The ShanghAI Lectures

人工智能

上海授课

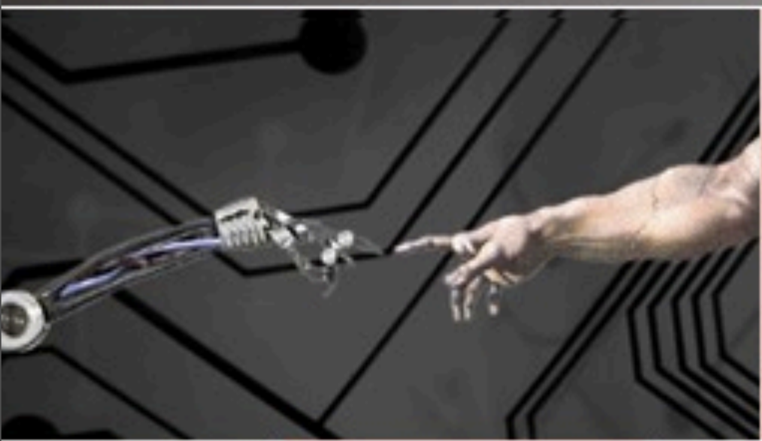
The ShanghAI Lectures
An experiment in global teaching

Rolf Pfeifer
National Competence Center Research in Robotics (NCCR Robotics)
Artificial Intelligence Laboratory
University of Zurich

Today from the University of Sussex, UK
(and other sites)

Brighton, 24 October 2013

EUCOG



European Network for the Advancement of Artificial Cognitive Systems, Interaction and Robotics

HOME | EVENTS › | PROJECT › | MEMBERS › | OUTREACH AREA | EDUCATION ›

Search...

Fourth EUCogIII Conference

- General Information
- Programme
- Conference Venue
- Travel Information
- Event Registration Form
- List of Participants
- Press Section

23-24 October 2013 - Fourth EUCogIII Members Confere Falmer/Brighton

Have a look at the online component of the meeting in Brighton http://robotsandyou.eucognition.org

"Social and Ethical Aspects of Cognitive Systems"

Organizers: Ron Chrisley (UOS) & Vincent C. Müller (Anatolia/ACT)

Local Organizer: Simon Bowes

Venue: University of Sussex, Falmer/Brighton, UK www.sussex.ac.uk.

Description: Cognitive systems research has a growing impact on society that will soon be very substantial s areas: e.g., industry; the workplace environment; transport; the employment economy; entertainment; healthca military; even our conception of ourselves and what it is to be human. However, there are different perceptions what these effects might be, and whether they will be beneficial, or a menace. This members meeting will pro forum for European cognitive systems researchers to discuss the potential social implications of their work. Bu than that, it will engage the wider community in a dialogue about cognitive systems research. How do di sectors of society hope to benefit from cognitive systems? What are their concerns about cognitive s technology? How might those hopes and concerns be used to inform cognitive systems research and the presents itself to society at large?

University of Sussex (geography)

more precisely:
Falmer,
just outside Brighton

Lecture 2

Cognition as computation: Successes and failures

The need for an embodied perspective on intelligence

24 October 2013

Last week's topics

- intelligence
- measuring intelligence: IQ and its issues
- Turing test and the "Chinese Room"
- artificial intelligence
- the synthetic methodology

Today's topics

- The classical approach: Cognition as computation
- Successes and failures of the classical approach
- Some problems of the classical approach
- The need for an embodied approach
- The "frame-of-reference" problem

“Birth” of AI, 1956

George A. Miller, Psychologist
“The Magical Number Seven Plus or Minus Two”

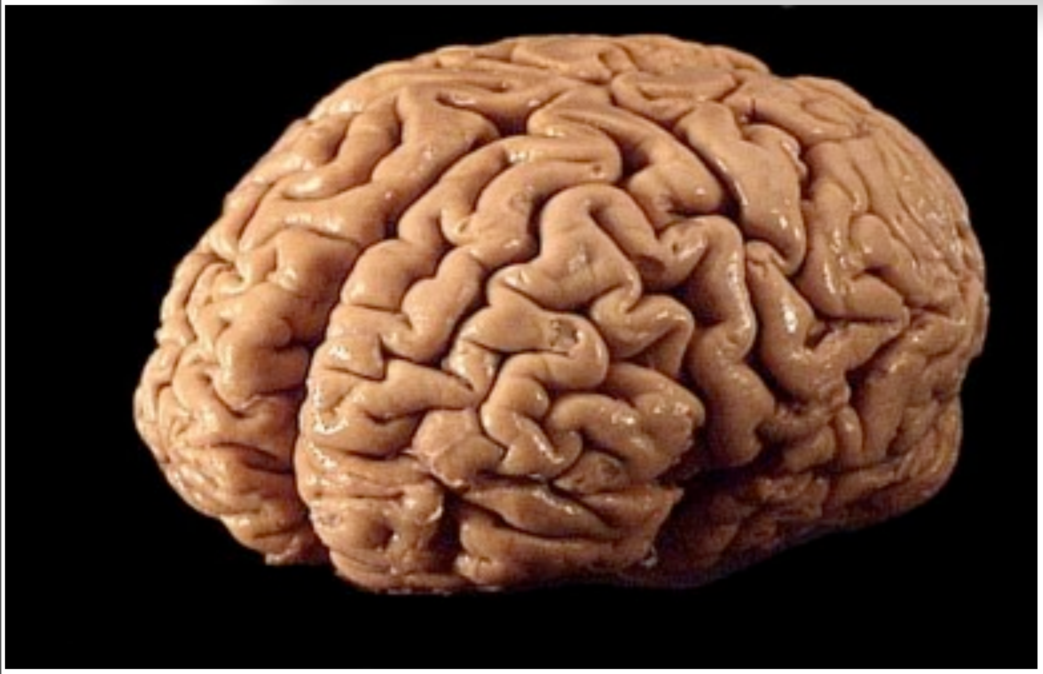
John McCarthy, Computer Scientist
Initiator of Artificial Intelligence

Herbert Simon
and Allen Newell
The “Logic Theorist”

Noam Chomsky, Linguist
“Syntactic Structures”

Conference in 1956: three seminal papers
Newell and Simon: For the first time in history, a thinking machine.
Noam Chomsky: Syntactic structures, a computational approach to language (transformational grammars)
George Miller: Capacity of STM - hard to measure in bits and bytes —> “chunks” (e.g. phone number in Germany: international access code, country code, area code, number).

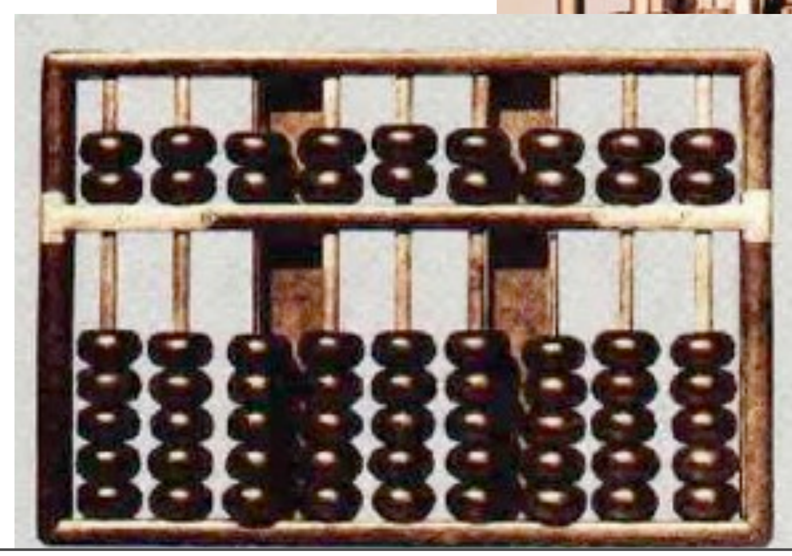
Functionalism and the “Physical Symbol Systems Hypothesis”



biological

electronic

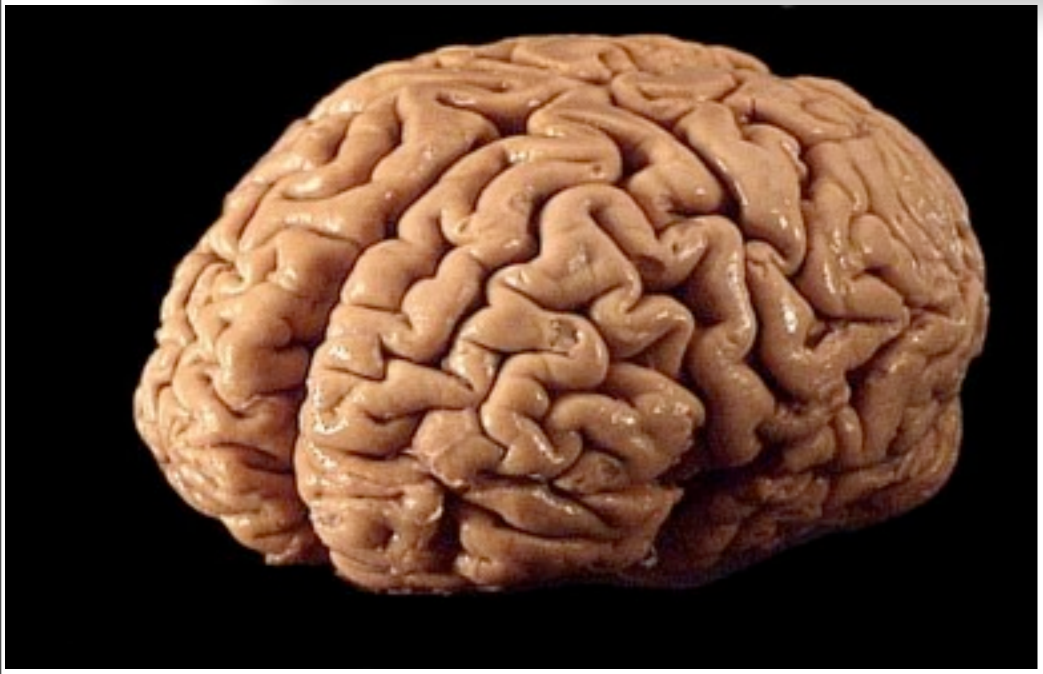
mechanical



Freitag, 1. November 13

The term functionalism is used differently in different contexts. Here we use Putnam’s notion, i.e. the idea that a (mental) process can be implemented in different physical media, mechanical, neural spikes in a biological systems, flip-flops in a computers, or “Swiss cheese”.

Functionalism and the “Physical Symbol Systems Hypothesis”

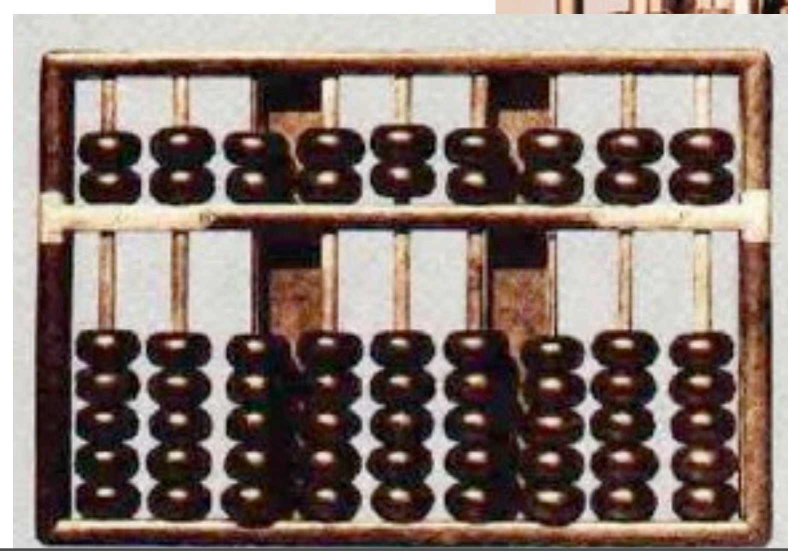


biological

electronic

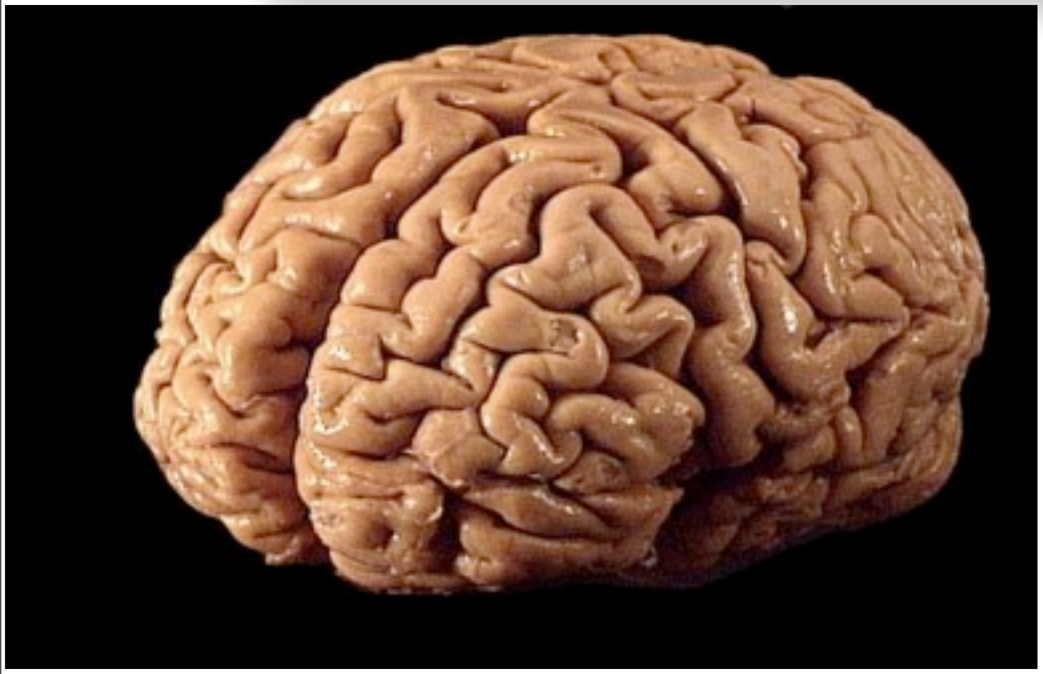
Swiss cheese
Hilary Putnam
(American
Philosopher)

mechanical



The term functionalism is used differently in different contexts. Here we use Putnam’s notion, i.e. the idea that a (mental) process can be implemented in different physical media, mechanical, neural spikes in a biological systems, flip-flops in a computers, or “Swiss cheese”.

Functionalism and the "Physical Symbol Systems Hypothesis"

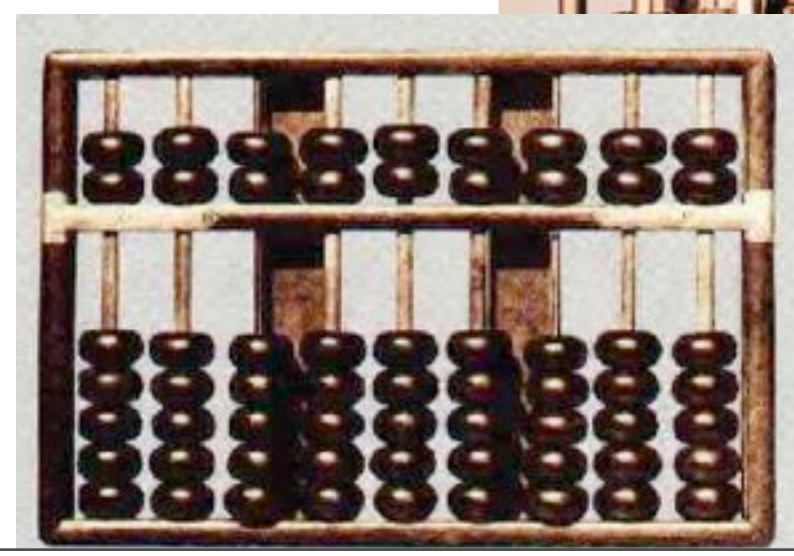


biological

electronic

Swiss cheese
Hilary Putnam
(American
Philosopher)

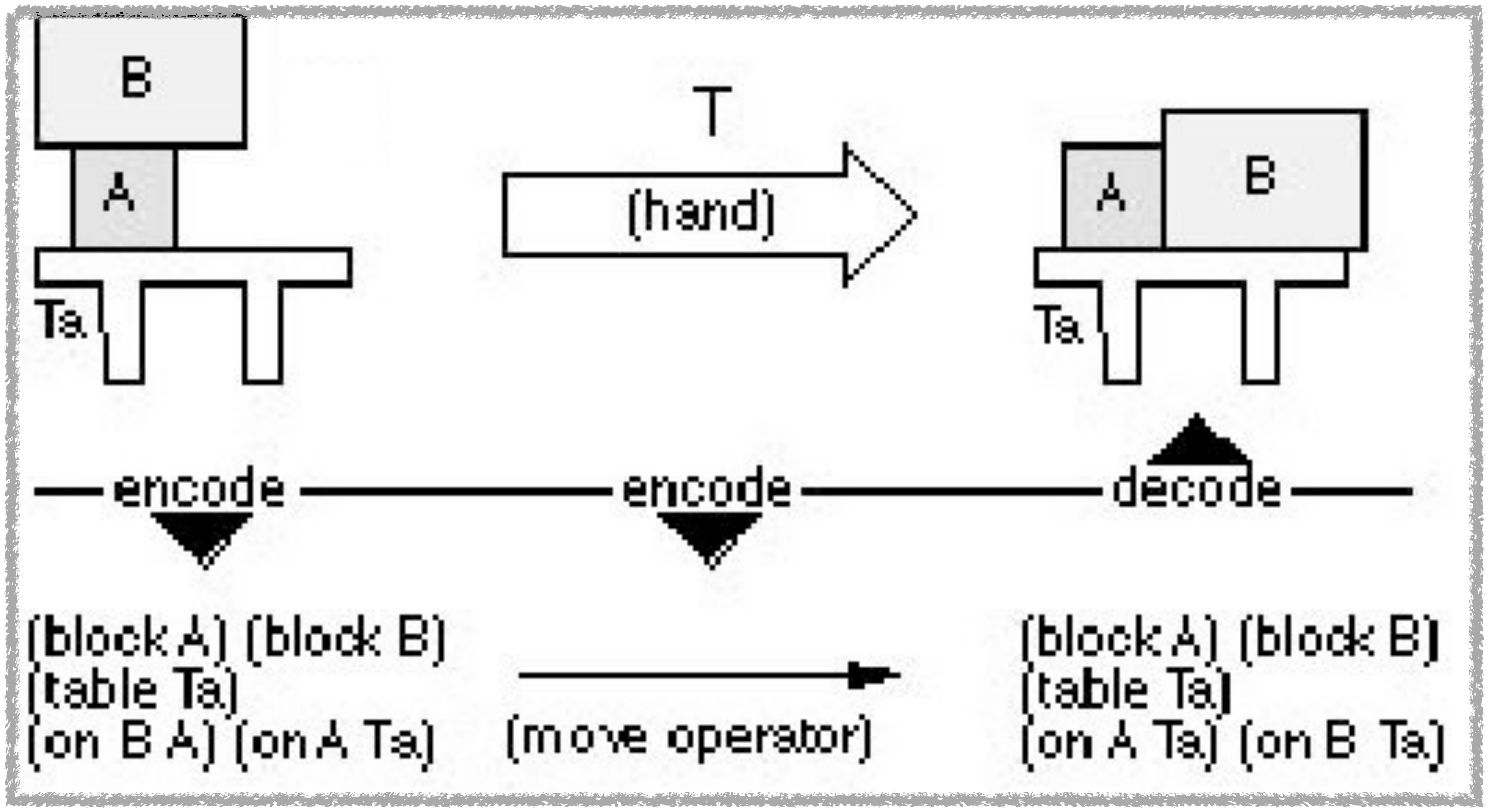
mechanical



The term functionalism is used differently in different contexts. Here we use Putnam's notion, i.e. the idea that a (mental) process can be implemented in different physical media, mechanical, neural spikes in a biological systems, flip-flops in a computers, or "Swiss cheese".

Functionalism and the “Physical Symbol Systems

Model/Representation:



University of Zurich[UZH]

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

Freitag, 1. November 13

The so-called physical symbol systems hypothesis (PSSH) by Allen Newell and Herbert Simon states that for “general intelligent action” it is a necessary and sufficient condition that the system be a “physical symbol system”, i.e., a system that can build and manipulate symbol structures and has a physical implementation (e.g., a brain or a computer).

GOFAI

G

O

F

A

I

University of Zurich[UZH]
robotics Swiss National Centre of Competence in Research
ai lab
人工智能 上海授课 The ShanghAI Lectures

From John Haugeland, “Artificial intelligence — the very idea.” MIT Press, 1993.

GOFAI

G ood

O ld

F ashioned

A rtificial

I ntelligence

From John Haugeland, “Artificial intelligence — the very idea.” MIT Press, 1993.

Classical AI: Research areas

- **problem solving**
- **knowledge representation and reasoning**
- **acting logically**
- **uncertain knowledge and reasoning**
- **learning and memory**
- **communicating, perceiving and acting**

(adapted from Russell/Norvig: Artificial intelligence, a modern approach)

Today's topics

- short recap
- The classical approach: Cognition as computation
- **Successes and failures of the classical approach**
- Some problems of the classical approach
- The need for an embodied approach

University of Zurich[UZH]

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

Classical AI: Successes

- search engines
- formal games (chess!)
- text processing systems/translation —> next week
- data mining systems
- restricted natural language systems
- appliances
- manufacturing

Classical AI: Successes

- search engines
- formal games (chess!)
- text processing systems/translation —> next week
- data mining systems
- restricted natural language systems
- appliances
- manufacturing

Indistinguishable from computer applications in general

Chess: New York, 1997

1 win   3 draws   2 wins

Classical AI: Failures

- recognizing a face in the crowd
- vision/perception in the real world
- common sense
- movement, manipulation of objects
- walking, running, swimming, flying
- speech (everyday natural language)

in general:
more natural forms of intelligence

Why is perception hard?

For humans perception is in fact very easy - it occurs without effort - whereas for artificial systems it has turned out to be a truly hard problem, as researchers tried to build artificial perception systems. They immediately realized that perception is much more than having a camera, a pattern matching algorithm, and a data base: it is a dynamic process requiring the physical interaction of an agent with the real world, a sensory-motor coordination.

Today's topics

- The classical approach: Cognition as computation
- Successes and failures of the classical approach
- **Some problems of the classical approach**
- The need for an embodied approach
- The "frame-of-reference" problem

University of Zurich[UZH]

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能上海授课

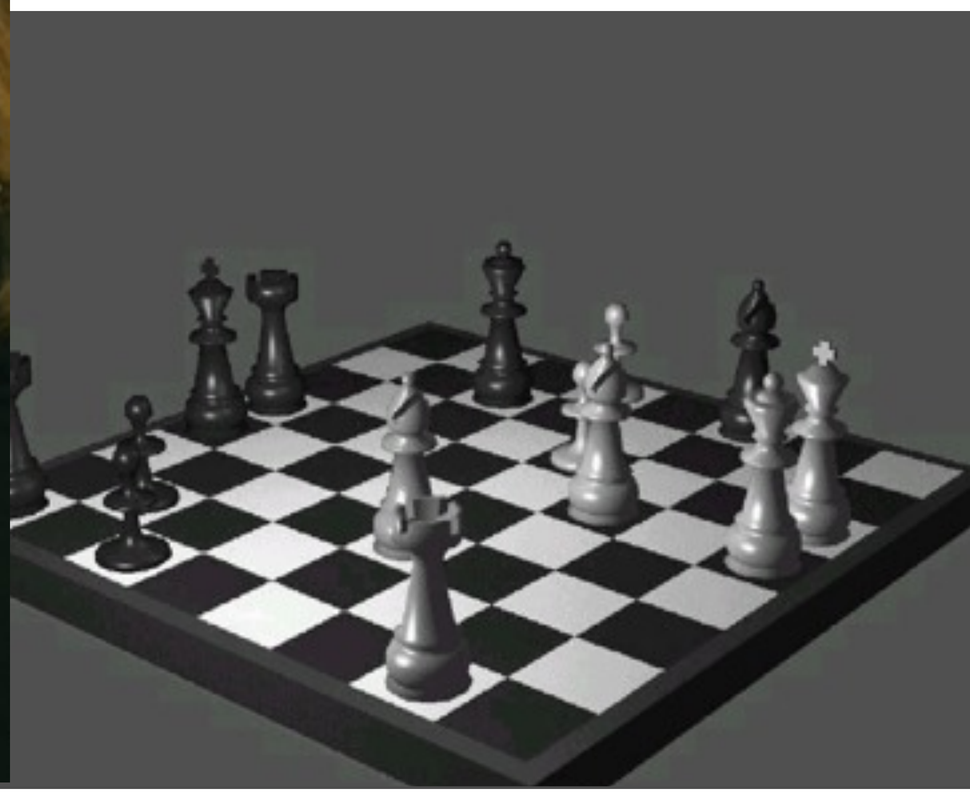
Fundamental problems of the classical approach

Monika Seps, chess master
former master student
AI Lab, Zurich

in general:
anything to do with real world interaction

fundamental differences: real — virtual

virtual, formal world



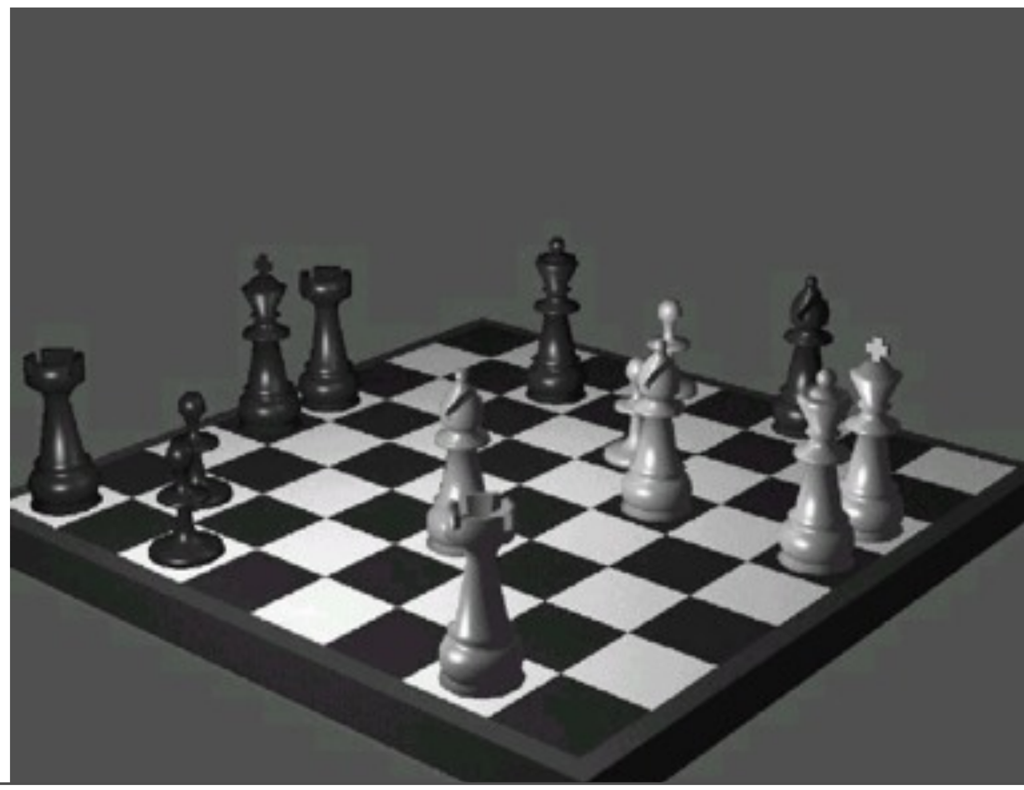
real world

Fundamental problems of the classical approach

in general:
anything to do with real world interaction

fundamental differences: real — virtual

virtual, formal world



real world

Differences
real vs. virtual worlds

examples:
- information acquisition takes time
- only limited information available (high level of uncertainty)
- physical devices subject to noise and malfunction
- no clearly defined, discrete states
- must do several things simultaneously
- real world has own dynamics - things change rapidly
- non-linear, limited predictability

Herbert Simon's “bounded rationality”

Differences
real vs. virtual worlds

examples:
- information acquisition takes time
- only limited information available (high level of uncertainty)
- physical devices subject to noise and malfunction
- no clearly defined, discrete states
- must do several things simultaneously
- real world has own dynamics - things change rapidly
- non-linear, limited predictability

Herbert Simon's "bounded rationality"

examples:
- information acquisition takes time
- only limited information available (high level of uncertainty)
- physical devices subject to noise and malfunction
- no clearly defined, discrete states
- must do several things simultaneously
- real world has own dynamics - things change rapidly
- non-linear, limited predictability
Herbert Simon's "bounded rationality"

Successes and failures of the classical approach

successes

applications (e.g. Google)

chess

manufacturing

(applications:“controlled” artificial worlds)

failures

foundations of behavior

natural forms of intelligence

interaction with real world

(scientific: “real worlds”)

Industrial environments vs. real world

industrial environments	real world environment
environment well-known	**limited knowledge and predictability**
little uncertainty	**rapidly changing**
predictability	**high-level of uncertainty**
(“controlled” artificial worlds)	(“real” worlds)

Industrial robots vs. natural systems

principles:

- **strong, precise, fast motors**
- **centralized control**
- **computing power**
- **optimization**

Industrial robots

Industrial robots vs. natural systems

principles:
- **low precision**
- **compliant**
- **reactive**
- **coping with uncertainty**

humans

→ no direct transfer of methods

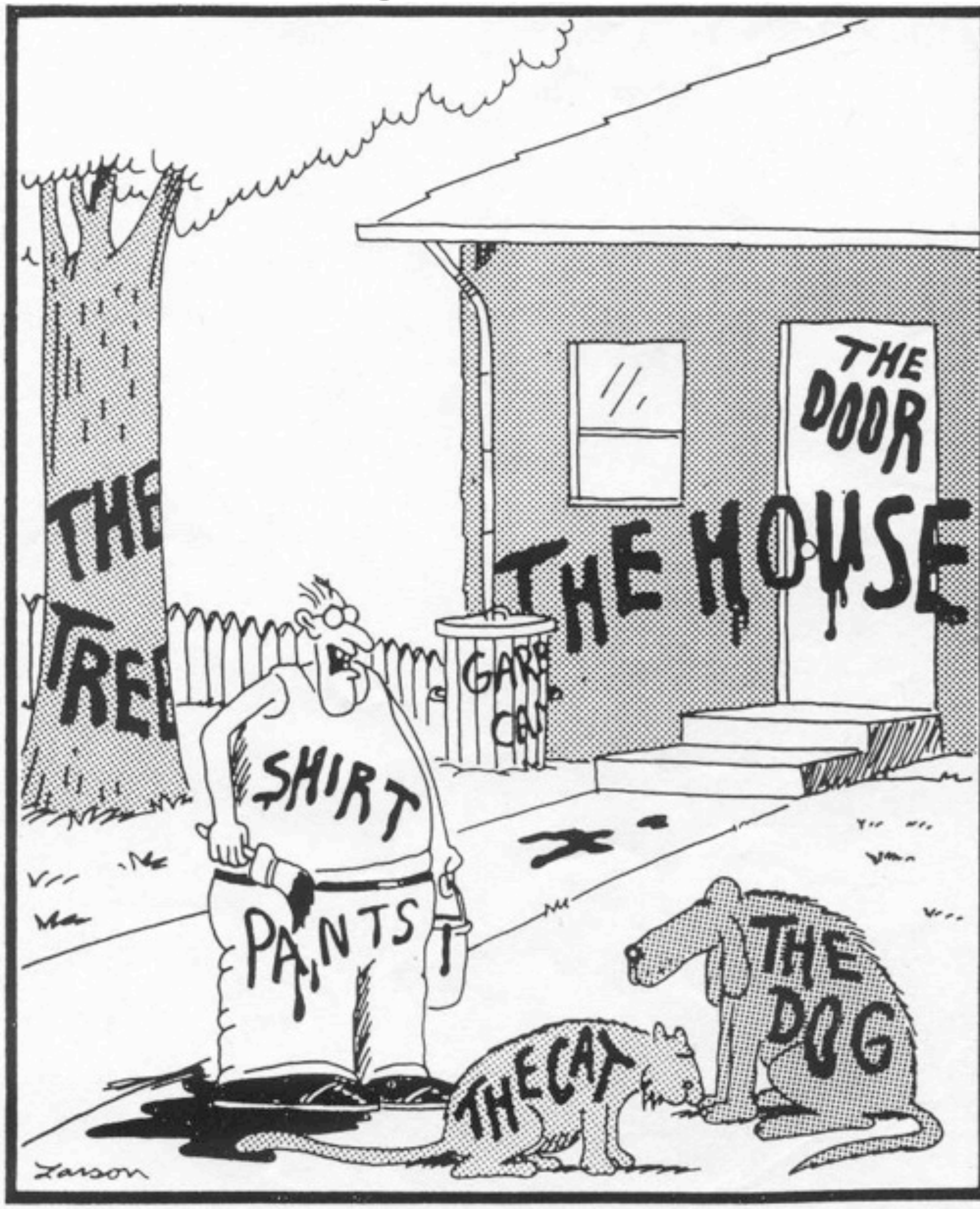
It is important to understand that the people who first built humanoids (the Asimo), the engineers at Honda, came from a background of control engineering. These methods work extremely well in an industrial environment, but not in the "real" world, where everything changes quickly and there is a high degree of uncertainty.

Fundamental problems of classical approach

- "symbol grounding problem"
- "frame problem"
- "homunculus problem"

The “symbol grounding” problem

real world: doesn’t come with labels ...



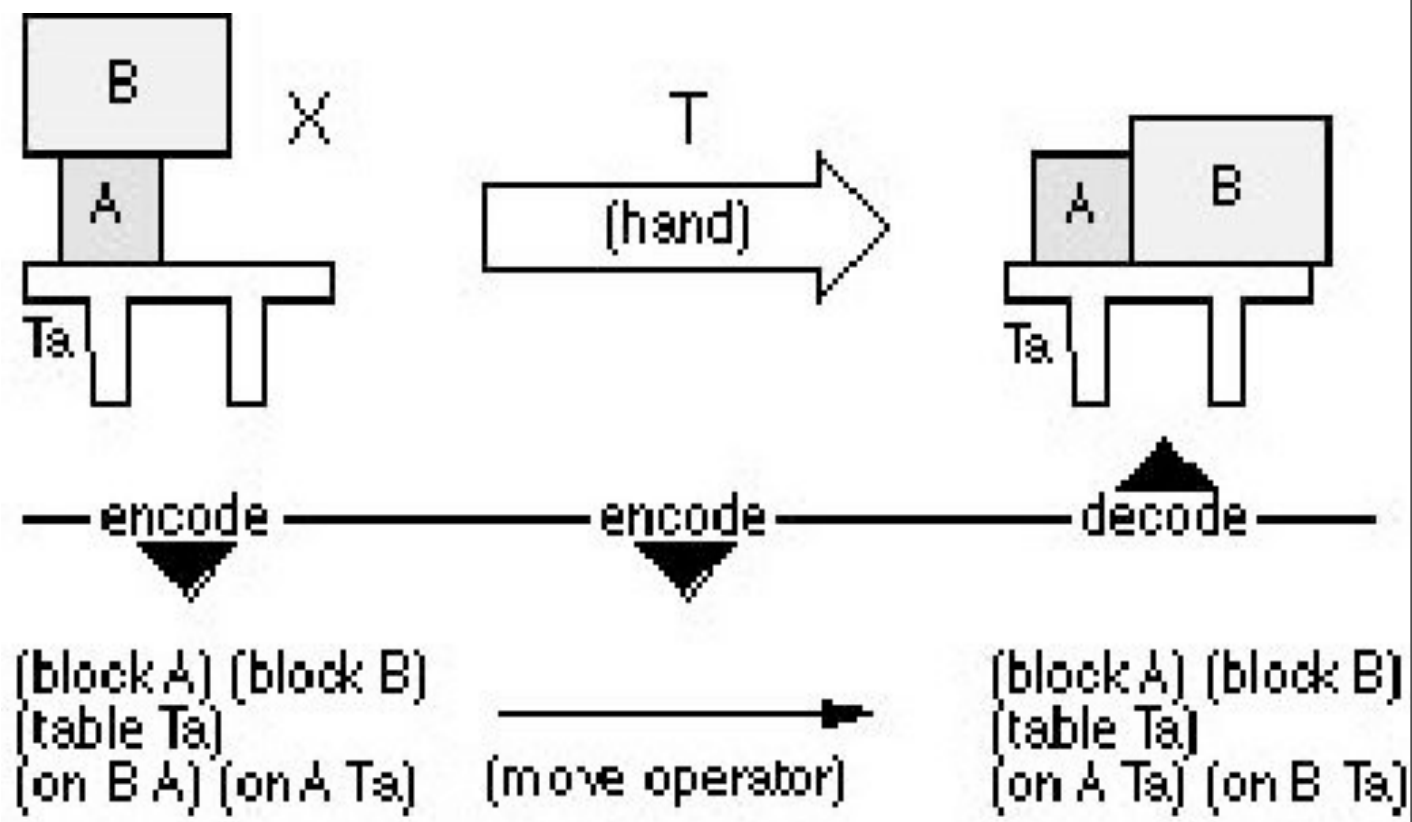
“Now! ... *That* should clear up a few things around here!”

Gary Larson

Gary Larsen’s point is that the world doesn’t come with labels. How can we associate symbols with a continuous stream of sensory stimulation?

The "frame problem": Maintaining model of real world

- **the more detailed the harder**
- **information acquisition**
- **most changes: irrelevant to current situation**



Freitag, 1. November 13

- if you leave the room, the color of the walls doesn't change
- the fact that there are millions of cars moving through the cities is at this moment irrelevant to our lecture
- however, if you have a both a battery and a bomb on a cart, and you move the cart out of the room to get rid of the bomb, the battery will move along with it; please read Dan Dennett's story about this.

The “frame problem”

Daniel Dennett, American philosopher (philosophy of mind)

R1: robot

R1D1: robot deducer

R2D1: robot relevant

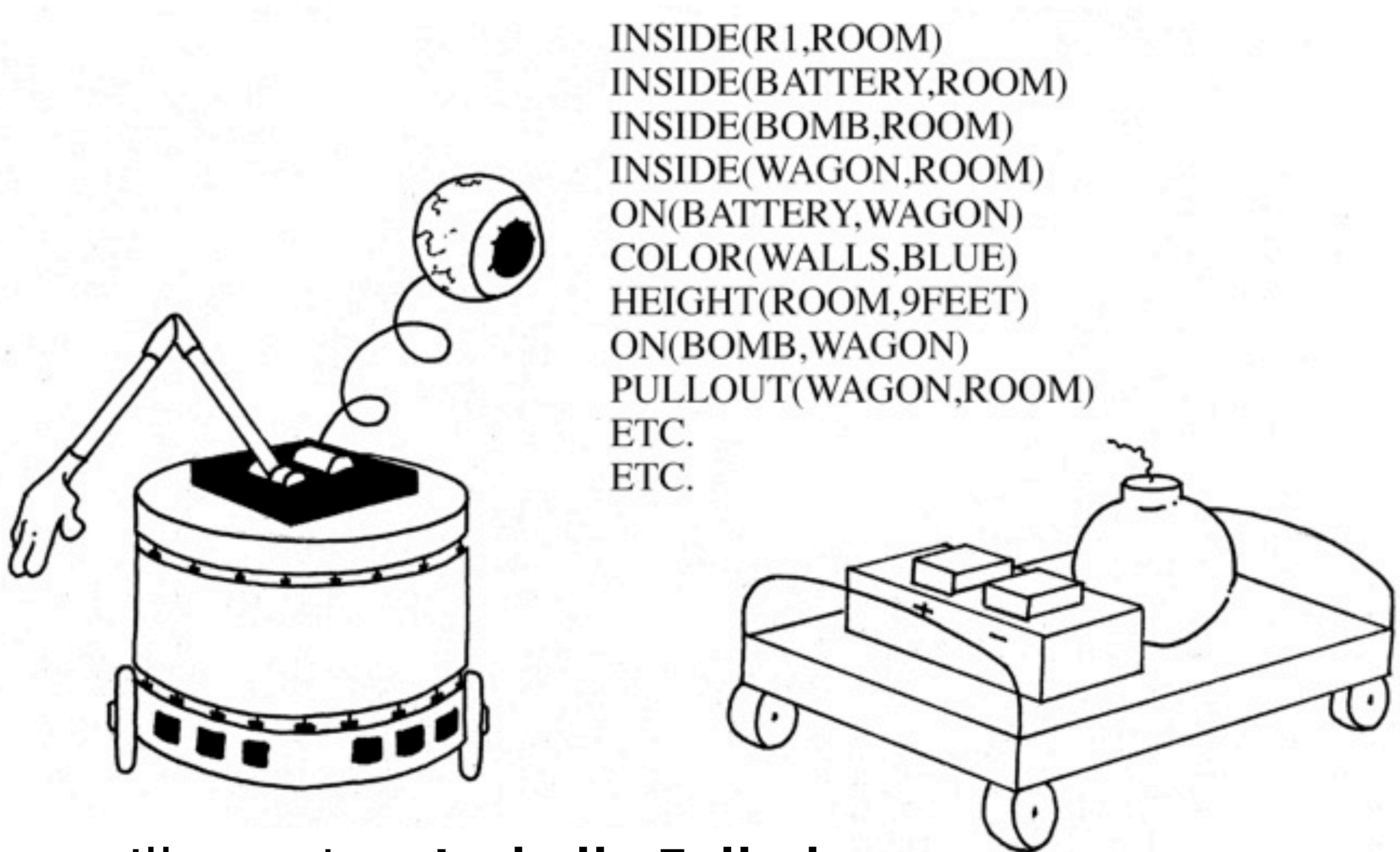


Illustration: **Isabelle Follath**

Once upon a time there was a robot, named R1 by its creators. Its only task was to fend for itself. One day its designers arranged for it to learn that its spare battery, its precious energy supply, was locked in a room with a time bomb set to go off soon.

R1 located the room, and the key to the door, and formulated a plan to rescue its battery. There was a wagon in the room, and the battery was on the wagon, and R1 hypothesized that a certain action which it called PULLOUT(WAGON, ROOM) would result in the battery removed from the room. Straightaway it acted, and did succeed in getting the battery out of the room before the bomb went off. Unfortunately, however, the bomb was also on the wagon. R1 knew that the bomb was on the wagon in the room, but didn't realize that pulling the wagon would bring the bomb out along with the battery. Poor R1 had missed that obvious implication of its planned act.

The “frame problem” (1)

Once upon a time there was a robot, named R1 by its creators. Its only task was to fend for itself. One day its designers arranged for it to learn that its spare battery, its precious energy supply, was locked in a room with a time bomb set to go off soon.

R1 located the room, and the key to the door, and formulated a plan to rescue its battery. There was a wagon in the room, and the battery was on the wagon, and R1 hypothesized that a certain action which it called PULLOUT(WAGON, ROOM) would result in the battery removed from the room. Straightaway it acted, and did succeed in getting the battery out of the room before the bomb went off. Unfortunately, however, the bomb was also on the wagon. R1 knew that the bomb was on the wagon in the room, but didn't realize that pulling the wagon would bring the bomb out along with the battery. Poor R1 had missed that obvious implication of its planned act.

Once upon a time there was a robot, named R1 by its creators. Its only task was to fend for itself. One day its designers arranged for it to learn that its spare battery, its precious energy supply, was locked in a room with a time bomb set to go off soon.

R1 located the room, and the key to the door, and formulated a plan to rescue its battery. There was a wagon in the room, and the battery was on the wagon, and R1 hypothesized that a certain action which it called PULLOUT(WAGON, ROOM) would result in the battery removed from the room. Straightaway it acted, and did succeed in getting the battery out of the room before the bomb went off. Unfortunately, however, the bomb was also on the wagon. R1 knew that the bomb was on the wagon in the room, but didn't realize that pulling the wagon would bring the bomb out along with the battery. Poor R1 had missed that obvious implication of its planned act.

The “frame problem” (2)

Back to the drawing board. ``The solution is obvious,'' said the designers. ``Our next robot must be made to recognize not just the intended implications of its acts, but also the implications about their side-effects, by deducing these implications from the descriptions it uses in formulating its plans.'' They called their next model, the robot-deducer, R1D1. They placed R1D1, in much the same predicament that R1 had succumbed to, and as it too hit upon the idea of PULLOUT(WAGON, ROOM) it began, as designed, to consider the implications of such a course of action. It had just finished deducing that pulling the wagon out of the room would not change the color of the room's walls, and was embarking on a proof of the further implication that pulling the wagon out would cause its wheels to turn more revolutions than there were wheels on the wagon---when the bomb exploded.

Once upon a time there was a robot, named R1 by its creators. Its only task was to fend for itself. One day its designers arranged for it to learn that its spare battery, its precious energy supply, was locked in a room with a time bomb set to go off soon.

R1 located the room, and the key to the door, and formulated a plan to rescue its battery. There was a wagon in the room, and the battery was on the wagon, and R1 hypothesized that a certain action which it called PULLOUT(WAGON, ROOM) would result in the battery removed from the room. Straightaway it acted, and did succeed in getting the battery out of the room before the bomb went off. Unfortunately, however, the bomb was also on the wagon. R1 knew that the bomb was on the wagon in the room, but didn't realize that pulling the wagon would bring the bomb out along with the battery. Poor R1 had missed that obvious implication of its planned act.

The “frame problem” (3)

Back to the drawing board. ``We must teach it the difference between relevant implications and irrelevant implications,'' said the designers, ``and teach it to ignore the irrelevant ones.'' So they developed a method of tagging implications as either relevant or irrelevant to the project at hand, and installed the method in their next model, the robot-relevant-deducer, R2D1 for short. When they subjected R2D1 to the test that had so unequivocally selected its ancestors for extinction, they were surprised to see it sitting, Hamlet-like, outside the room containing the ticking bomb ... ``Do something!'' they yelled at it. ``I am,'' it retorted. ``I'm busily ignoring some thousands of implications I have determined to be irrelevant. Just as soon as I find an irrelevant implication, I put it on the list of those I must ignore, and ...'' the bomb went off.

(from Dennett, D. (1987). Cognitive wheels: the frame problem in AI. In C. Hookway (Ed.), Minds, Machines, and Evolution: Philosophical Studies. Bantain: Bean Books.)

Once upon a time there was a robot, named R1 by its creators. Its only task was to fend for itself. One day its designers arranged for it to learn that its spare battery, its precious energy supply, was locked in a room with a time bomb set to go off soon.

R1 located the room, and the key to the door, and formulated a plan to rescue its battery. There was a wagon in the room, and the battery was on the wagon, and R1 hypothesized that a certain action which it called PULLOUT(WAGON, ROOM) would result in the battery removed from the room. Straightaway it acted, and did succeed in getting the battery out of the room before the bomb went off. Unfortunately, however, the bomb was also on the wagon. R1 knew that the bomb was on the wagon in the room, but didn't realize that pulling the wagon would bring the bomb out along with the battery. Poor R1 had missed that obvious implication of its planned act.

Summary of Dennett's points

- obvious to humans, not obvious to robots (robot only has symbolic model/representation of world)
- vast number of potential side effects, mostly irrelevant
- distinction between relevant and irrelevant inferences → must test all

Today's topics

- short recap
- The classical approach: Cognition as computation
- Successes and failures of the classical approach
- Some problems of the classical approach
- **The need for an embodied approach**

University of Zurich UZH

robotics Swiss National Centre of Competence in Research

ai lab

Two views of intelligence

classical:
cognition as computation

embodiment:
cognition emergent from sensory-motor and interaction processes

Freitag, 1. November 13

Cognition as computation: many problems, e.g.
- perception
- manipulation
- movement
- locomotion (walking, running)

has lead to the problems mentioned (symbol grounding, frame problem, etc.).
In general: neglect of interaction with real world

The need for an embodied perspective

- "failures" of classical AI
- fundamental problems of classical approach
- Wolpert's quote:

Are we just doing something different now? Movement, locomotion? Where is the "thinking", the "intelligence"? — > Wolpert's quote

The need for an embodied perspective

"Why do plants not have brains?"

The need for an embodied perspective

“Why do plants not have brains? The answer is actually quite simple — they don’t have to move.” Lewis Wolpert, UCL

evolutionary perspective on development of intelligence/cognition

University of Zurich[UZH]

robotics Swiss National Centre of Competence in Research

ai lab

人工智能 上海授课 The ShanghAI Lectures

Evolutionary selectionist pressure on development of brain: from the need to move, to locomote, and to orient in space.
And the evolution of the brain (and what we now call cognition or intelligence) has always taken place as part of a complete organism that had to interaction, survive, and reproduce in the real world. In other words, the brain and the body have co-evolved.

The need for an embodied perspective

- **“failures” of classical AI**
- **fundamental problems of classical approach**
- **Wolpert’s quote: Why do plants not …?**
- **Interaction with environment: always mediated by body**
- **brain/neural system: always part of complete organism**

Today’s topics

- short recap
- The classical approach: Cognition as computation
- Successes and failures of the classical approach
- Some problems of the classical approach
- The need for an embodied approach
- **The “frame-of-reference” problem**

The "frame-of-reference" problem — introduction

Video "Heider and Simmel"

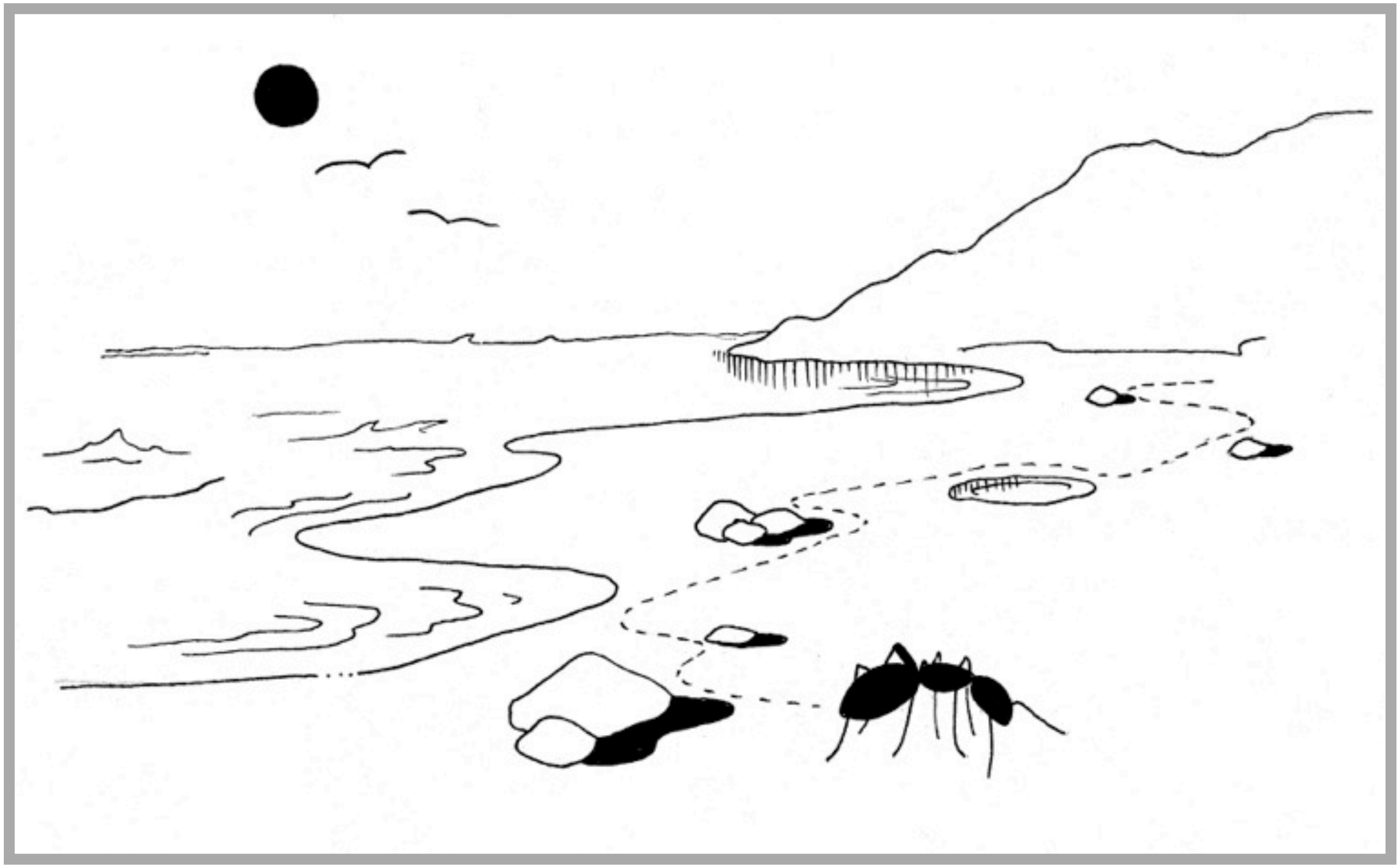
More than sixty years ago (1944) the psychologists Fritz Heider and Marianne Simmel conducted an experimental study, which can be seen as the starting point of attribution theory research. Moving symbols were shown during a short animated cartoon which subjects unanimously described as living objects (mostly people).
"Anthropomorphization, the incurable disease", David McFarland, Oxford University

The "frame-of-reference" problem — introduction

Video "Heider and Simmel"

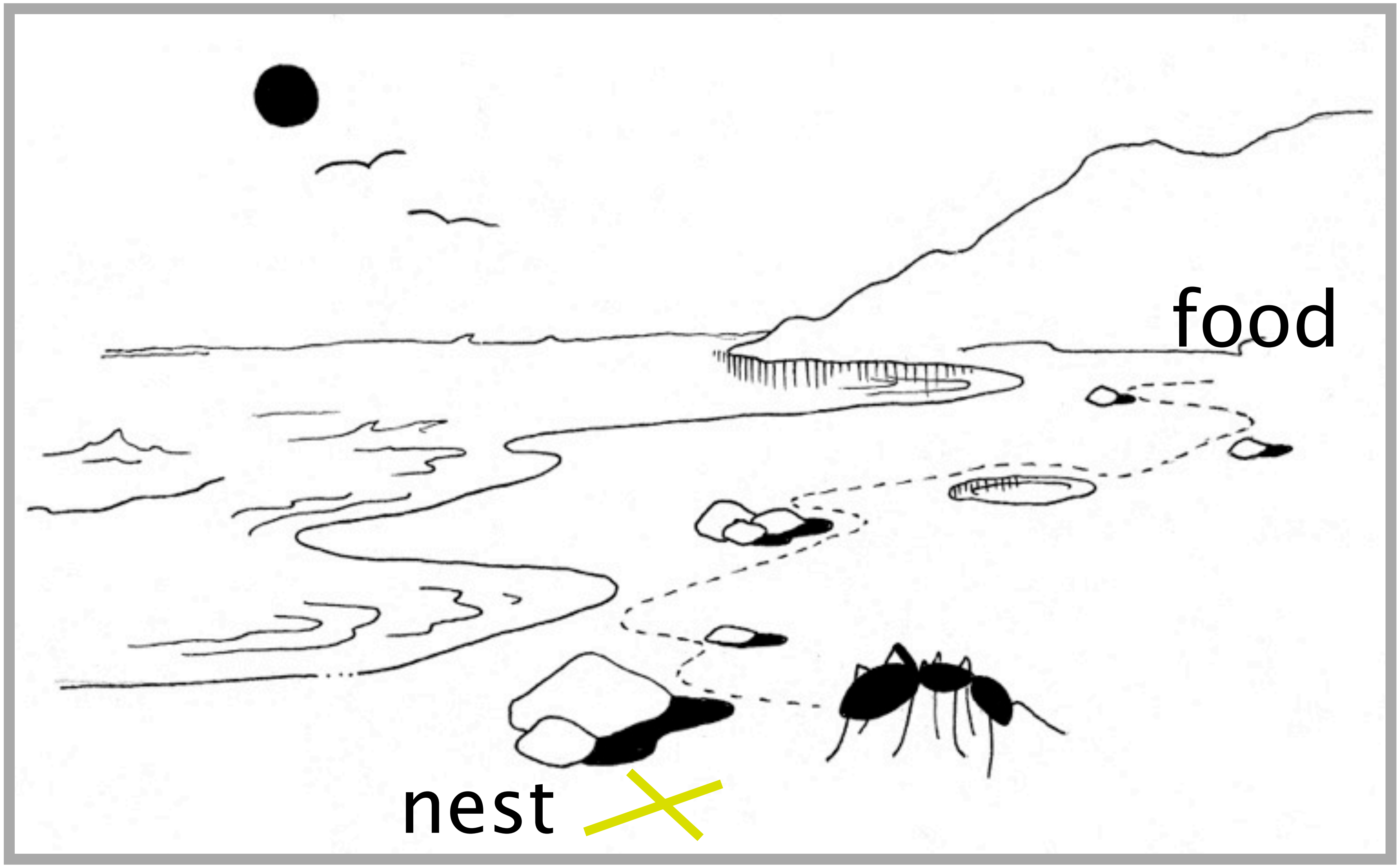
comment on video?

“Frame-of-reference” Simon’s ant on the beach

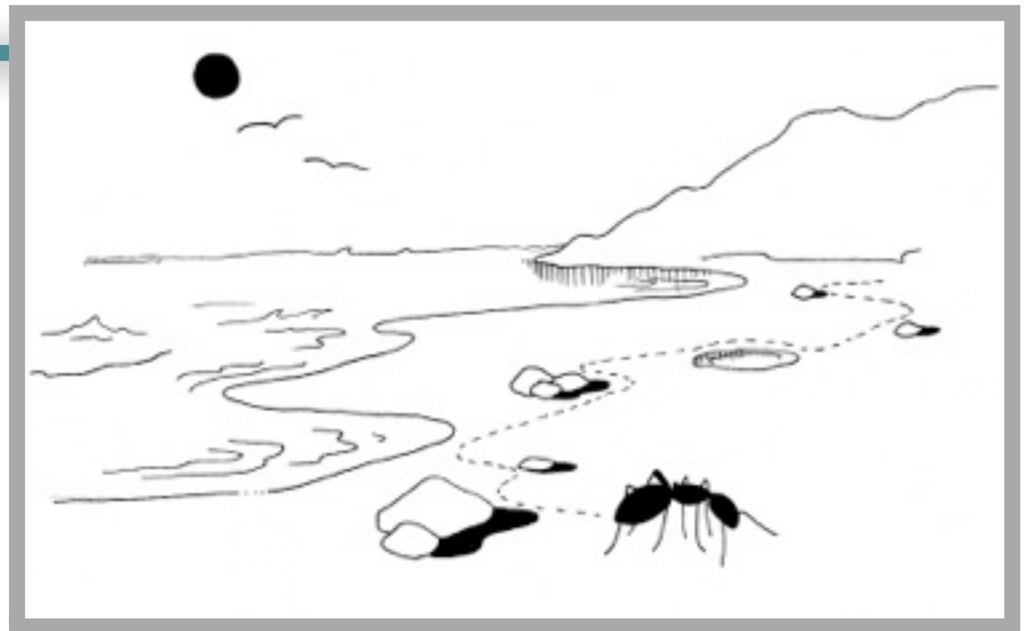


From Herbert A. Simon: The sciences of the artificial. MIT Press.

"Frame-of-reference" Simon's ant on the beach



“Frame-of-reference”
Simon’s ant on the beach

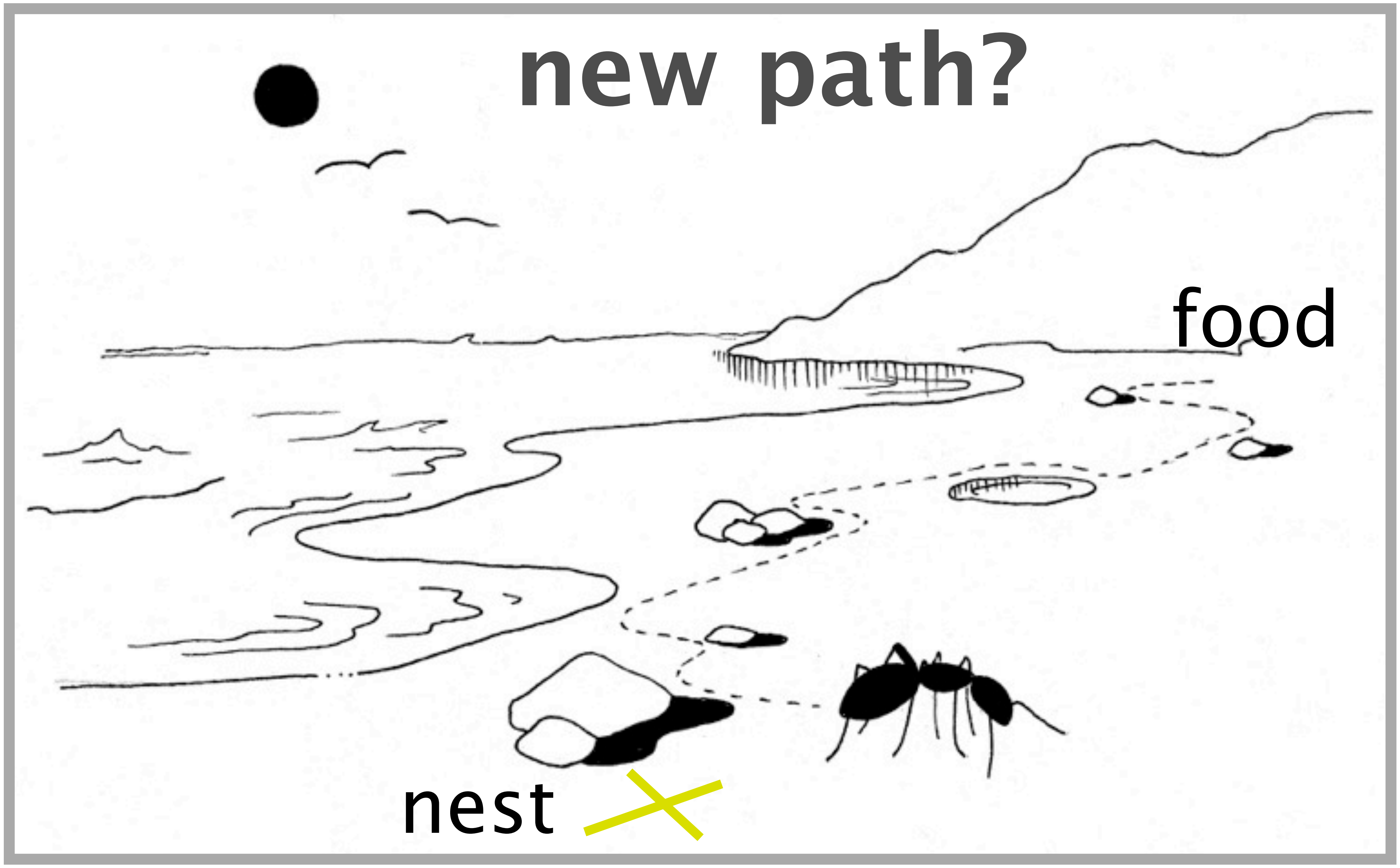


- **simple behavioral rules**
- **complexity in interaction,**
  not — necessarily — in brain

- thought experiment:
  increase body by factor of 1000
  everything else the same

Simple behavioral rules, e.g.:
- obstacle on left, turn right
- obstacle on right, turn left
- otherwise go straight

“Frame-of-reference” Simon’s ant on the beach



New path? more or less straight line.

“Frame-of-reference” F-O-R

- **perspectives issue**
- **behavior vs. mechanism issue**
- **complexity issue**

F-O-R competition
Will Fabio continue the contest?

?

Swiss Chocolate

Champagne

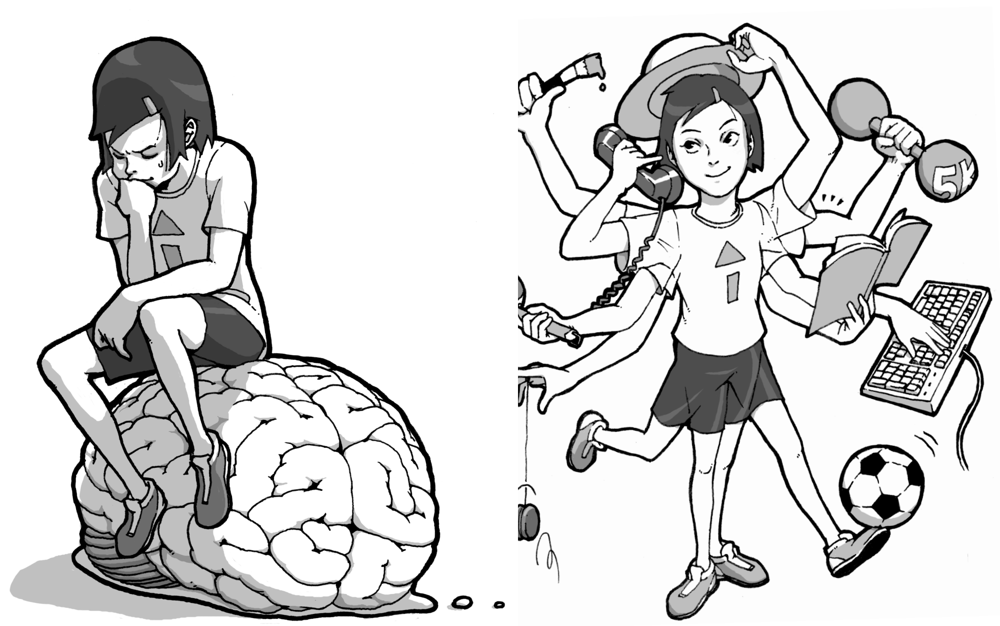
Freitag, 1. November 13

Should there be a day in the ShanghAI lectures where the FOR problem is not mentioned, the first to discover will get either a box of Swiss chocolate or a bottle of champagne. If there is one idea from the class that everyone should remember, it's the FOR problem.

End of lecture 2

Thank you for your attention!

back to Fabio Bonsignorio

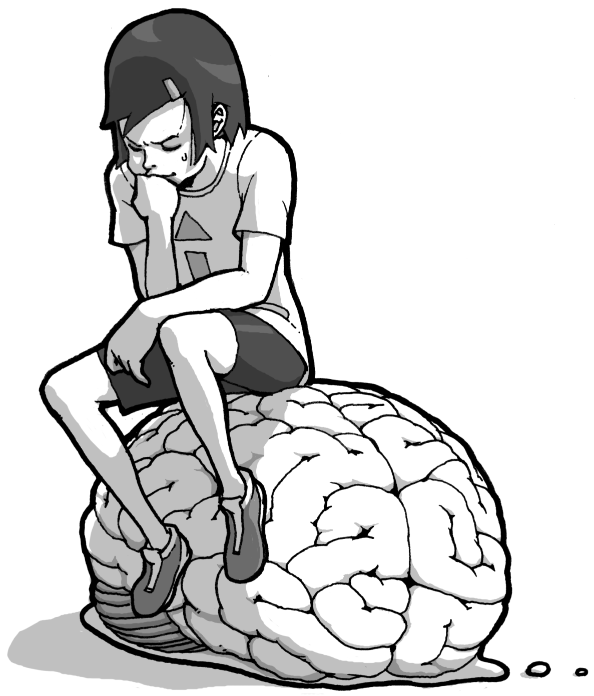


End of lecture 2

Thank you for your attention!

stay tuned for lecture 3

“Towards a theory of intelligence”



University of Zurich[UZH]

Additional slide materials for self-study

The "homunculus problem"

"Homunculus" literally means "little man"; as used here, it designates a "little man in the head."

The homunculus problem, or the homunculus fallacy, as it is also called, refers to circular accounts of psychological processes. These processes are circular because they ascribe to some internal mechanism (the homunculus) the very psychological properties being investigated in the first place. For example, a theory of vision might postulate that there is within the brain a mechanism - the "homunculus" that scans, views, or inspects images on the retina. Such a theory would be vacuous, however, since scanning, viewing, and inspecting are all instances of the very visual processes the theory was supposed to illuminate in the first place (Gregory 1987, p. 313). In other words, the theory has assumed the very things it set out to explain.

Gregory, R.L (1987). The Oxford companion to the mind. Oxford, UK: Oxford University Press.

Problems to think about: Meaning?

Imagine that you are using an electronic train schedule to figure out a train from Shanghai to Beijing. What does the program know about trains and train schedules? What is the input to the program and what is its output? How come, you will finally end up in the right train? Where does the meaning come from?

What happens if you take the human out of the loop (in which case you have an autonomous system)?

End of additional slide materials for self-study