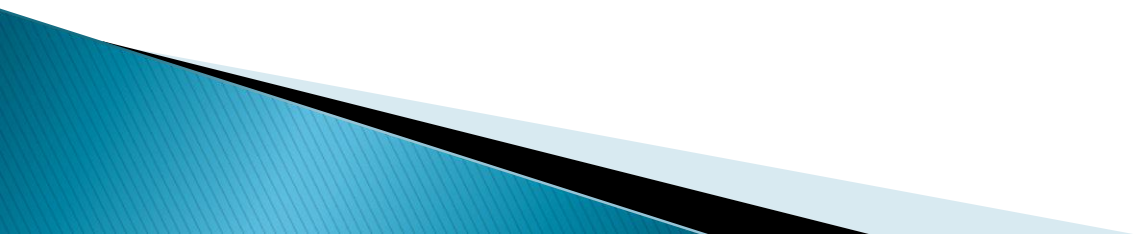
Artificial Intelligence

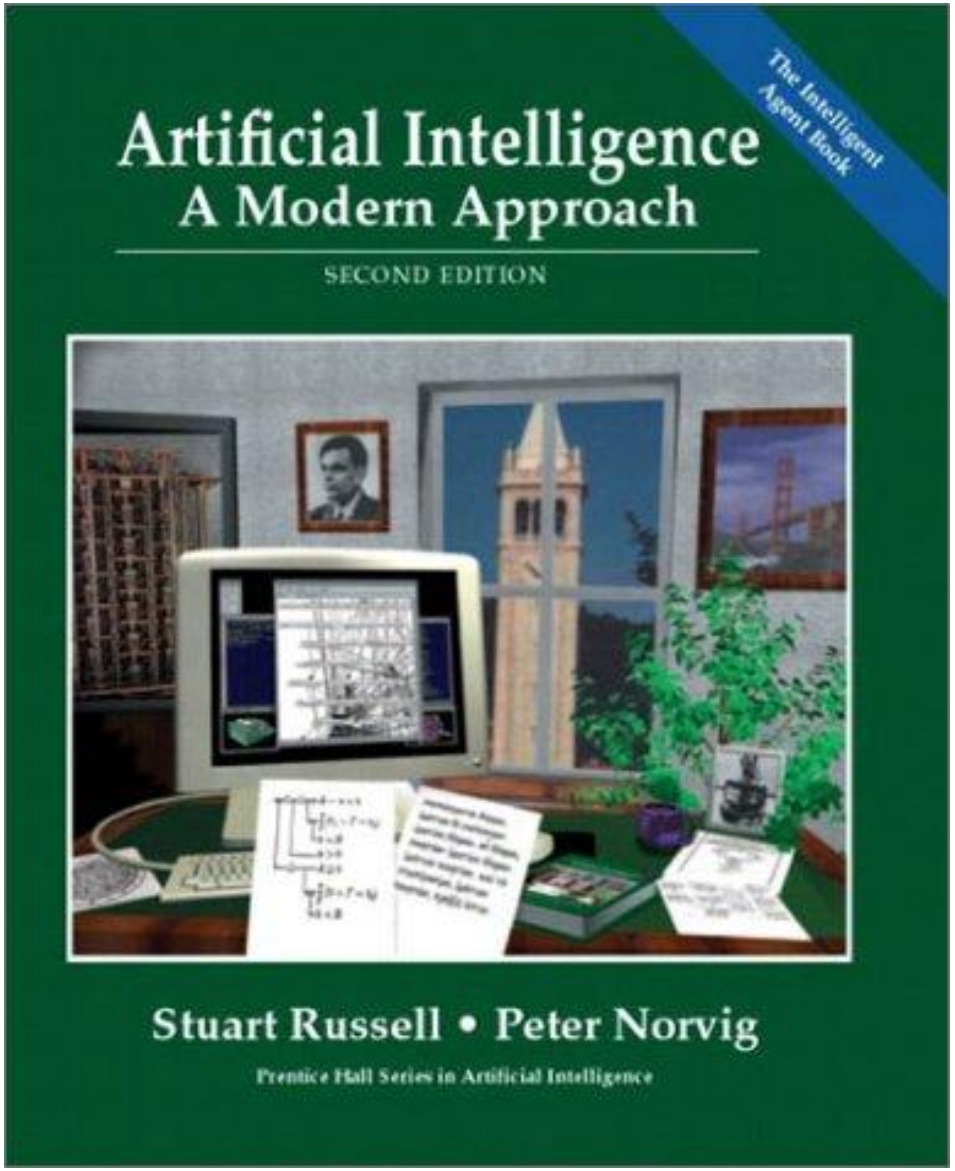
Intelligent Systems Group
Dept of Measurement and Information Systems,
Budapest University of Technology and Economics



Course info

- Course site https://www.mit.bme.hu/eng/oktatas/targyak/VIMIAC00
- Lecturers Péter Antal, antal@mit.bme.hu
  Tadeusz Dobrowiecki, tade@mit.bme.hu
  Péter Sárközy, psarkozy@mit.bme.hu
  György Strausz, strausz@mit.bme.hu
- Schedule
  - Monday 14.15-15.45, Room IE.224, Bld. I, Wing E, 2nd floor
  - Odd academic weeks: Thursday 8.30-10.00, Room IE.224
- Contact hours
  - By appointment, BME IE.412
- Book
  - S. Russell and P. Norvig: Artificial Intelligence: A Modern Approach, Prentice Hall, 2nd Ed.
- Slides
  - Based on AIMA slides from S.Russel/T.Leanert/H.Ng (At course site)

Course info



http://aima.cs.berkeley.edu/

Homework, midterms, …grading

- Grading:
  - Obligatory midterm tests,
    - 0-55 points, min.40%,
    - at the 8th and the last week.
  - Major one-person homework
    - divided into 3 parts/topic assignments,
    - deadlines later,
    - max 12 points per assignment,
    - at least 2 assignments should be over 40% (5-5 points).
  - Overall
    - Weights: midterm: 55%, major: 45%
- Midterm tests are closed-book exams.

Course outline

- Artificial Intelligence – what it is all about?
- Problem solving with search
- Knowledge and logic
- Working with uncertainty
- Machine learning
- Cooperative intelligence

Overview of today lecture

- What is intelligence and what may be artificial intelligence?
- Trends in computer science
  - The knowledge era and the data-intensive age
  - The age of online learning with autonomy
- Autonomous agents

Browsing for „intelligent„

(Google: about 403,000,000 results in 0.61 seconds)

~ house, household, robot, airplane, car, ship, vacuum cleaner, telephone, washing machine, …

~ city, sensor, material, label, traffic system, form filler, news searcher, company, …

~ motor, glasses, hand tool, accu, surgical scalpel, intercom, camera, coffee machine, electric meter, traction control, lamp pylon, AC adapter, …

~ bandage, pill, medicine, pen, paper, map, pram, bulb, connector, protective coat, roof system, key cabinet, …

~ plasticine, cream cake, vine bottle, fountaine, toothbrush, washing powder, washing, …

Intelligent pill?

May do something better.
But what should it do better?
How should do it better?
Perhaps it shoud do also
something else?

Intelligent systems – a bit more seriously

What is the meaning of being „intelligent”?

Why we should need something „intelligent”?

How to realize to be „intelligent”?

Intelligent system (probably) are:

smart, know more, can learn, adaptive, robust, autonomous, can speak, understand others, stick together, solve problems, …

What is (human) intelligence?

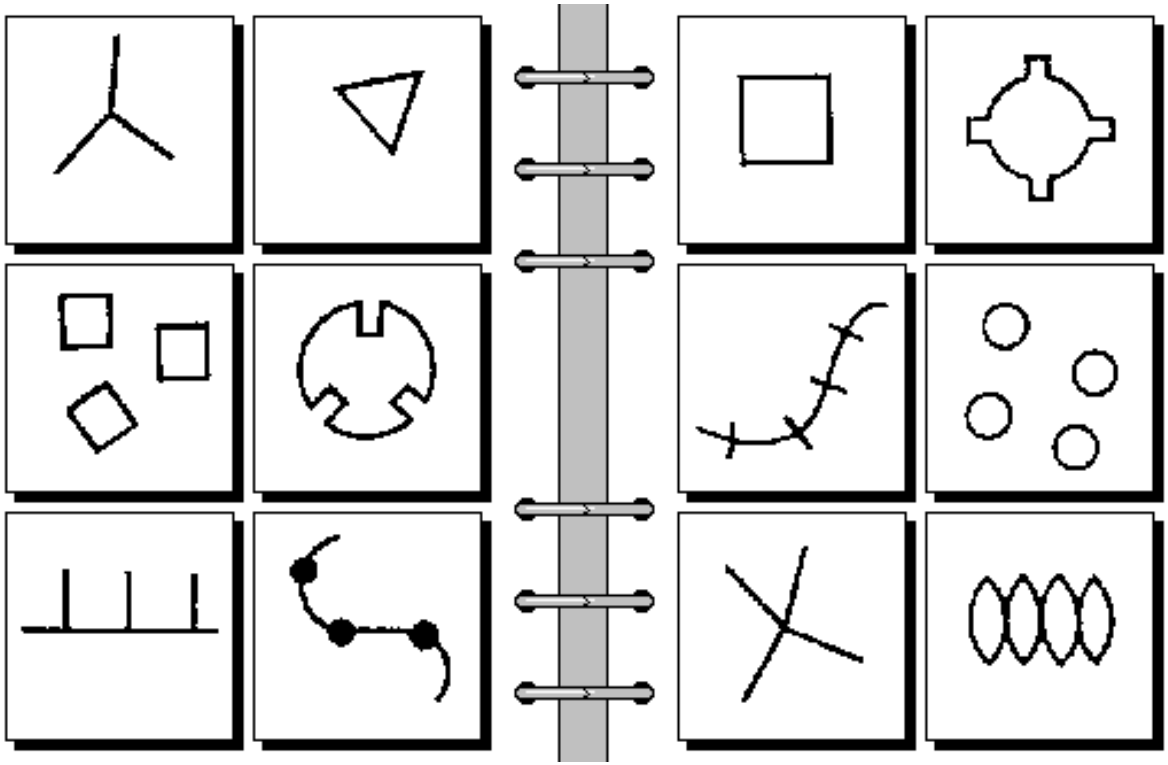
In a human – resultant of multiple factors.

Intelligence (generally) is the ability to (pose and to) realize goals in the world.

It can be found in humans in many forms and levels, in many animals, and in some machine systems.

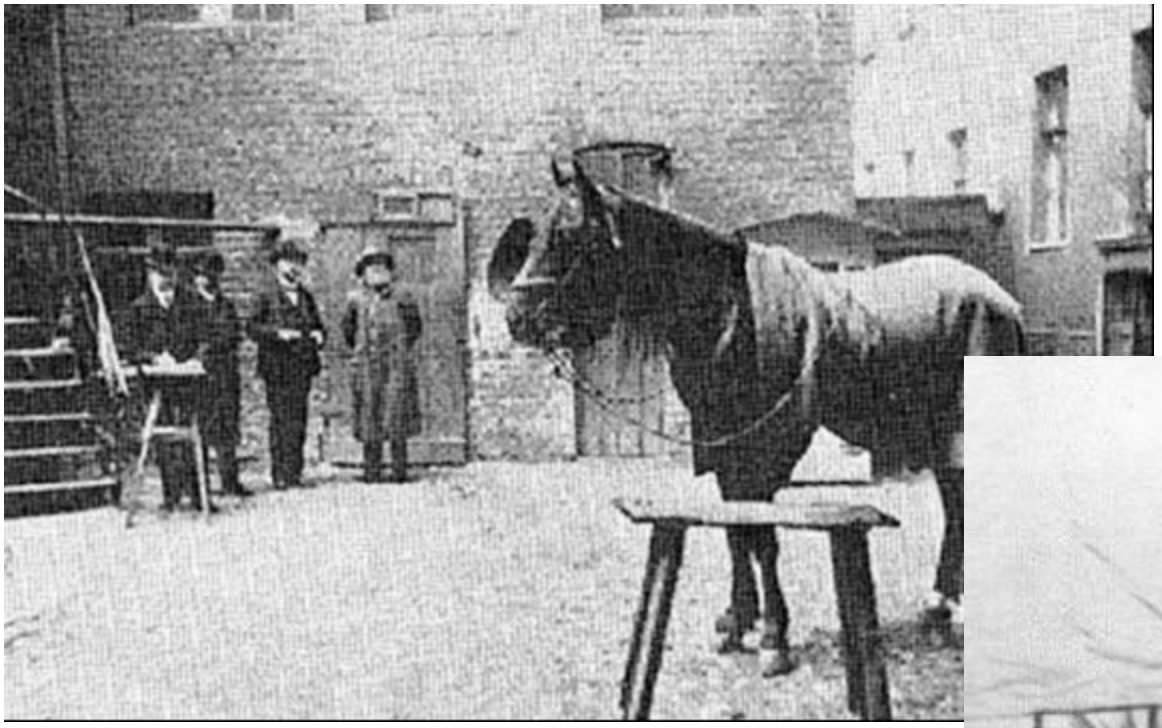
Artificial intelligence typically takes from here **learning**, **reasoning**, **problem solving**, and **language**.

Verbal/Linguistic
The capacity to use language to express what's on your mind and to understand other people.
Existential
To exhibit the proclivity to pose and ponder questions about life, death and ultimate realties.
Logical/Mathematical
The ability to understand the underlying principles of some kind of causal system.
Interpersonal
The ability to understand other people.
Multiple Intelligences
Intrapersonal
Having an understanding of yourself, of knowing who you are, what you can do, etc.
Visual/Spatial
The ability to present the spatial world internally in your mind.
Musical/Rhythmic
The capacity to think in music, to be able to hear patterns, recognize them, and perhaps maniipulate them.
Naturalist
The ability to discriminate among living things as well as sensitivity to other features of the natural world.
Bodily/Kinesthetic
capacity to use your whole or parts of your body, to solve problem, make something, or put on a production.
M.I. Smart! jcp
1999

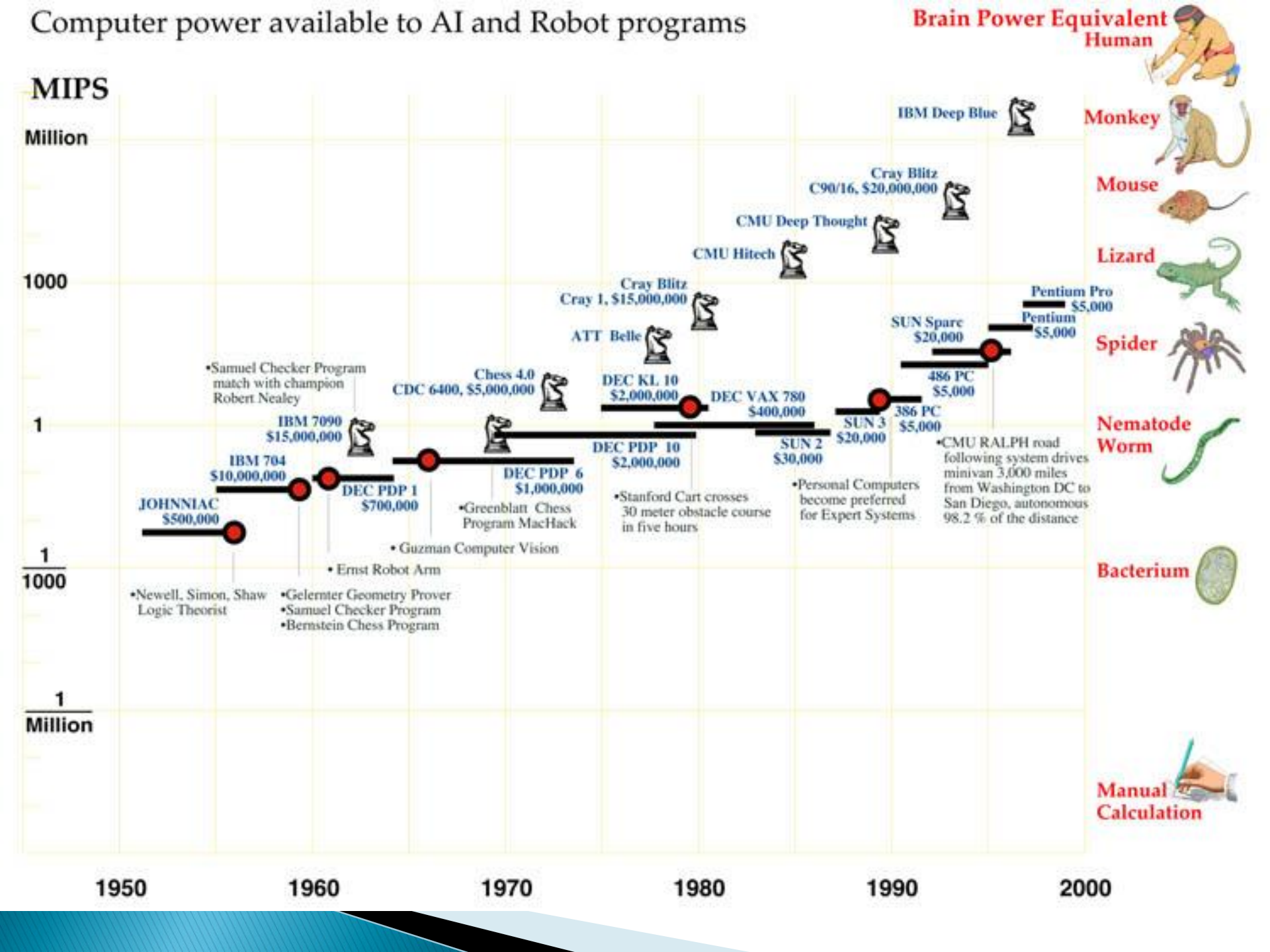


Robot, go toward „more”!

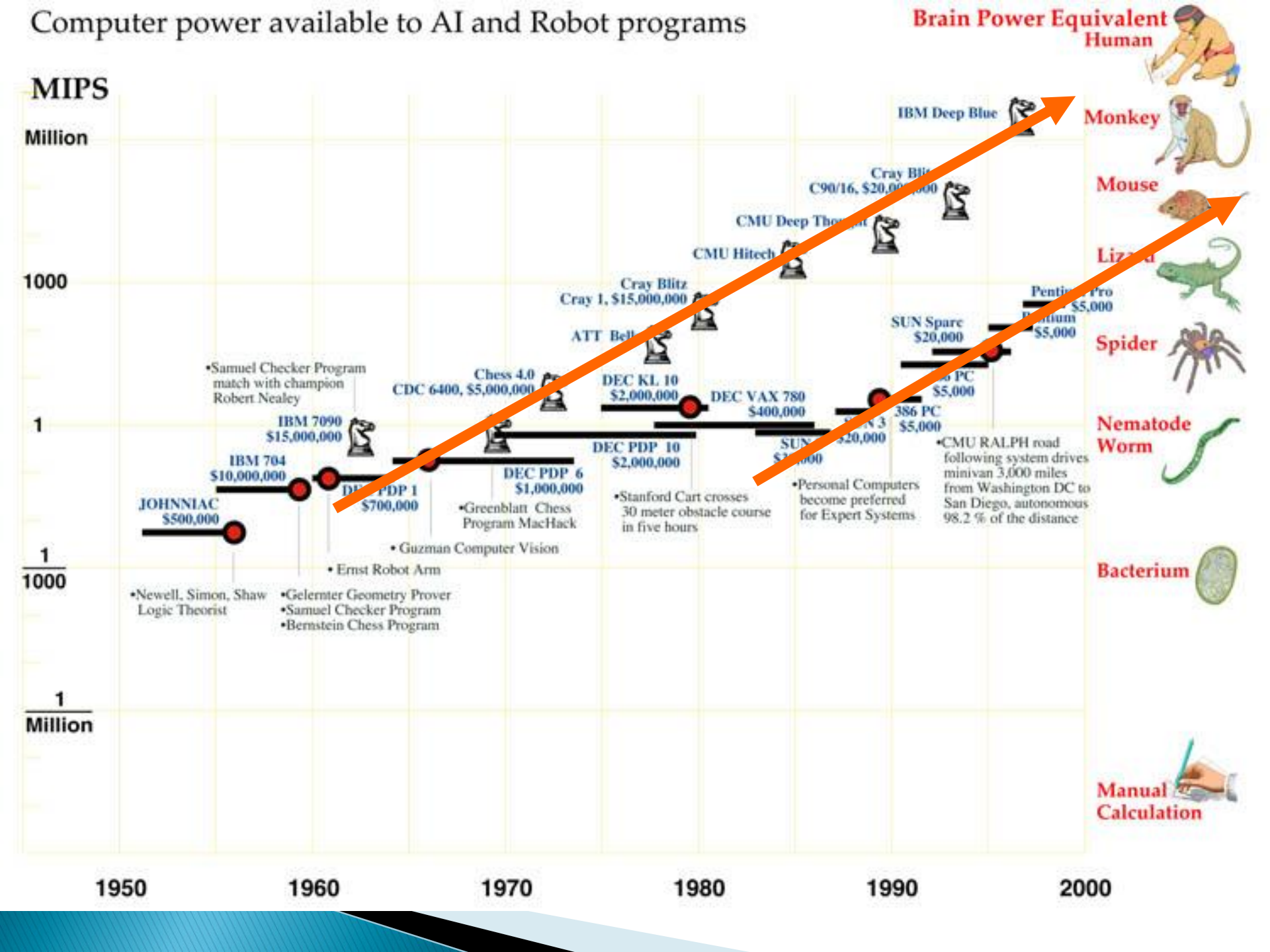
(Bongard problems,
https://en.wikipedia.org/wiki/Bongard_problem)



Computer power available to AI and Robot programs



Computer power available to AI and Robot programs



Factors behind intelligence explosion

- Computation
  - Moore's law
- Data
  - Big data age
- Knowledge
  - Publications, knowledge bases,…
- Technologies
  - Artificial intelligence? Language understanding?
  - Machine learning? Deep learning?

Why do we need AI at all?

- Understanding human cognition
- Supporting and complementing human expertise
- Supporting, complementing, extending human abilities
- No choice: data & knowledge exceeded the scope of human cognition
- Instead of human experts, it is
  - slightly cheaper ;-), scalable, multiplicable, durable, ..
- **Curiosity(!) + optimism(?)**

What is then the AI?

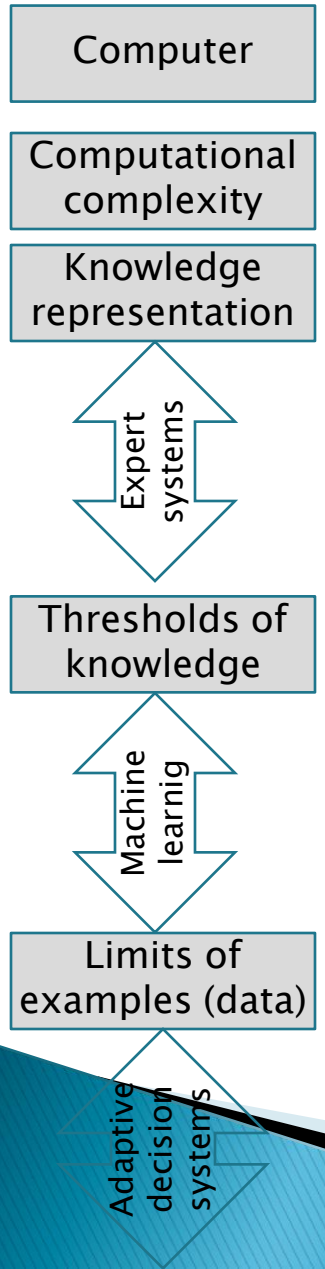
Science of building intelligent machines,
particularry intelligent computer programs.

(For us) the AI is a part of the Computer Science and Computer Engineering.

It studies the possible computer realization of such abilities as sensing, reasoning, learning, furthermore it attempts to formulate principles how to build computer systems exposing such abilities.

Is AI a „true” intelligence? Here the philosophers are fighting, but an engineer can work undisturbed.

Milestones and phases in AI



- ~1930: Zuse, Neumann, Turing..: „instruction is data”:
  - Laws of nature can be represented, „executed”/simulated, learnt
  - Knowledge analogously: representation, execution, adaptation and learning
- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1956 Dartmouth meeting: the term "Artificial Intelligence”
- 1950s Early AI programs (e.g. Newell & Simon's Logic Theorist)
- **The Symbolic system hypothesis: search**
- 1965 Robinson's complete algorithm for logical reasoning
- 1966-73 AI discovers computational complexity
  Neural network research almost disappears
- 1969-79 Early development of knowledge-based systems

- **The knowledge system hypothesis: knowledge is power**
- 1986- Neural networks return to popularity
- 1988- Probabilistic expert systems
- 1995- Emergence of machine learning
- **The „big data” hypothesis: let data speak**
- 2005/2015-- Emergence of autonomous adaptive decision systems („robots”, agents)**The autonomy hypothesis??**

What is then AI (minimalistic experience)?

Dealing only with **feasible** and **solvable** questions:

- how we should analyze **problems demanding intelligence**, and how to **specify** them for the machine solution

- how to **describe** and **manage** the **knowledge formally** (with mathematical tools)

- what system architectures to use to realize machine problem solving based on the knowledge management

- central concepts **knowledge representation**, and **search**.

State of the art: ☺

- Deep Blue defeated the reigning world chess champion Garry Kasparov in 1997
- Proved a mathematical conjecture (Robbins conjecture) unsolved for decades
- No hands across America (driving autonomously 98% of the time from Pittsburgh to San Diego)
- During the 1991 Gulf War, US forces deployed an AI logistics planning and scheduling program that involved up to 50,000 vehicles, cargo, and people
- NASA's on-board autonomous planning program controlled the scheduling of operations for a spacecraft
- Proverb solves crossword puzzles better than most humans
- Google search/car/face recognition/…

AI approaches?

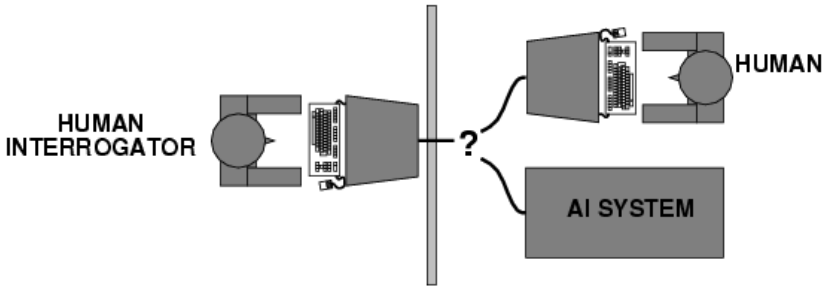
Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

More about

Textbook, Chapter 1.1, What Is AI?

Acting humanly: Turing Test

- Turing (1950) "Computing machinery and intelligence":
- "Can machines think?" → "Can machines behave intelligently?"
- Operational test for intelligent behavior: the Imitation Game



- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in following 50 years
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

Thinking humanly: cognitive modeling

- 1960s "cognitive revolution": information-processing psychology
- Requires scientific theories of internal activities of the brain
- -- How to validate? Requires

  1) Predicting and testing behavior of human subjects (top-down), or

  2) Direct identification from neurological data (bottom-up)

- Both approaches (roughly, Cognitive Science and Cognitive Neuroscience) are now distinct from AI, but ➔
- Hassabis, Demis, et al. "Neuroscience-inspired artificial intelligence." *Neuron* 95.2 (2017): 245-258.

Thinking rationally: "laws of thought"

- Aristotle: what are correct arguments/thought processes?
- Several Greek schools developed various forms of *logic*: *notation* and *rules of derivation* for thoughts; may or may not have proceeded to the idea of mechanization
-
- Direct line through mathematics and philosophy to modern AI
-
- Problems:
  1. Not all intelligent behavior is mediated by logical deliberation
  2. What is the purpose of thinking? What thoughts should I have?

➔ **(Symbolic) reasoning is mainly for collaborative thinking**!

Acting rationally: rational agents

- Rational behavior: doing the right thing
- The right thing: that which is expected to maximize goal achievement, given the available information
- Doesn't necessarily involve thinking – e.g., blinking reflex – but thinking should be in the service of rational action

Rational agents

- An agent is an entity that perceives and acts
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions:

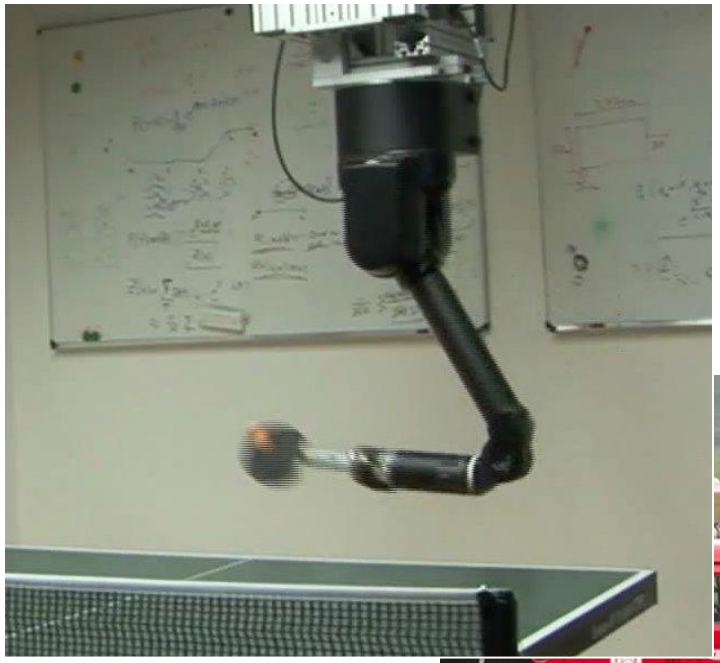
$$[f: P^* \rightarrow A]$$

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat: computational limitations make perfect rationality unachievable

→ design best program for given machine resources

Rational agents

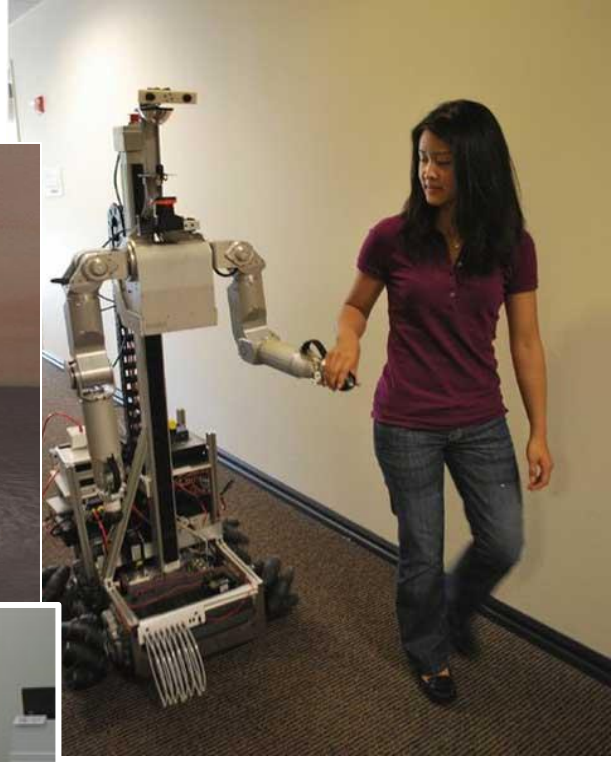
let see some …



DARPA
GRAND CHALLENGE

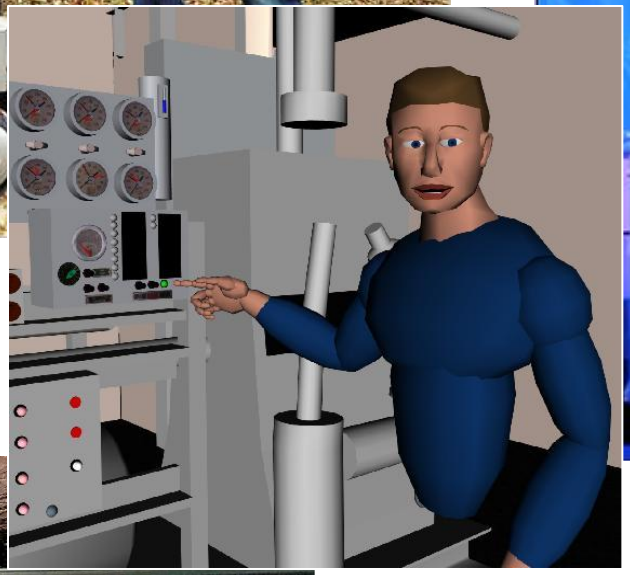
APPLANIX
REDTEAM

PVAMC
MGH



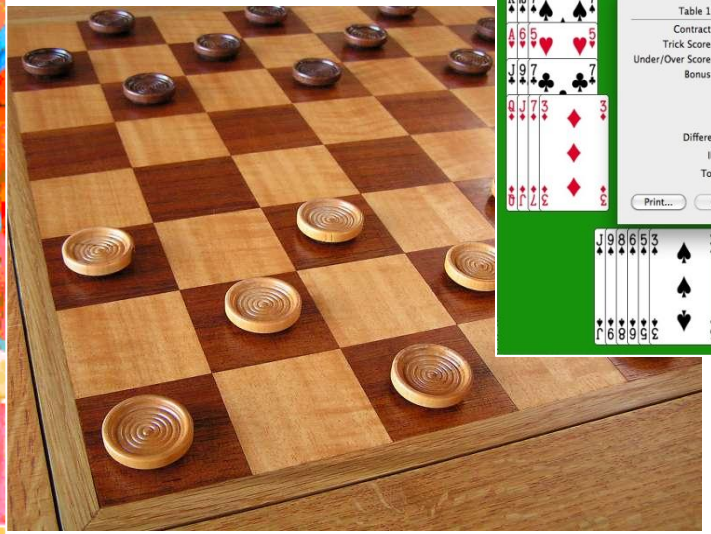
RIBA II
YouTube HD

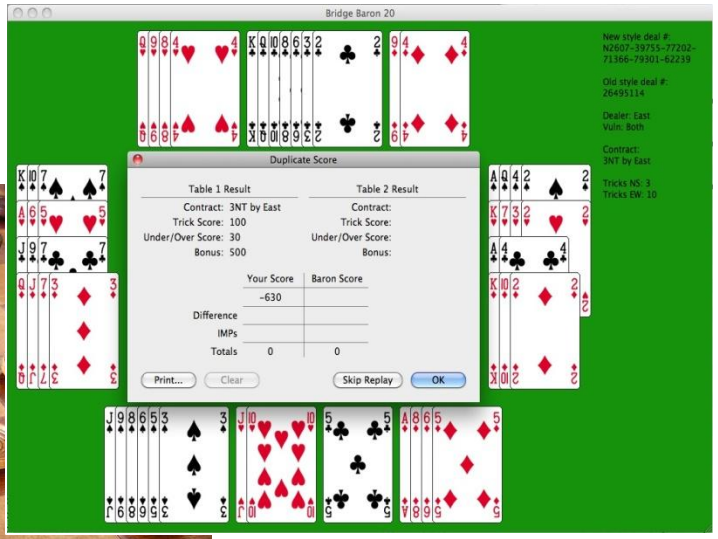
ADE Geriatric Oncology
Adele
Case
M. Jones
Role
Physician
Pre-clinical
View Chart
View Patient
View Progress
References
Notebook
Exit
Doctor, I have a lump on my chest...

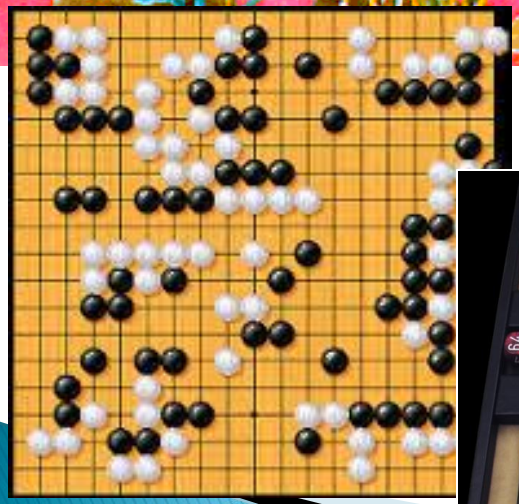


$18,200
$21,440
$5,600
WATSON
BRAD
contempt 97%
contemn 14%
Despised Icon 10%

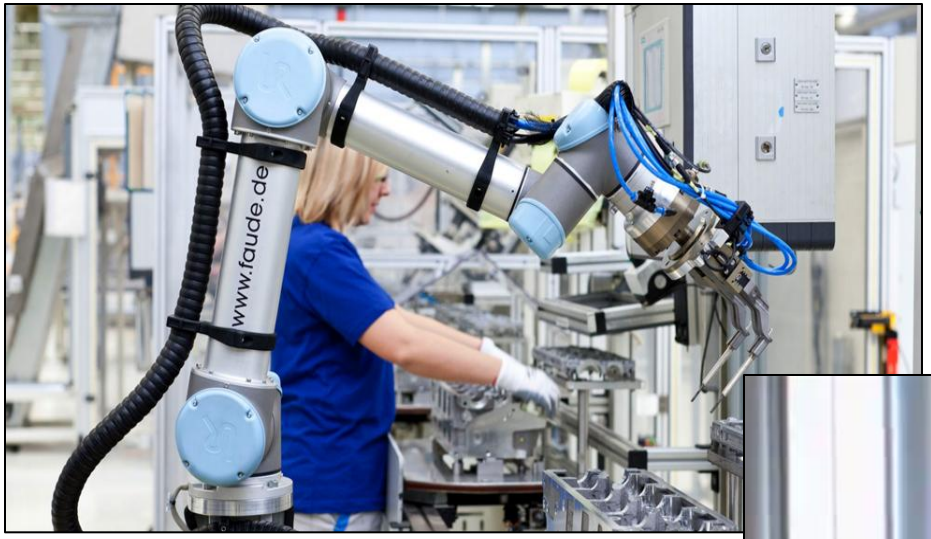
TOP STRATEGY GAMES PC



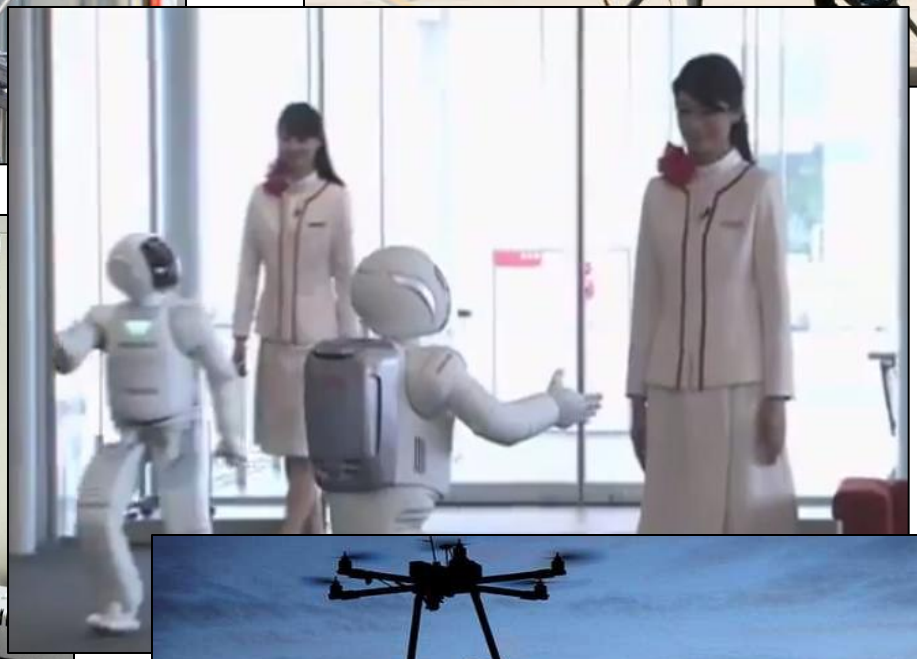

Bridge Baron 20
Duplicate Score
Table 1 Result
Contract: 3NT by East
Trick Score: 100
Under/Over Score: 30
Bonus: 500
Table 2 Result
Contract:
Trick Score:
Under/Over Score:
Bonus:
Your Score
Baron Score
-630
Difference
IMPs
Totals 0 0
Print...
Clear
Skip Replay
OK
New style deal #:
N2607-39755-77202-71366-79301-62239
Old style deal #:
26495114
Dealer: East
Vuln: Both
Contract:
3NT by East
Tricks NS: 3
Tricks EW: 10



IBM
GARRY
SPAROV
DEEP
BLUE


www.faude.de

U.S. AIR FORCE



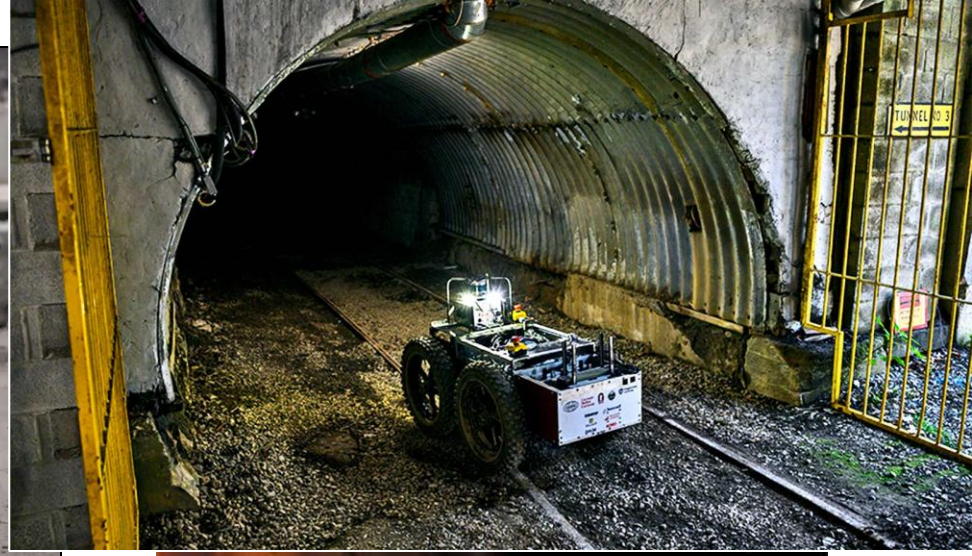
Caliper
IX

SEP 15 2001
11:23:47 PM

TUD
TUD
TUD

Roomba
Discovery
iRobot

Poseidon
DIRECT

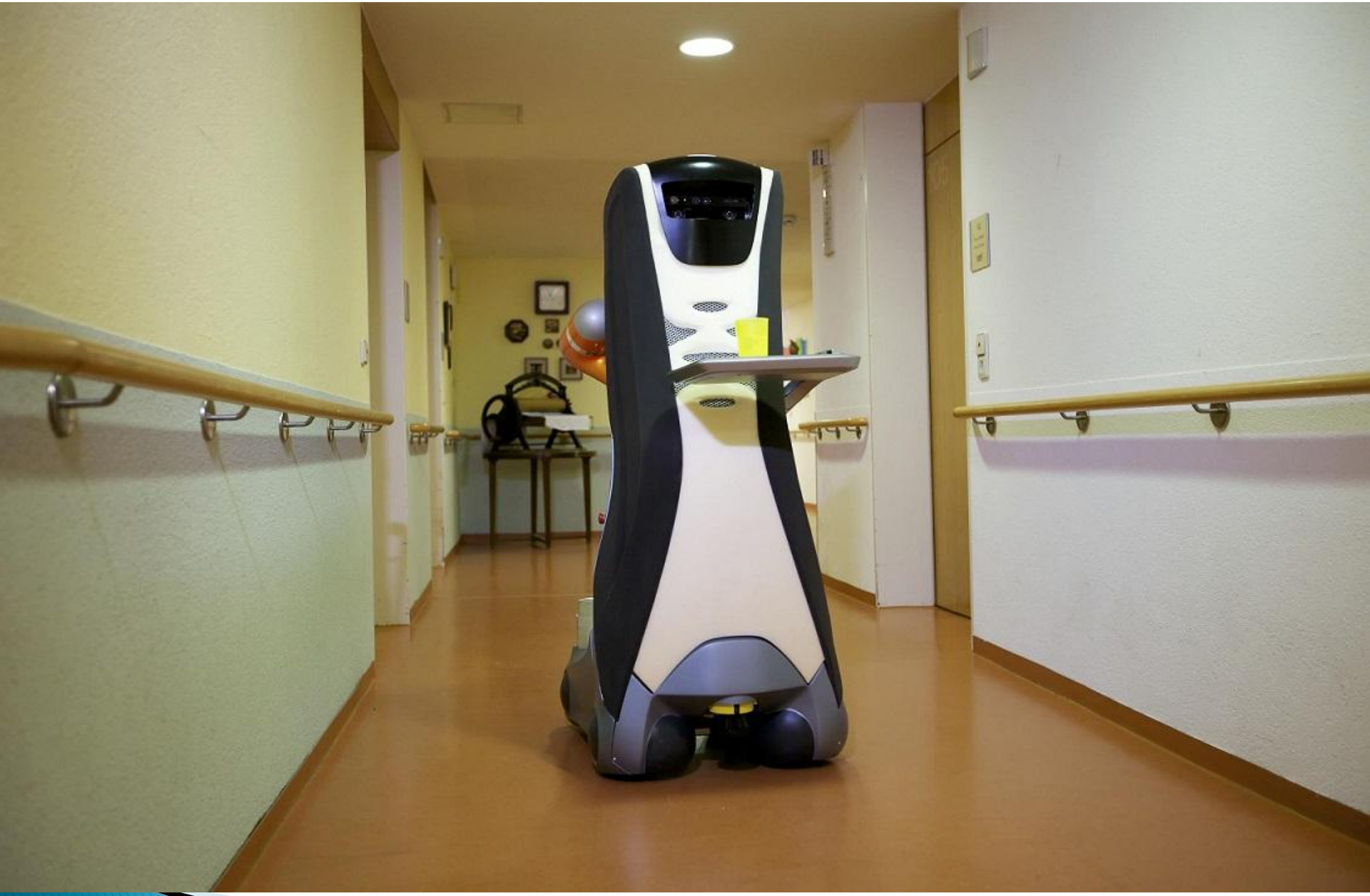

TUNNEL NO 3

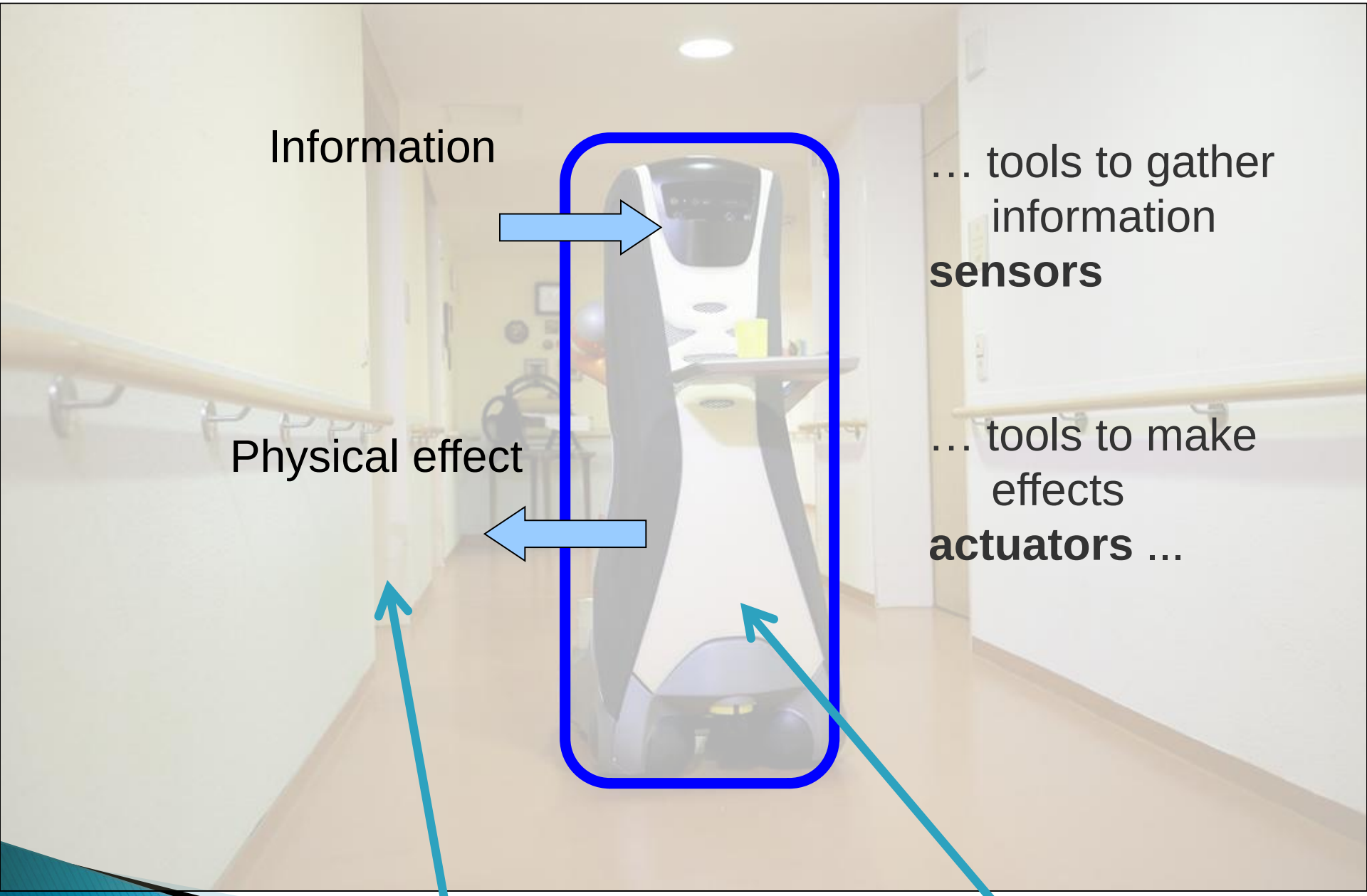
INTOUCH HEALTH

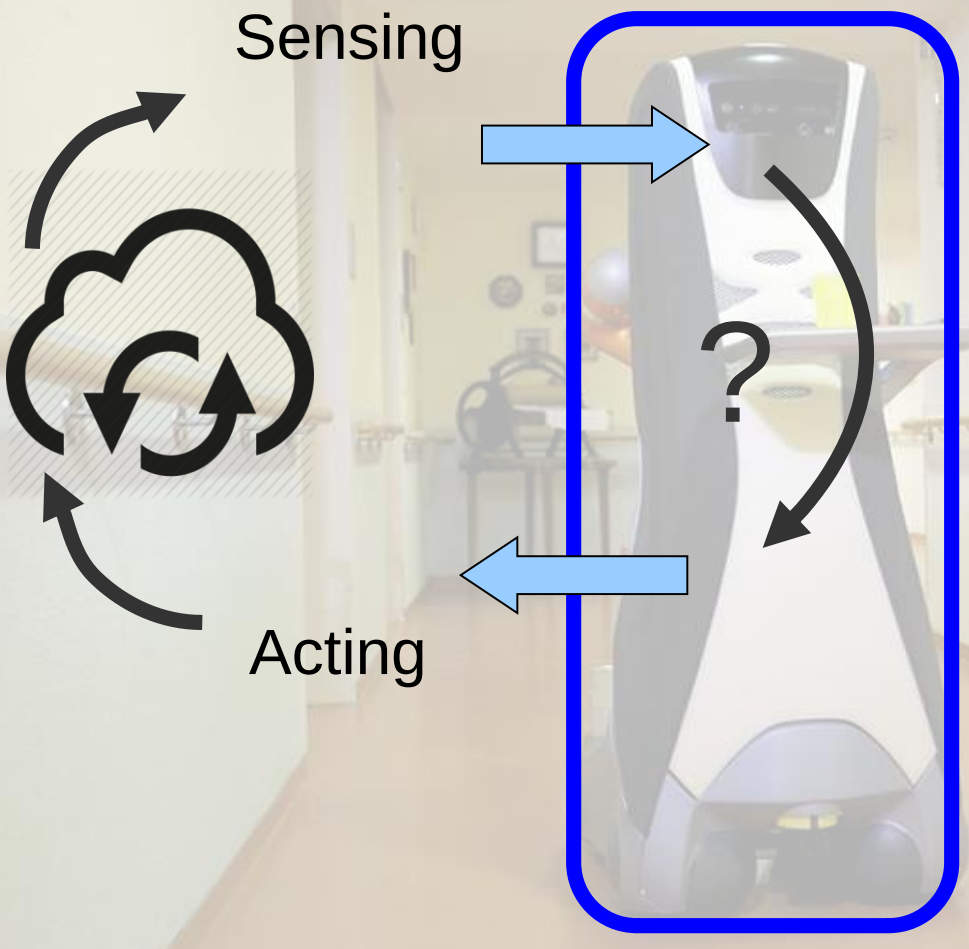
Rational agents

let look inside …

Care-O-bot 3, Support of elderly people at home
https://www.care-o-bot.de/en/care-o-bot-3.html




Information
Physical effect
… tools to gather information
sensors
… tools to make effects
actuators ...
Outside: an **environment**
Inside: a **system**



Agent

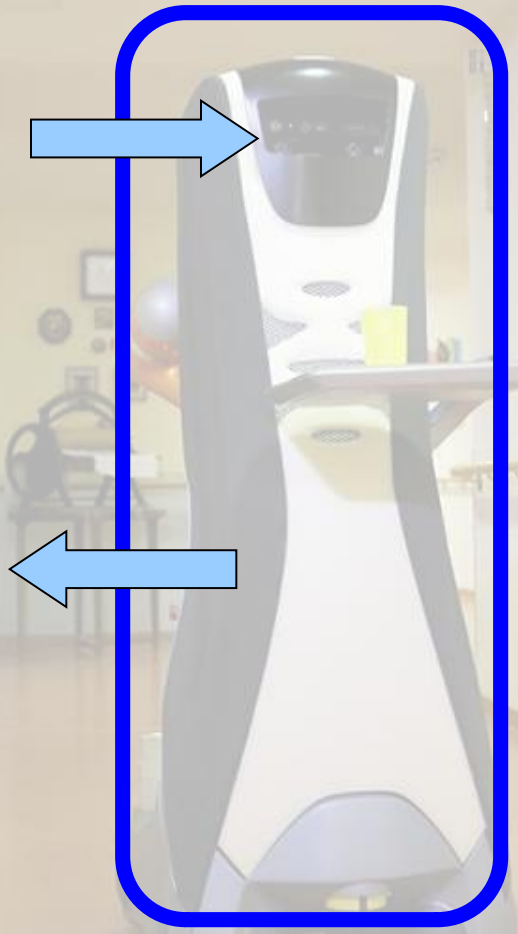
A system embedded (physically) in its environment, with which it continuously interacts, it senses the environment with sensors and influences it with actuators ...

States of agent's environment

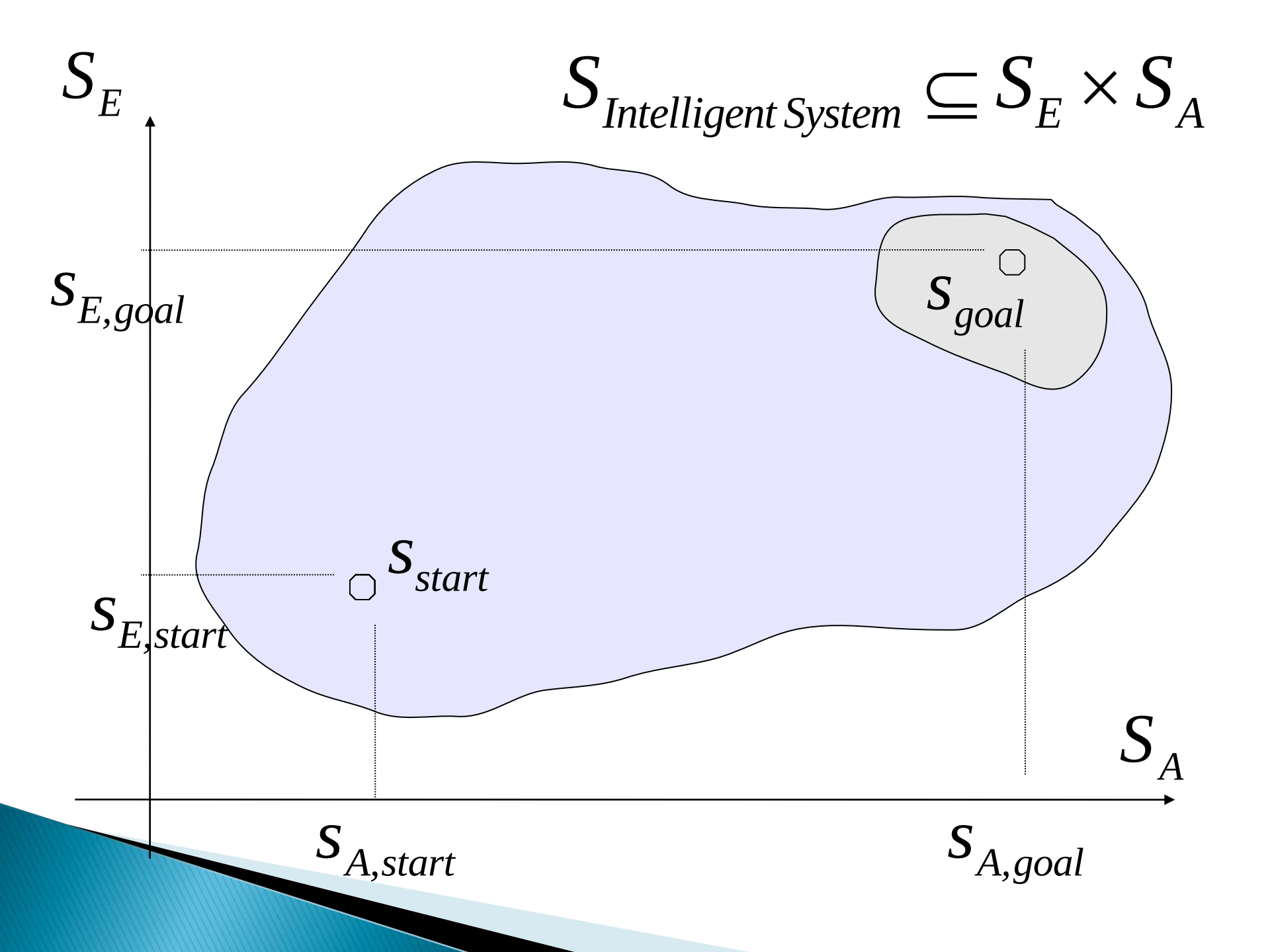
$$s_E(t) \in S_E$$

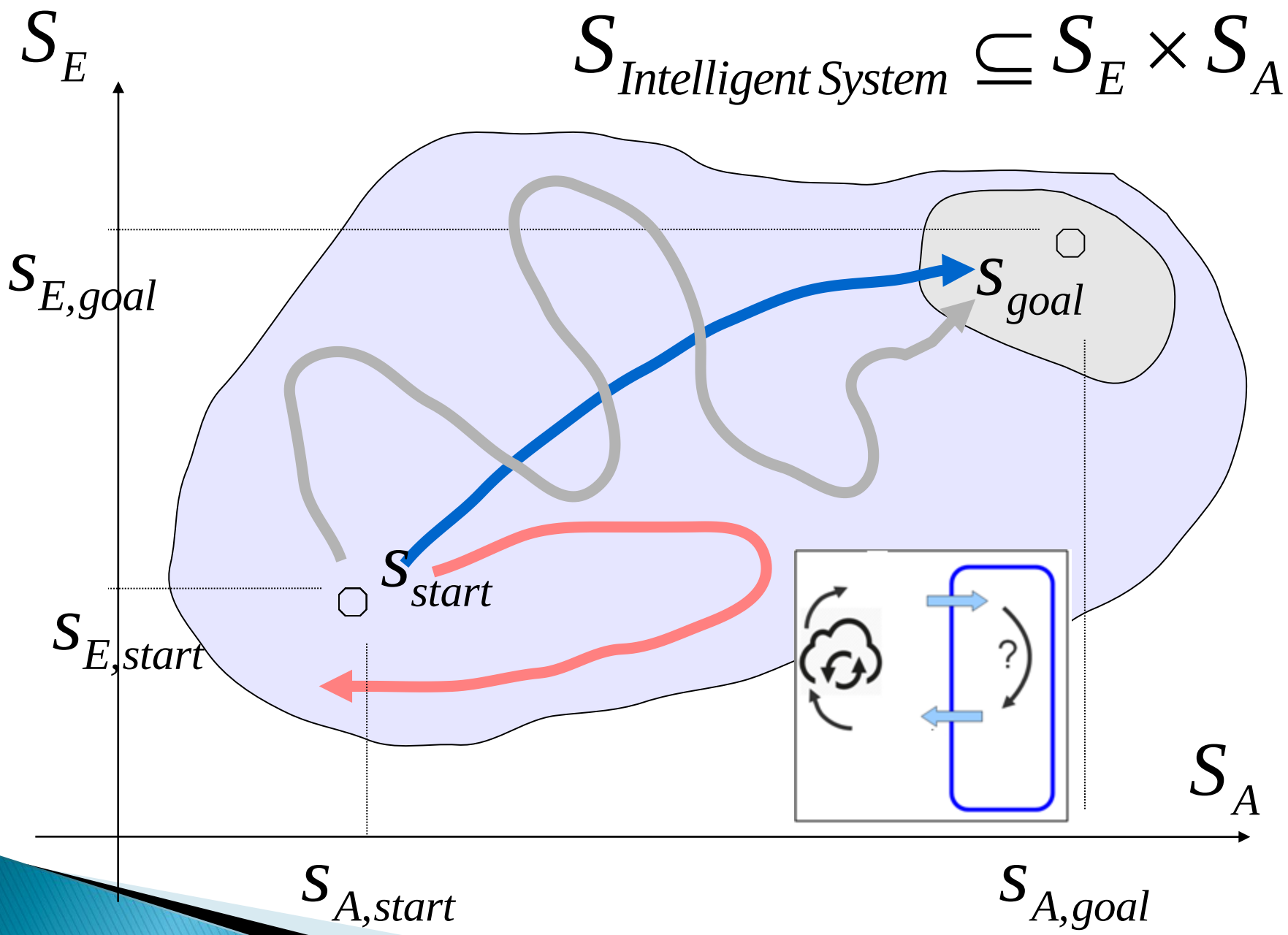
States of agent

$$s_A(t) \in S_A$$



$$S_{Intelligent\ System} \subseteq S_E \times S_A$$


$S_{Intelligent\,System} \subseteq S_E \times S_A$
S_E
$S_{E,goal}$
S_{goal}
S_{start}
$S_{E,start}$
S_A
$S_{A,start}$
$S_{A,goal}$


$S_{Intelligent\ System} \subseteq S_E \times S_A$
S_E
$S_{E,goal}$
$S_{E,start}$
S_{goal}
S_{start}
?
S_A
$S_{A,start}$
$S_{A,goal}$

Some simple comments

The „**goal**” of an agent is to reach or to „make” such specific state of its environment, which is useful or desirable for the agent.

Executing loop of sensing and acting agent moves in the state-space along a **trajectory**. Trajectories may differ. An agent, which realizes a good (efficient, …) trajectory, is „smarter” than other agents.

Rational action = action toward the goal.

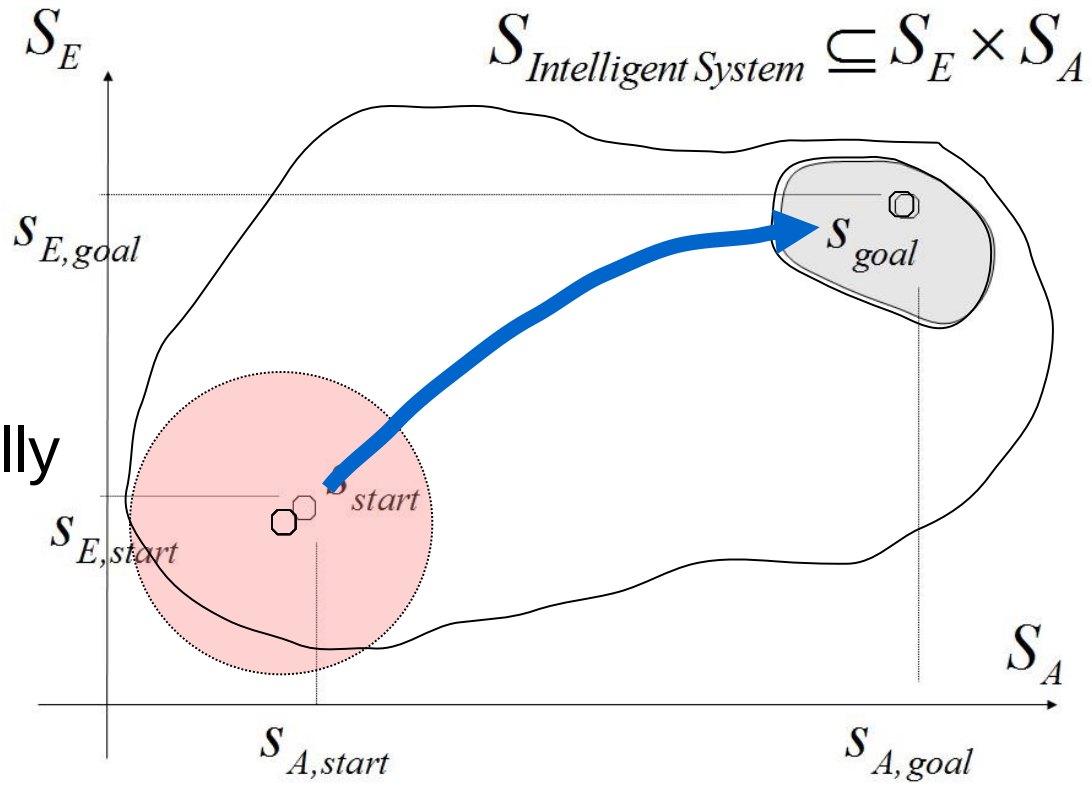
Intelligent a**gent –** chooses its actions rationally and reaches its goals, acting against environmental changes and difficulties.

The ideal rationality is impossible, the computing demands are too high. **Bounded rationality** - acting properly, when there is no time to do all the computations.

Mechanism to keep an agent on its trajectory, as long as it senses a difference between the actual and the goal state. Agent's task: to „**compute**” the **action** from the **sensing**.

Problem: „distance”

Local radius of an action
Global goal distance

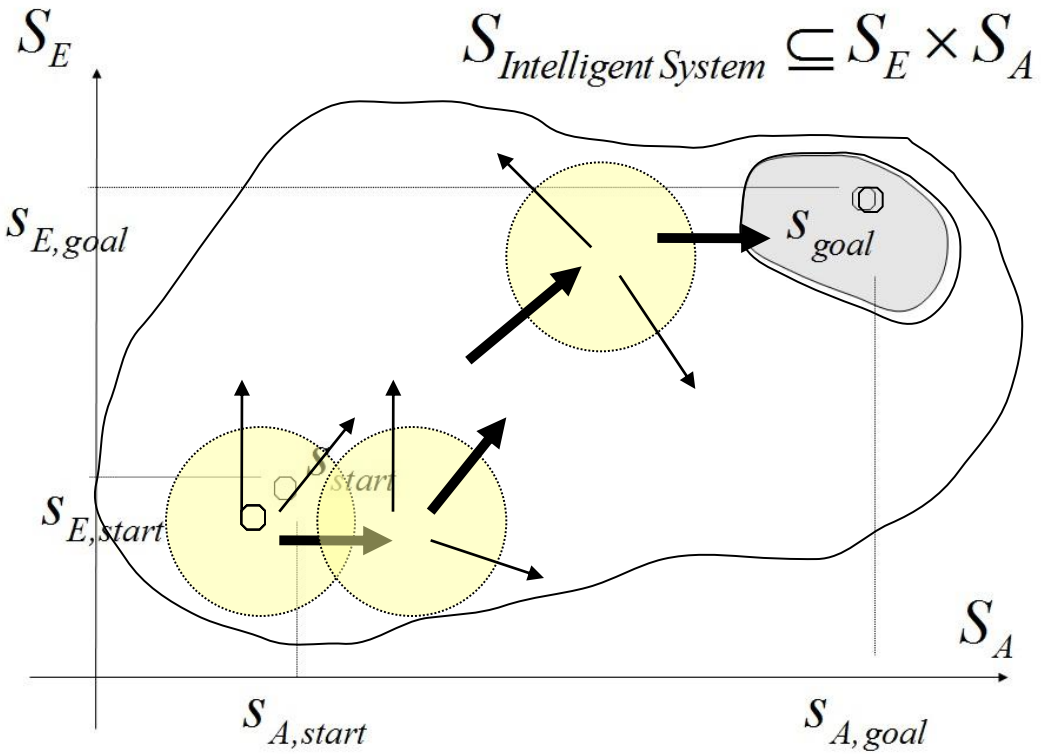


A single action generally is not enough to solve the problem

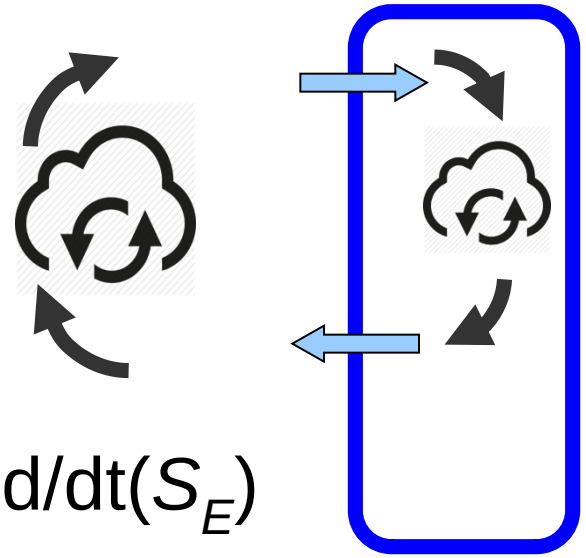
Problem: „correct” direction

Universal search algorithm
what would happen, if I move …,
and then what would happen, if I move further
and then ...

What is the „direction” of an action?



Problem: dynamics



If „deteriorate” is faster than „improve”?

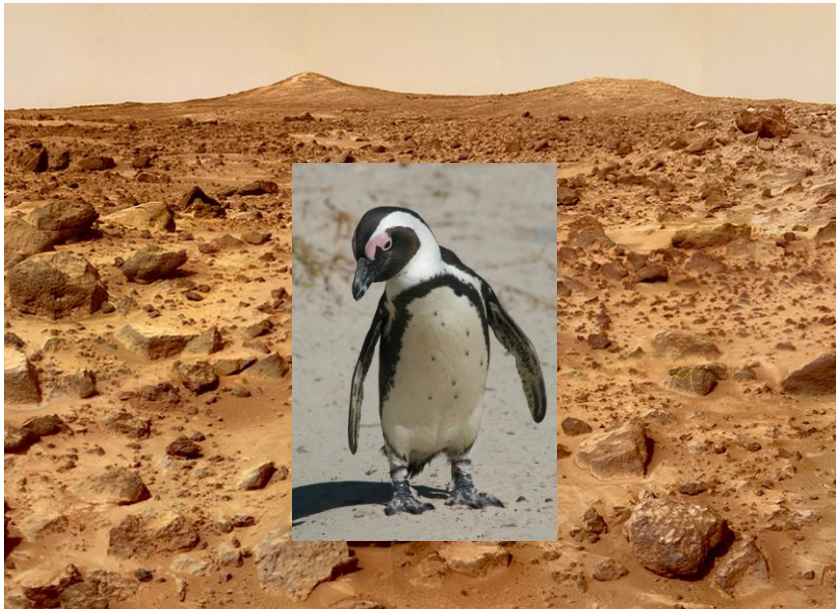
Who is faster?

The environment? The adaptation, the learning of an agent never will be successful.

The agent? It will succeed despite the changing environment.

Same? A true problem when more agents are around. Their mutual learning won’t converge!

„Pinguin principle”
– a system can be truly intelligent only in its „native” environment.



Personal view!

„Ferrari principle”
– a system is truly intelligent, if it is adequate to the demands of the task. Less and more Intelligence - disadvantage.

IT (Information Technology) perspective:

Intelligence is a designable and scalable system property.

With intelligence we realize demanding and new system services.

A computer science engineer should be able to manage the intelligence level of the designed systems.

We will show, how a (part of) intelligence can be:
expressed with maths,
coupled to algorithms, and
embedded into suitable architectures.

Agent = (Agent)Architecture + (Agent)Program

Starting observations:

Environments can be classified into types
(Environment types)

Problems appearing in the environments can be classified into types (Problem types)

Programs can be classified into types
(Agent functions and Agent types)

.

More about

Textbook, Chapter 2, Intelligent Agents

Environments

accessible	not accessible
deterministic	not deterministic
episodic	not episodic
static	dynamic
discrete	continuous
single agent	multiple agent

Problems

Single state problem: physical state = belief state

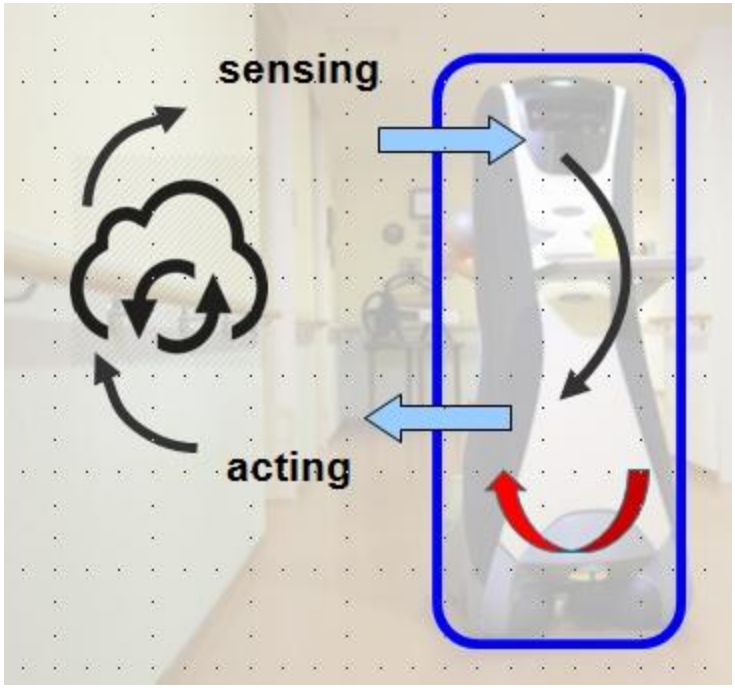
Multiple state problem: physical state set = belief state

Contingency problem:

(finite) deficit in information, contingency plans

Discovery problem

Essential information deficit, learning critical situations



Agent function

a = action
o = observation
KB = knowledge base
T = {1, …, t-1}

Mathematically an agent:

$$a_t = f(o_T , a_T , KB_{t-1})$$

$a_t = f(o_{t-1})$
$a_t = f(o_T)$
$a_t = f(o_{t-1} , a_{t-1})$
$a_t = f(KB_{t-1})$
$a_t = f(o_{t-1} , a_{t-1} , KB_{t-1})$
…

Which agent function means more intelligence?

Agent types (abstract architectures)

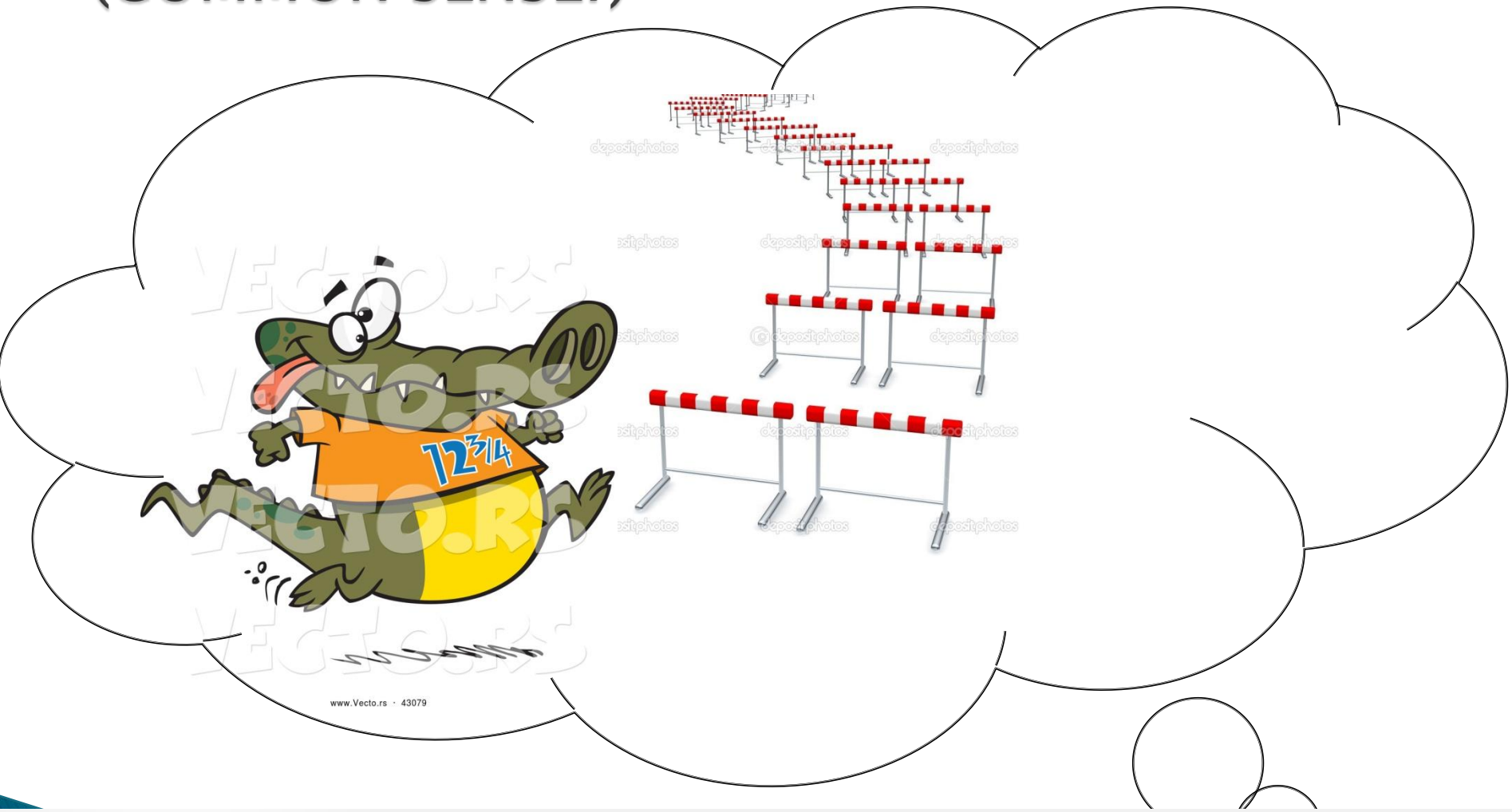
Simple reflex agent
Model based agent
Goal-oriented agent
Utility-oriented agent
Learning agent

Hybrid (layered, modular) agent
MAS (Multiple Agent Systems) agent

Artificial Intelligence

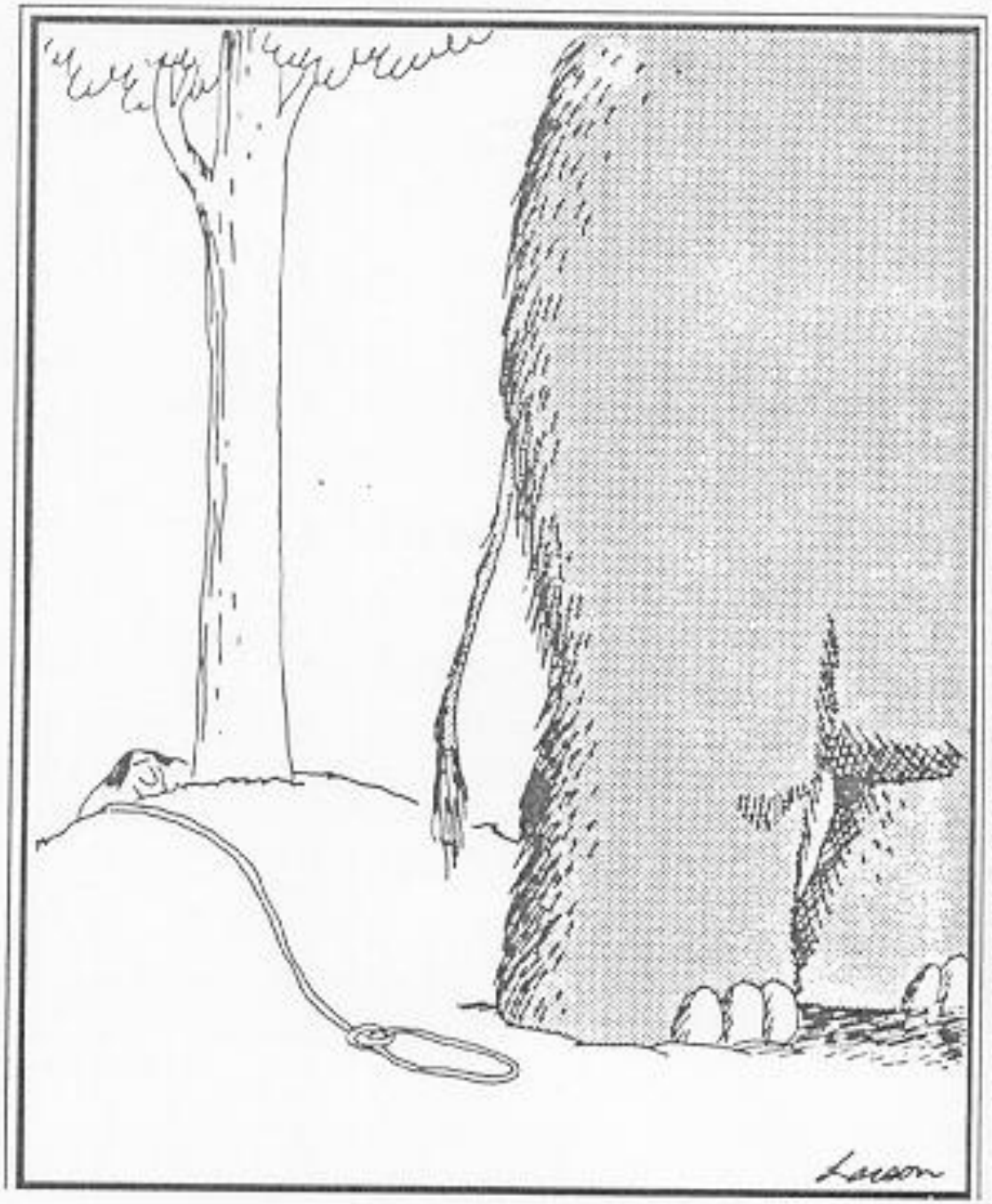
Optimistic visions,
Pessimistic visions,
Fears, and
Some tough questions

WHY CAN'T MY COMPUTER UNDERSTAND ME? (COMMON SENSE?)



Google

Can an alligator run the hundred-metre hurdles?


Lucas

Safe?
Legal?
Etical?
Pays taxes?

Misused?
Not interested
in our goals if
superintelligent?

"da Vinci Robot Allegedly Marketed to Less-skilled Doctors„
https://www.lawyersandsettlements.com/legal-news/da-Vinci-robot/davinci-lawsuit-robot-2-18655.html

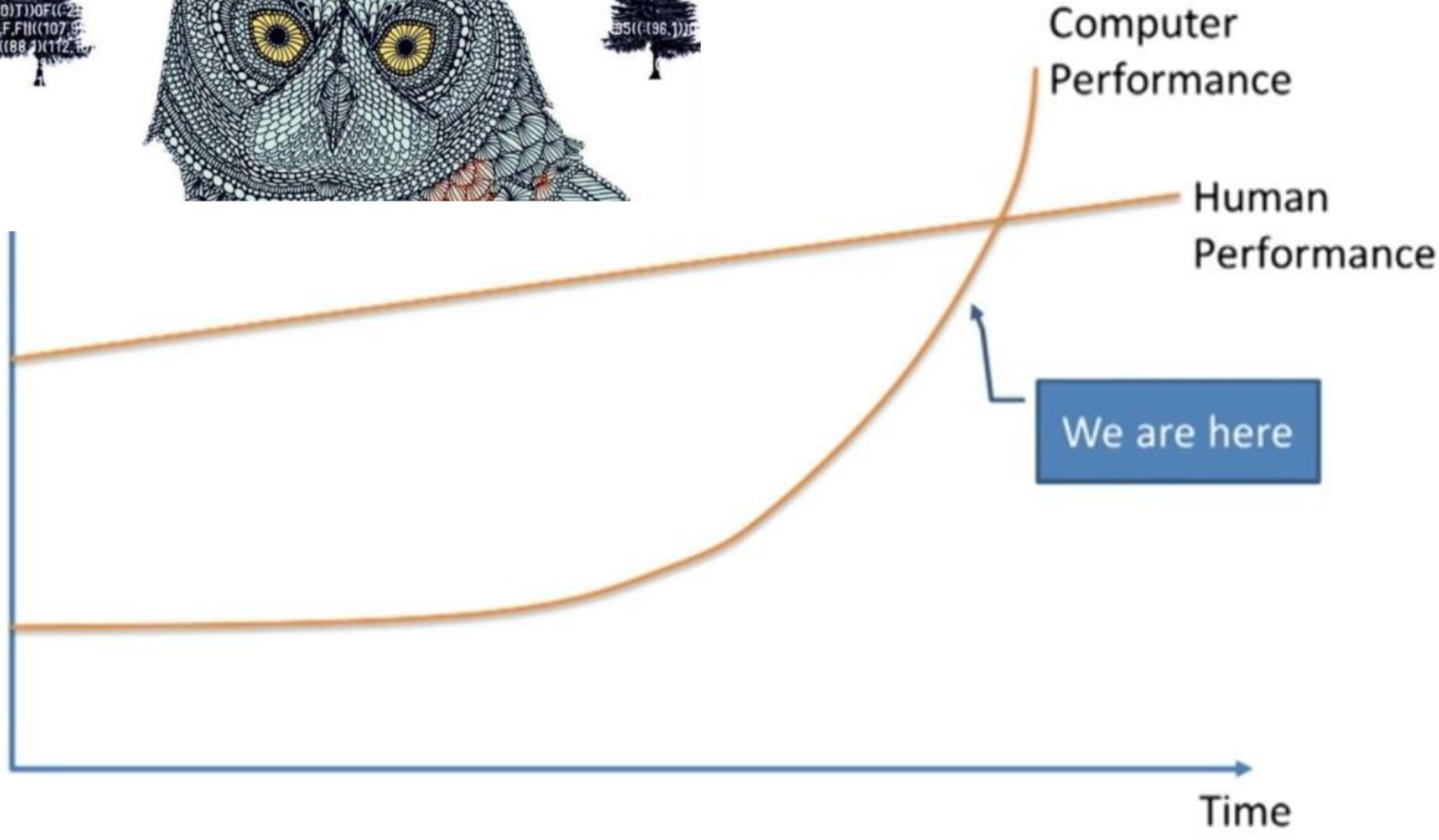
„AI Cited for Unlicensed Practice of Law”
https://www.wired.com/2007/03/ai_cited_for_un/

„ List of self-driving car fatalities”
https://en.wikipedia.org/wiki/List_of_self-driving_car_fatalities

„Automated to Death”
http://spectrum.ieee.org/computing/software/automated-to-death

SUPERINTELLIGENCE

Paths, Dangers, Strategies



Myth:
Superintelligence by 2100 is inevitable

Myth:
Superintelligence by 2100 is impossible

Fact:
It may happen in decades, centuries or never: AI experts disagree & we simply don't know

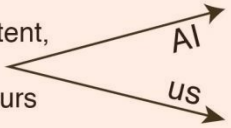
Myth:
Only Luddites worry about AI

Fact:
Many top AI researchers are concerned

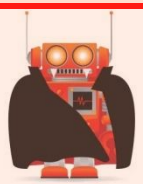
Mythical worry:
AI turning evil

Mythical worry:
AI turning conscious

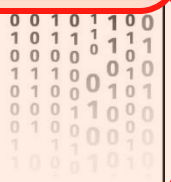
Actual worry:
AI turning competent, with goals misaligned with ours



Myth:
Robots are the main concern



Fact:
Misaligned intelligence is the main concern: it needs no body, only an internet connection



Myth:
AI can't control humans

Fact:
Intelligence enables control: we control tigers by being smarter

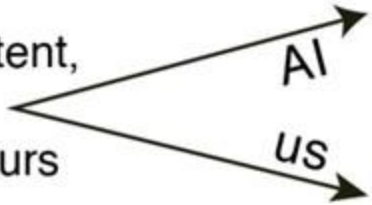
Myth:
Machines can't have goals

Fact:
A heat-seeking missile has a goal

Mythical worry:
Superintelligence is just years away

Actual worry:
It's at least decades away, but it may take that long to make it safe

Actual worry:
AI turning competent, with goals misaligned with ours



Artificial Narrow Intelligence (ANI) (Weak AI): √

Artificial Narrow Intelligence is AI that specializes in one area. There's AI that can beat the world chess champion in chess, but that's the only thing it does. Ask it to figure out a better way to store data on a hard drive, and it'll look at you blankly.

Artificial General Intelligence (AGI) (Strong, Human-level AI): ?

Artificial General Intelligence refers to a computer that is as smart as a human across the board—a machine that can perform any intellectual task that a human being can. Ability to reason, plan, solve problems, think abstractly, comprehend complex ideas,

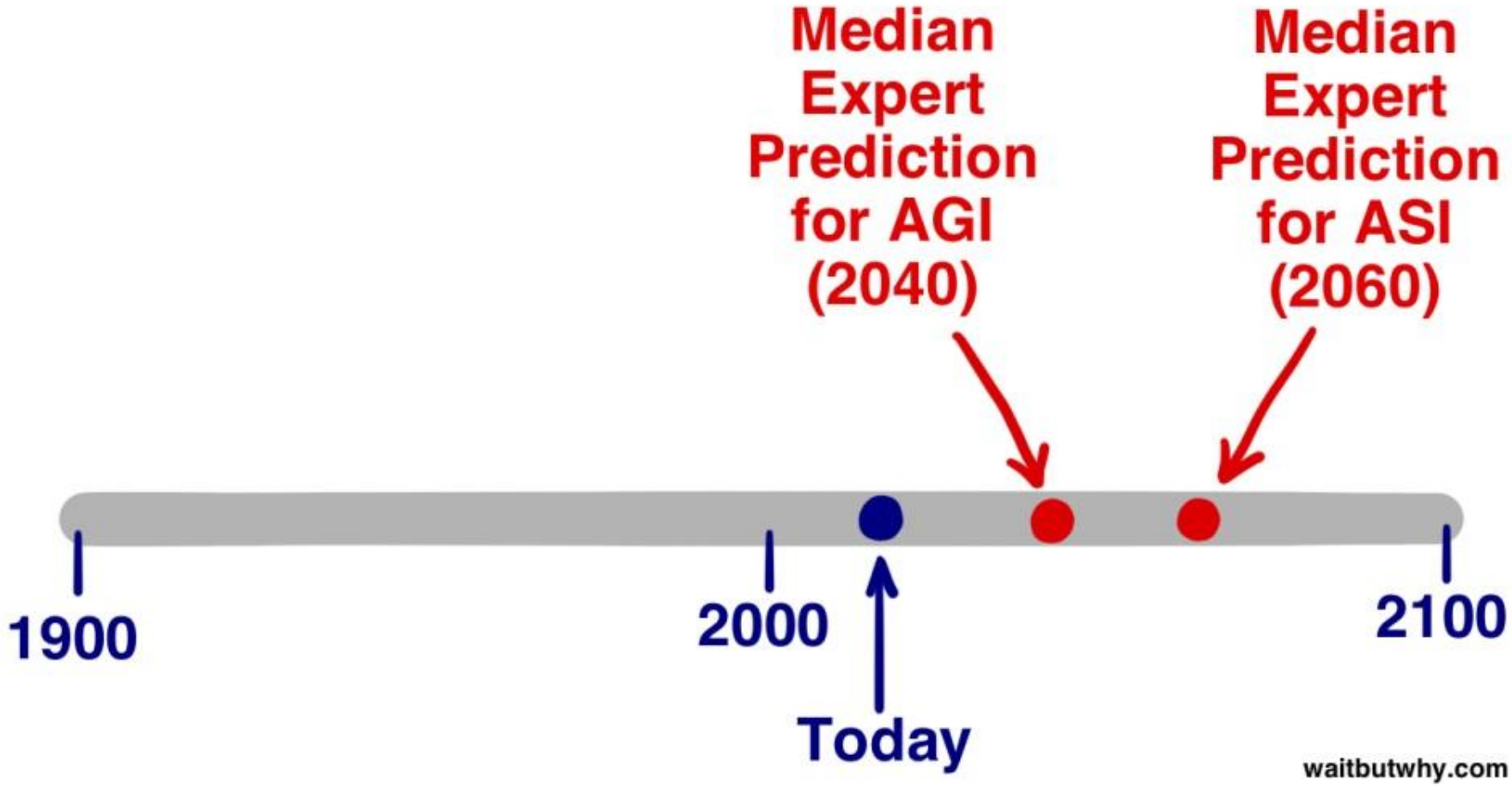
learn quickly, and learn from experience. AGI would be able to do all of those things as easily as you can.

Artificial Superintelligence (ASI): … ??

An oracle, which answers nearly any question posed to it with accuracy, including complex questions that humans cannot easily answer.

A genie, which executes any high-level command it's given, and then awaits its next command.

A sovereign, which is assigned a broad and open-ended pursuit and allowed to operate in the world freely, making its own decisions about how best to proceed.

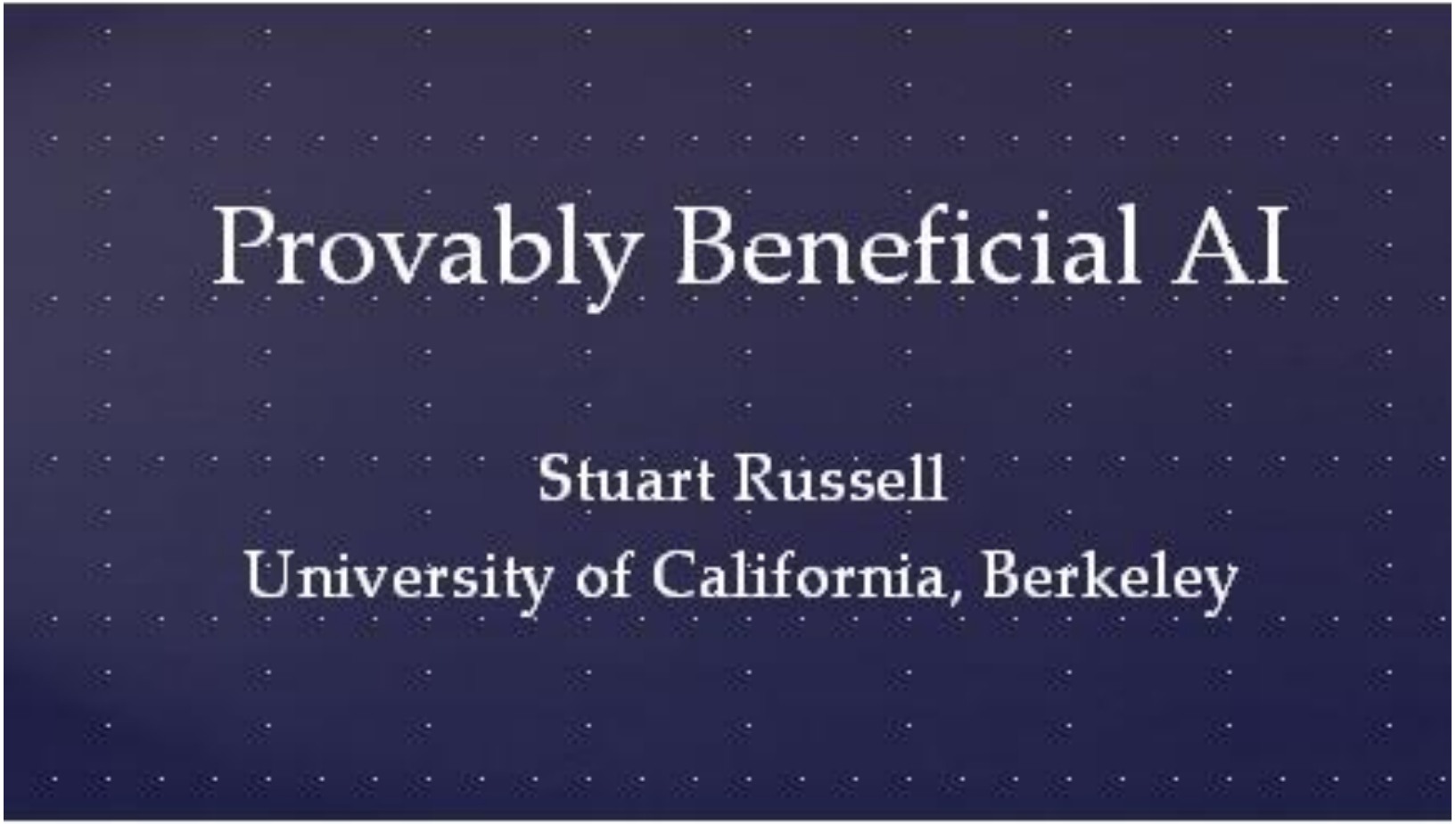

Median Expert Prediction for AGI (2040)
Median Expert Prediction for ASI (2060)
1900
2000
Today
2100
waitbutwhy.com

Technological singularity

a hypothetical point in the future when technological growth becomes uncontrollable and irreversible, resulting in unfathomable changes to human civilization.

According to the most popular version of the singularity hypothesis, called intelligence explosion, an upgradable intelligent agent (such as a computer running software-based artificial general intelligence) would enter a "runaway reaction" of self-improvement cycles, with each new and more intelligent generation appearing more and more rapidly, causing an intelligence explosion and resulting in a powerful superintelligence that would, qualitatively, far surpass all human intelligence.

When? How to prepare? How to avoid? How to control? How to …?


Provably Beneficial AI
Stuart Russell
University of California, Berkeley

(excerpts)

Summary

- Four approaches to AI
- **Rational decisions: autonomous agents**
- Recent applications of AI
- And some questions and outlooks

Further readings, clips, etc.:

DARPA Grand Challenge, https://en.wikipedia.org/wiki/DARPA_Grand_Challenge
DARPA Urban Challenge, http://archive.darpa.mil/grandchallenge/
Bridge Baron, http://www.cs.umd.edu/~nau/bridge/bridge.html
Chinook, http://webdocs.cs.ualberta.ca/~chinook/project/
Robbins algebra, https://www.cs.unm.edu/~mccune/papers/robbins/
Otter: An Automated Deduction System, http://www.cs.unm.edu/~mccune/otter/, http://www.cs.unm.edu/~mccune/prover9/
AARON Cybernetic Artist, http://www.kurzweilcyberart.com/
Google-glasses, http://www.technologyreview.com/featuredstory/532691/google-glass-is-dead-long-live-smart-glasses/
„AI Cited for Unlicensed Practice of Law", http://www.kurzweilai.net/ai-cited-for-unlicensed-practice-of-law
Turing test, http://plato.stanford.edu/entries/turing-test/
The Loebner Prize in Artificial Intelligence, http://www.loebner.net/Prizef/loebner-prize.html
"da Vinci Robot Allegedly Marketed to Less-Skilled Doctors", http://www.prweb.com/releases/2013/5/prweb10710140.htm, http://www.davincisurgery.com/
Automated to Death, http://spectrum.ieee.org/computing/software/automated-to-death
Home robots, http://www.care-o-bot.de/en/care-o-bot-3.html
Remote Agent, http://ti.arc.nasa.gov/tech/asr/planning-and-scheduling/remote-agent/
Titan Mare Explorer, http://www.nasa.gov/pdf/580675main_02_Ellen_Stofan_TiME_.pdf
Robonaut 2, http://robonaut.jsc.nasa.gov/
Test Arenas, http://www.nist.gov/el/isd/testarenas.cfm
Robotic Search and Rescue, http://wiki.robocup.org/wiki/Robot_League
Ambiens Intelligencia (Ambient Intelligence), https://en.wikipedia.org/wiki/Ambient_intelligence
Ambient Assisted Living, http://www2.csd.uoc.gr/hy564/files/material/03/survey on aal tools for older adults.pdf
Care-O-bot 3 informational video, https://www.youtube.com/watch?v=s9CraxEzZLw
Service robots in nursing homes: Care-O-bot 3 and CASERO, https://www.youtube.com/watch?v=dx0zxr3D_zU
https://www.youtube.com/watch?v=ABpOtvLzh2U
First Movement of Robonaut 2 on ISS, https://www.youtube.com/watch?v=glLX_sKTU2I
IBIS PIAP, https://www.youtube.com/watch?v=c-X6tHyNpf8
AI vs. AI. Two chatbots talking to each other, https://www.youtube.com/watch?v=WnzlbyTZsQY
Creepily realistic robot, https://www.youtube.com/watch?v=IhVu2hxm07E
Robot doctors helping Canadians with limited access to health care, https://www.youtube.com/watch?v=D3PMmzIS_Kk
A Day in the Life of a KR[illegible] Robot, https://www.youtube.com/watch?v=6KRjuuEVEZs
NIST Robotics Test Facility, https://www.youtube.com/watch?v=iyEWJszoqZU

https://www.edge.org/conversation/ray_kurzweil-the-singularity
https://futureoflife.org/beneficial-agi-2019/
https://futureoflife.org/background/aimyths/
https://www.youtube.com/watch?v=EBK-a94IFHY
https://www.youtube.com/watch?v=GYQrNfSmQ0M
http://intelligence.org/files/AIPosNegFactor.pdf
https://nickbostrom.com/ethics/artificial-intelligence.pdf
https://futureoflife.org/2019/05/09/state-of-ai/
http://futureoflife.org/AI/open_letter_autonomous_weapons
D.J. Chalmers: The Singularity: A Philosophical Analysis, http://consc.net/papers/singularity.pdf
R. Kurzweil: How to Create a Mind: The Secret of Human Thought Revealed, http://www.amazon.ca/How-Create-Mind-Thought-Revealed/dp/0670025291
The Off-Switch Game, https://people.eecs.berkeley.edu/~dhm/papers/off_switch_AAAI_ws.pdf