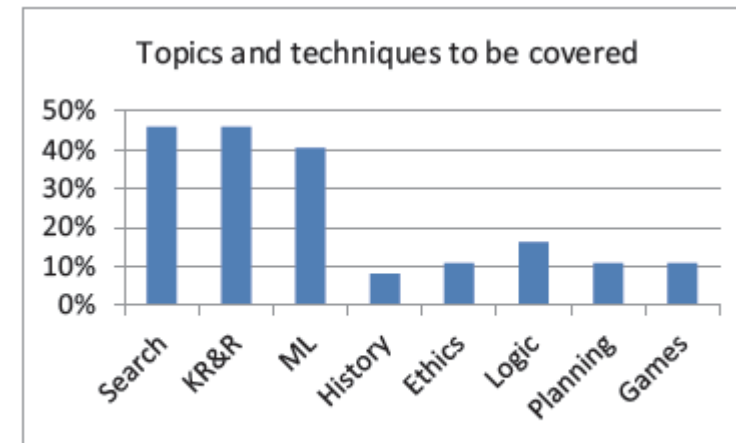
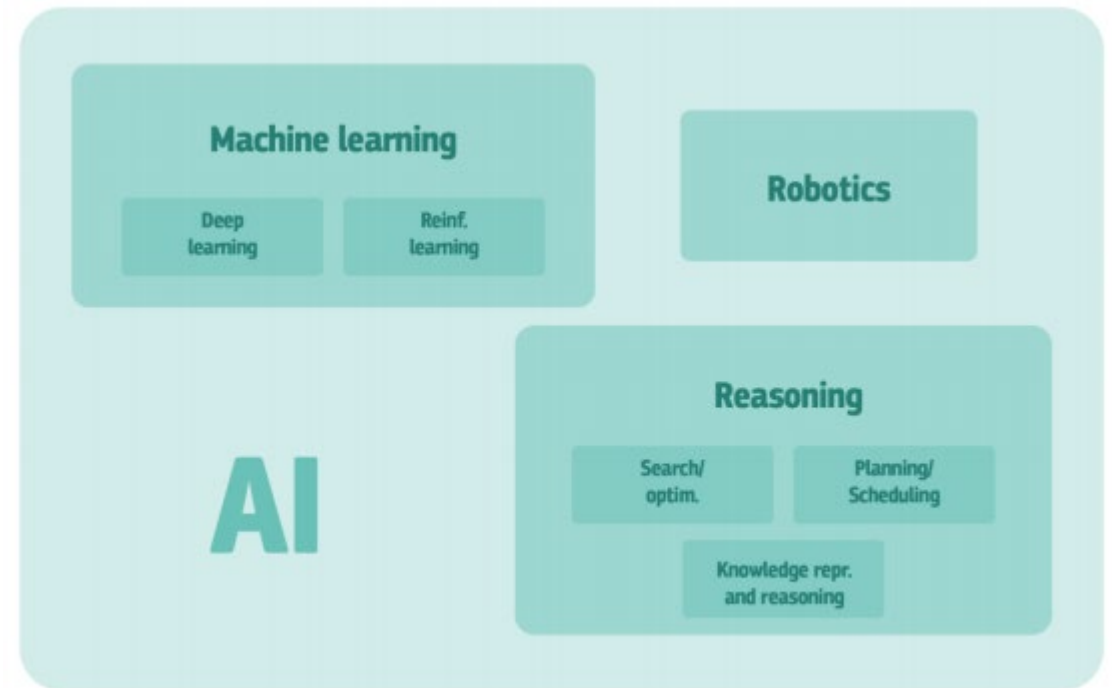


Figure 4: Topics and techniques that should be covered

What is AI?

- › EC High-level expert group on AI
 - Definition, scope etc
- <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>



AI Ethics and Regulation

- Ethics guidelines
<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- EU Regulatory framework proposal on artificial intelligence
<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

US has 'moral imperative' to develop AI weapons, says panel

Draft Congress report claims AI will make fewer mistakes than humans and lead to reduced casualties



Syllabus

- › Problem Solving:
 - Searching
 - › Uninformed, Informed
 - › Local
 - › Adversarial
 - Constraint Satisfaction Problems
- › Reinforcement Learning
 - MDPs, RL, Multi-agent systems
- › Ethics
- › Intelligence from Computation
 - Search
 - Planning
 - CSP
- › Intelligence from Data
 - Bayes nets
 - ML

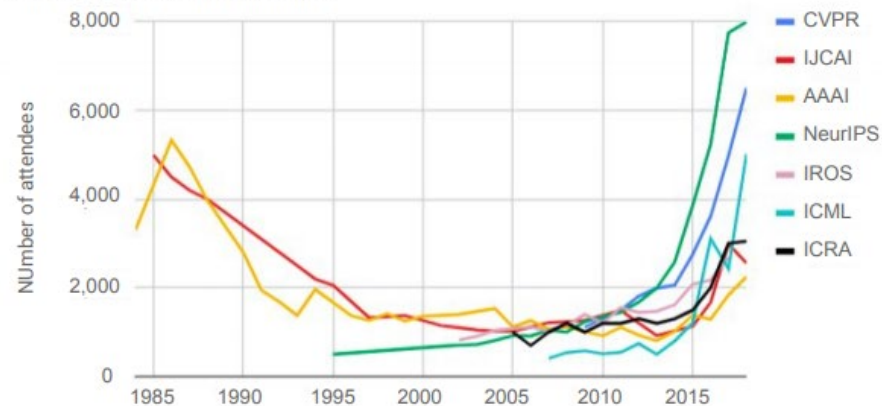
AI trends

› AI index annual report

- <http://cdn.aiindex.org/2018/AI%20Index%202018%20Annual%20Report.pdf>

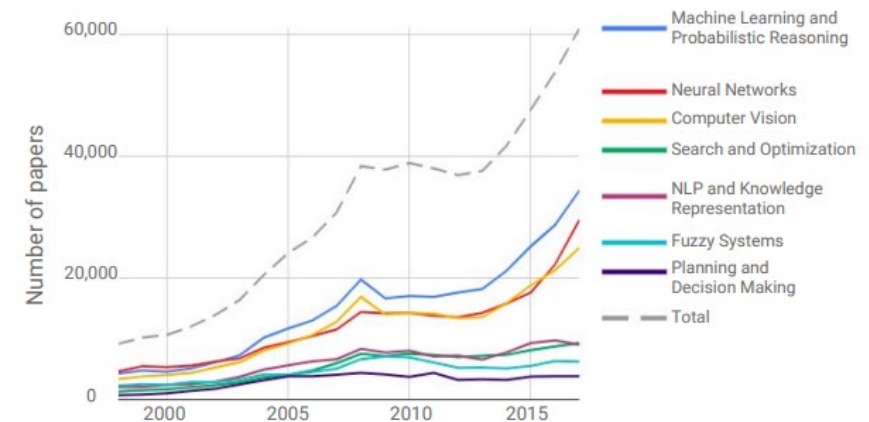
Attendance at large conferences (1984–2018)

Source: Conference provided data

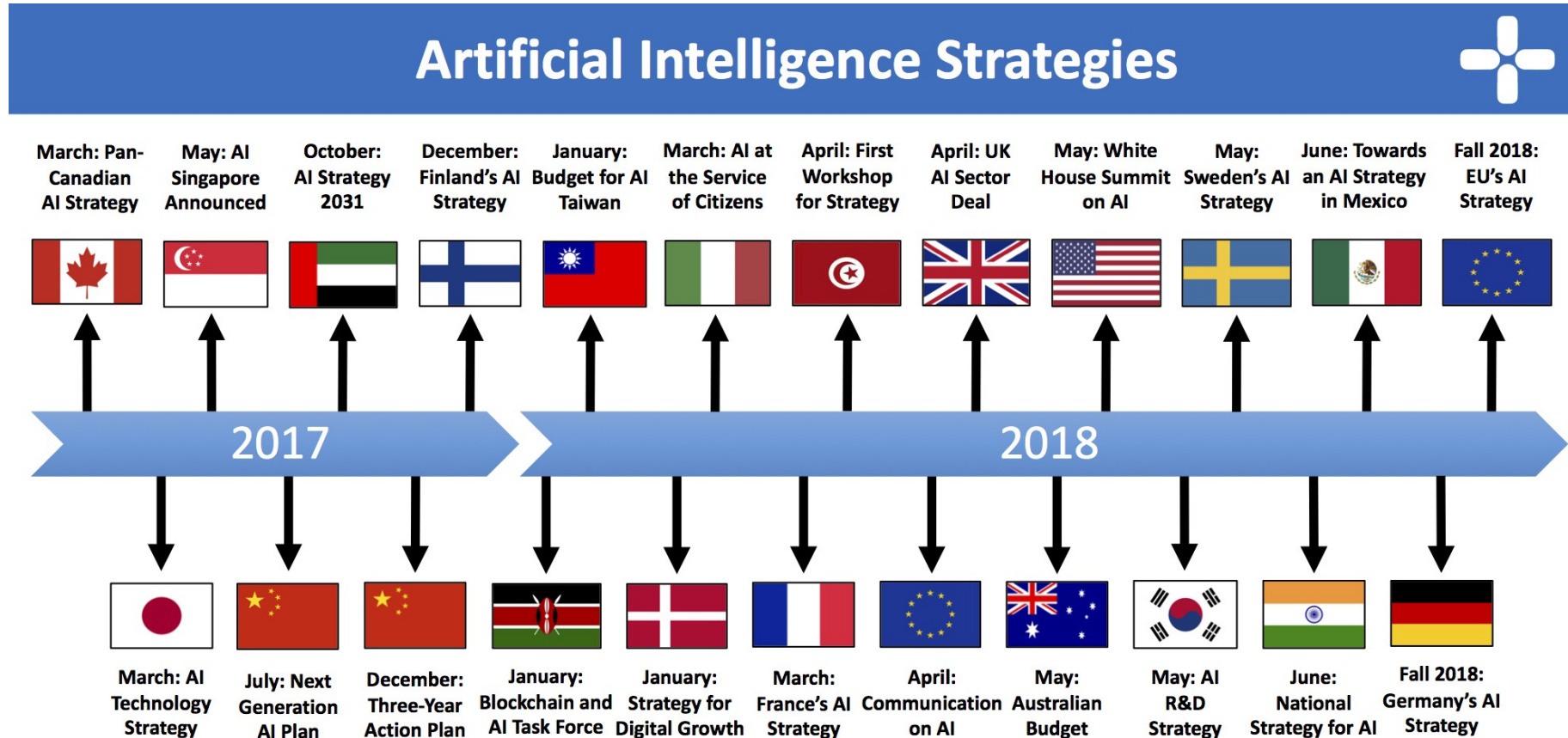


Number of AI papers on Scopus by subcategory (1998–2017)

Source: Elsevier



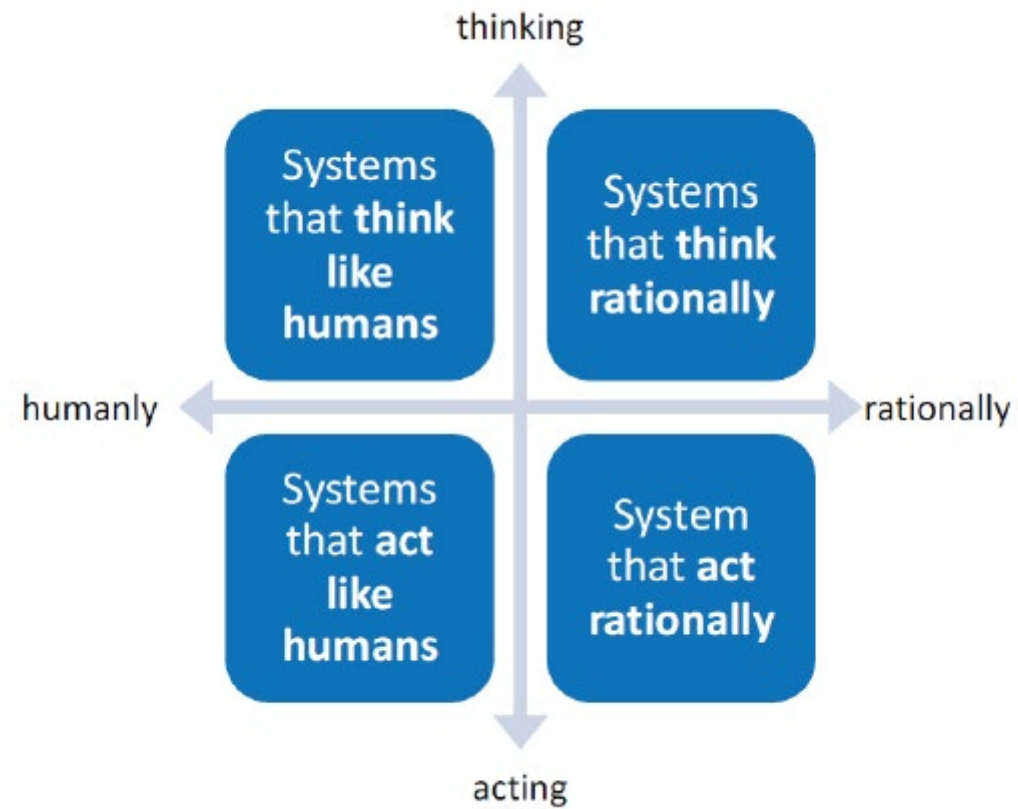
AI trends – national initiatives



AI predictions and challenges - 2022

- › <https://research.aimultiple.com/future-of-ai/>
- › Some examples:
 - › Ethics and legislation
 - › Explainability and accountability - XAI
 - › Balance between surveillance and security
 - › Training on small data sets
 - › Transfer learning
 - › Generalization of algorithms to multiple domains
 - › Model-free not enough – increase in self-supervised and model-based RL
 - › Auto ML

What is AI?



Acting Humanely approach (1 of 2)

- › Turing test approach
 - A computer passes the test if a human interrogator cannot tell whether the responses come from a person or computer
- › However, more important to study underlying principles of intelligence than exactly duplicate the exemplar (ie a human)
- › Progress in following research areas:
 - Natural language processing
 - Knowledge representation
 - Automated reasoning
 - Machine learning
 - Optional: computer vision, robotics
- › Microsoft twitter bot
- › GPT-3 **Generative Pre-trained Transformer 3** - language model that uses deep learning to produce human-like text
- › DALL-E that creates images from text captions

Acting Humanely approach (2 of 2)

› The Winograd Schema Challenge

- The city councilmen refused the demonstrators a permit because they feared violence.
 - The city councilmen refused the demonstrators a permit because they advocated violence.
- › Does the pronoun "they" refers to the city councilmen or the demonstrators? switching between the two instances of the schema changes the answer

Thinking Humanely approach

- › Cognitive science/cognitive neuroscience – understand how humans think
- › Issues:
 - Requires scientific theories of internal activities of the brain
 - Also, humans often don't think (or act) in ways we consider intelligent



A prior track record of exceptional achievement in engineering **is** required, but **no** prior experience working on the brain or human physiology is required. We will teach you what's known about the brain, which is not much tbh.

6:19 PM · Sep 19, 2020



Thinking Rationally approach

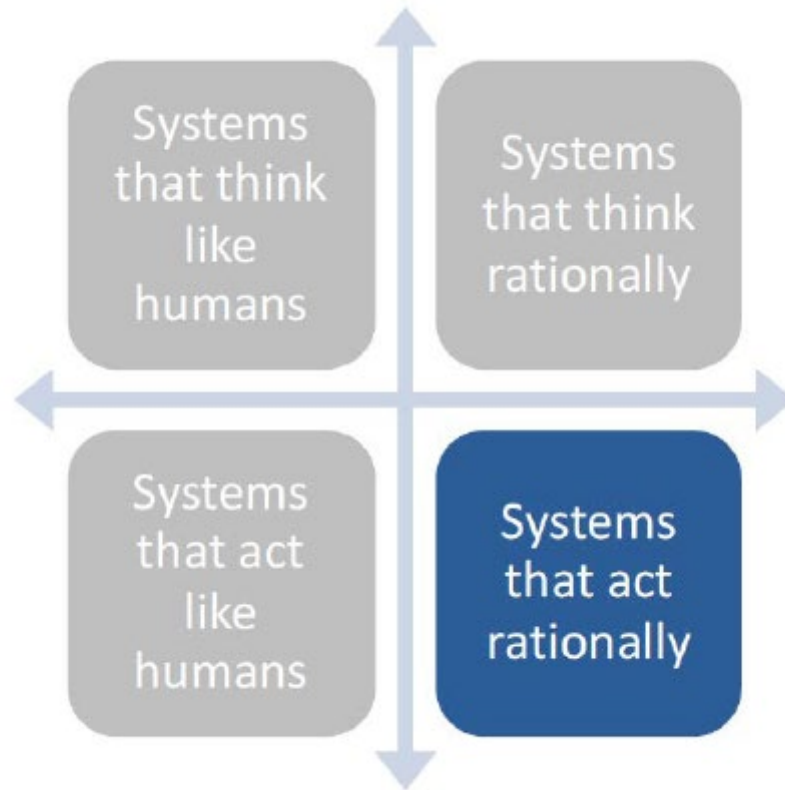
- Logic – patterns for argument structures
- Issues:
 - › How to represent all knowledge using logical notation, especially uncertain knowledge
 - › Solving a problem in theory vs in practice

Acting Rationally approach

- › Maximizing your expected utility/outcome
- › Rational agent
 - Focus of AI today – general principles of rational agents and components for constructing them

“AI is the field that studies the synthesis and analysis of computational agents that act intelligently” (Poole & Mackworth, Artificial Intelligence: Foundations of Intelligent Agents)

What is AI?



Rational Agents

- › An agent is an entity that perceives and acts in an environment
- › An agent acts intelligently if:
 - its actions are appropriate for its goals and circumstances
 - it is flexible to changing environments and goals
 - it learns from experience
 - it makes appropriate choices given perceptual and computational limitations (finite memory and limited time)

Rational agents

- › This course is about designing rational agent
- › Abstractly, an agent is a function from percept histories to actions:

$$f: \mathcal{P}^* \rightarrow \mathcal{A}$$

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- › Computational limitations make perfect rationality unachievable
 - design best program for given machine resources
- › Goal of AI: understand the principles that make intelligent behaviour possible

Short history of AI

- › 1940-1950: Early days
 - 1943: McCulloch & Pitts: Boolean circuit model of brain
 - 1950: Turing's "Computing Machinery and Intelligence"
- › 1950—70: Excitement: Look, Ma, no hands!
 - 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
 - 1956: Dartmouth meeting: "Artificial Intelligence" adopted
 - 1965: Robinson's complete algorithm for logical reasoning
- › 1966-1973 – AI discovers computational complexity; neural networks research almost disappears
- › 1970—90: Knowledge-based approaches
 - 1969—79: Early development of knowledge-based systems
 - 1980—88: Expert systems industry booms
 - 1988—93: Expert systems industry busts: "AI Winter"
- › 1988 – the return of neural networks
- › 1990— 2012: Statistical approaches + subfield expertise
 - Resurgence of probability, focus on uncertainty
 - General increase in technical depth
 - Agents and learning systems... "AI Spring"?

(When) will there be another AI winter?

Recent history of AI

- › 2011 — IBM's Watson wins at Jeopardy! IBM Watson: The inside story of how the Jeopardy-winning supercomputer was born, and what it wants to do next (Tech Republic, September 9, 2013)
- › 2012 —Google unveils a “deep learning” systems that recognized images of cats
- › 2015 —Image recognition systems outperformed humans in the ImageNet challenge
- › 2016 — AlphaGo defeats world Go champion Lee Sedol: “In Two Moves, AlphaGo and Lee Sedol Redefined the Future” (Wired, March 16, 2016)
- › 2018 —Self-driving cars hit the road as Google's Waymo launched (a very limited) self-driving taxi service in Phoenix, Arizona

(<https://mindmatters.ai/2020/01/so-is-an-ai-winter-really-coming-this-time/>)

AI state of the art: what can AI do well in 2022?

Solve humanity's greatest problems...

...such as finding Waldo in 4.45 seconds, better than most 5 year olds

https://www.youtube.com/watch?time_continue=10&v=-i7HMPpxB-Y

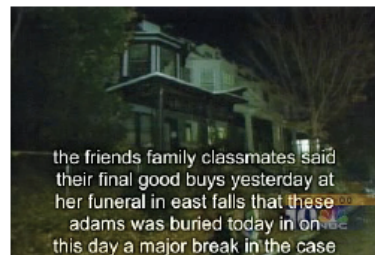
Extensive list per application area:

<https://deepindex.org/>

Natural Language

Speech technologies (e.g. Siri)

- Automatic speech recognition (ASR)
- Text-to-speech synthesis (TTS)
- Dialog systems



Language processing technologies

- Question answering
- Machine translation

"Il est impossible aux journalistes de rentrer dans les régions tibétaines"

Bruno Philip, correspondant du "Monde" en Chine, estime que les journalistes de l'AFP qui ont été expulsés de la province tibétaine du Qinghai "n'étaient pas dans l'illégalité".

Les faits Le dala-lama dénonce l'"enfer" imposé au Tibet depuis sa fuite, en 1959

Vidéo Anniversaire de la rébellion

"It is impossible for journalists to enter Tibetan areas"

Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

Facts The Dalai Lama denounces the "hell" imposed since he fled Tibet in 1959

Video Anniversary of the Tibetan rebellion: China on guard



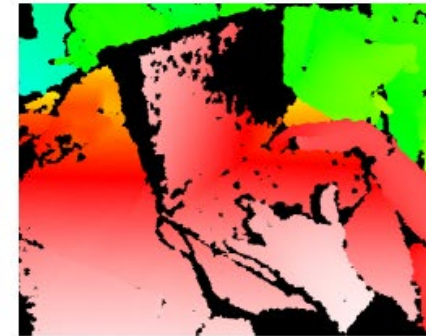
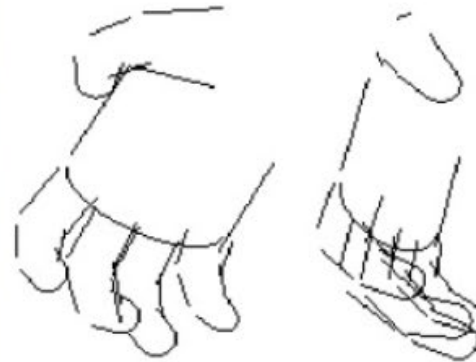
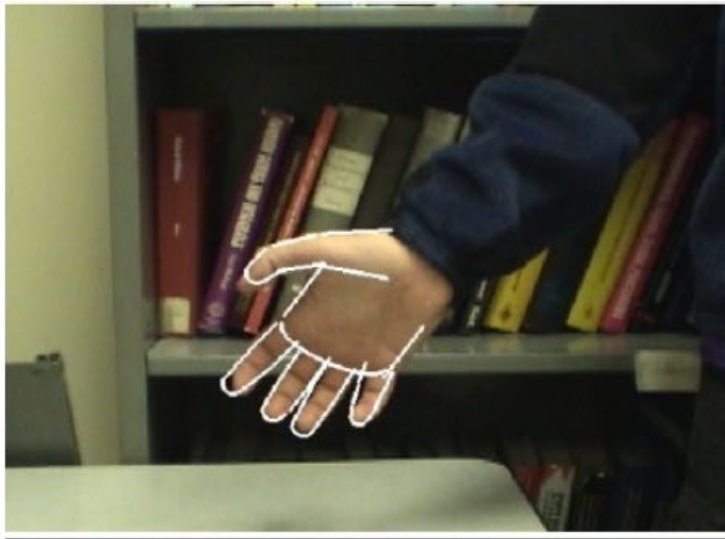
- Web search
- Text classification, spam filtering, etc...

Vision (Perception)

Object and face recognition

Scene segmentation

Image classification



Robotics

Robotics

- Part mech. eng.
- Part AI
- Reality much harder than simulations!

Technologies

- Vehicles
- Rescue
- Soccer!
- Lots of automation...



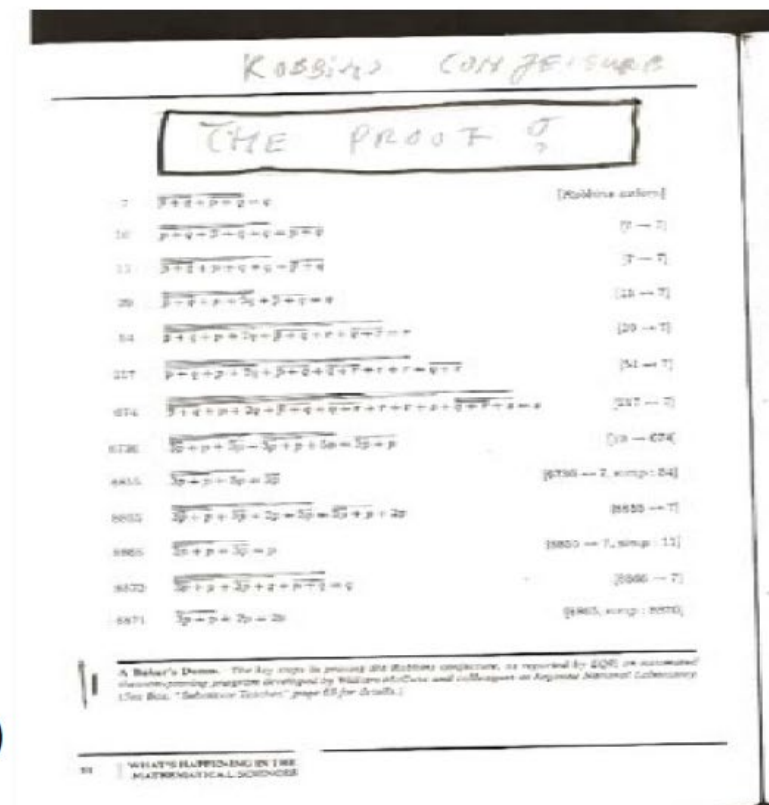
Logic

Logical systems

- Theorem provers
- NASA fault diagnosis
- Question answering

Methods:

- Deduction systems
- Constraint satisfaction
- Satisfiability solvers (huge advances!)



Game playing

Classic Moment: May, '97: Deep Blue vs. Kasparov

- First match won against world champion
- “Intelligent creative” play
- Examining 200 million moves per second, or 50 billion positions in the three minutes allocated for a single move in a chess game
- Humans understood 99.9 of Deep Blue's moves
- Can do about the same now with a PC cluster

1996: Kasparov Beats Deep Blue

- “I could feel --- I could smell --- a new kind of intelligence across the table.”

1997: Deep Blue Beats Kasparov

- “Deep Blue hasn't proven anything.”

Open question:

- How does human cognition deal with the search space explosion of chess?
- Or: how can humans compete with computers at all??

Huge game-playing advances recently, e.g. in Go!



AlphaGo documentary

<https://www.youtube.com/watch?v=WXuK6gekU1Y>

Decision making

Applied AI involves many kinds of automation

- Scheduling, e.g. airline routing, military
- Route planning, e.g. Google maps
- Medical diagnosis
- Web search engines
- Spam classifiers
- Automated help desks
- Fraud detection
- Product recommendations
- ... Lots more!

AI state of the art: examples of what is AI not so good at in 2022?

- › IBM Watson delivered ‘unsafe and inaccurate’ cancer recommendations

<https://www.massdevice.com/report-ibm-watson-delivered-unsafe-and-inaccurate-cancer-recommendations/>

- › Writing movie scripts

<https://www.youtube.com/watch?v=LY7x2lhqjmc>

- › Or jokes

<http://aiweirdness.com/post/174691534037/why-did-the-neural-network-cross-the-road>

- › Or songs

<https://techcrunch.com/2018/09/17/mumford-sons-beware-an-ai-can-now-write-indie-music/>

- › Consumer AI? - I've seen the future of consumer AI, and it doesn't have one

https://www.theregister.co.uk/2018/09/05/consumer_ai_ifa_2018_roundup/

- › Chatbots

<https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>

- Reliable non-biased facial recognition – increasing bans on use of facial recognition

➤ <https://www.bbc.com/news/technology-51148501>