The ShanghAI Lectures

上海授课

人工智能

TTD: 11 Oct 2012

- check videos with Nathan
- short intro to Shanghai JiaoTong University (by Prof. Weidong Chen)
- check data on machine translation system and produce slider slides)

Video clips

- Didabots (30s to 1.10)
- Didabots heap building (first minute)
- “Puppy” (original version)
- “Mini-dog” (entire video, 15s)

The ShanghAI Lectures
by the University of Zurich
An experiment in global teaching

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

Today from Shanghai Jiao Tong University, Shanghai, China

欢迎您参与

“来自上海的人工智能系列讲座”

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab


Shanghai
CHINA
RUSSIAN FEDERATION
KAZAKHSTAN
MONGOLIA
KYRGYZSTAN
TAJIKISTAN
PAKISTAN
INDIA
NEPAL
BHUTAN
BANGLADESH
MYANMAR
LAOS
THAILAND
VIETNAM
PHILIPPINES
TAIWAN
NORTH KOREA
SOUTH KOREA
JAPAN
Sea of Japan (East Sea)
Yellow Sea
PACIFIC OCEAN
South China Sea
Bay of Bengal
Lake Khanka
Gobi Desert
Tarim Pendi
Kunlun Shan
Qaidam Pendi
Bayan Har Shan
Qilian Shan
Quingzang Gaoyuan (Plateau of Tibet)
HIMALAYA
Fengshui Shan 1398 m
Taiping Ling 1712 m
Altun Shan 5798 m
Muz Tag 7723 m
Ngaglong Kangri 6596 m
BEIJING
Forbidden City (Palace Museum)
Temple of Heaven
Hanging Temple
Terra Cotta Warriors
Red Hill
Flaming Mountain
Potala Palace
Giant Pandas
Stone Forest
Seven Star Park
Elephant Trunk Hill
The Lingering Garden
Hulun Buir
Qiqihar
Harbin
Ulanhot
Jilin
Changchun
Xilinhot
Fuxin
Shenyang
Anshan
Dalian
Yantai
Qingdao
Tongshan
Hohhot
Jining
Yinchuan
Lanzhou
Xining
Tianshui
Xi'an
Chengdu
Chongqing
Guiyang
Kunming
Liuzhou
Nanning
Guangzhou (Canton)
Macau
Hong Kong
Haikou
Nanchang
Gullin
Hangzhou
Altay
Karamay
Yining
Ürümqi
Turpan
Aksu
Korla
Kashi
Hotan
Lhasa
Lop Nur
Qinghai Hu
Nu Jiang
Huang He
Tropic of Cancer
90° E
105° E
120° E
135° E
75° E
45° N
30° N
15° N
LEGEND
Country Boundary
Major Road
River
Seaport
Major City
Country Capital
Airport
Point of Interest
Lake
Great Wall of China
0 150 300 Km
0 75 150 Miles
Copyright © 2012-13 www.mapsofworld.com
(Created on 21st Sept. 2012)

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY 电子信息与电气工程学院 School of Electronic, Information and Electrical Engineering

《上海人工智能系列讲座（The ShanghAI Lectures）

short intro by Prof. Weidong Chen

Calling on:
Please ask questions any time!

- Xi'an —> The Chinese Room Experiment (student pres.)
- Chiba —> Simon's ant on the beach
- Budapest —> Didabot video: comment
- Salford —> what are the robots doing?
- Shanghai —> requirements for clustering; Didabots
- Madrid —> change of morphology
- Berlin (Humboldt) —> changing the number of Didabots
- skku/isrc —> pros and cons of natural language/mathemtatics for theory formulation
- Moscow —> why emergence in "Puppy"?
- KIT (Karlsruhe): Guest lecture by Prof. Alex Waibel
- ACT (Thessaloniki): Guest lecture by Prof. Vincent C Müller

Lecture 3

Towards a theory of intelligence

11 October 2012

University of Zurich

Simultaneous translation

- most advanced translation system
- Prof. Alex Waibel, Carnegie-Mellon University and Karsruhe Institute of Technology

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

Today's topics

- short recap
- the "Chinese Room Thought Experiment" (student presentation from NPU, Xi'an)
- Braitenberg vehicles (for reading)
- the "Swiss Robots"
- prerequisites for a "theory of intelligence"
- intelligent systems: properties and principles (mostly lecture 4)
- Guest lectures by Alex Waibel (CMU and KIT) and Vincent Müller (American College, Tessaloniki, and Oxford Univer.)

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

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 上海人工智能授课

Some problems:
- symbol grounding problem
- frame problem
- homunculus problem

The need for an embodied approach:
previous lecture: Louis Wolpert’s quote: why do plants not have brains?
Also evolutionary argument by Rodney Brooks: (“The embodied turn”)
(see later)

The “embodied turn” Rodney Brooks, MIT

Single-cell entities arose out of the primordial soup roughly 3.5 billion years ago. A billion years passed before photosynthetic plants appeared. After almost another billion and a half years—around 550 million years ago—the first fish and vertebrates came into being, and 100 million years later insects emerged. Let us quote directly from Brooks’s argument:

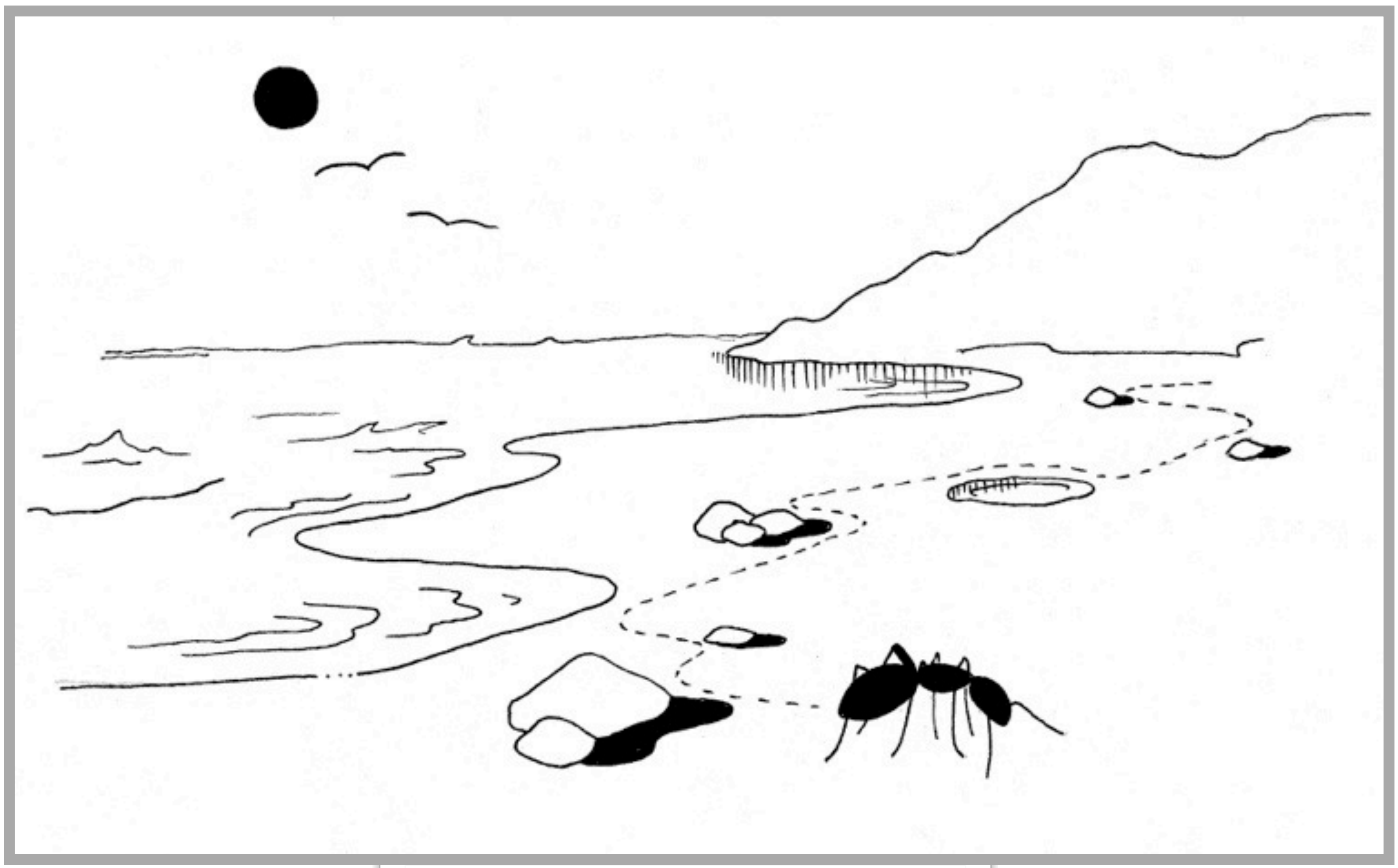
“Then things started moving fast. Reptiles arrived 370 million years ago, followed
by dinosaurs at 330 and mammals at 250 million years ago. The first primates appeared 120 million years ago and the immediate predecessors to the great apes a mere 18 million years ago. Man arrived in roughly his present form 2.5 million years ago. He invented agriculture a mere 19,000 years ago, writing less than 5,000 years ago and “expert” knowledge only over the last few hundred years.” (Brooks, 1990, p. 5)

Brooks’ argument: time from primordial soup to insects - much longer than from insects to humans. We must start to study insects first.

Because of this interest in insects, walking and locomotion in general became important research topics in artificial intelligence and robotics research.

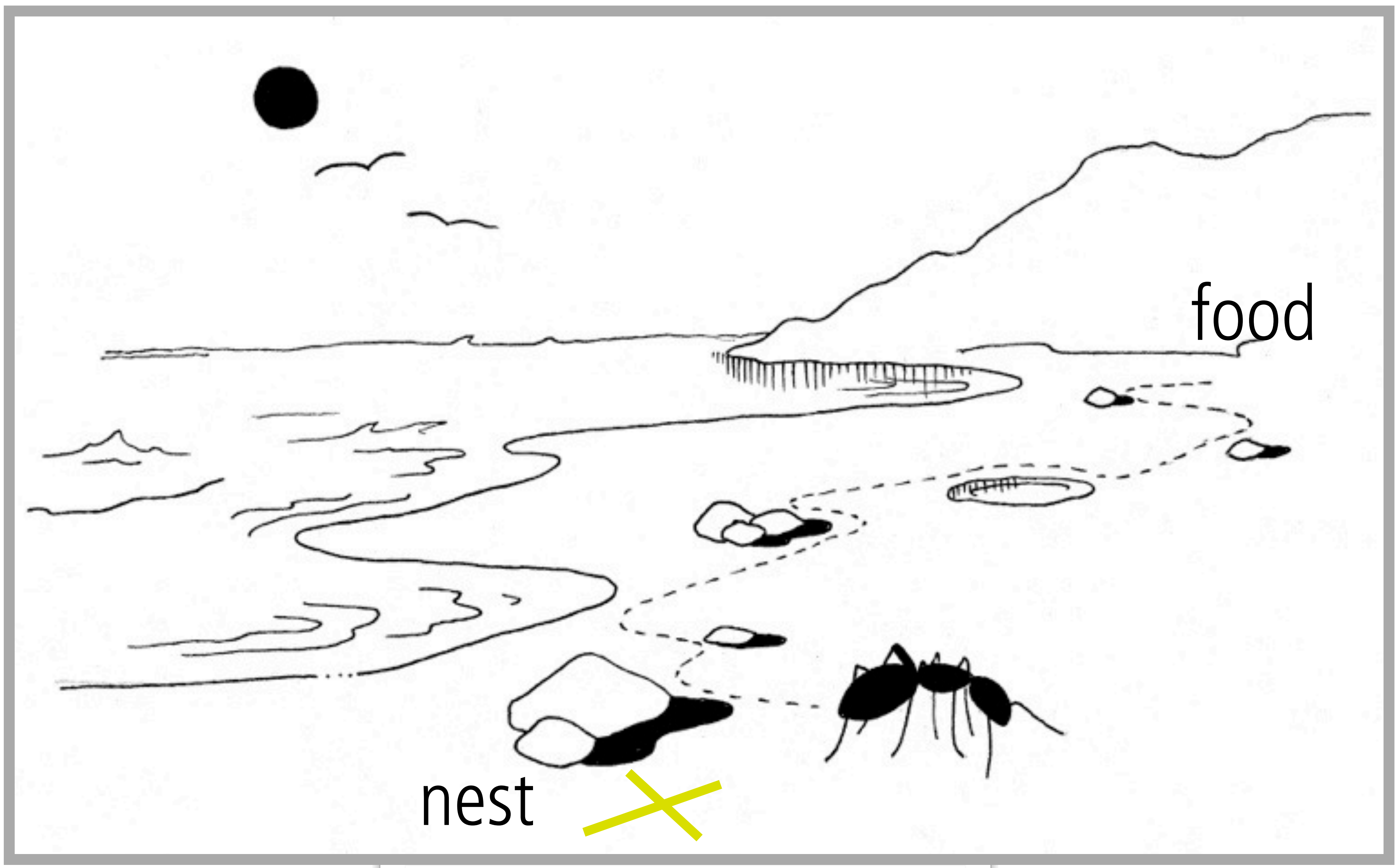
Brooks, R. A. (1990). Elephants don’t play chess. In P. Maes, ed. Designing autonomous agents: Theory and practice from biology to engineering and back. Cambridge, MA: MIT Press, 3-15.

"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



University of Zurich

robotics Swiss National Centre of Competence in Research

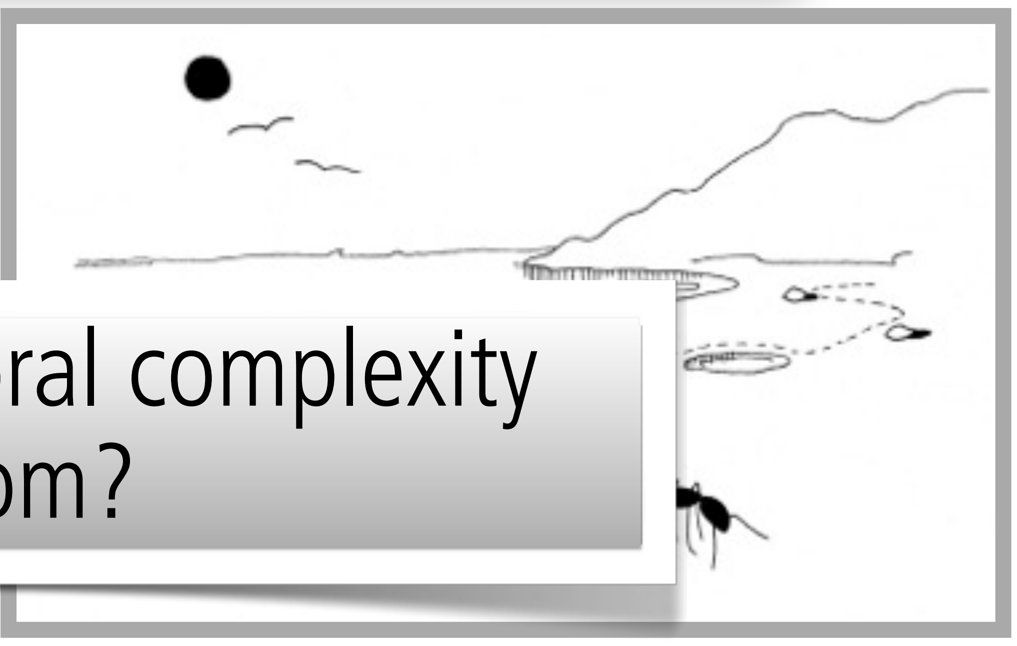
ai lab

The ShanghAI Lectures 人工智能 上海授课

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

- **simple behavioral rules**



—> **Chiba?**

where is the behavioral complexity coming from?

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

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

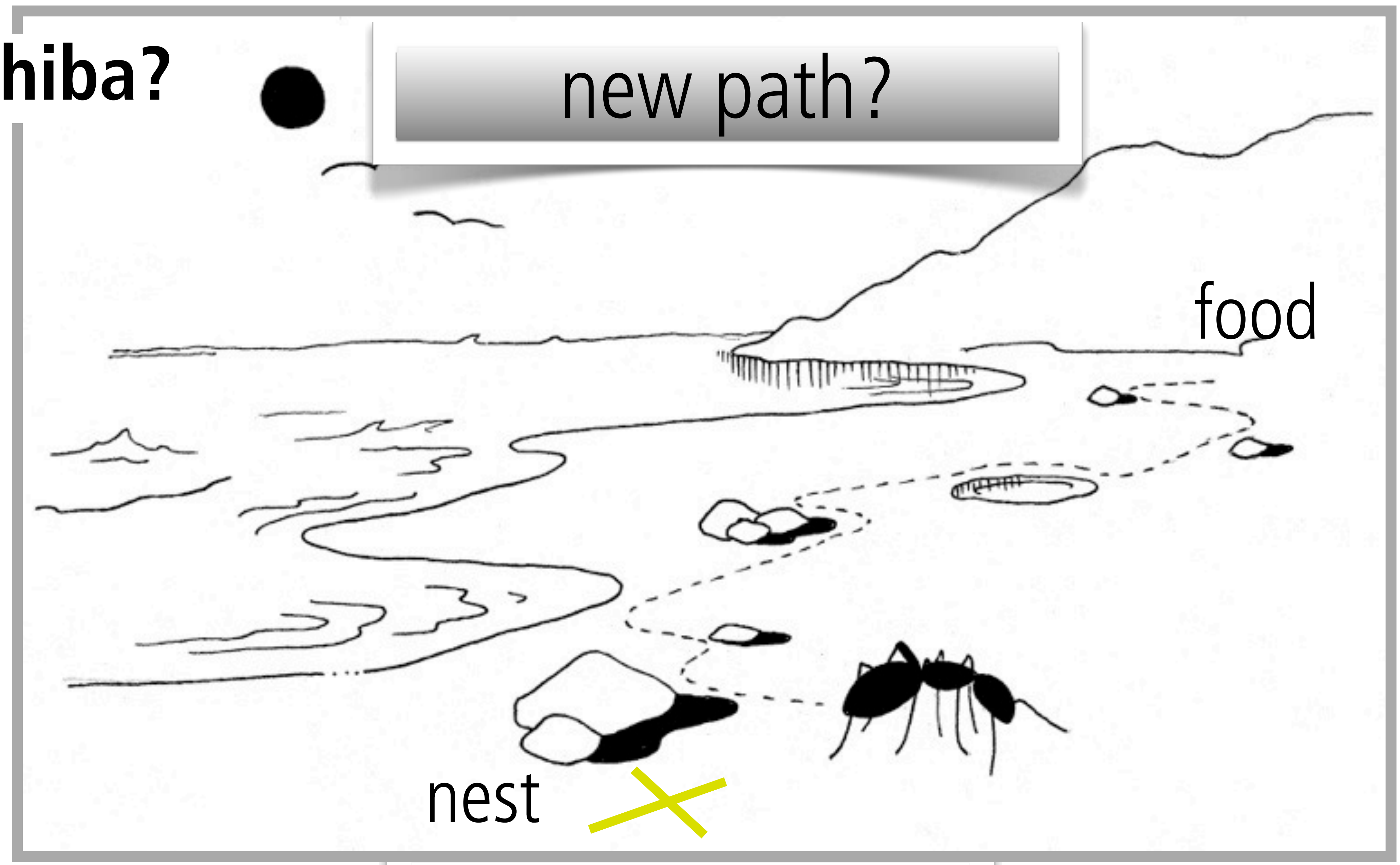
The ShanghAI Lectures 人工智能 上海授课

complexity emerges from interaction of ant with environment - it's neither the environment alone, nor the ant along.

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

—> Chiba?

new path?



"Frame-of-reference" F-O-R

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

F-O-R competition

?

Swiss Chocolate

Champagne

Should I forget to mention the FOR problem in any of the lectures, 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.

Today's topics

- short recap
- the "Chinese Room Thought Experiment" (student presentation from NPU, Xi'an)
- Braitenberg vehicles (for reading)
- the "Swiss Robots"
- prerequisites for a "theory of intelligence"
- intelligent systems: properties and principles (mostly lecture 4)
- Guest lectures by Alex Waibel (CMU and KIT) and Vincent Müller (American College, Tessaloniki, and Oxford Univer.)

The "Chinese Room Thought Experiment"

Student presentation from Xi'an

University of Zurich

robotics Swiss National Centre of Competence in Research

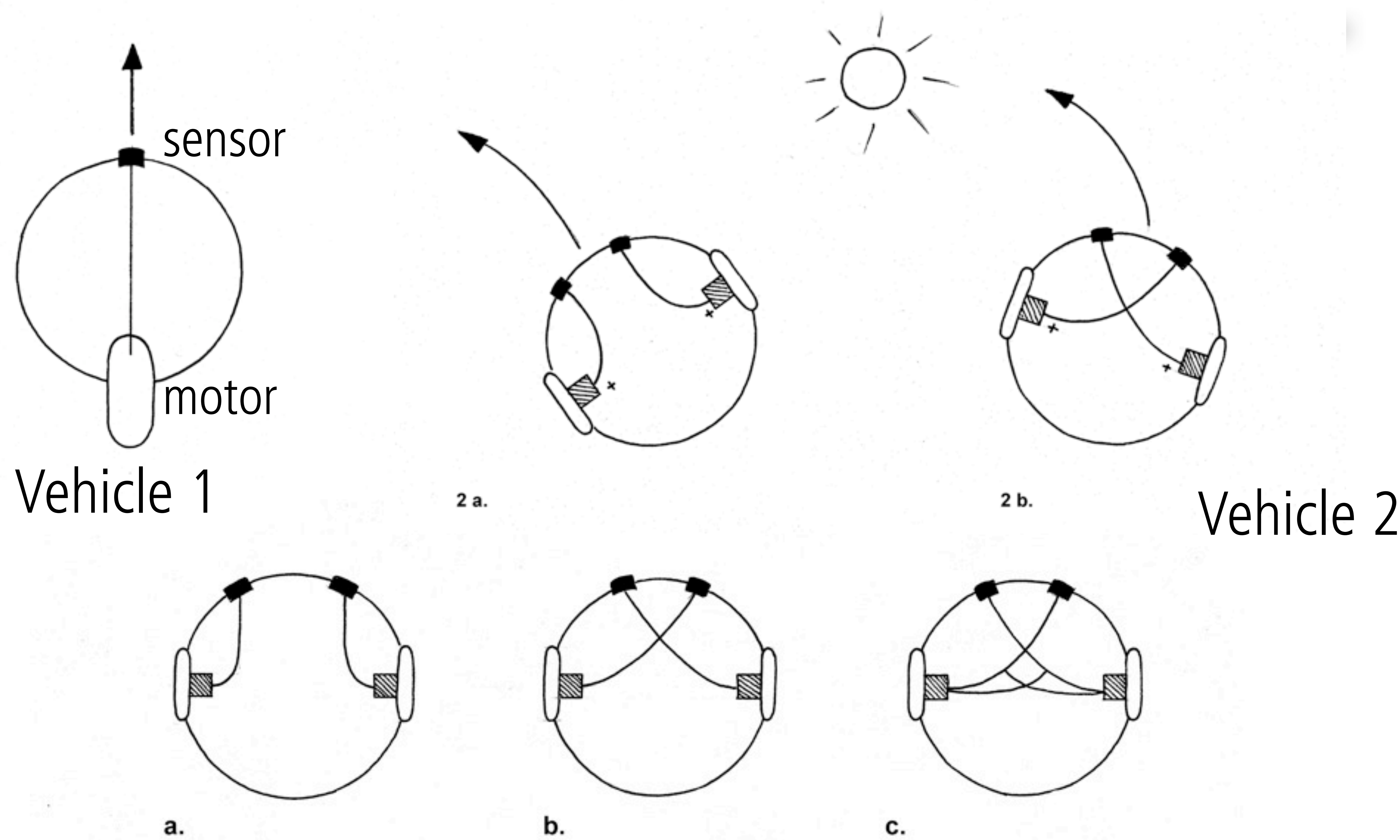
ai lab

人工智能 上海授课 The ShanghAI Lectures

Today’s topics

- short recap
- the “Chinese Room Thought Experiment” (student presentation from NPU, Xi’an)
- Braitenberg vehicles (for reading)
- the “Swiss Robots”
- prerequisites for a “theory of intelligence”
- intelligent systems: properties and principles (mostly lecture 4)
- Guest lectures by Alex Waibel (CMU and KIT) and Vincent Müller (American College, Tessaloniki, and Oxford Univer.)

Braitenberg vehicles



Please read about the Braitenberg vehicles:
Braitenberg, V. (1986). Vehicles: Experiments in synthetic psychology. MIT Press.
or
Chapter 6 in “Understanding intelligence”.

Braitenberg vehicles represent a series of agents of increasing complexity. Although some are purely reactive, others include learning mechanisms, and thus have their own history.

In the simplest vehicles, it is quite obvious what they do. As matters get slightly more involved, predicting their behavior turns out to be very difficult, even in purely reactive systems, because the mechanisms generating the vehicles' behavior interact in interesting ways. Even if we have complete knowledge of the vehicle's insides, it still proves difficult to control it. Its interaction with itsenvironment adds complexity, which leads to some degree of unpredictability, even if the driving mechanisms are entirely deterministic -- in physics, there are always fluctuations. Let us examine the Braitenberg vehicles one by one. As always, we pay attention to the frame-of-reference problem. In examining this series of vehicles, it is always a good idea to imagine how they move around under various conditions. This process of imagination is best complemented with computer simulations or with experiments on real robots.

Vehicle 1
The first Braitenberg vehicle has one sensor, for one particular quality, and one motor. The sensor and the motor are connected very simply: The more there is of the quality to which the sensor is tuned, the faster the motor goes. If this quality is temperature, it will move fast in hot regions and slow down in cold regions. An observer might get the impression that such a vehicle likes cold and tries to avoid heat. The precise nature of thisquality does not matter; it can be concentration of chemicals, temperature, light, noise level, or any other of a number of qualities. The vehicle always moves in the direction in which it happens to be
pointing.

If we introduce friction into the vehicle's environment, its behavior gets interesting, because friction is always a bit asymmetric. The vehicle eventually deviates from its straight course, and in the long run, is seen to move in a complicated trajectory, curing one way or another without apparent (to the observer!) good reason. Perturbations

Observation exercise

Video "Didabots"

what's happening, what's going on?

mind the "frame-of-reference"

University of Zurich

Swiss National Centre of Competence in Research

- Didabot video

Observation exercise

Video "Didabots"

what's happening, what's going on?

mind the "frame-of-reference"

—> Budapest

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

- Didabot video

Today's topics

- short recap
- the "Chinese Room Thought Experiment" (student presentation from NPU, Xi'an)
- Braitenberg vehicles (for reading)
- **the "Swiss Robots"**
- **prerequisites for a "theory of intelligence"**
- **intelligent systems: properties and principles (mostly lecture 4)**
- **Guest lectures by Alex Waibel (CMU and KIT) and Vincent Müller (American College, Tessaloniki, and Oxford Univer.)**

A robot experiment

6x6m arena with Styrofoam cubes

Arena filled with Styrofoam cubes, initially randomly distributed. Put robots in there, they start working.

Sequence

frames: 2-3min

Sequence

frames: 2-3min

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

Sequence

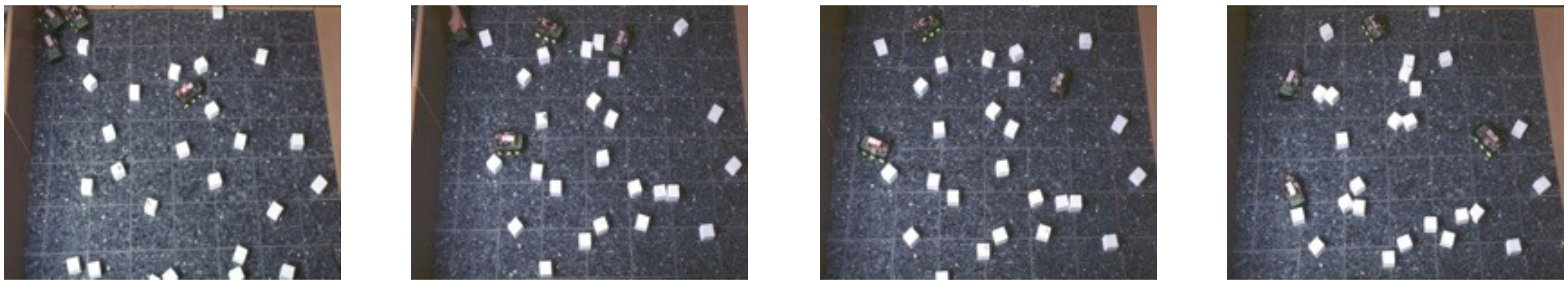
frames: 2-3min

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

Sequence



frames: 2-3min

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

Sequence

frames: 2-3min

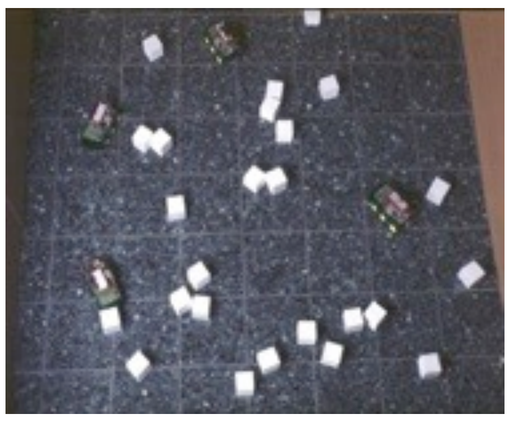
University of Zurich

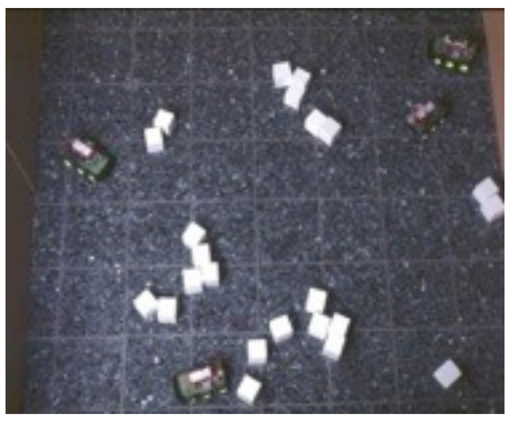
robotics Swiss National Centre of Competence in Research

ai lab

Sequence





frames: 2-3min

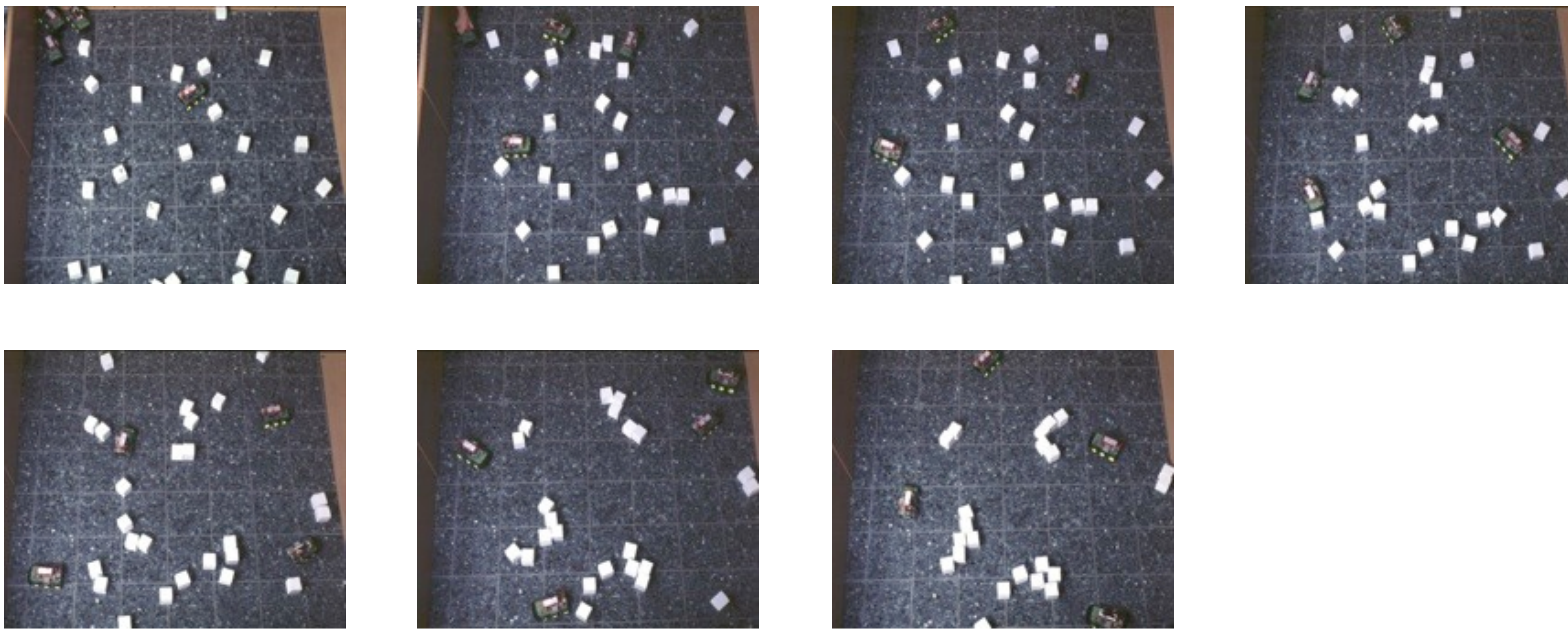
University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

人工智能 上海授课 The ShanghAI Lectures

Sequence



frames: 2-3min

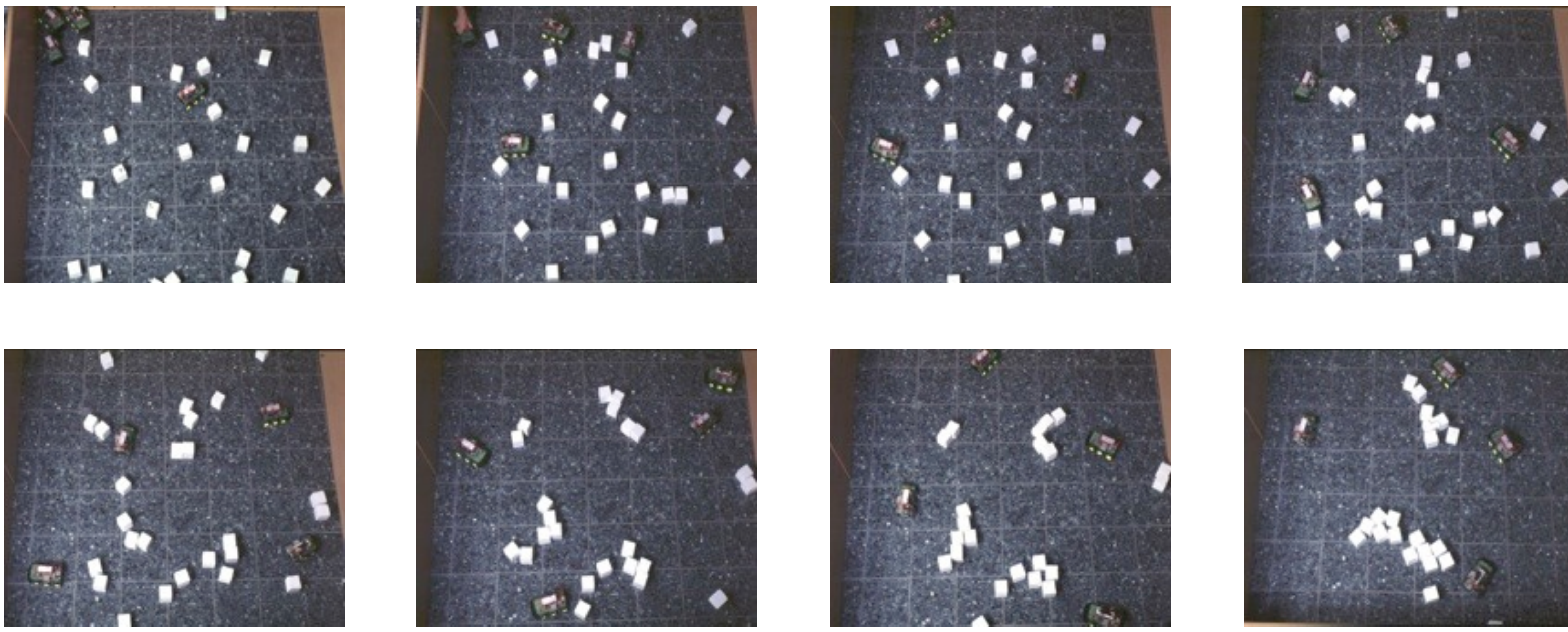
University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

Sequence



entire process: ～20min
frames: 2-3min

The frames are taken about every two to three minutes; the whole process takes roughly twenty minutes. Although the clusters and exact positions of the cubes vary, the final configurations are all qualitatively similar, i.e. a few clusters, and some cubes distributed along the walls.

What are the robots doing?

—> Salford

- forming clusters
- moving cubes together
- making free space

What are the robots doing?

- forming clusters
- moving cubes together
- making free space
- cleaning up

→ the "Swiss Robots"

observer's perspective

What are the robots really doing?

mechanisms underlying their behavior

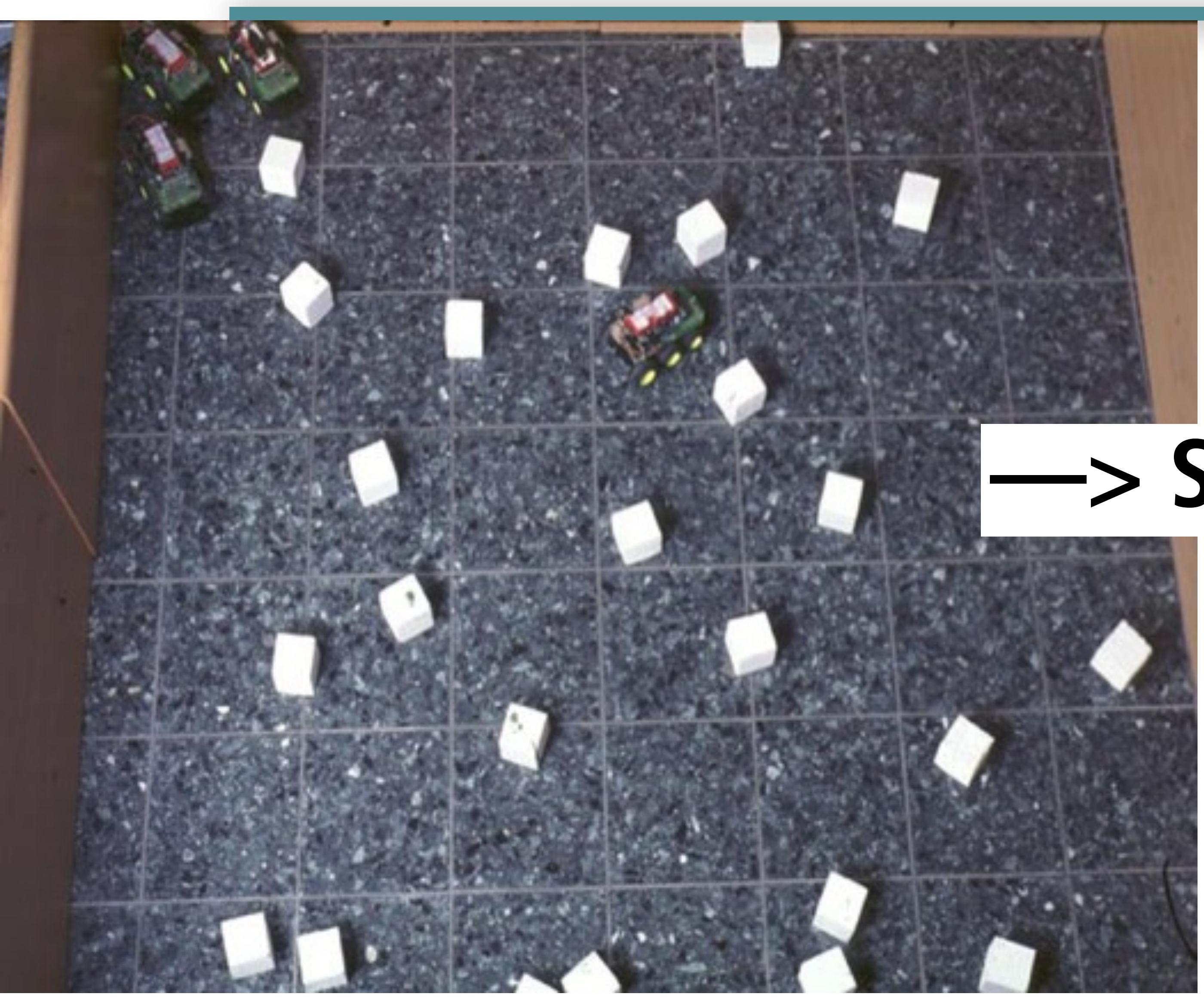
Activities involved in clustering: standard solution

6x6m arena with Styrofoam cubes

Requirements for design of robots?

Look for cube, if possible, the nearest one). Pick up cube (somehow). Look for nearest cluster. Go to cluster. Deposit cube. Look for new cube, etc.
Requires quite sophisticated perceptual skills on the part of the robots (recognizing a cube from different angles and distances, recognizing a cluster). Some elementary motor skills for grasping and dropping. locomotion skills.

Activities involved in clustering: standard solution



Requirements for design of robots?

—> Shanghai

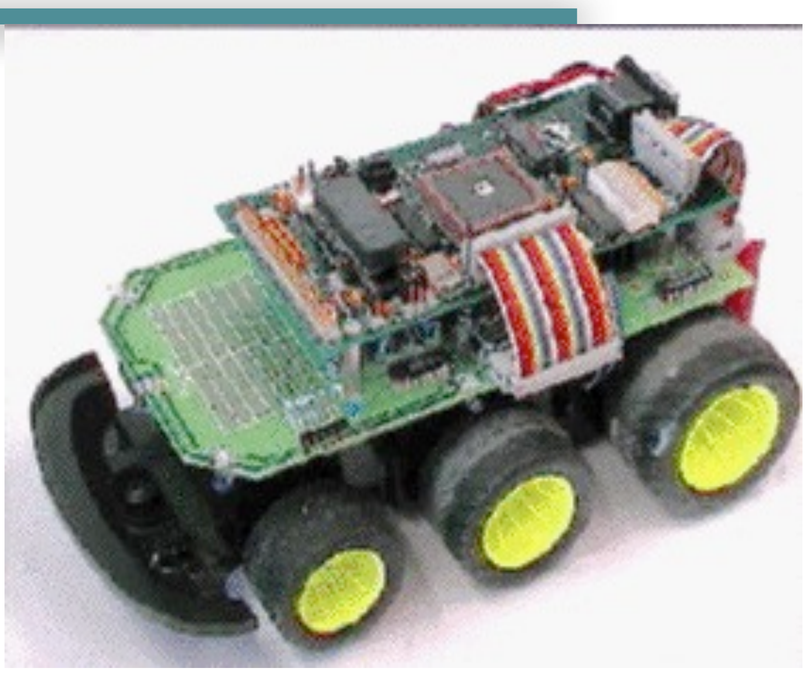
6x6m arena with Styrofoam cubes

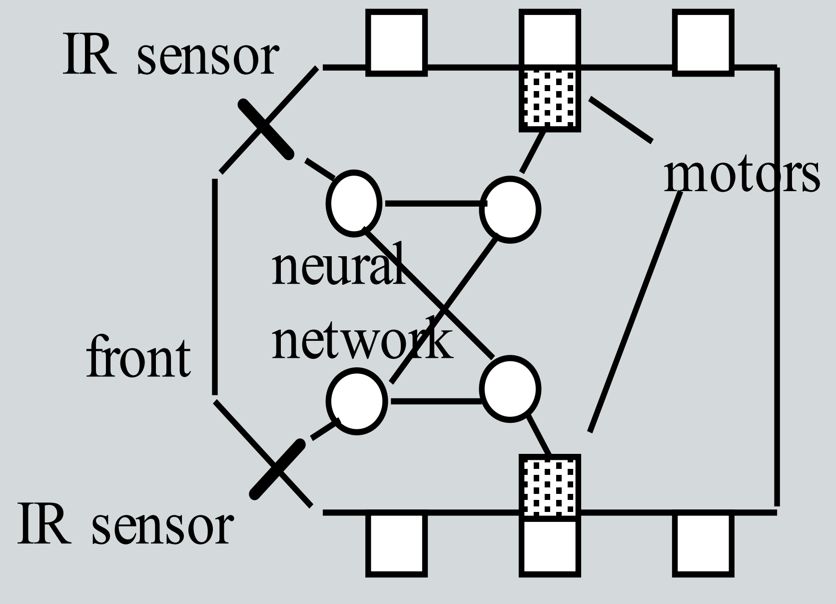
Look for cube, if possible, the nearest one). Pick up cube (somehow). Look for nearest cluster. Go to cluster. Deposit cube. Look for new cube, etc.
Requires quite sophisticated perceptual skills on the part of the robots (recognizing a cube from different angles and distances, recognizing a cluster). Some elementary motor skills for grasping and dropping. locomotion skills.

The Swiss robots

6x6m arena with Styrofoam cubes

Didabot
simple robot
for didactial
purposes





University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

What are the robots really doing?

behavioral rule:

sensory stimulation on left: turn right

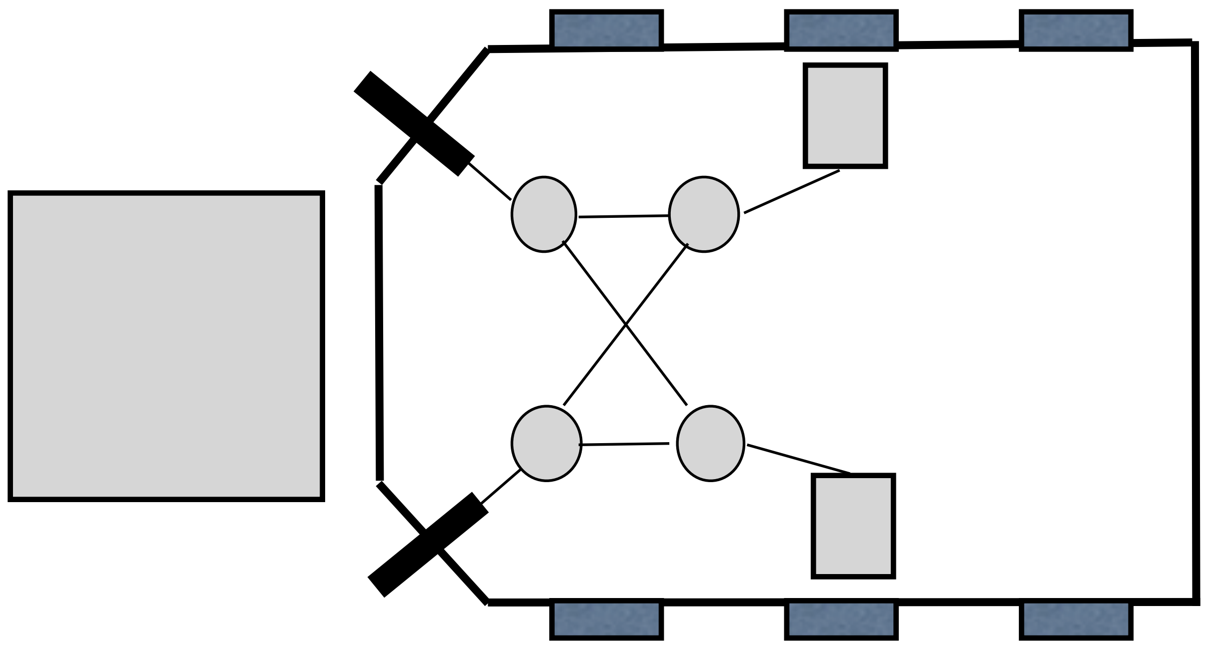
sensory stimulation on right: turn left

(obstacle avoidance)

→

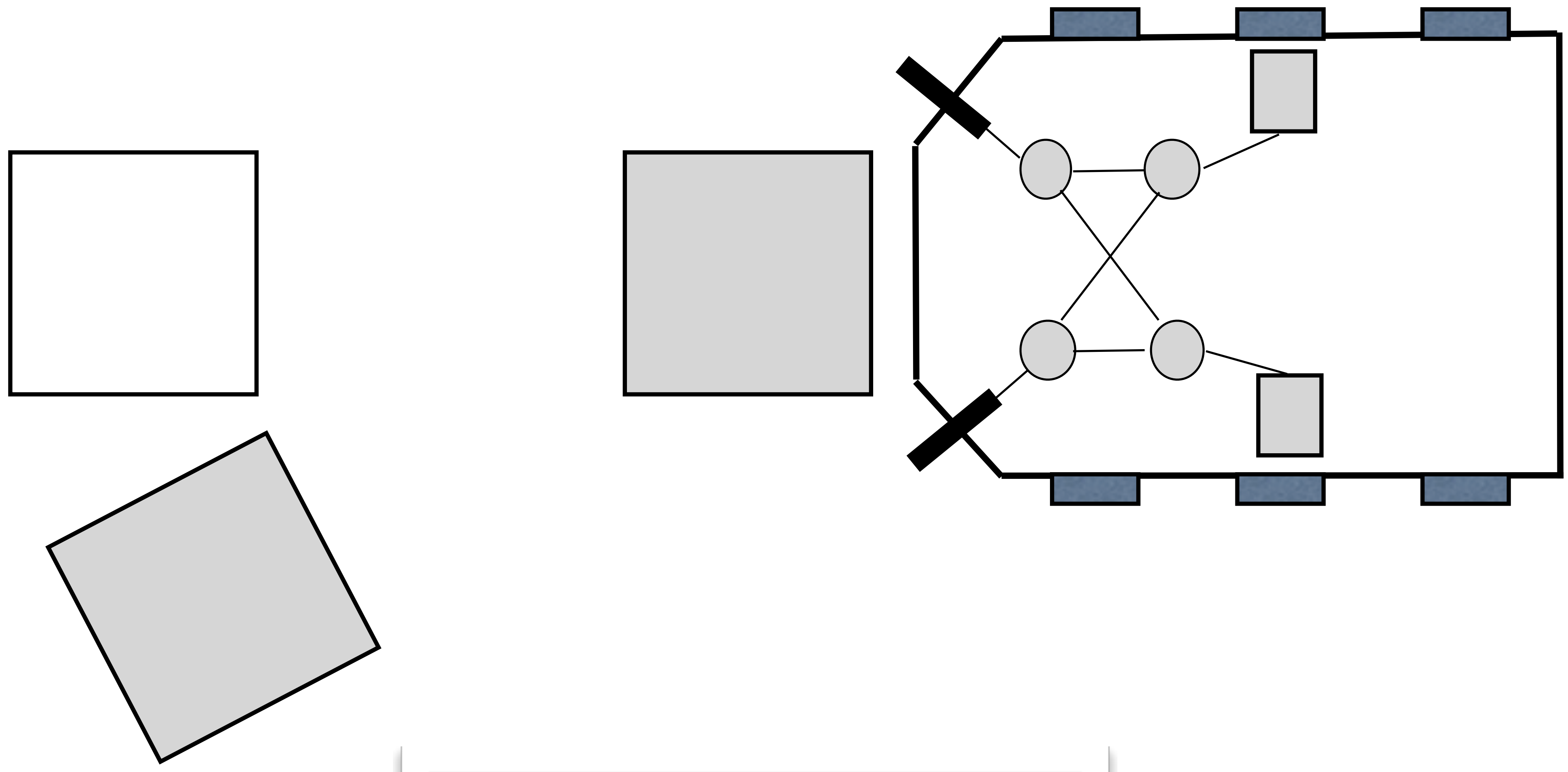
situated perspective (from the agent's point of view)

Cluster formation



Encountering a cube head-on by chance —> no stimulation of the sensors. What happens?
How far will it push the cube?

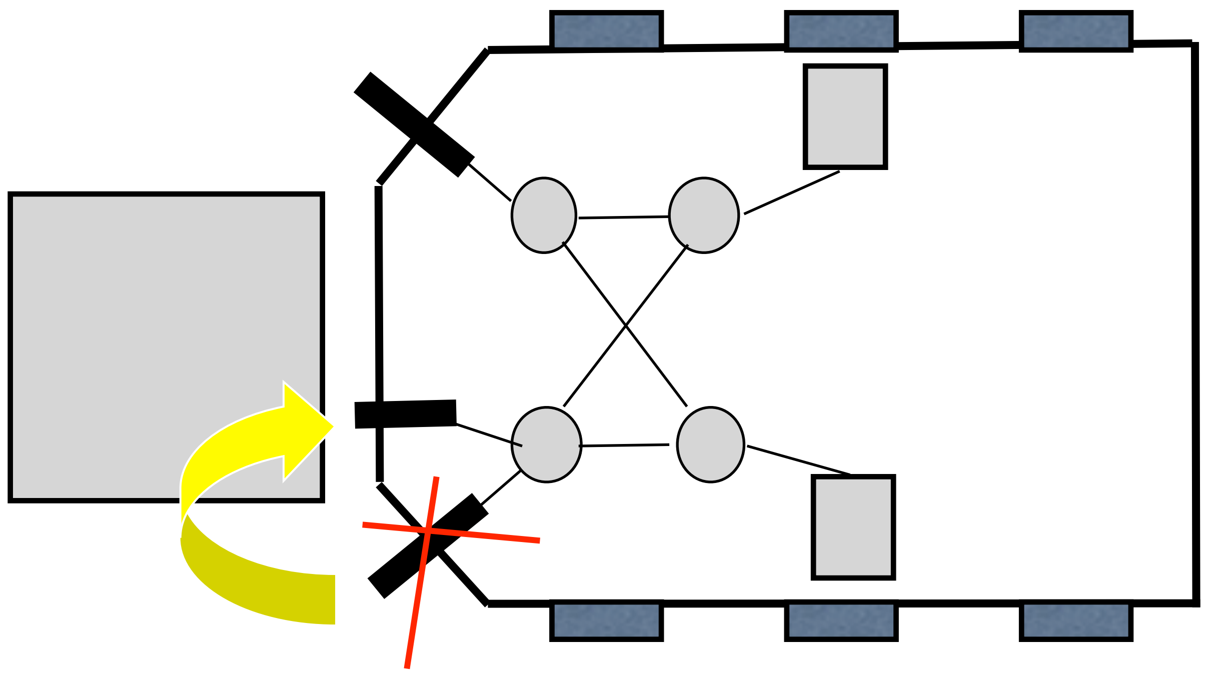
Cluster formation



until it encounters an object with one of its sensors - here the one on the left; it will then turn right, leaving two cubes together.

Change of morphology

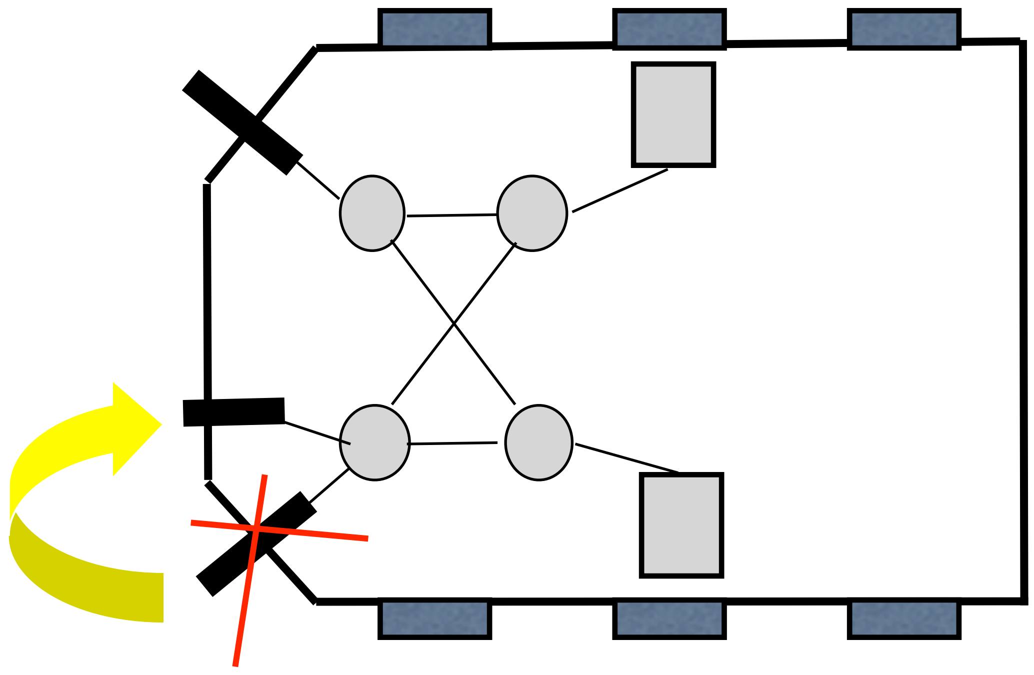
What happens if everything stays the same except that the position of one of the sensors is changed as shown in the figure?



Change of morphology

What happens if everything stays the same except that the position of one of the sensors is changed as shown in the figure?

—> Madrid



"Swiss Robots" summary

- frame-of-reference
- self-organization and emergence
- embodiment: interdependence morphology — behavior
- exploitation of ecological niche: "cheap design" (see later)

Video "Didabots heap building"

University of Zurich

robotics Swiss National Centre of Competence in Research

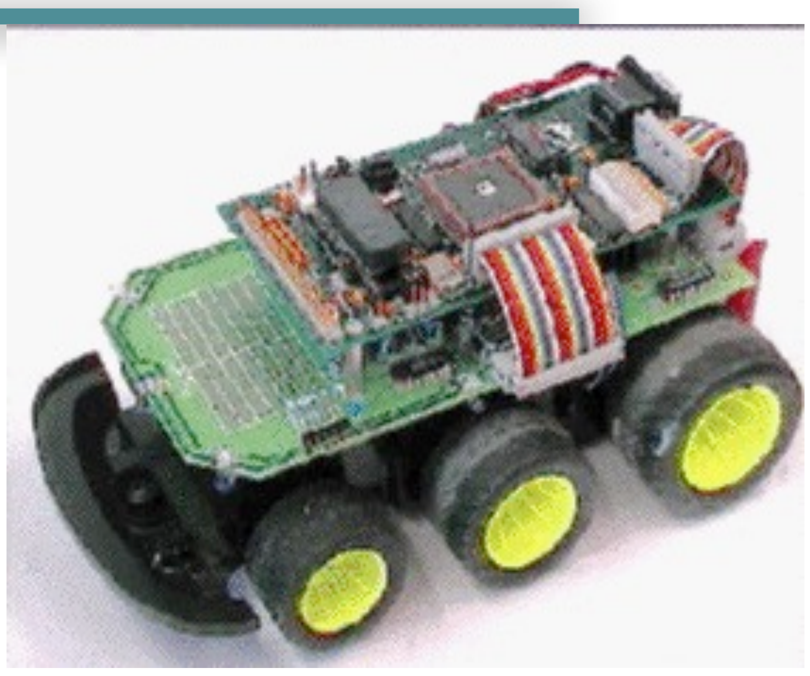
ai lab

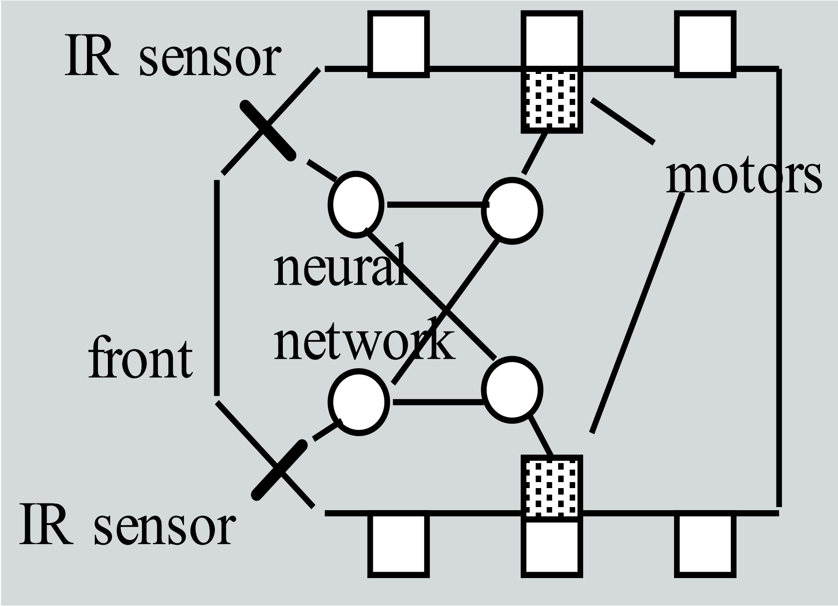
The ShanghAI Lectures 人工智能 上海授课

Daniel Dennett
(philosophy of mind)

6x6m arena with Styrofoam cubes

Didabot
simple robot
for didactial
purposes





University of Zurich
robotics Swiss National Centre of Competence in Research
ai lab
The ShanghAI Lectures 上海 人工 智能 授课

These robots are cleaning up, but that's not what they think they are doing (joke — can they think?)

Change of number of robots

What happens if instead of four to five, only one robot is used?

—> Humboldt (Berlin)

The "Swiss Robots"

"building clusters"

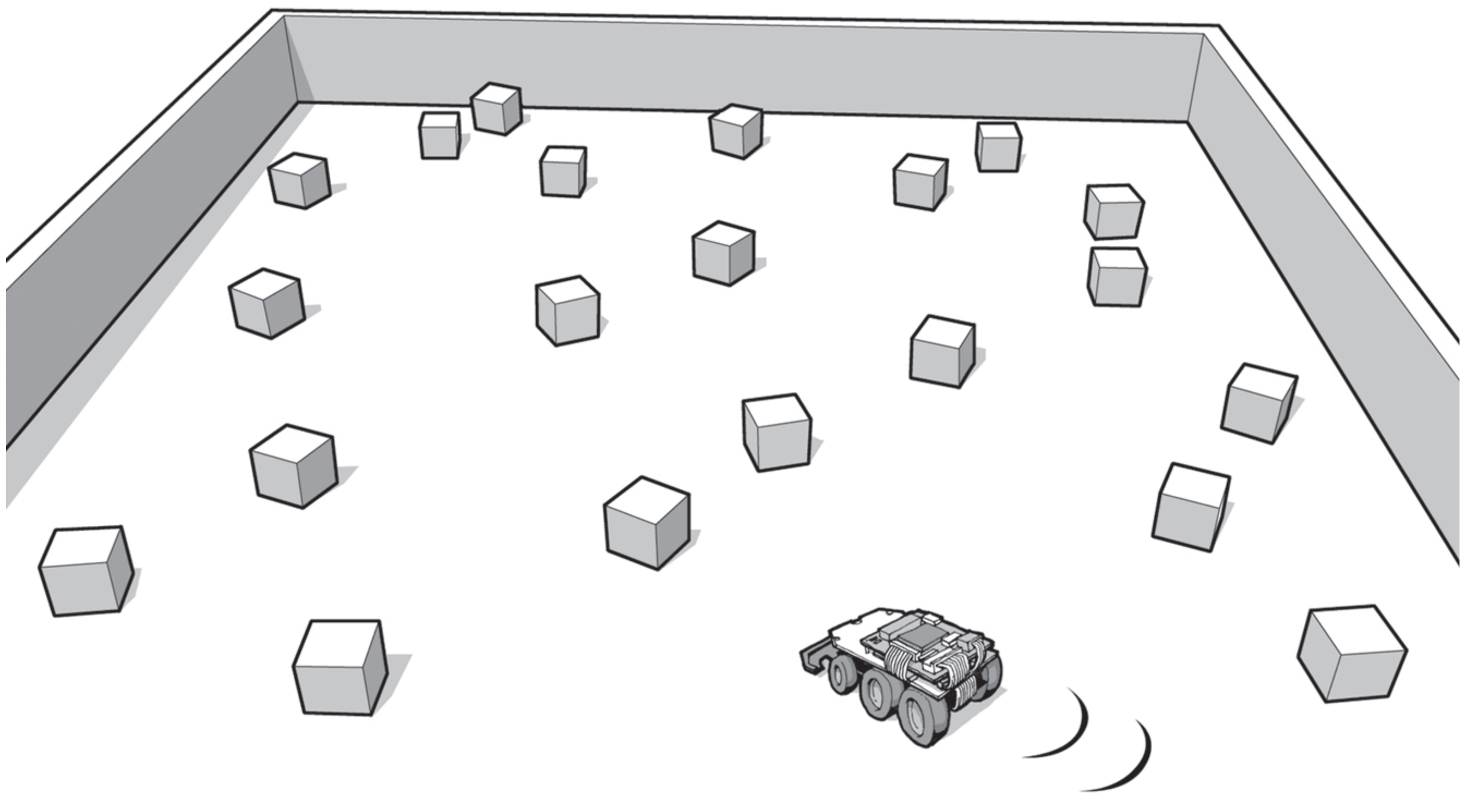
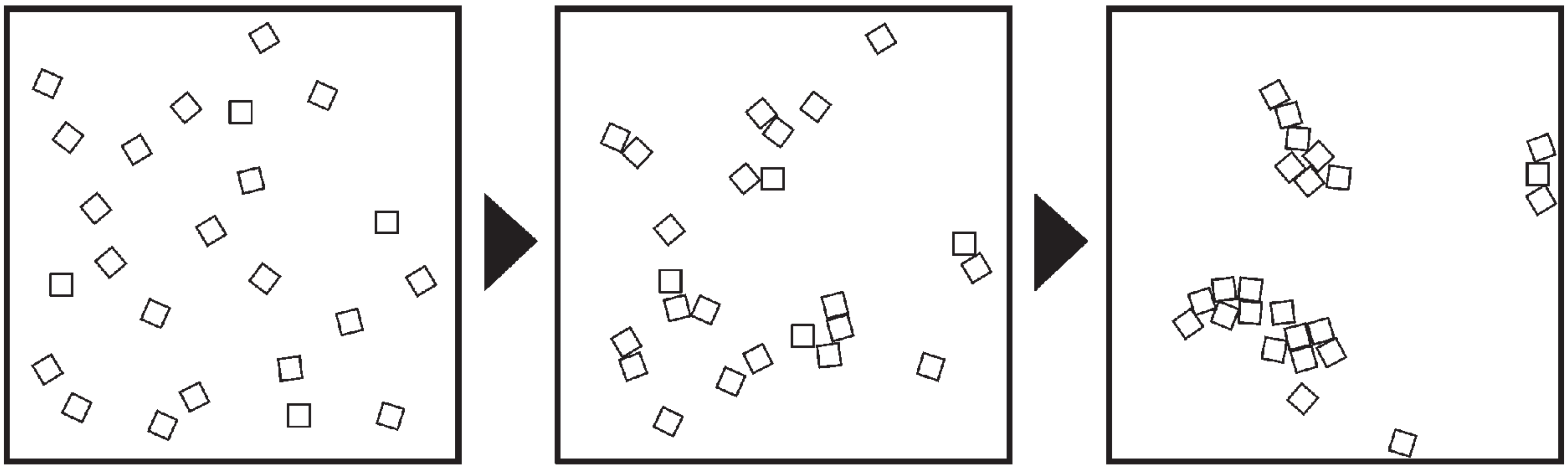


Figure 3.1 from "How the body …"



University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

The "Swiss Robots": cluster formation — morphology change

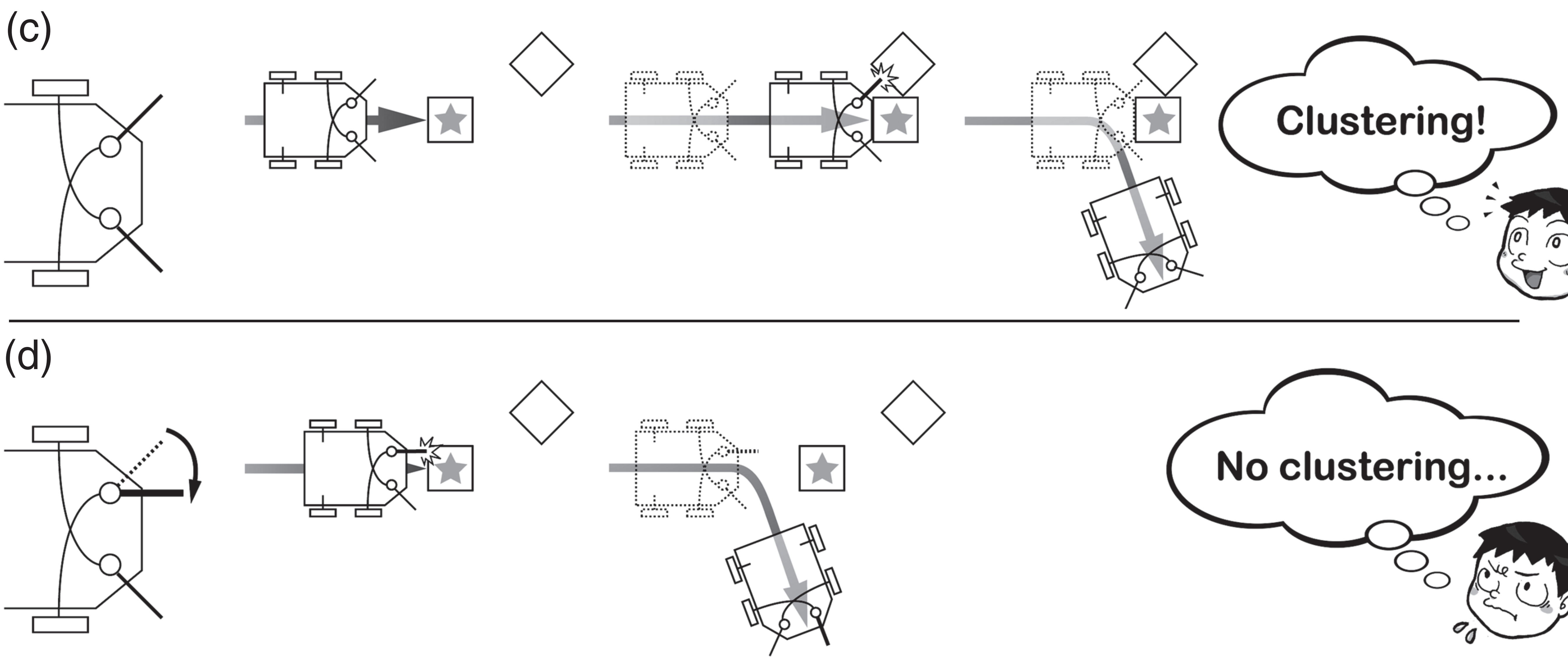


Figure 3.1 from "How the body ..."

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

Today’s topics

- short recap
- the “Chinese Room Thought Experiment” (student presentation from NPU, Xi’an)
- Braitenberg vehicles (for reading)
- the “Swiss Robots”
- **prerequisites for a “theory of intelligence”**
- **intelligent systems: properties and principles (mostly lecture 4)**
- **Guest lectures by Alex Waibel (CMU and KIT) and Vincent Müller (American College, Tessaloniki, and Oxford Univer.)**

Prerequisites for a "theory of intelligence"

- form of theory
- frame-of-reference
- synthetic methodology
- time perspectives
- emergence

Form of theory

- philosophy of science, not one answer, subjective, preferences
- verbal: low precision
- mathematical: rigorious (will use dynamical systems)
- algorithmic: GOFAI
- intelligence: not only "understanding", but also "building" —> synthetic methodology

design principles: joining engineering and science

What form the theory should have is not a scientific questions but on of philosophy of science.

While natural language has high expressivity it has, for the formulation of a scientific theory, a number of disadvantages:
- vague, lack of rigor
- ambiguous
- easy to gloss over details

Form of theory

- philosophy of science, not one answer, subjective, preferences
- verbal: low precision
- mathematical: rigorious (will us
- algorithmic: GOFAI
- intelligence: not only “understanding”, but also “building” —> synthetic methodology

—> verbal/ mathematical theories: isrc/skku, Seoul, Korea

design principles: joining engineering and science

University of Zurich

robotics Swiss National Centre of Competence in Research

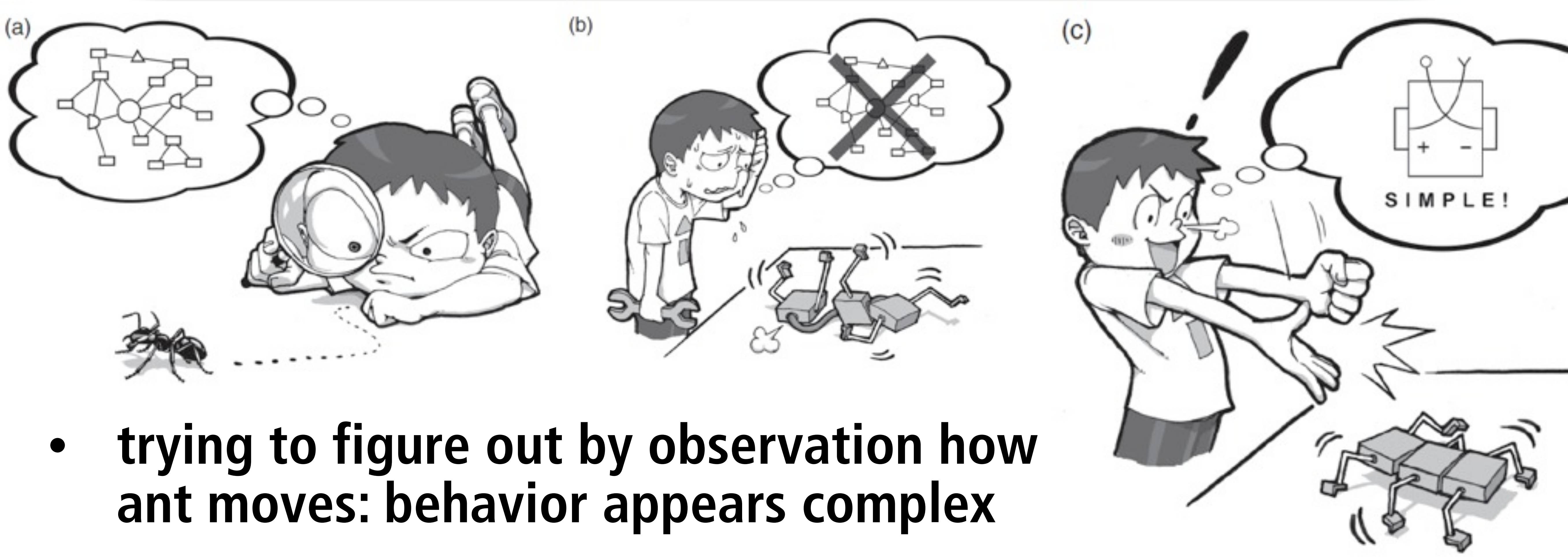
ai lab

The ShanghAI Lectures 人工智能 上海授课

What form the theory should have is not a scientific questions but on of philosophy of science.

While natural language has high expressivity it has, for the formulation of a scientific theory, a number of disadvantages:
- vague, lack of rigor
- ambiguous
- easy to gloss over details

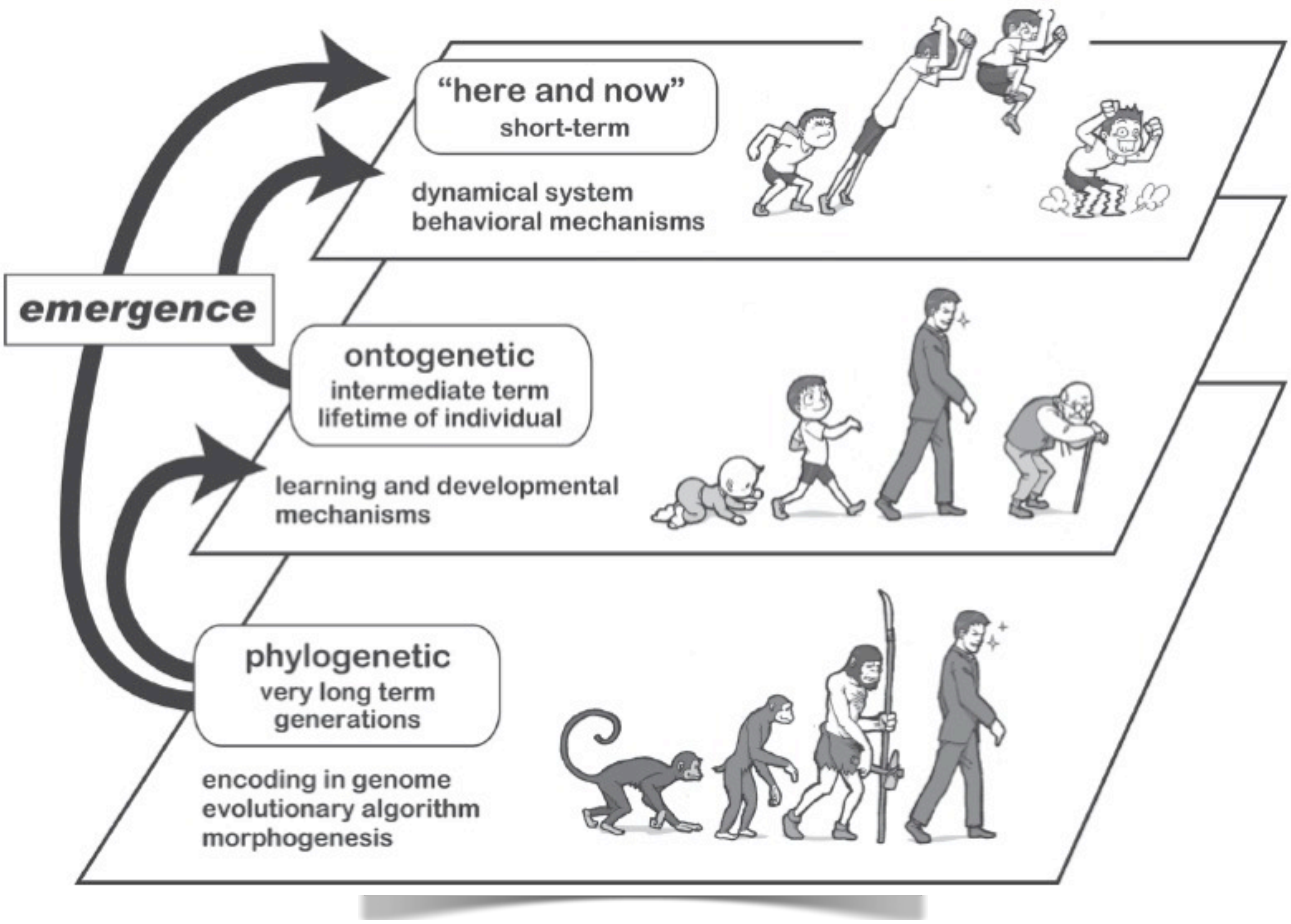
Synthetic methodology and frame-of-reference



- **trying to figure out by observation how ant moves: behavior appears complex**
- **first design fails**
- **complex behavior from simple behavioral architecture**

cf. Braitenberg vehicles

Time perspectives



Time perspectives in understanding and design

state-oriented **"hand design"**	"here and now" perspective
learning and development **initial conditions, learning and developmental processes**	"ontogenetic" perspective
evolutionary **evolutionary algorithms, morphogenesis**	"phylogenetic" perspective

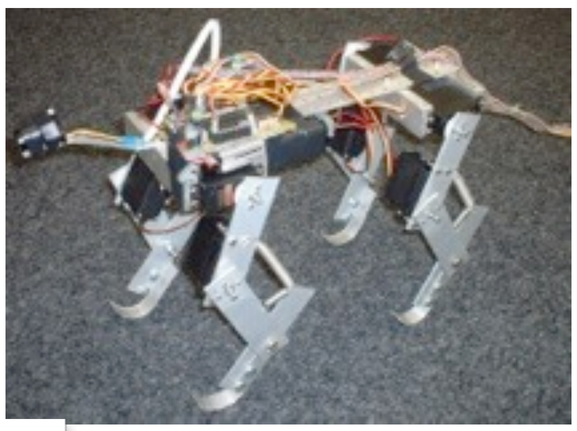
Understanding: **all three perspectives requires**
Design: **level of designer commitments, relation to autonomy**

Emergence

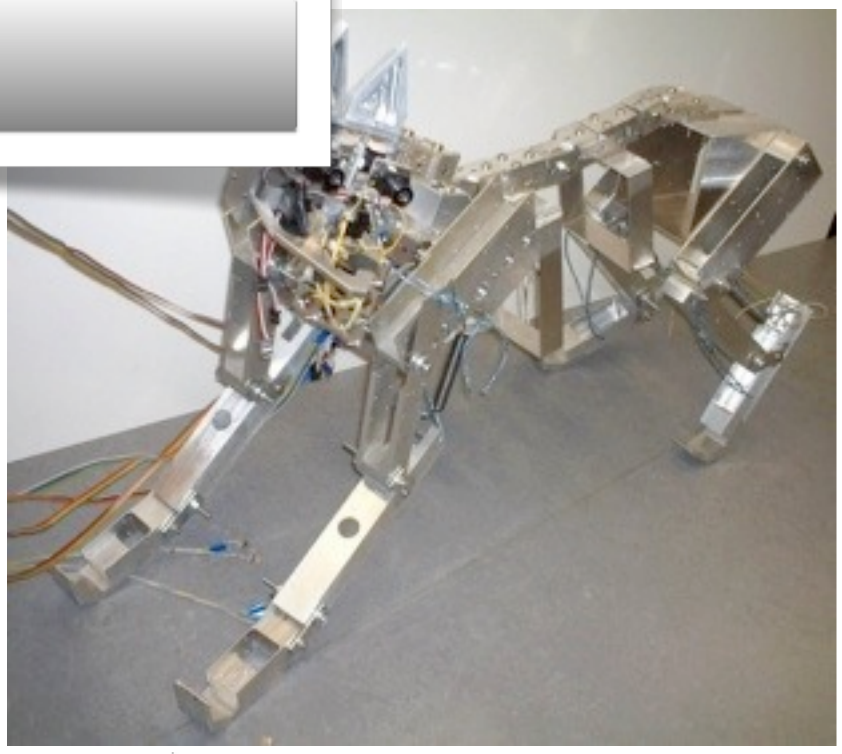
- collective behavior: global patterns from local interactions (e.g. “Swiss Robots”, bird flocks, clapping) (see chapter 7)
- behavior of individual: emergent from interaction with environment
- from time scales (see later)

Emergence of behavior: the quadruped "Puppy"

rapid locomotion in biological systems



Video "Puppy" standard



Design and construction:
Fumiya Iida, AI Lab, UZH and ETH-Z

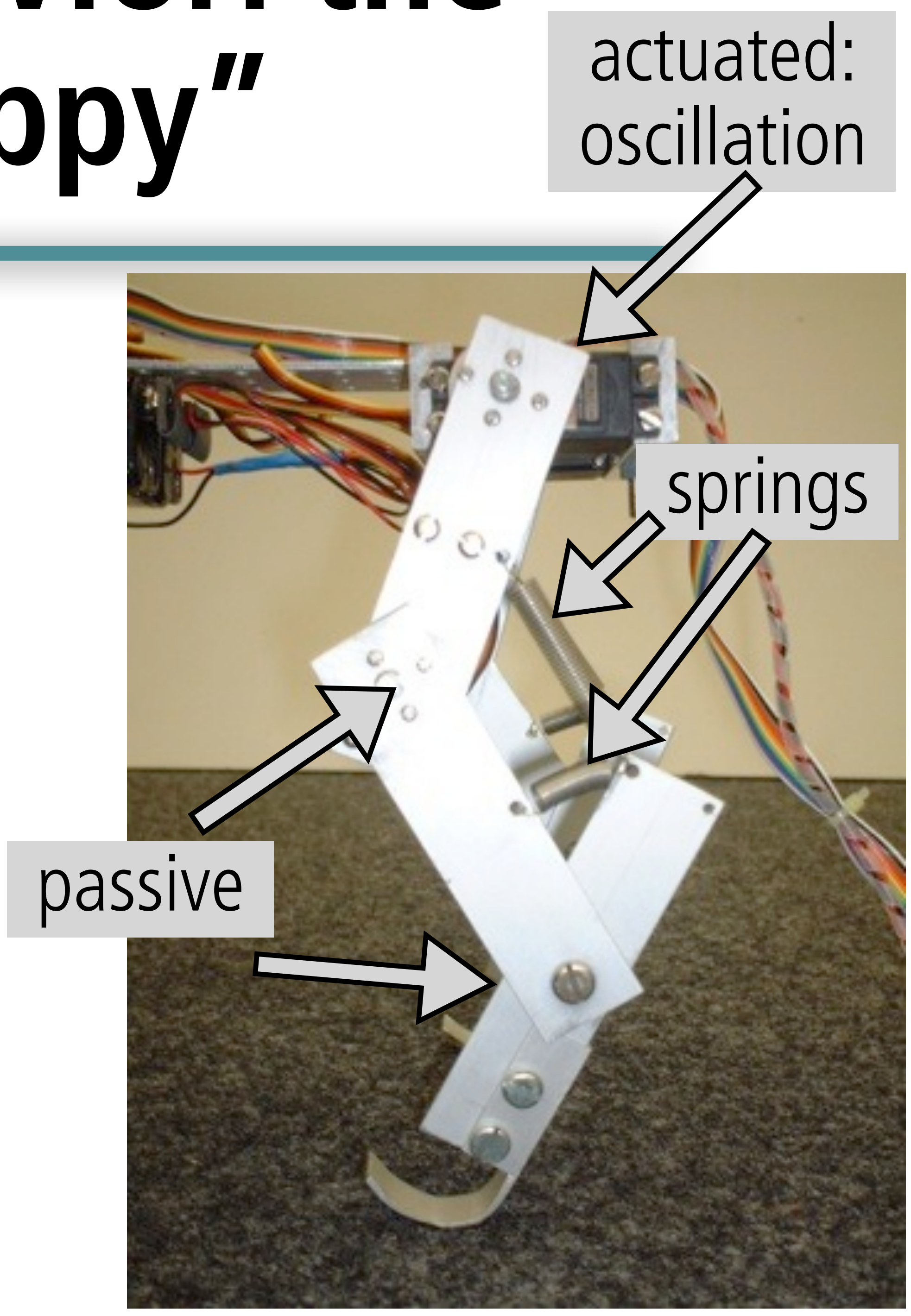
University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

Emergence of behavior: the quadruped "Puppy"

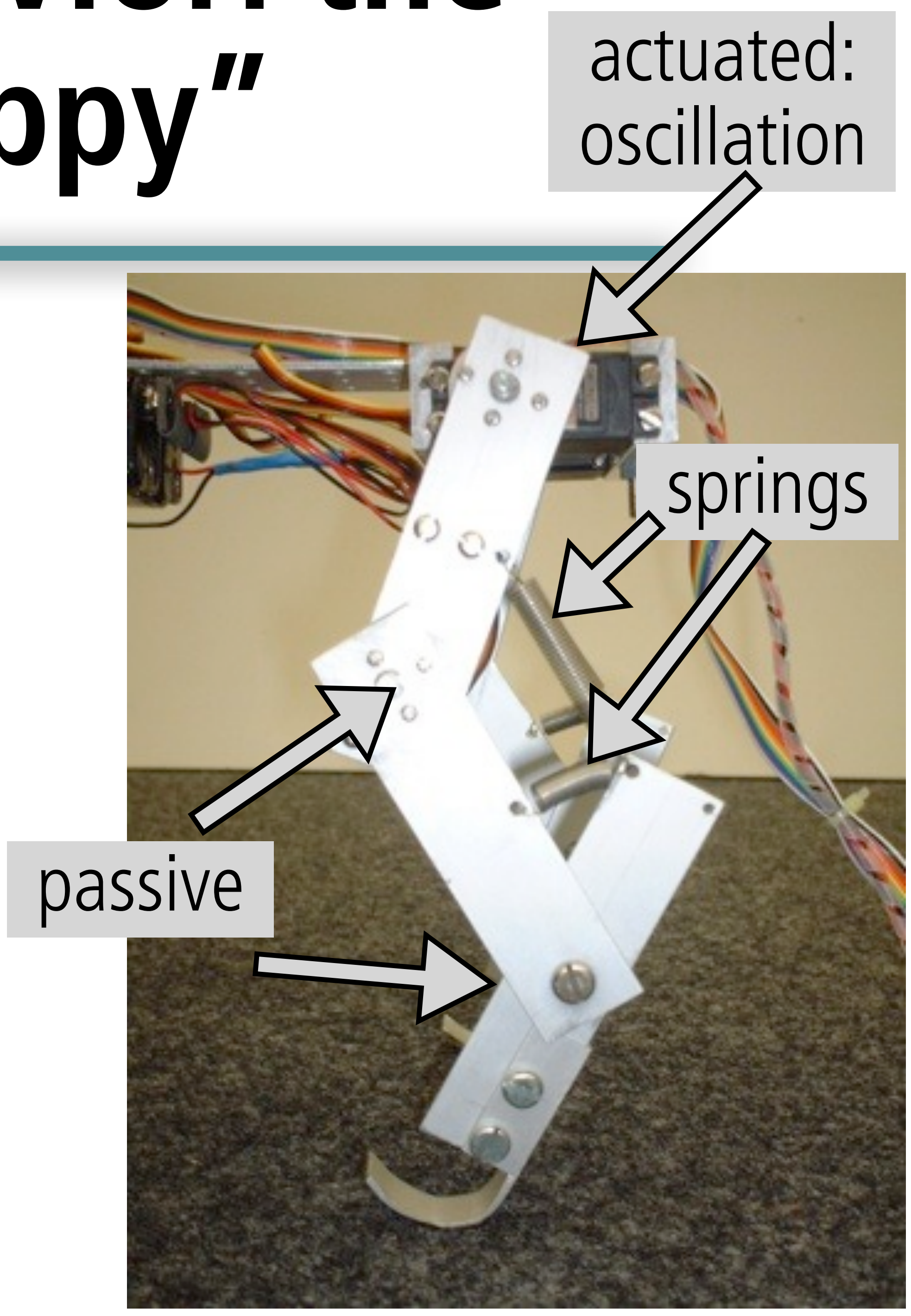
- simple control (oscillations of "hip" joints)
- spring-like material properties ("under-actuated" system)
- self-stabilization, no sensors
- "outsourcing" of functionality



Emergence of behavior: the quadruped "Puppy"

- **simple control (oscillations of "hip" joints)**
- **spring-like material properties ("under-actuated" system)**
- **self-stabilization, no sensors**
- **"outsourcing" of functionality**

→ morphological computation



University of Zurich

robotics Swiss National Centre of Competence in Research

36

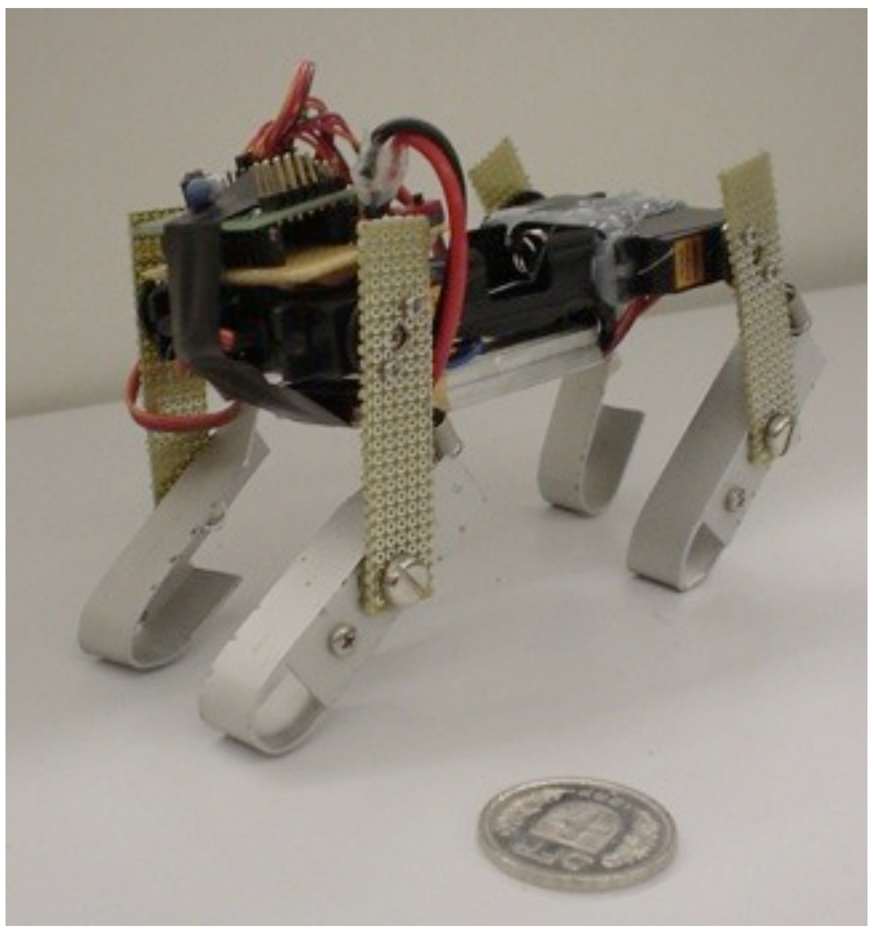
ai lab

The ShanghAI Lectures 人工智能 上海授课

In this example, only the "shoulder" or "hip" joint is actuated, the others are passive but connected through springs.

The concept of morphological computation is about outsourcing functionality to morphological and material characteristics of the organism, e.g. elastic damped springs, arrangement of arrays of light sensors, passively deformable tissue, etc. Many examples will be given during the lectures.

"Puppy" video



Video "Mini-dog"

Design and construction: **Fumiya Iida, ETH-Z (formerly AI Lab, University of Zurich)**

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

Emergence of behavior in "Puppy"

- control: oscillation
- only "hip joints" driven, other joints passive
- trajectories of passive joints: not programmed into robot (but they do the right thing)

→ exploitation of passive dynamics

joints self-organize into proper trajectory (without being directly controlled). The fact that the joint is not directly controlled doesn't imply that it's not doing the right thing.

Emergence of behavior in "Puppy"

why emergence?
—> RGGU Moscow

ation

- only "hip joints" driven, other joints passive
- trajectories of passive joints: not programmed into robot (but they do the right thing)

→ exploitation of passive dynamics

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课

Today's topics

- short recap
- the "Chinese Room Thought Experiment" (student presentation from NPU, Xi'an)
- Braitenberg vehicles (for reading)
- the "Swiss Robots"
- prerequisites for a "theory of intelligence"
- **intelligent systems: properties and principles (mostly lecture 4)**
- **Guest lectures by Alex Waibel (CMU and KIT) and Vincent Müller (American College, Tessaloniki, and Oxford Univer.)**

Intelligent systems: properties and principles

"How the body …", chapter 4

- **complete agents: embodied, situated, autonomous, self-sufficient**
- **real worlds vs. virtual worlds (see lecture 2)**
- **dynamical systems**
- **properties of complete agents**
- **agent design principles**

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

The Japanese psychologist Masanao Toda introduced the so-called "Fungus Eaters", creatures that were sent to a distant planet to collect uranium ore. Because it was so distant, it could not be directly controlled, it had to be autonomous. It had to learn about the environment from its own perspective, i.e. it had to be situation. Because there was nobody to replace the batteries or to repair it, it had to be self-sufficient. Finally, because it had to survive on the real planet and collect ore, it obviously had to be embodied.

Real worlds vs. virtual worlds have been discussed in lecture 2

Agents can be viewed as complex dynamical systems, because they are physical systems. This implies that we can apply the concepts and the terminology of the field, which provides a highly productive and intuitive metaphor for talking about intelligent systems.

Complete agents have a number of properties, see p. 95 of "How the body …".

For agent design principles, see chapter 4.

Design principles for intelligent systems

Principle 1: Three constituents principle

Principle 2: Complete agent principle

Principle 3: Parallel, loosely coupled processes

Principle 4: Sensory-motor coordination/ information self-structuring

Principle 5: Cheap design

Principle 6: Redundancy

Principle 7: Ecological balance

Principle 8: Value

Assignments for next week

- Read chapter 4 of "How the body …"
- Read materials for self-study (Braitenberg vehicles)

University of Zurich

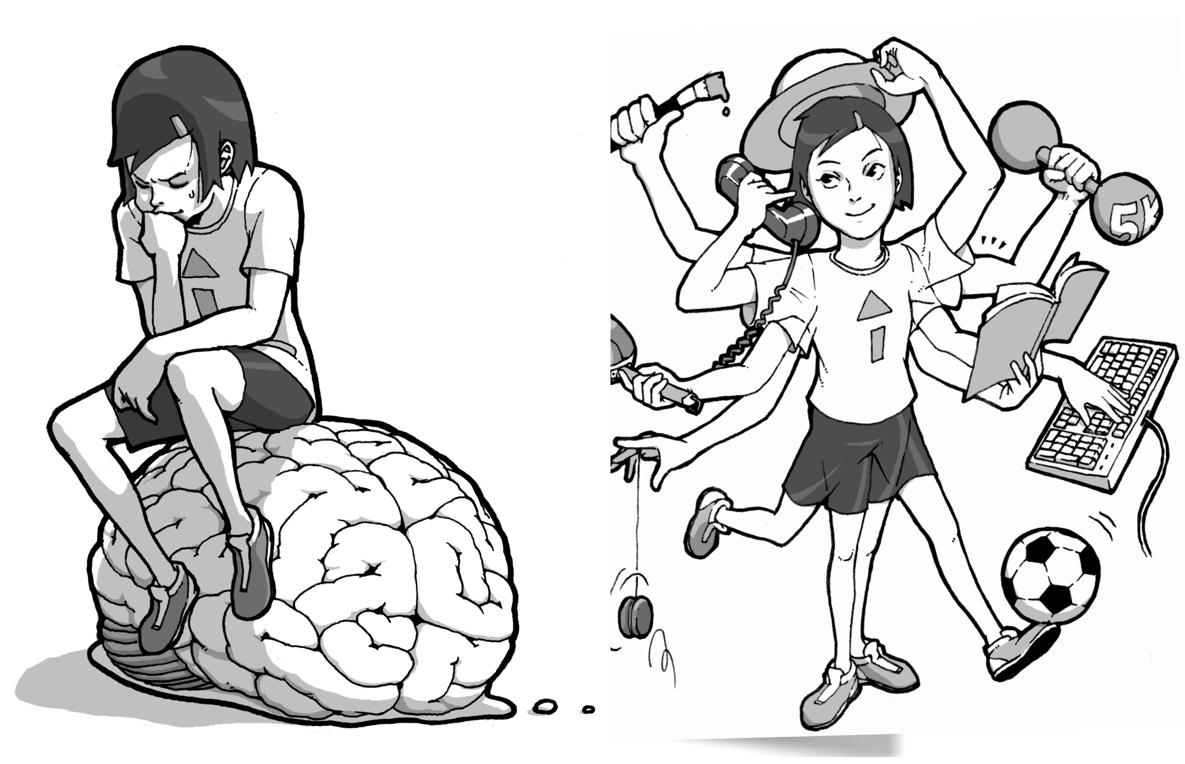
End of lecture 3

Thank you for your attention!

stay tuned for guest lectures

by Prof. Alex Waibel and Prof. Vincent C. Müller

University of Zurich



lab

Guest speaker

from Karlsruhe Institute of Technology and CMU

Prof. Alex Waibel
(Automatic speech translation)

Today, 10.00 CET (9.00 GMT)

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 上海人工智能授课

Prof. Alex Waibel is the world champion in speech recognition and machine translation. He is also a very successful businessman, who has founded several companies.

Dr. Alexander Waibel is a Professor of Computer Science at Carnegie Mellon University, Pittsburgh and at the Karlsruhe Institute of Technology (KIT) Germany. He directs InterACT, the International Center for Advanced Communication Technologies at both Universities with research emphasis in speech recognition, language processing, speech translation, multimodal and perceptual user interfaces. At Carnegie Mellon, he also serves as Associate Director of the Language Technologies Institute and holds joint appointments in the Human Computer Interaction Institute and the Computer Science Department.

Guest speaker

from American College, Tessaloniki Greece

Prof. Vincent C. Müller
"Computers Can Do Almost Nothing – Except Cognition (Perhaps)"

Today, 10.30 CET (9.30 GMT)

University of Zurich

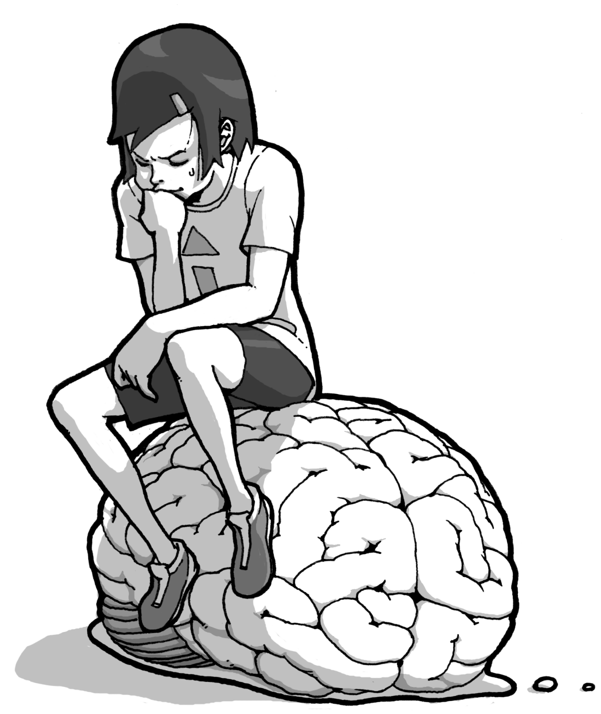
Prof. Mueller is coordinating the EU-Cog network and is interested, among other things, in the philosophy and theory of artificial intelligence.

End of lecture 3

Thank you for your attention!

stay tuned for lecture 4

“Design principles for intelligent systems, Part I”

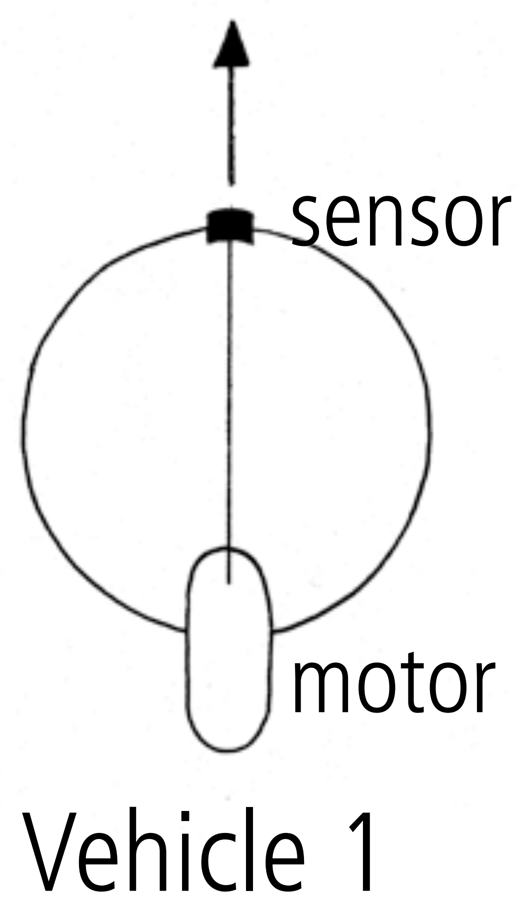


University of Zurich

lab

Materials for self-study

Braitenberg vehicle 1



quality: e.g. temperature

medium: water

wire: the higher the temperature the faster the motor

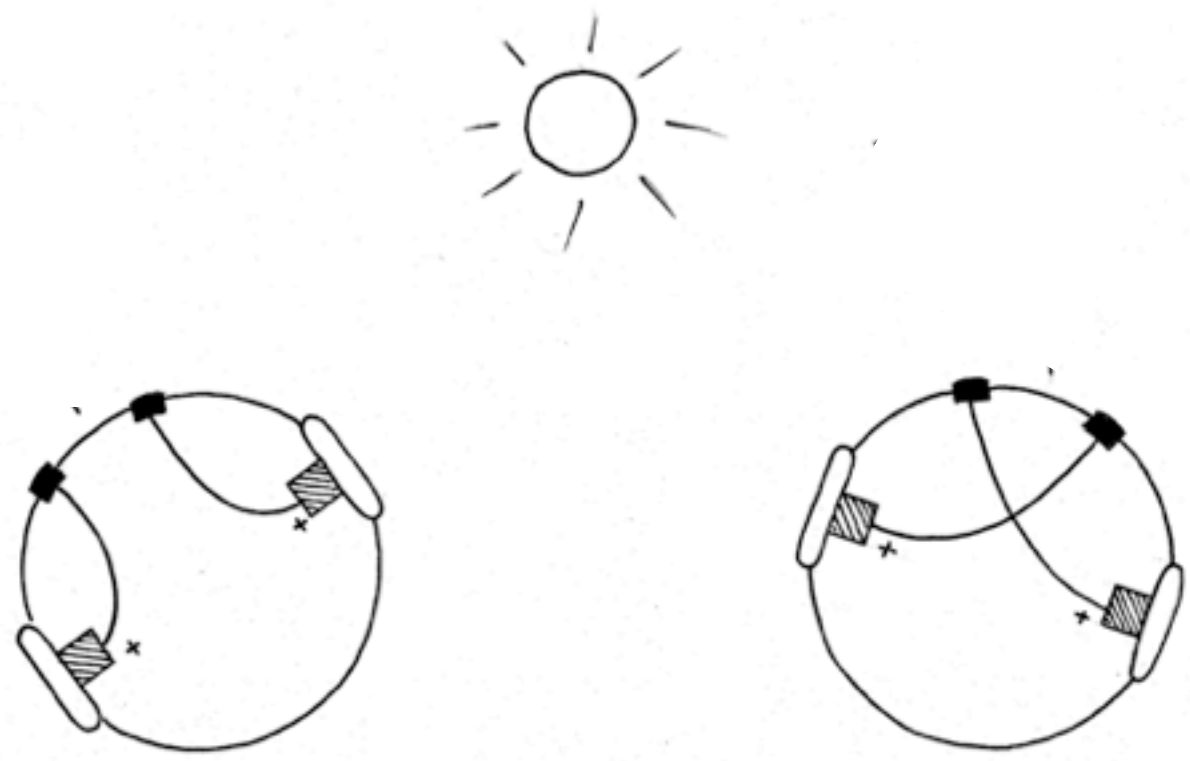
If the vehicle functions in the water and the quality is temperature, how will the vehicle behave?

Braitenberg vehicle 2

quality: light intensity

medium: flat ground

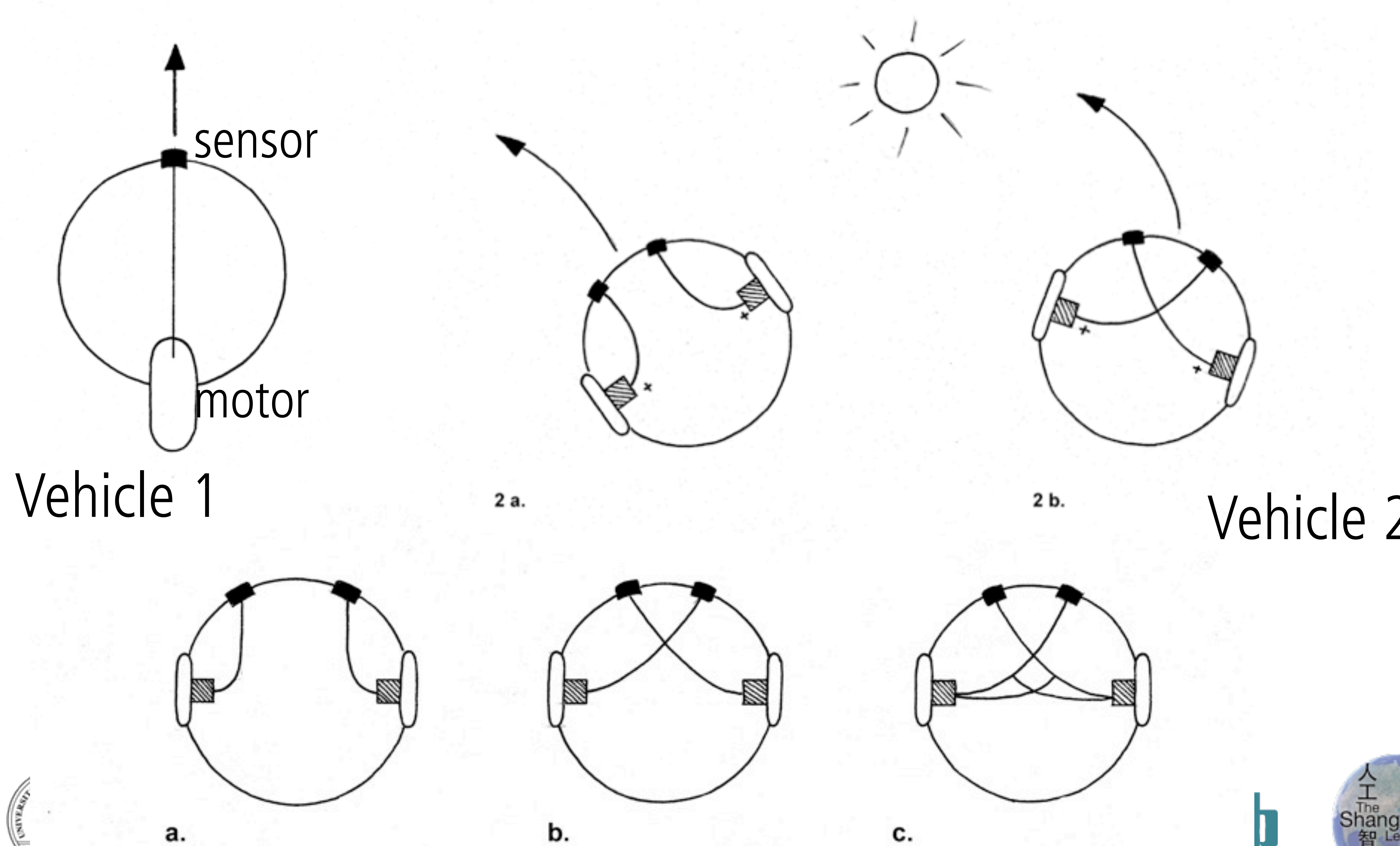
wires: brighter
—> faster



2 a. 2 b.

In the arrangement shown in the slides, what will the vehicles do?

Braitenberg vehicles



Please read about the Braitenberg vehicles:
Braitenberg, V. (1986). Vehicles: Experiments in synthetic psychology. MIT Press.
or
Chapter 6 in “Understanding intelligence”.

Braitenberg vehicles represent a series of agents of increasing complexity. Although some are purely reactive, others include learning mechanisms, and thus have their own history.

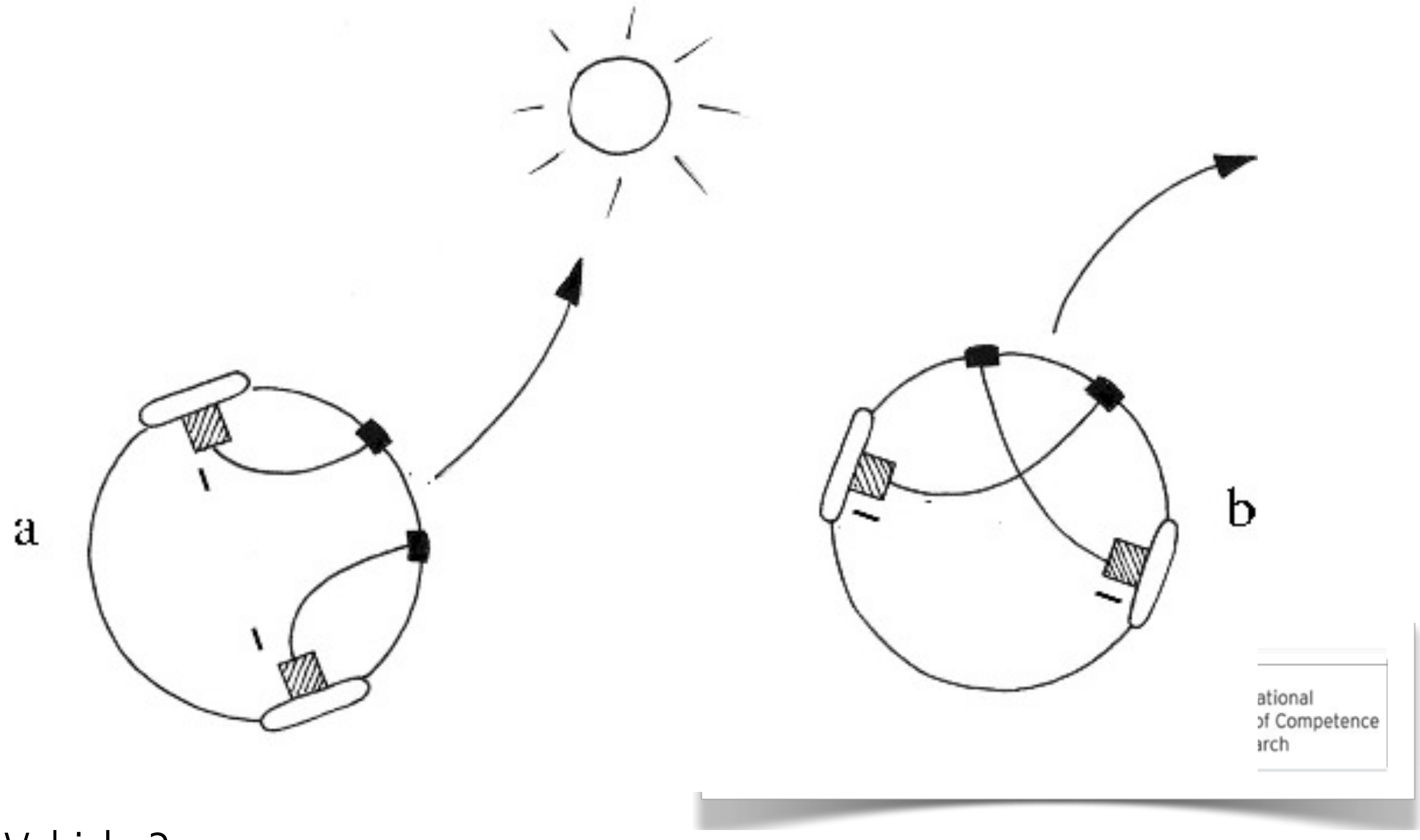
In the simplest vehicles, it is quite obvious what they do. As matters get slightly more involved, predicting their behavior turns out to be very difficult, even in purely reactive systems, because the mechanisms generating the vehicles' behavior interact in interesting ways. Even if we have complete knowledge of the vehicle's insides, it still proves difficult to control it. Its interaction with itsenvironment adds complexity, which leads to some degree of unpredictability, even if the driving mechanisms are entirely deterministic -- in physics, there are always fluctuations. Let us examine the Braitenberg vehicles one by one. As always, we pay attention to the frame-of-reference problem. In examining this series of vehicles, it is always a good idea to imagine how they move around under various conditions. This process of imagination is best complemented with computer simulations or with experiments on real robots.

Vehicle 1
The first Braitenberg vehicle has one sensor, for one particular quality, and one motor. The sensor and the motor are connected very simply: The more there is of the quality to which the sensor is tuned, the faster the motor goes. If this quality is temperature, it will move fast in hot regions and slow down in cold regions. An observer might get the impression that such a vehicle likes cold and tries to avoid heat. The precise nature of thisquality does not matter; it can be concentration of chemicals, temperature, light, noise level, or any other of a number of qualities. The vehicle always moves in the direction in which it happens to be
pointing.

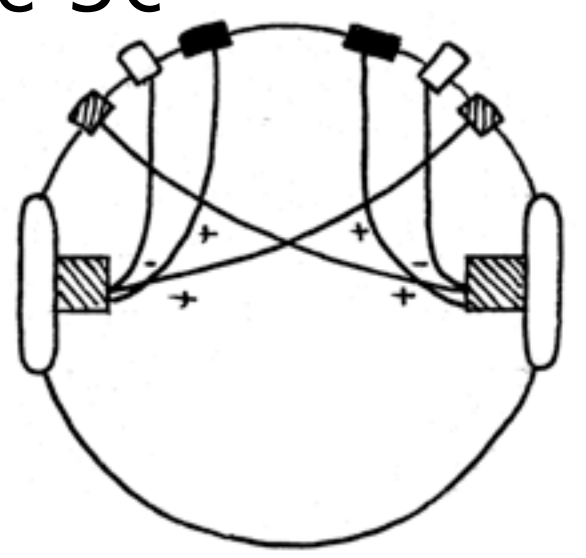
If we introduce friction into the vehicle's environment, its behavior gets interesting, because friction is always a bit asymmetric. The vehicle eventually deviates from its straight course, and in the long run, is seen to move in a complicated trajectory, curing one way or another without apparent (to the observer!) good reason. Perturbations

Braitenberg vehicles

Vehicle 3



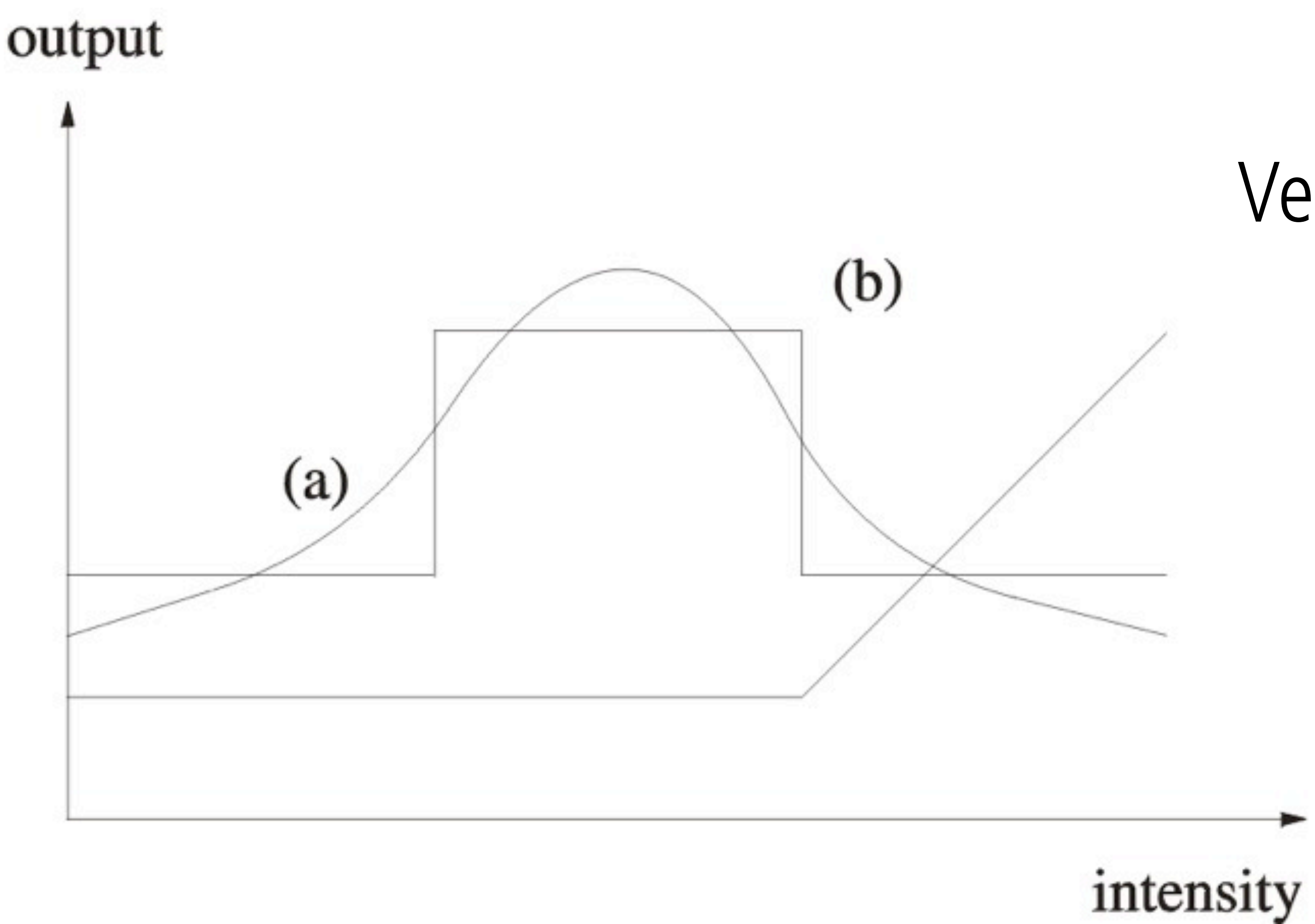
multisensorial
vehicle 3c



Vehicle 3
The first two Braitenberg vehicles have only excitation: the more stimulation at the sensors, the more the motors are powered. Let usnow introduce inhibition: the more stimulation, the less power is delivered to the motors. This principle is incorporated in Braitenberg vehicle 3.

The behaviors involved are fairly obvious. Vehicle 3a ends up facing, say, a light source, whereas vehicle 3b turns away from it but also remains near the source, unless there is a disturbance, like another source. Additional sensors can also be introduced, and each stimulus can be connected either to the motor on the same or the opposite side, and can be excitatory or inhibitory (see figure 6.6). Stimuli to which the sensors are attuned could be light, oxygen concentration, temperature, concentration of organic molecules (food), or similar things. The vehicle has a tendency to stay longer in certain areas than in others because when its sensors are activated by the presence of a stimulus, its motor and thus its movement are inhibited. We cannot help admitting that the vehicle appears to have a set of ``values'' and that it incorporates them in some way that we would want to call ``knowledge.'' ``Knowledge'' in this context does not mean ``stored representations;'' that is, it is not--as in the classical AI view--stored in an explicit form to be manipulated by the agent (the vehicle) itself. Rather, it is attributed to the vehicle as a whole by an outside observer. Attributing ``knowledge'' to an agent is a way of describing its behavior--it has nothing to do with the agent's internal structure.

Braitenberg vehicles



Vehicle 4

Vehicle 4 has non-linear collections from sensors to motors. For example, up to s certain light intensity, the vehicle is still, and if the light intensity then further increases, the vehicle starts to move (as if it were waking up). Still, it is a purely reactive system: it does not have its own history; that is, it does not change over time. Nevertheless, it looks very much like an autonomous agent. If it has many sensors and they are connected in complex ways to the motors, it would in fact be very difficult to control the agent's behavior.

Vehicle 5
We can now add arbitrary complexity by introducing threshold devices. In chapter 5 we called these ``devices'' nodes or model neurons. The kinds of nodes suggested here are of the linear threshold variety, but they could also be of the sigmoid type. They can either be interposed between sensors and motors or connected to each other in various ways. A vehicle possessing these devices is of Braitenberg type 5. These kinds of models are also called artificial neural networks and will be further discussed in lecture 5.

Braitenberg vehicles: Vehicles 6 and beyond

- **Vehicles 6: evolution**
- **Vehicles 7 to 14: increasing the internal neural connections**

Vehicle 6:Evolution
Suppose we put a number of vehicles that we have built on a table containing light sources, sounds, smells, and so forth and let them move around. We pick out one vehicle, the model, make a copy of it, and put both the model and the copy back on the table. We pick out another, and repeat the process indefinitely. Of course, we do not choose vehicles that have fallen on the floor, because they are obviously incapable of coping with this particular environment. We produce vehicles at a pace that roughly matches the rate at which vehicles fall off the table.

If we play this game in a hurry, we are likely to make mistakes now and then. A well-tested vehicle might still fall off the table.
Particularly shrewd variations might also be introduced unwittingly into the pattern of connections with the result that our copy survives, whereas the original may turn out to be unfit for survival after all. If the imperfect copying results simply from sloppiness, the chances that something interesting will emerge because of the mistakes in copying are small. However, a ``better'' sort of error would involve creating new combinations of partial mechanisms, and structures such as IR sensors, cameras, motors, or wheels, each of which has not been disrupted in its own well-tested functionality. Such errors have a much greater chance of transcending the intelligence of the original plan. If these ``lucky'' incidents live forever, they will have many descendants, because they and their descendants will frequently be chosen for copying simply because they stay on the table all the time.

This is, of course, a model of Darwinian evolution. It reminds usof the metaphor of the blind watchmaker, created by Richard Dawkins (1988) to describe evolution. Vehicles created in such a scenario are said to be Braitenberg type 6. We may, by accident, create vehicles whose behavior is extraordinary without understanding why they behave as they do, because building something that works is typically much easier than analysis: Braitenberg called this the ``law of uphill analysis and downhill invention.'' Indeed in evolutionary approaches you can quite often get the agents to do what they should do, but it is usually hard to understand why they do what they do. (Evolutionary methods are discussed in detail in chapter 8.)

Other Vehicles
We discuss the remaining vehicles, vehicle 7 to 14, only briefly, because for our purposes, the simple vehicles are more interesting: They illustrate the sensory-motor couplings and how they lead to remarkable behavior. The later ones, especially vehicles 7 and above, have a cognitivistic flavor and are therefore bound to run into the problems the classical approac to AI.

End of materials for self-study

University of Zurich

robotics Swiss National Centre of Competence in Research

ai lab

The ShanghAI Lectures 人工智能 上海授课