UNIVERSITÄT DES SAARLANDES

# Artificial Intelligence

# Introduction

Podcasts are available with these slides:
Login to the CMS, go to
Information > Materials > Lecture Podcasts

**Prof. Dr. habil. Jana Koehler**

Summer 2020

# Recommended Reading

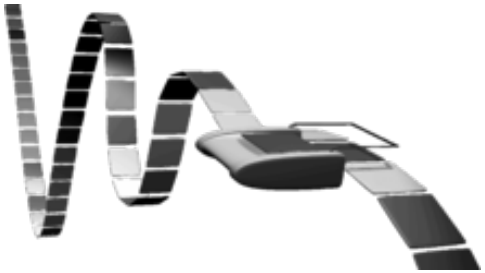
- AIMA Chapter 1: Introduction
  - 1.1 What Is AI?
  - 1.2 The Foundations of Artificial Intelligence
  - 1.3 The History of Artificial Intelligence
  - 1.4 The State of the Art
  - 1.5 Summary, Bibliographical and Historical Notes, Exercise

- For further reading, see slides at the end of this deck

# The Beginnings of Artificial Intelligence

## Alan Turing: Can Machines Think?

A. M. Turing (1950)
Computing Machinery and Intelligence
Mind 49: 433-460



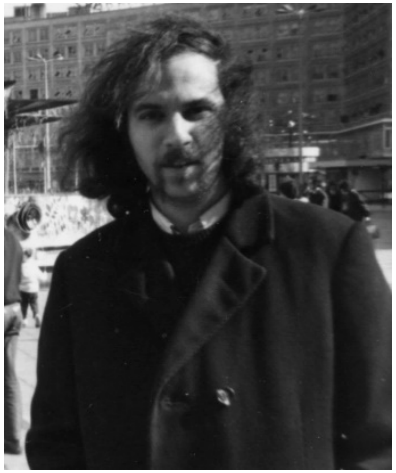
# 1956 Dartmouth College Summer Workshop

- «Complex Computer Applications»
- «Artificial Intelligence»

**John McCarthy**
4.9. 1927 - 23. 10. 2011

"Stanford School"
thinking rationally

***„Symbolic“***



**Marvin L. Minsky**
9.8.1927 - 24.1.2016

"MIT School"
thinking humanly

***„Subsymbolic“***

# Symbolic Representations

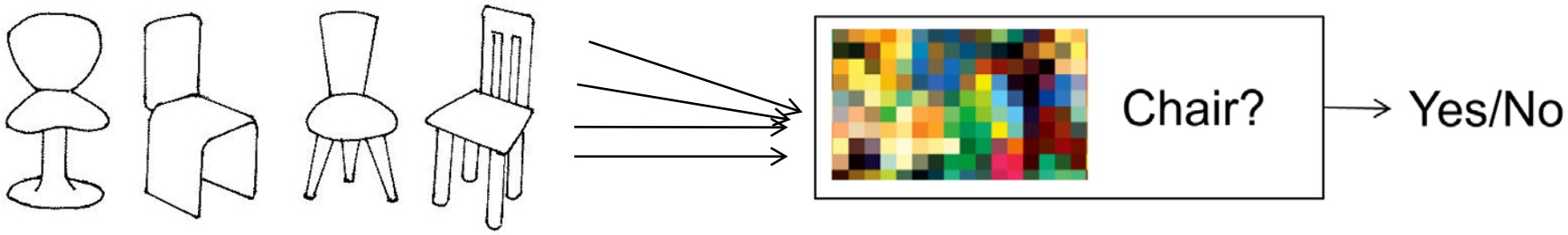
A chair

- is a portable object
- has a horizontal surface at a suitable height for sitting
- has a vertical surface suitably positioned for leaning against

Find a definition

- using symbols, concepts, rules, some formalism
- apply automated reasoning procedures

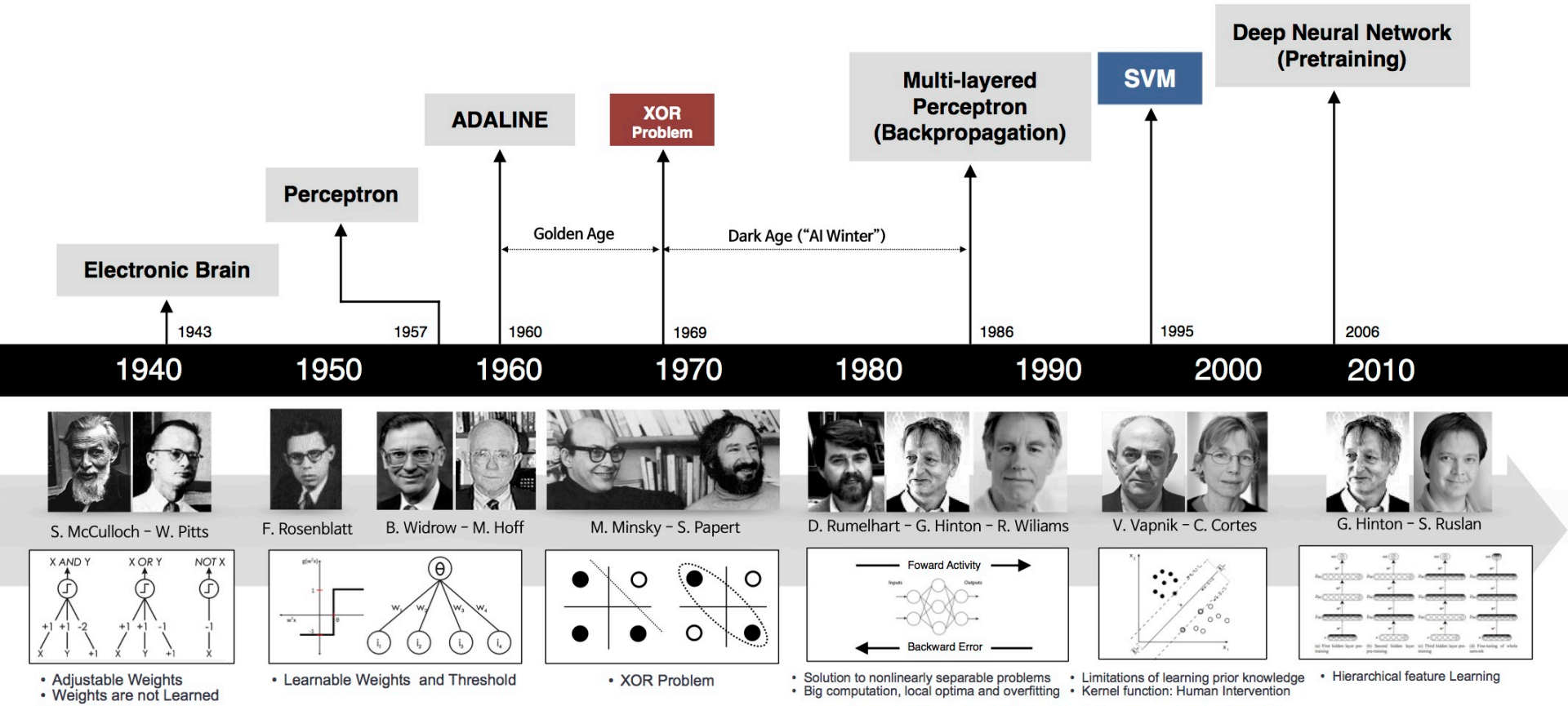
# Subsymbolic Representations




- Use many different (arbitrary) features to describe the object
  - low-level inputs bits, encoding of neurons
  - show examples to the system and let it learn a generalization pattern

- If the pattern is correct, the system has «learned» the concept without using an explicit definition

# Dramatic Progress over the Last Decades – Neural Networks




*https://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html*

# Dramatic Progress over the Last Decades – Search

Systematic Search
Alpha Beta Pruning
$10^5/10^3$

Temporal Difference Learning
$10^{144}$

Heuristic Search and Knowledge
$10^{123}$

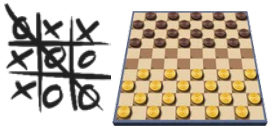
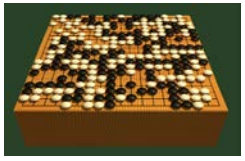
Conflict Learning
Monte-Carlo Tree Search

Reinforcement Learning
Deep Learning
$10^{360}$

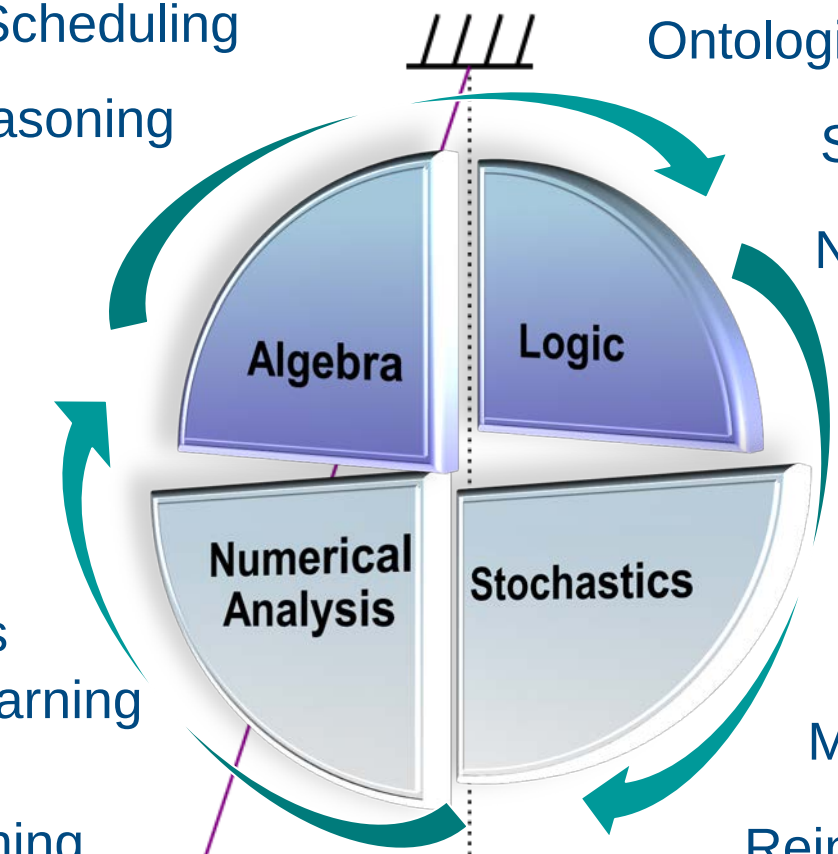
Discounted Regret Minimization
Imperfect Information Games

Population-based/ Multi-Agent RL
Long-/Shortterm Planning

Policy Distillation

| 1940s-1970s | 1992 | 1996/97 | 2012 | 2016 | 2017 | 2019 |
|---|---|---|---|---|---|---|
|  | TD-Gammon  IBM | Deep Blue  IBM | Watson  IBM Watson | AlphaGo  > 20 People Google | Libratus  Carnegie Mellon University Computer Science Department | AlphaStar  > 40 People DeepMind |
| TicTacToe/Checkers  | Backgammon  | Chess  | Jeopardy  | Go  | Poker  | Starcraft II  |

# The PENDULUM of AI Methods




Planning/Scheduling

Constraint Reasoning

Mathematical Optimization

Computational Science

Neural Networks & Deep Learning

Machine Learning

Ontologies/Knowledge Graphs

Satisfiability

Nonmonotonic Reasoning

Reasoning about Action

Statistics

Bayesian Networks

Markov Models

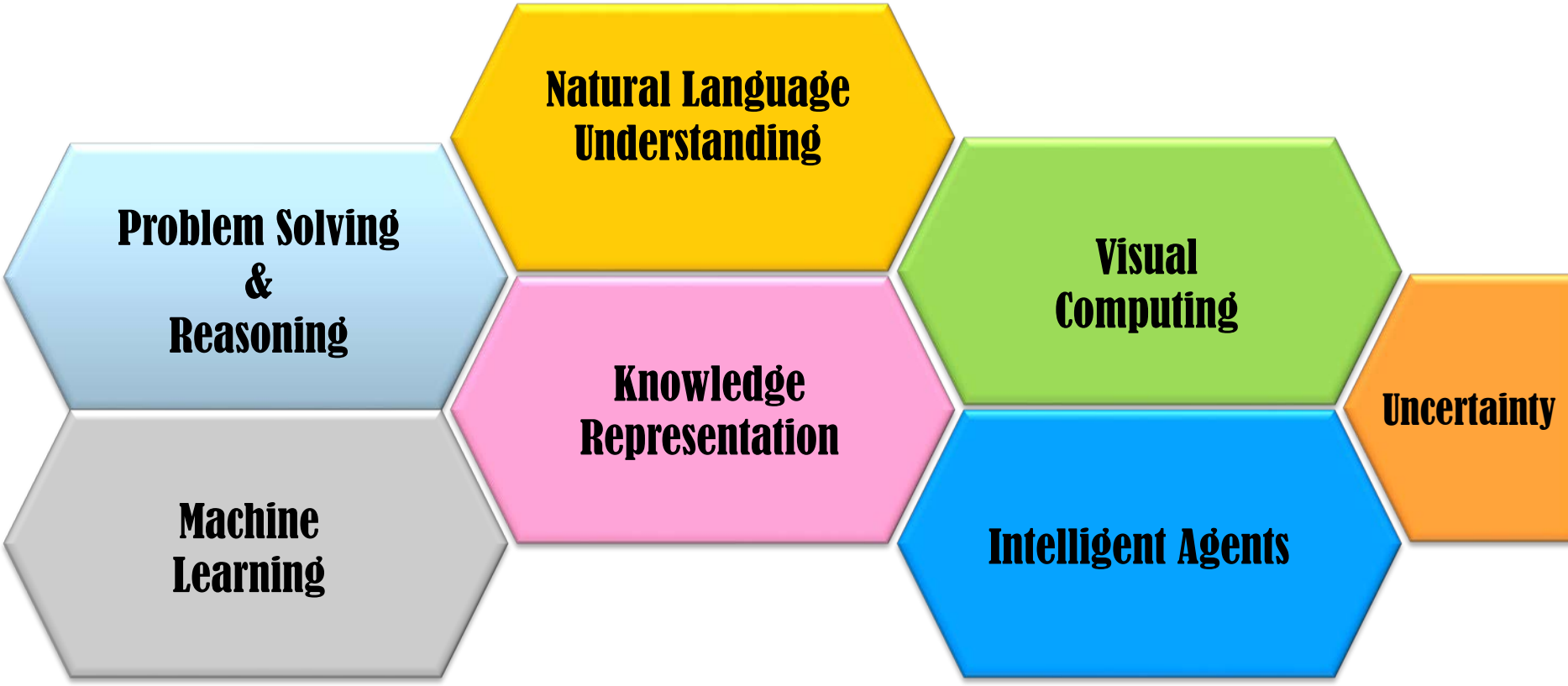
Reinforcement Learning

# AI – An Interdisciplinary Research Area rooted in Computer Science – Dartmouth Proposal

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

# Important Research Areas

- *1956: automatic computers (simulations), use language, neuron nets, theory of the size of a computation, self-improvement, abstractions, randomness and creativity*



# Artificial Intelligence vs. Cognitive Science

***Build Intelligent Software Systems***

- algorithms to achieve intelligent behavior
- problems that only humans can solve are solved by computers
- do not mimic/replicate human intelligence

➢ part of computer science, links to mathematics, economics

***Understand Human Intelligence***

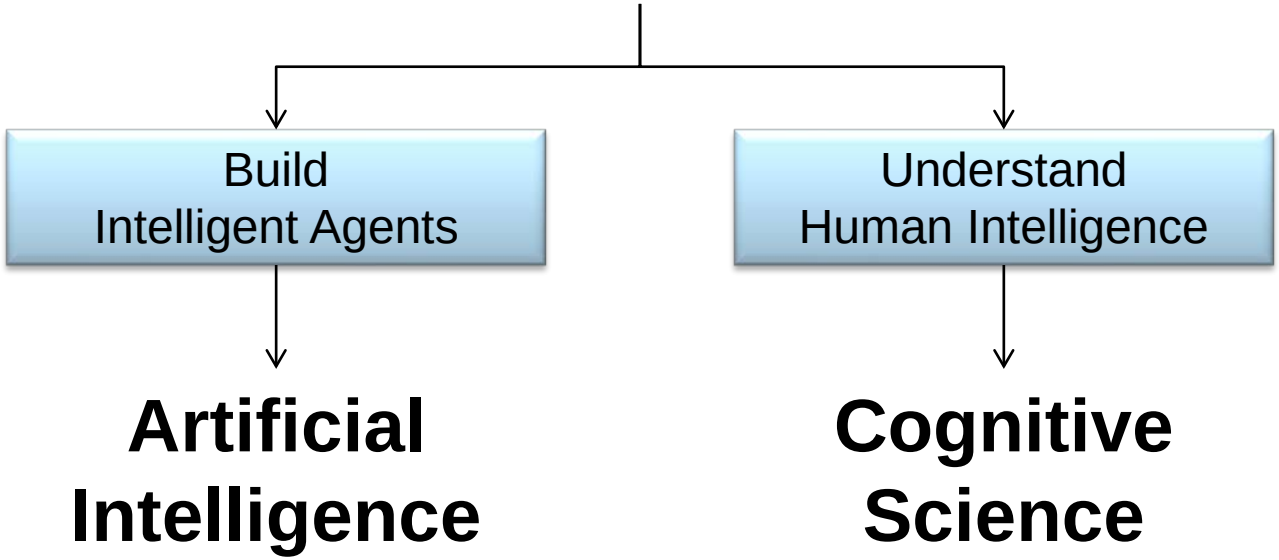
- what constitutes (human) intelligence?
- how do people solve problems?
- build models of human intelligence/the brain

➢ part of neuro & brain sciences, links to psychology

# My Working Definition

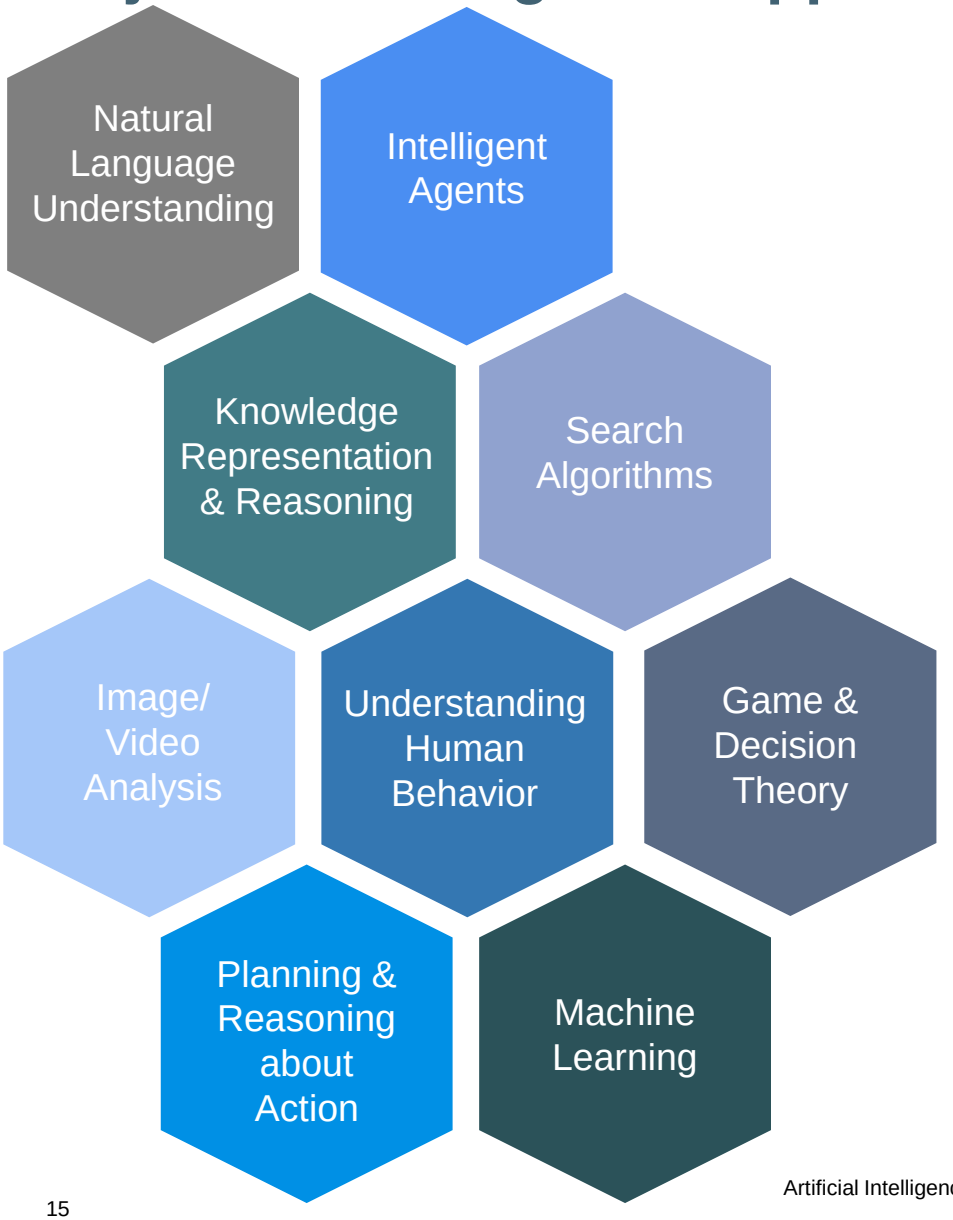
**Intelligence**

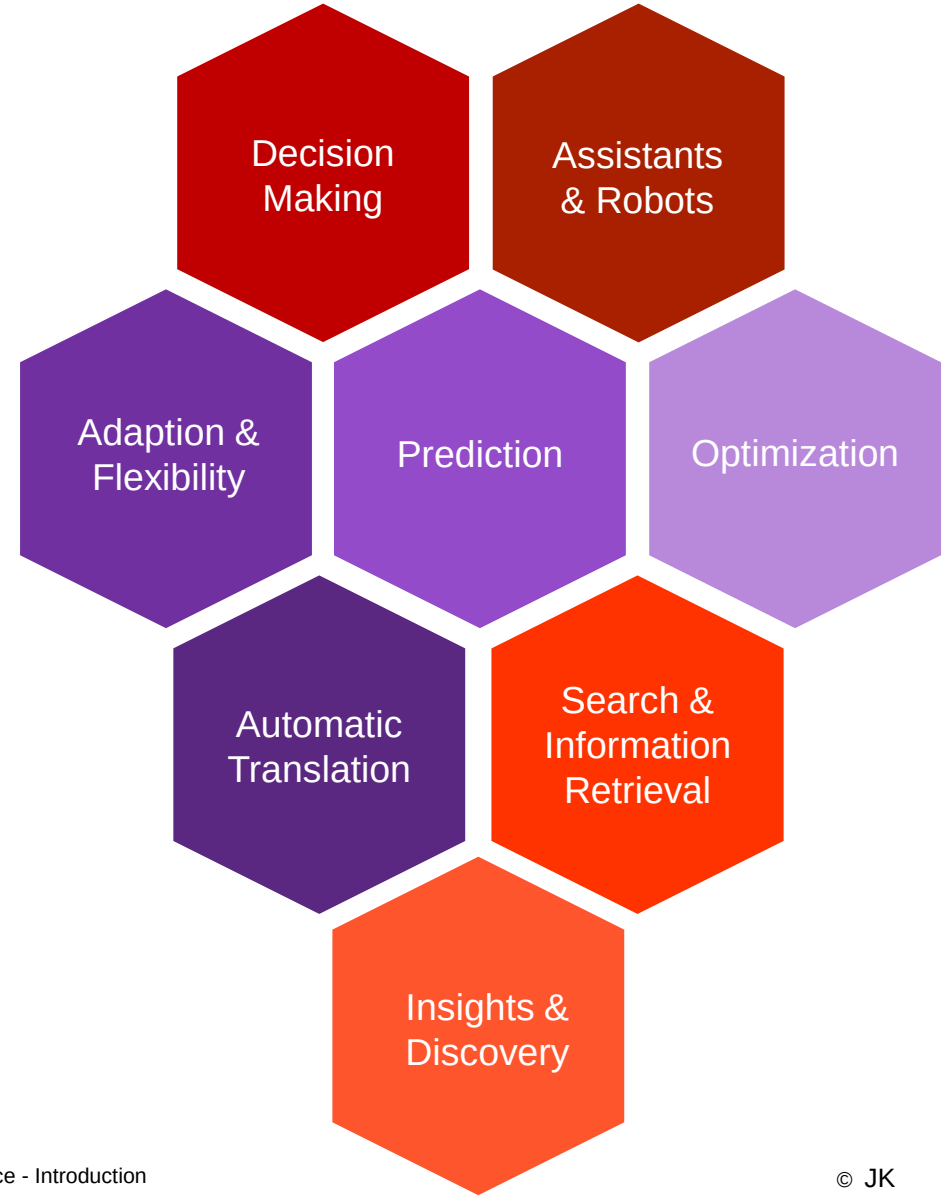
… is the ability to define and set goals and to develop behavior to achieve these goals …



# AI is about

# Major Technologies & Application Areas



# What is Artificial Intelligence?

- The attempt to make computers more "intelligent"?
- The attempt to better understand human intelligence?
- Four approaches:
  - is it about thought thinking. . .
  - or acting?
  - oriented towards a human model (with all its defects). . .
  - or normative (how should a rational being think/act)?

# Selected Definitions

| Thinking Humanly | Thinking Rationally |
| --- | --- |
| "The exciting new effort to make computers think . . . machines with minds, in the full and literal sense." (Haugeland, 1985)<br><br>"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978) | "The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985)<br><br>"The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992) |
| **Acting Humanly** | **Acting Rationally** |
| "The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)<br><br>"The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991) | "Computational Intelligence is the study of the design of intelligent agents." (Poole et al., 1998)<br><br>"AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998) |

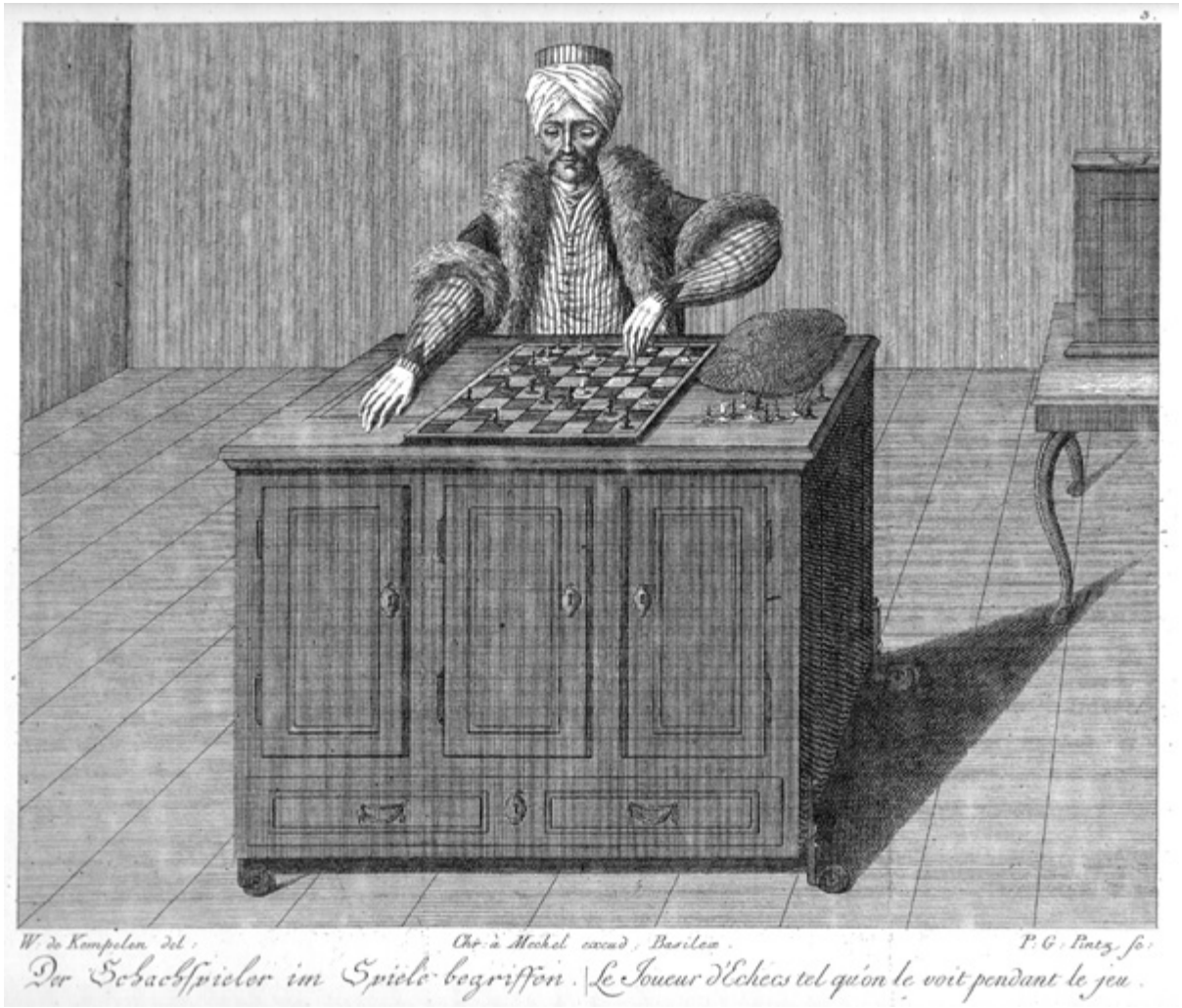
# Thinking Humanly

- What cognitive capabilities are necessary to produce intelligent performance?
- Not important: Being able to solve problems correctly
- Important: Being able to solve problems like a human would

- **Cognitive science and cognitive psychology**
  - **Introspection (observing ourselves)**
  - **Psychological experiments**
  - **Brain research**

- . . . will not be discussed in this course

# Acting Humanly: The Mechanical Turk 1770 - 1854




Wolfgang von Kempelen

# Thinking Rationally

*"Socrates is a man,
all men are mortal,
Socrates is mortal."*

- What are the laws of thought?
- How should we think?

- Represent problems using a formal notation
- Use computational laws to derive conclusions

- Early greek philosophers, e.g. Aristoteles
  - Drawing correct conclusions using logical rules
  - Logical formalisms play a major role in AI although commonsense and everyday knowledge cannot be captured in logic alone

# Acting Rationally

- Rational agents (or rational actors)
- A rational agent acts so as to achieve its given goals, under the assumption that its impressions of the world and its convictions are correct
- Rational thinking is a prerequisite for rational acting, although it is not a necessary condition
  - what to do, for example, when we must make a decision faced with insufficient information?

➢ We look at the concept of a rational agent in more detail later …

# The Turing Test

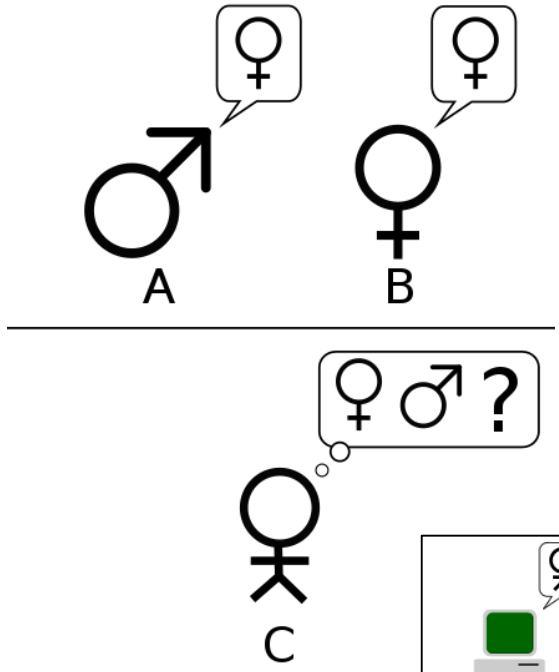
*“The new form of the problem can be described in terms of a game which we call the 'imitation game.´ It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart front the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman.*

*It is A's object in the game to try and cause C to make the wrong identification. The object of the game for (B) is to help the interrogator.*

*… We now ask the question, "What will happen when a machine takes the part of A in this game?" …These questions replace our original, "Can machines think?"*

*The new problem has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man.”*

A man trying to convince the interrogator that he is a woman, is replaced by a computer trying to convice the interrogator that it is a human (or woman?)



*Ex Machina Trailer*

https://www.youtube.com/watch?v=gyKqHOgMi4g

# A Bit of AI History

# Major Phases

- 1943 – 1955
  - neural networks, basic learning mechanisms, genetic algorithms

- 1956 Darthmouth Conference
  - "Artificial Intelligence" as a research programme
  - AI as part of computer science
    - duplicating human skills such as creativity, self-improvement, use of language
    - building machines that will function autonomously in complex and changing environments

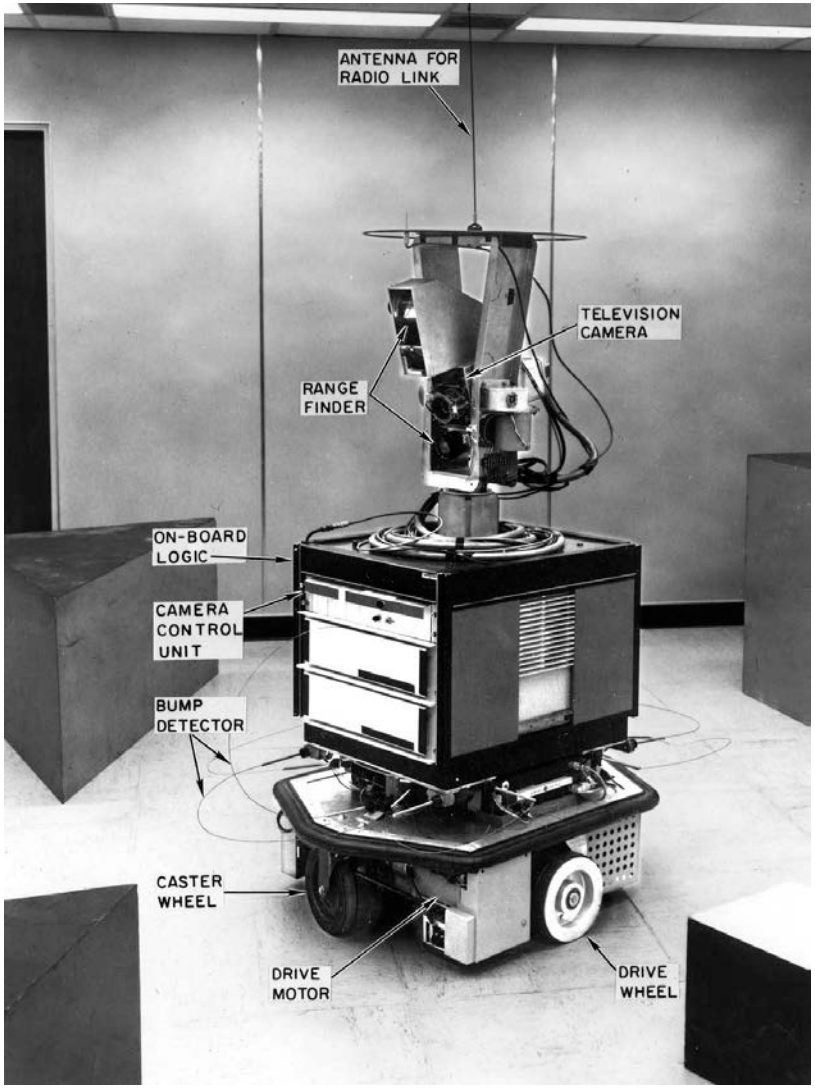
# Early Successes

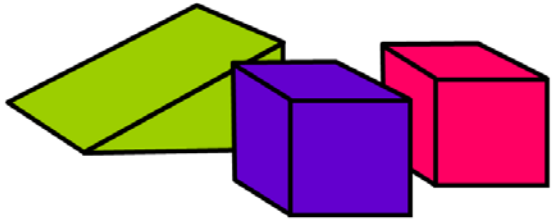
- 1952 – 1969
  - logical theorem proving, game playing programms,formal models of the world, various logic-based inference mechanisms
  - **Physical symbol system hypothesis**

"a physical symbol system has the necessary and sufficient means for general intelligent action"

  - **Microworlds**: small (toy) domains on which these systems operated

# Shakey the Robot - Stanford Research Institute 1966 - 1972



"blocksworld"

# Early Enthusiasm

*"It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create.*

*Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which the human mind has been applied."*

1957

Herbert Simon
Nobel prize 1978

"for his pioneering research into the decision-making process within economic organizations"

# A Dose of Reality 1966 – 1973

- Pure syntatic manipulations led to infamous failures in language translation
  - "the spirit is willing, but the flesh is weak" = the vodka is good, but the meat is rotten
- Intractability of many problems
  - exhaustive enumeration of variants does not scale in exponential state spaces – ***combinatorial explosion***
- Limitations of many mechanisms only slowly understood
  - perceptrons can learn everything they can represent – but they can represent very little

# LISP (1958 John McCarthy)

```
;; Add 2 and 3 and 4:
(+ 2 3 4)

;; Set the variable p to the value
3.1415:
(setf p 3.1415)

;; Define a function that squares
its argument:
(defun square (x)
  (* x x))

;; Square the number 3:
(square 3)
```

# PROLOG (1972 Alain Colmerauer)

```
% Prologue text with facts:
father(adam, tobias).
father(tobias, frank).
father(tobias, ulrike).

% Prologue text with rule:
grandfather(X, Y) :-
    father(X, Z),
    father(Z, Y).

grandfather(X, Y) :-
    father(X, Z),
    mother(Z, Y).


?- grandfather(adam, ulrike).
yes.
?- grandfather(X, frank).
X=adam
```

# Knowledge-Based Systems 1969 – 1985 – the Key to Power?

- Adding more domain knowledge to an AI system
  - learn from human experts
- Explicit representations of rules, plans, goals, states, objects, confidence/uncertainty
  - "frames" – concept logics – ontologies
- From "weak" to "strong" methods
- AI technology companies

# Mycin 1973

**An Artificial Intelligence Program to Advise Physicians Regarding Antimicrobial Therapy***

EDWARD H. SHORTLIFFE†, STANTON G. AXLINE, BRUCE G. BUCHANAN, THOMAS C. MERIGAN, AND STANLEY N. COHEN

*Stanford University, Stanford, California 94305*

SAMPLE RULE

*LISP*

```
PREMISE:
        ($AND   (SAME (VAL CNTXT GRAM) GRAMPOS)
                (SAME (VAL CNTXT MORPH) COCCUS)
                (SAME (VAL CNTXT CONFORM) CHAINS)
                3)
ACTION:
        (CONCLUDE CNTXT IDENT STREPTOCOCCUS TALLY .7)
```

*ENGLISH TRANSLATION*

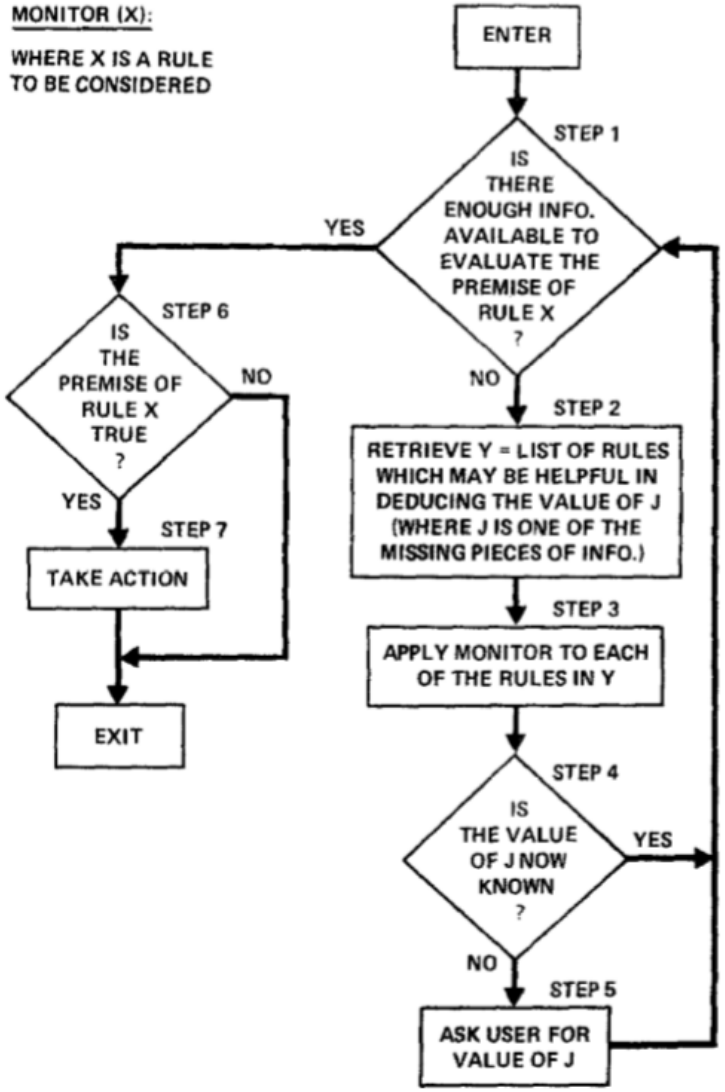
IF:

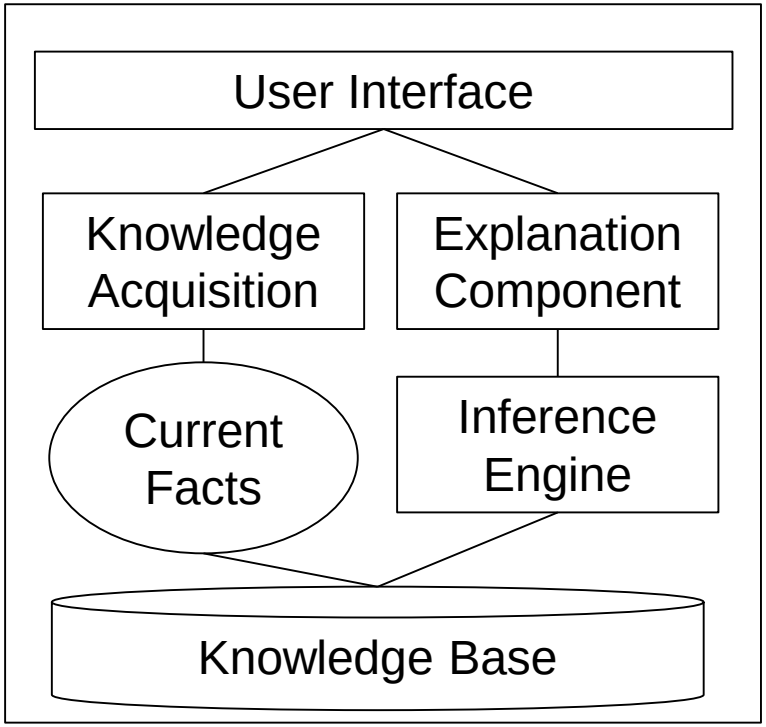
THE GRAMSTAIN OF THE ORGANISM IS GRAMPOS, AND THE MORPHOLOGY OF THE ORGANISM IS COCCUS, AND THE GROWTH CONFORMATION OF THE ORGANISM IS CHAINS

THEN:

CONCLUDE THAT THE IDENTITY OF THE ORGANISM IS STREPTOCOCCUS (MODIFIER: THE CERTAINTY TALLY FOR THE PREMISE TIMES .7)

# Mycin Inference Engine



# Knowledge-Intensive ("strong") versus Knowledge-Sparse Approaches ("weak")

- Knowledge-sparse: axioms, some facts + strong inference mechanisms
  - Google search uses indexing + the page rank
  - Heuristic and stochastic search in Deep Blue Chess Computer
  - Theorem proving in Logic theorist

- Knowledge-intensive: domain ontologies and encoded knowledge bodies
  - Chess opening book
  - Expert systems, e.g. Mycin
  - Google user profiles

# AI Winter (1987–93)

- Subsymbolic (connectionist) approaches remained vague
- Understanding Speech remains a challenge

- LISP machine market collapsed

- Expert systems did not scale and could not handle contradictory knowledge and beliefs

- Japan's 5$^{th}$ generation computer systems project never managed to provide concurrent logic programming on special hardware

# Data-Driven AI (since 2000)

- Significant progress in all AI subfields
  - *better understanding of problems and complexity properties*
  - *increased mathematical sophistication*
- Scalable, robust algorithms and well-understood expressive representation formalisms
- Maturity of AI programming frameworks
- AI services and solutions by major IT vendors (e.g. Amazon AWS, Facebook PyTorch, Google Keras/TensorFlow)

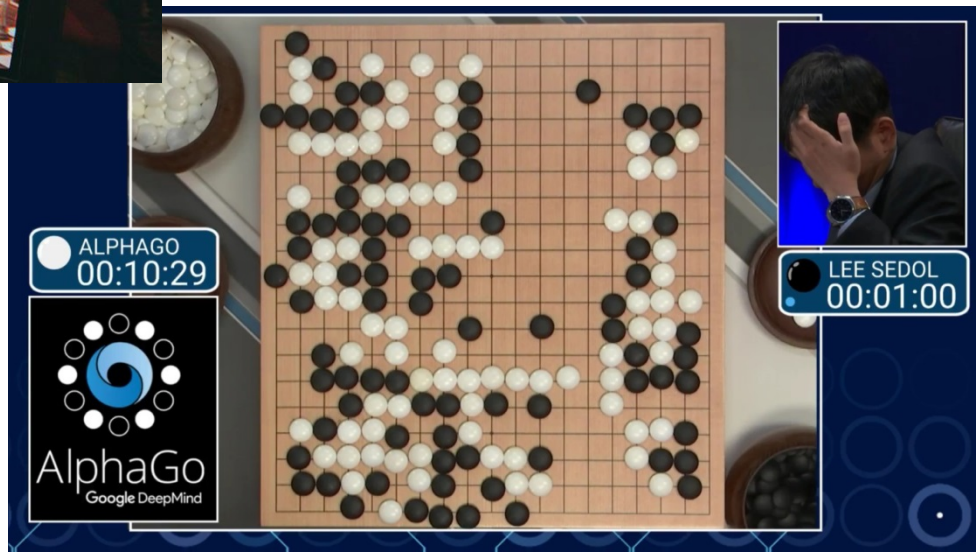
# Digital Transformation paves Ground for AI Revolution

- Internet access everywhere for everybody
- Digital transformation generates a lot of data
- Applications with billions of users
- Cloud services & elastic computing
- Maturity of software development

# Deep Blue and AlphaGo – Important Milestones

1997

2016



# Deep Blue 1996/97

- Winner against world champion Garri Kasparow
- Brute force computing power, written in C
- Massively parallel, RS/6000 SP Thin P2SC-based system with 30 nodes each node containing a 120 MHz P2SC, enhanced with 480 special purpose VLSI chess chips
- Capable of evaluating 200 million positions per second, 11.38 GFLOPS
- 259th on TOP500 supercomputer list in June 1997

# Deep Blue's Opening Book and Evaluation Function

- Evaluation function was initially with many to-be-determined parameters, optimal values for these parameters were determined by the system itself by analyzing thousands of master games

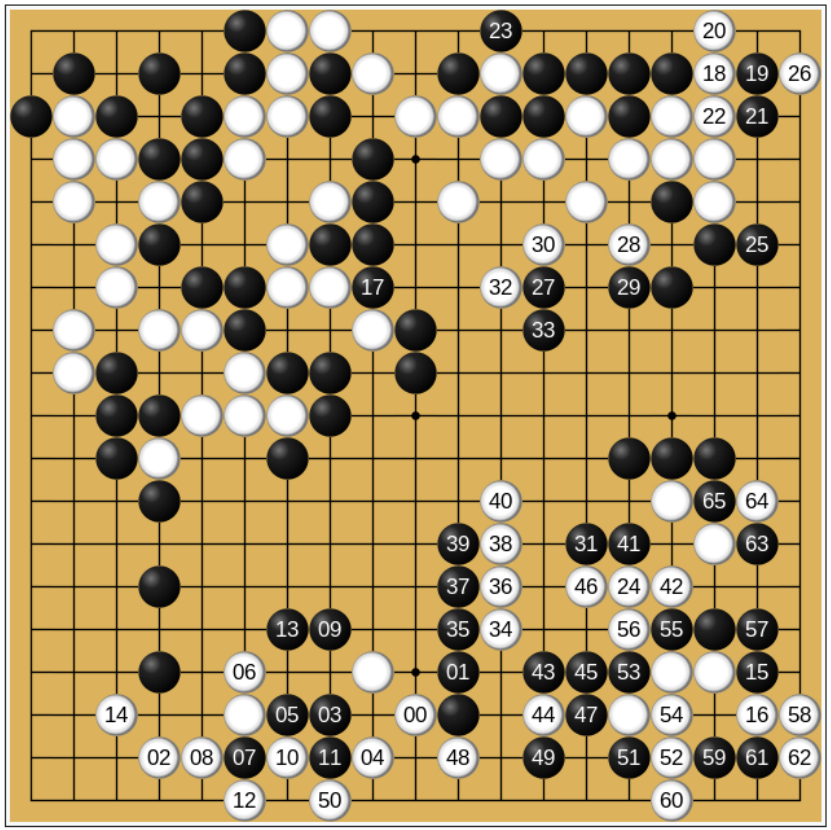
- Opening book contained over 4,000 positions and 700,000 grandmaster games
  - provided by grandmasters Miguel Illescas, John Fedorowicz, Nick de Firmian
  - Chess knowledge finetuned by grandmaster Joel Benjamin

# AlphaGo and AlphaZero 2016/2017

- Game tree complexity $10^{360}$ (chess $10^{123}$)
- No known heuristic methods to evaluate a situation on a Go game board

➢ Learn evaluation from analyzing millions of Go games

➢ Use stochastic search to determine promising next moves




*https://de.wikipedia.org/wiki/AlphaGo*

# IBM Watson Triggers Renewed Interest in AI in 2012


https://www.youtube.com/watch?v=WFR3lOm_xhE

# Hype Hurts: Steering Clear of Dangerous AI Myths

**Published:** 03 July 2017 **ID:** G00324274

**Analyst(s):** Tom Austin | Alexander Linden | Mike Rollings

https://www.gartner.com/smarterwithgartner/steer-clear-of-the-hype-5-ai-myths/

- Myth 1: Buy an AI to Solve Your Problems
- Myth 2: Everyone Needs an AI Strategy or a Chief AI Officer
- Myth 3: Artificial Intelligence Is Real
- Myth 4: AI Technologies Define Their Own Goals
- Myth 5: AI Has Human Characteristics
- Myth 6: AI Understands (or Performs Cognitive Functions)
- Myth 7: AI Can Think and Reason
- Myth 8: AI Learns on Its Own
- Myth 9: It's Easy to Train Applications That Combine DNNs and NLP
- Myth 10: AI-Based Computer Vision Sees Like we Do (Or Better)
- Myth 11: AI Will Transform Your Industry — Jump Now and Lead
- Myth 12: For the Best Results, Standardize on One AI-Rich Platform Now
- Myth 13: Maximize Investment in Leading-Edge AI Technologies
- Myth 14: AI Is an Existential Threat (or It Saves All of Humanity)
- Myth 15: There Will Never Be Another AI Winter

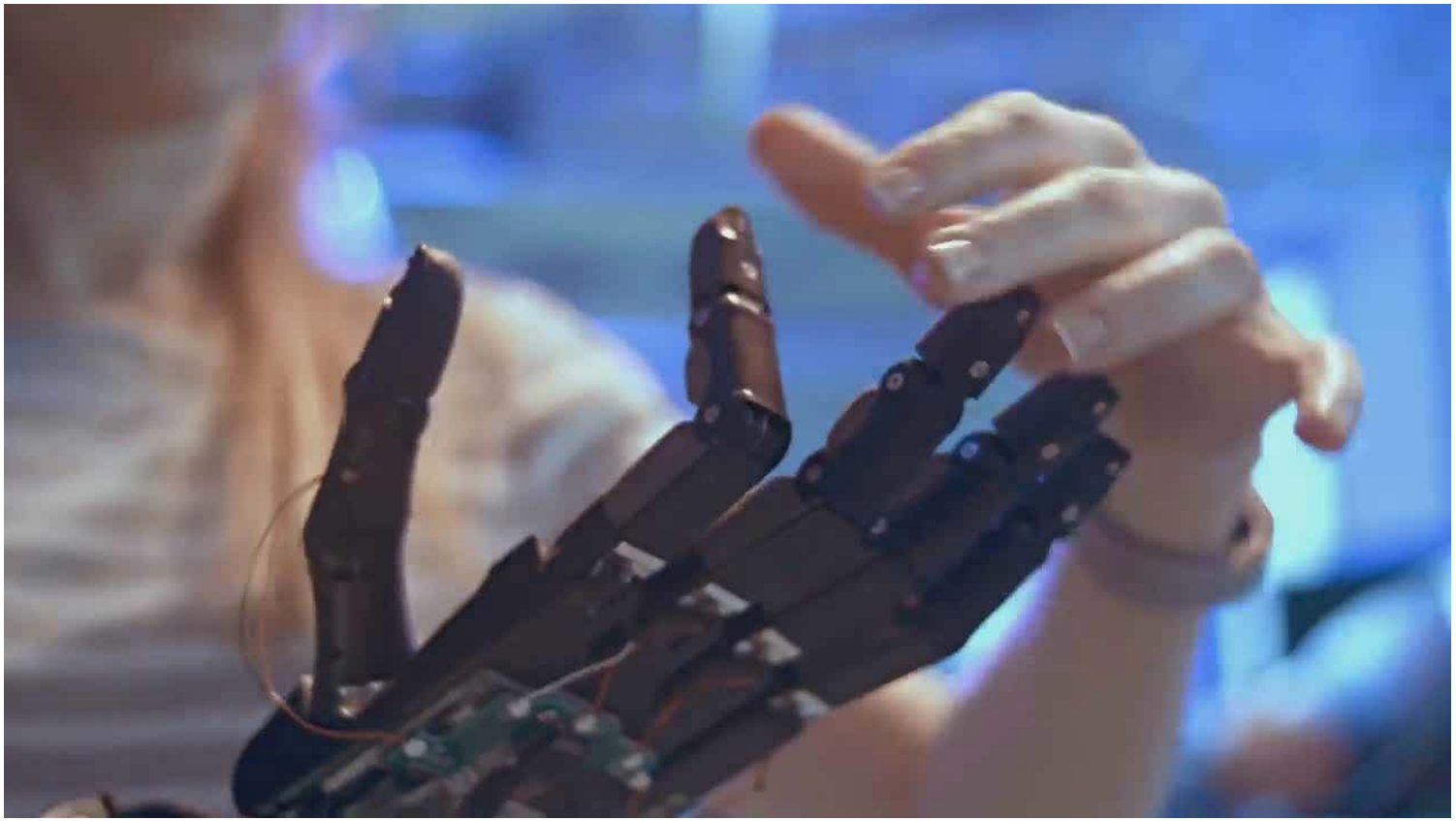
# Selected Current Challenges in AI

# Transfer Learning

https://www.youtube.com/watch?v=iaF43Ze1oeI

https://research.google/teams/brain/robotics/

# Reinforcement Learning



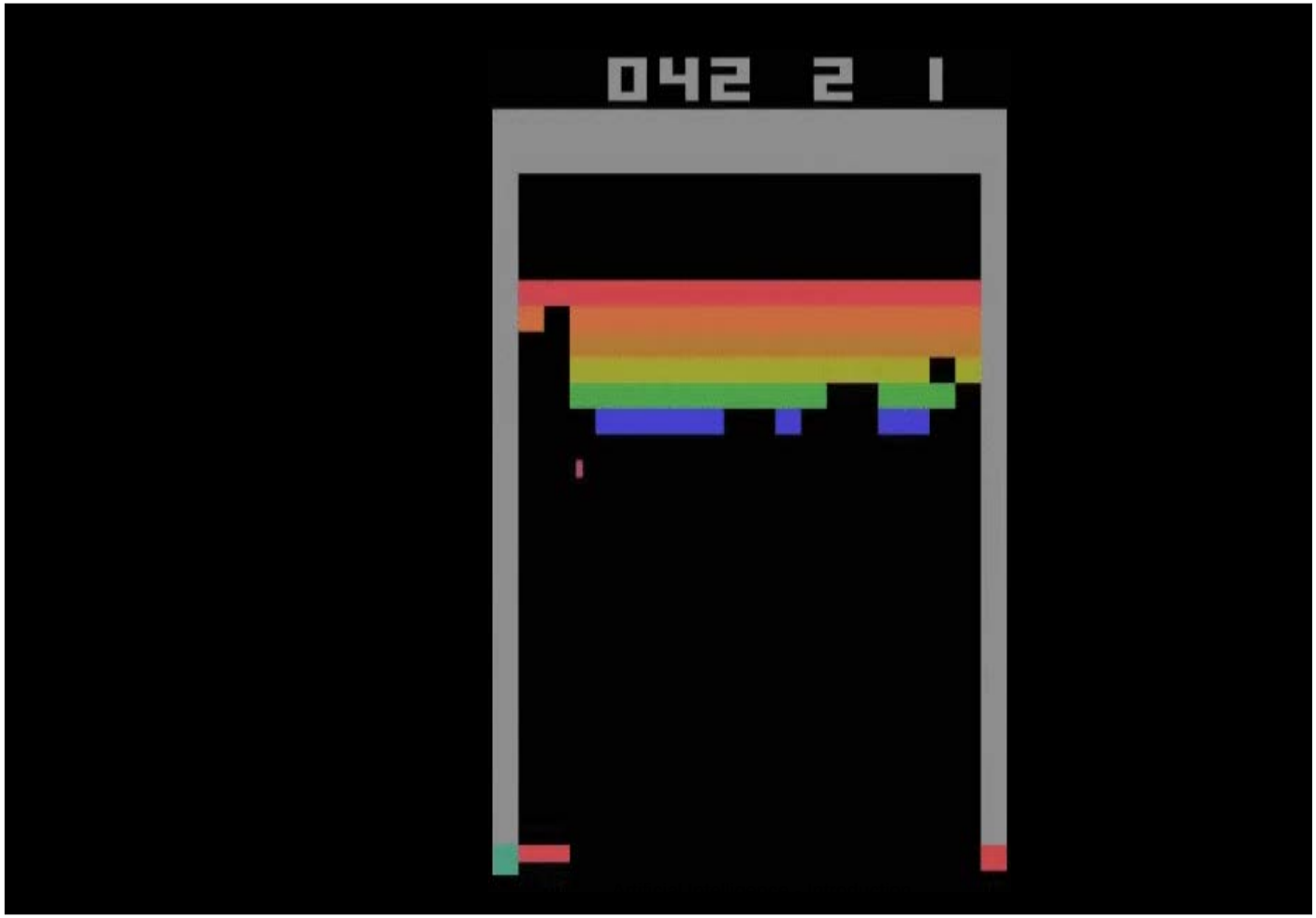
- 20 % success rate for arbitrary cube
- 60 % success rate for 15 moves problem
- 80 % dropped cubes

https://openai.com/blog/solving-rubiks-cube/

# «Magic» in the Learning of Optimal Strategies

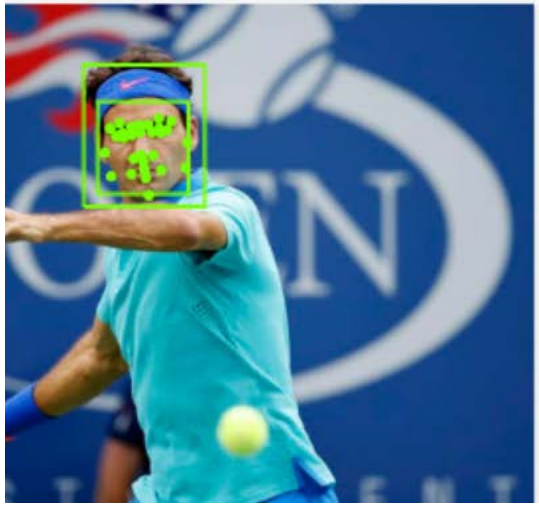
Google DeepMind's Deep Q-learning playing Atari Breakout after 240 minutes of training

https://www.youtube.com/watch?v=V1eYniJ0Rnk



# «Understanding» Images with Deep Learning

| | |
|---|---|
| Sports | 94% |
| Tennis | 93% |
| Tennis Player | 89% |
| Football Player | 88% |
| Ball Game | 86% |
| Racquet Sport | 82% |



Need to resolve:

- Hidden dependency on training data
- Non-calibrated confidence values
- Intransparancy of learned model

Headwear Very Likely

# Ethical Questions

- Is a genderless voice enough?
- Do we need machine-specific voices such that we can immediately recognize if we talk to a human or to a machine?

MATT SIMON SCIENCE 03.11.19 11:45 AM

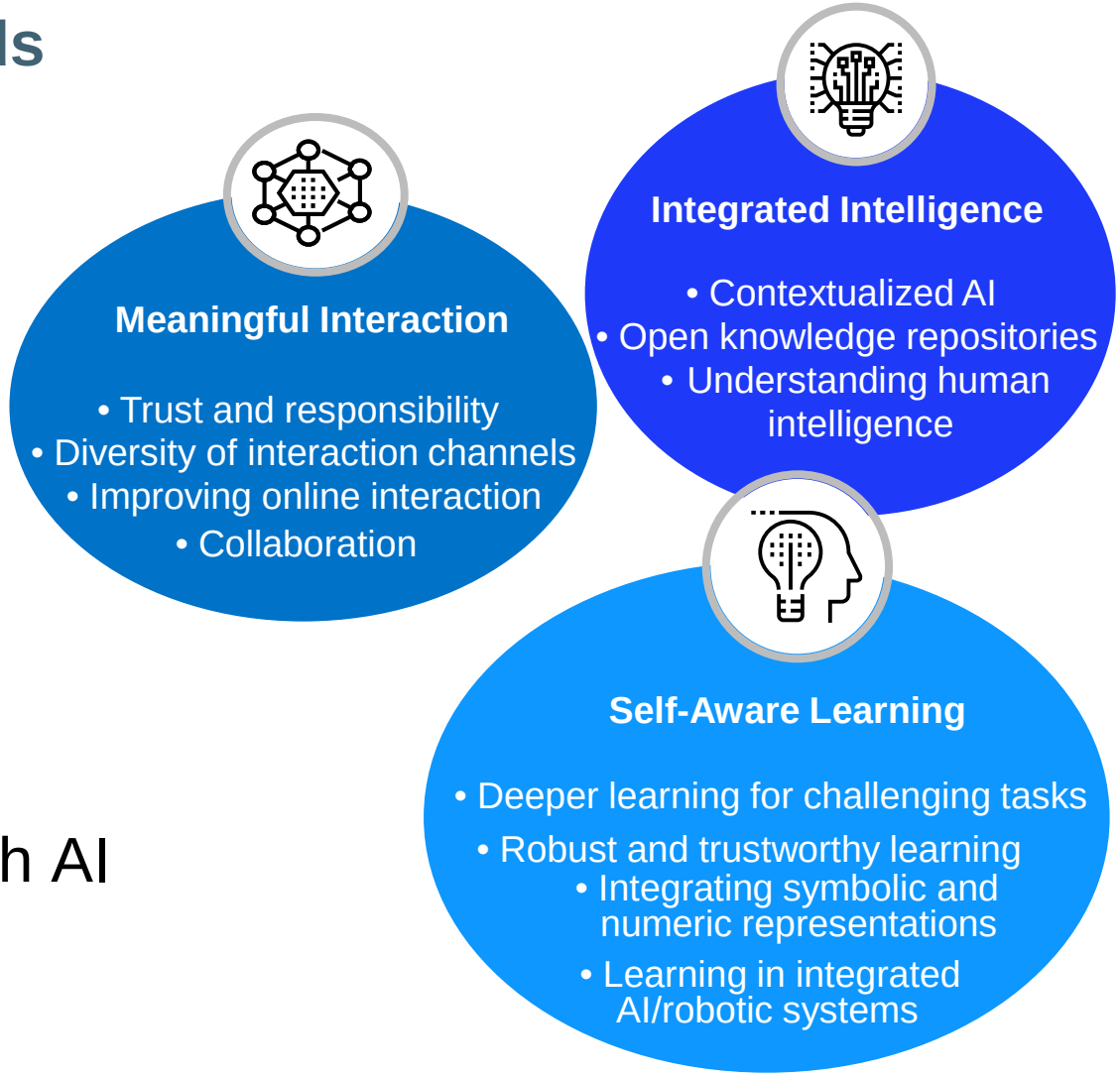
THE GENDERLESS DIGITAL VOICE THE WORLD NEEDS RIGHT NOW

*https://www.wired.com/story/the-genderless-digital-voice-the-world-needs-right-now/*

https://www.genderlessvoice.com/

# Current Research Trends

- Multi-AI Architectures
- Mastering Context
- Embodiment
- Scientific Discovery with AI




A 20-Year Community
Roadmap for Artificial Intelligence Research in the US

Yolanda Gil (USC) and Bart Selman (Cornell), co-chairs

# Summary

- Artificial Intelligence is a vivid research area with many different communities, opinions, and technologies
- The history of AI is marked by big successes, but also hypes and failures
- The increase in computing power, the availability of data, and the progress in all AI subfields has led to remarkable achievements over the last decade
- Each AI technology has its limitations and risk management as well as ethical considerations are very important when bringing AI into applications
- The next challenge for AI will be to master the „Four Big „A“: abstraction, analogy, argumentation, arbitrary common sense

# *Steven Spielberg: Artificial Intelligence, 2001 (Introductory Scene)*


Initial part available via youtube: https://www.youtube.com/watch?v=nqlhYcPx4K0
Watch final part of the introductory scene via the CMS: Information > Materials > Videos

# Working Questions

1. What is Artificial Intelligence about?

2. What do we mean by a symbolic or subsymbolic representation?

3. What are weak (knowledge-sparse) and strong (knowledge-intensive) methods?

4. What characterizes the main historic phases of AI research?

5. Discuss risks and opportunities of AI technology in the light of your own opinion.

# Further Reading

- A. M. Turing (1950): Computing Machinery and Intelligence, Mind 49: 433-460
  - E.g. available here https://www.csee.umbc.edu/courses/471/papers/turing.pdf
- Patrick Hayes, Kenneth Ford: Turing Test Considered Harmful
  - https://www.researchgate.net/profile/Kenneth_Ford/publication/220813820_Turing_Test_Considered_Harmful/links/09e4150d1dc67df32c000000.pdf
- McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955), A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence
  - http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf
- Podcast in German: https://www.zeit.de/digital/2019-07/kuenstliche-intelligenz-algorithmus-dfki-jana-koehler

# Further Reading

- Rodney Brooks: The Seven Deadly Sins of AI Predictions, MIT Technology Review 2017
  - https://www.technologyreview.com/s/609048/the-seven-deadly-sins-of-ai-predictions/

- Daniel Crevier: The Tumultuous History of the Search for Artificial Intelligence, Harper Collins, 1993.

- Nils Nilsson: The Quest for Artificial Intelligence – A History of Ideas and Achievements, Cambridge University Pres, 2010.

- Jerry Kaplan: Artificial Intelligence - What Everyone Needs to Know, Oxford University Press, 2016.