



# **The ShanghAI Lectures**

## **An experiment in global teaching**

Rolf Pfeifer

Today from Shanghai Jiao Tong University, China

上海交通大学

Shanghai, 23 October 2014



# Shanghai - Geography



# Shanghai - Geography



# Shanghai - Geography





# Lecture 2

---

**Cognition as computation:  
Successes and failures**

**The need for an embodied perspective on  
intelligence**

**23 October 2014**



# 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

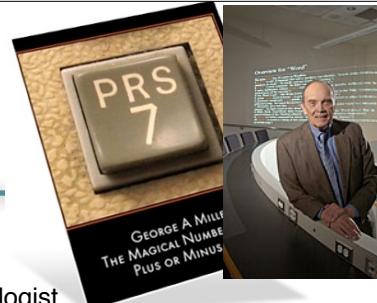
# “Birth” of AI, 1956



Herbert Simon  
and Allen Newell  
The “Logic Theorist”



Noam Chomsky, Linguist  
“Syntactic Structures”



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

John McCarthy, Computer Scientist  
Initiator of Artificial Intelligence



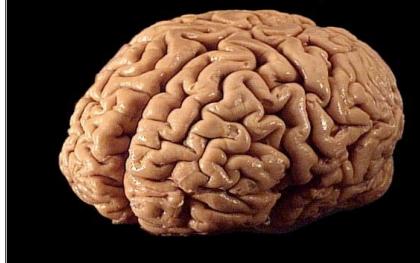
Conference in 1956: three seminal papers

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”



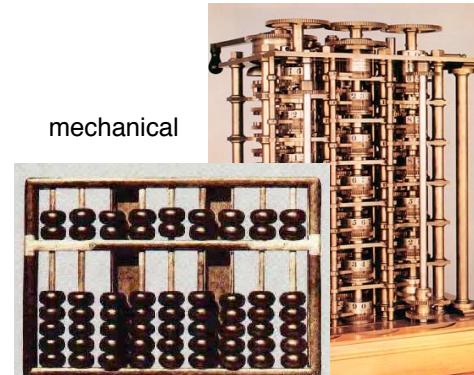
biological



electronic



Swiss cheese  
Hilary Putnam  
(American  
Philosopher)

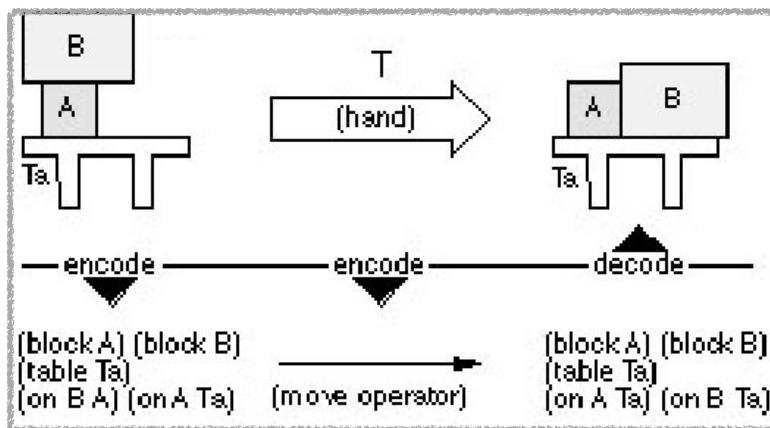


mechanical

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”

## Model/Representation:



11

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

**(John Haugeland, Philosopher)**

12

From John Haugeland, "Artificial intelligence – the very idea." MIT Press, 1993.



# GOFAI

---

**G ood  
O ld  
F ashioned  
A rtificial  
I ntelligence  
(John Haugeland, Philosopher)**

13

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
- The “frame-of-reference” problem



# Classical AI: Successes

---

- search engines
- formal games (chess!)
- machine learning
- data mining systems - “big data”
- restricted natural language systems
- manufacturing, control systems

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

18

The classical approach has not performed well on those tasks that are easy and effortless for humans such as walking, talking (conversation), perception, manipulation, reacting to rapid changes in the real-world environments, etc. It has been extremely successful at those tasks that are generally considered hard for humans such as abstract problem solving, logic, analysing large amounts of data, repetitive tasks, etc.



# Why is perception hard?

---

19

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.

Take an object, a pencil, a bottle, or a mobile phone, turn it, view it from different perspectives and distances, and under different lighting conditions.



# 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

# Fundamental problems of the classical approach



Monika Seps, chess master  
former master student  
AI Lab, Zurich, 2012



Photo: Rüdiger (c) www.kosterle.com

in general:  
anything to do with real world interaction

**fundamental differences: real –  
virtual**

virtual, formal world

real world



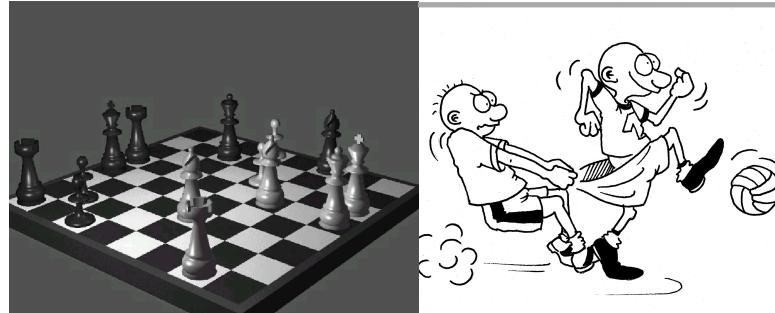
# Fundamental problems of the classical approach

in general:  
anything to do with real world interaction

**fundamental differences: real – virtual**

virtual, formal world

real world



# Differences real vs. virtual worlds

---



23

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"

# Successes and failures of the classical approach

successes

**search engines**

**chess**

**manufacturing**

**“Big Data”**

(applications:  
“controlled”artificial worlds)

failures

**foundations of behavior**

**natural forms of  
intelligence**

**interaction with real world**

(scientific: “real worlds”)

# industrial environment



- high predictability
- programmability



industrial robots (“hard”)

# real-world environment

- low predictability
- coping with high levels of uncertainty



humans (“soft”)



ai lab

26

electrical motors, metal, plastic

humans: skin, tissue, passively deformable, etc.



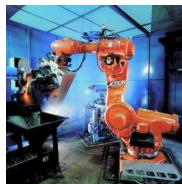
University of  
Zurich

robotics<sup>o</sup>  
Swiss National Institute of Technology in Research

# industrial environment



- high predictability
- programmability



industrial robots (“hard”)

# real-world environment

- low predictability
- coping with high levels of uncertainty



humans (“soft”)



ai lab

27

humans: 85% “soft”



University of  
Zurich



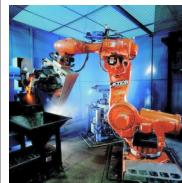
electrical motors, metal, plastic

humans: skin, tissue, passively deformable, etc.

# industrial environment



- high predictability
- programmability



industrial robots (“hard”)

—> no direct transfer of methods

# real-world environment

- low predictability
- coping with high levels of uncertainty



humans (“soft”)



ai lab

28

Industrial environment:  
electrical motors, metal, plastic  
principles:

- high predictability and programmability
- strong, precise, fast motors
- centralized control
- high computing power requirements
- optimization

humans: skin, tissue, passively deformable, etc.  
principles:

- low predictability
- high levels of uncertainty
- low precision
- compliance
- reactive



# Fundamental problems of classical approach

- “symbol grounding problem”
- “frame problem”
- “homunculus problem”

29

These fundamental issues are entailed by the classical symbol-processing approach to intelligence.

# The “symbol grounding” problem

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

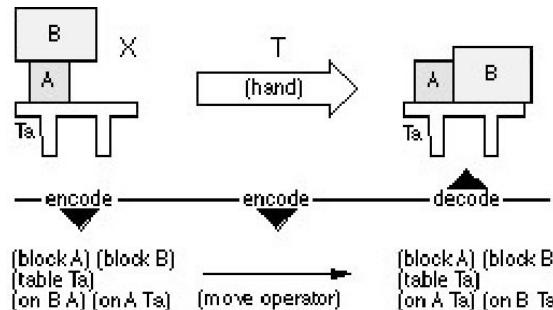
Gary Larson



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

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



31

- 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 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.

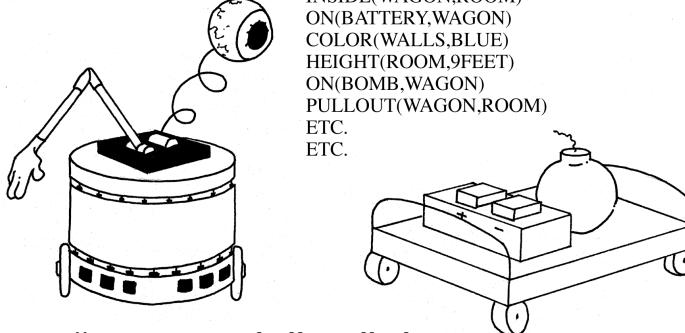
# The “frame problem”

**Daniel Dennett, American philosopher  
(philosophy of mind)**

**R1: robot**

**R1D1:  
robot deducer**

**R2D1:  
robot  
relevant**



INSIDE(R1,ROOM)  
INSIDE(BATTERY,ROOM)  
INSIDE(BOMB,ROOM)  
INSIDE(WAGON,ROOM)  
ON(BATTERY,WAGON)  
COLOR(WALLS,BLUE)  
HEIGHT(ROOM,9FEET)  
ON(BOMB,WAGON)  
PULLOUT(WAGON,ROOM)  
ETC.  
ETC.

Illustration: Isabelle Follath



University of  
Zurich UZH

32

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.

## 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-deductor, 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.

## 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.)*

## 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

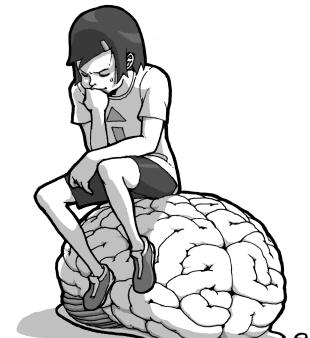
---

- 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**
- The “frame-of-reference” problem

# Two views of intelligence

classical:  
**cognition as computation**

embodiment:  
**cognition emergent from sensory-motor and interaction processes**



38

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:

39

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

41

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 interact, 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”

44

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



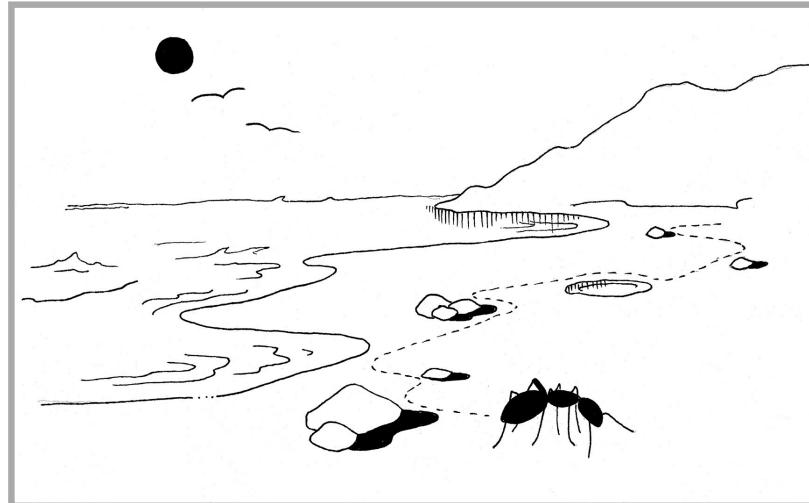
# The “frame-of-reference” problem – introduction

Video “Heider and Simmel”

**comment on video?**



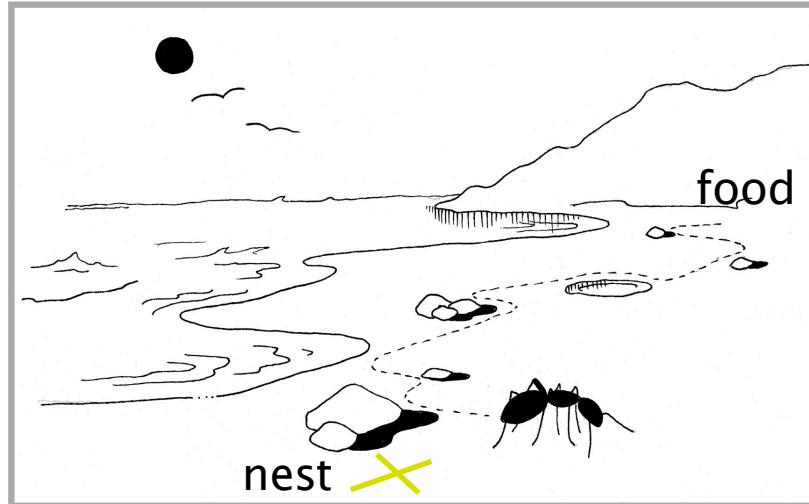
# “Frame-of-reference” Simon's ant on the beach



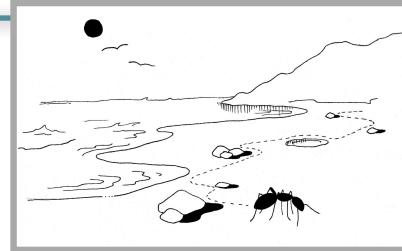
47

From Herbert A. Simon: The sciences of the artificial. MIT Press.  
(highly recommended reading)

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



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



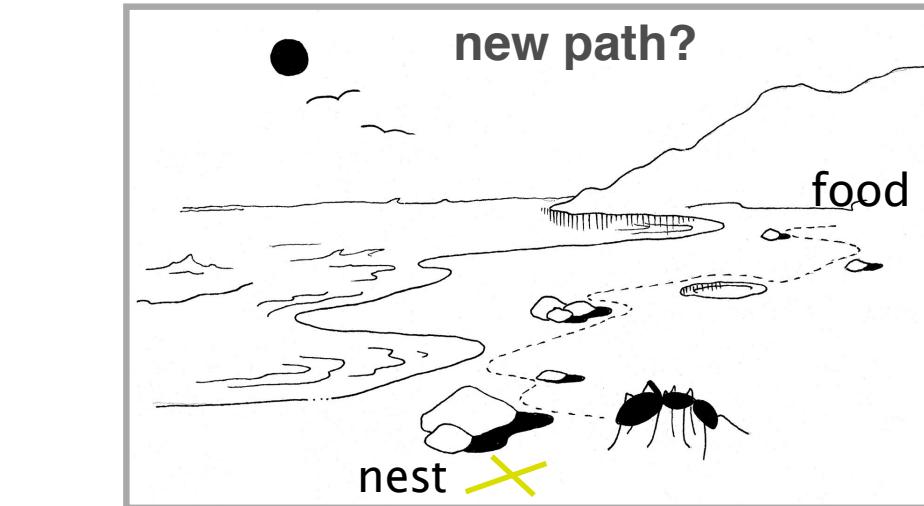
- simple behavioral rules
- complexity in interaction,  
not — necessarily — in brain
- thought experiment:  
**increase body by factor of 1000**  
**everything else the same**

49

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



50

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 ?



Swiss Chocolate



Champagne



52

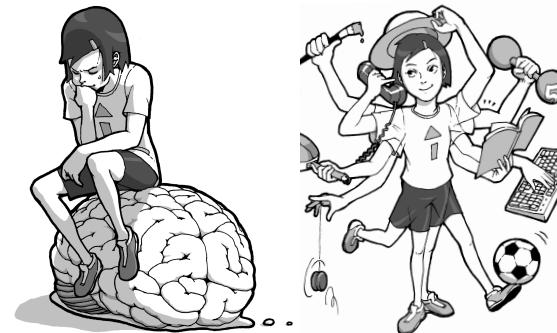
Should there be a day in the ShangAI lectures where the FOR problem is not mentioned, the first to discover will get either a box of Swiss chocolate or a bottle of champagne. If there is one idea from the class that everyone should remember, it's the FOR problem.

# End of lecture 2

---

Thank you for your attention!

*back to Fabio Bonsignorio*



# 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.

## Issue of infinite regress

55

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)?**

56

The reason to use autonomous systems in the study of intelligence is that precisely the human is removed from the loop, thereby forcing the artificial system to interact with the environment on its own, without human help.

# End of additional slide materials for self-study

---

