Introduction to Cognitive Science

(8: Modules and Architectures)

(Ch. 8)

Overall organization of the mind

- Architectures for the artificial agent
  - Russell and Norvig (excellent book)
- Fodor's modularity of mind
- Massive modularity hypothesis

Architectures

- AI systems are built from subsystems
- Three types of agents are interesting:
  - Reflex agent
  - Goal-based agent
  - Learning agent
- Not all artificial agents are intelligent agents!
- Sensors to see, effectors to interact
- 20 prototypes

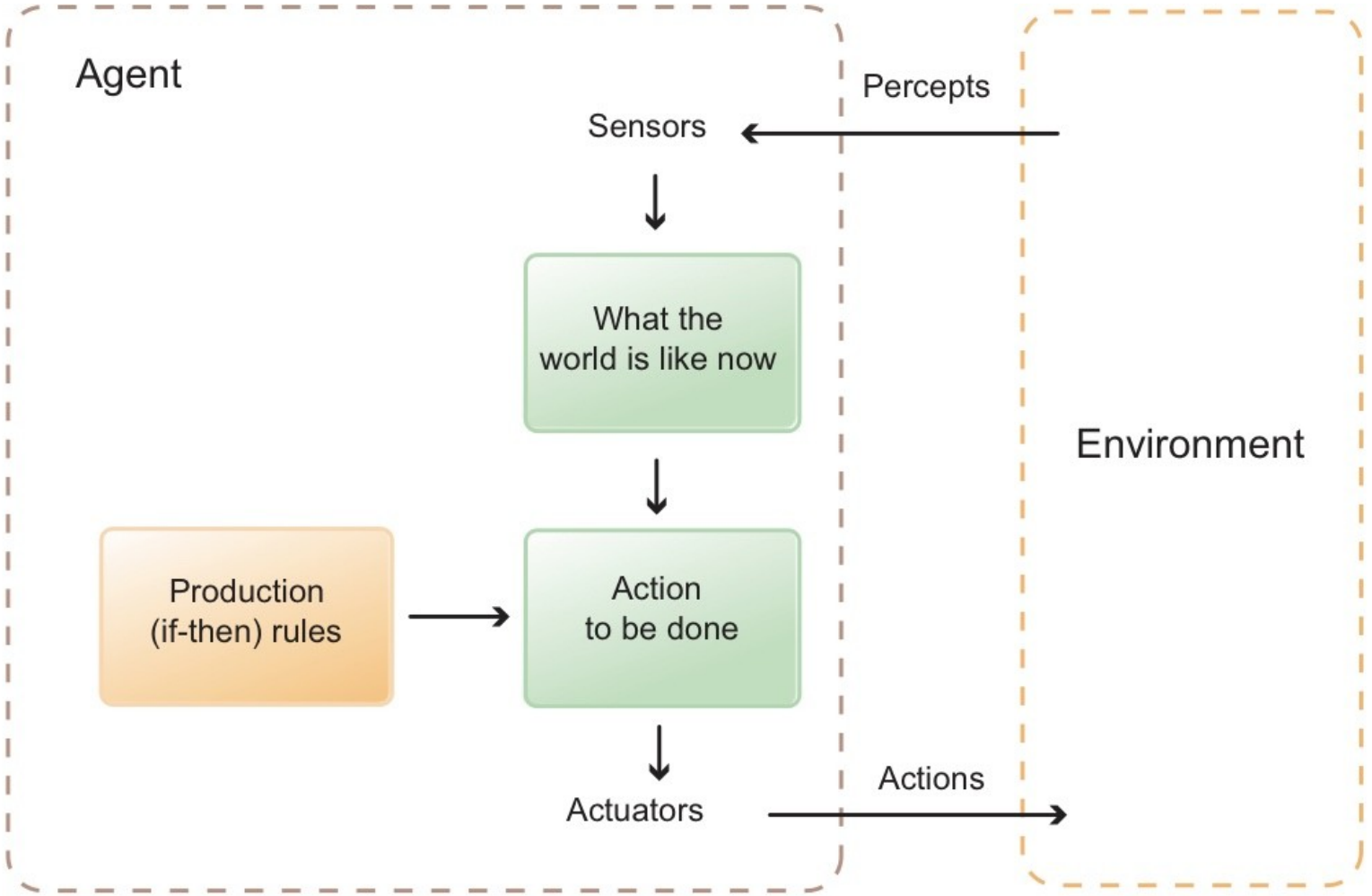

Figure 8.1 The architecture of a simple reflex agent. Production rules are all that intervene between sensory input and motor output. (Adapted from Russell and Norvig 2009)

Reflex agents

- They are not cognitive systems
- They react invariantly
- It is commonly taken that a feature of cognitive systems is that they can react differently in the same scenario depending on „internal feelings”

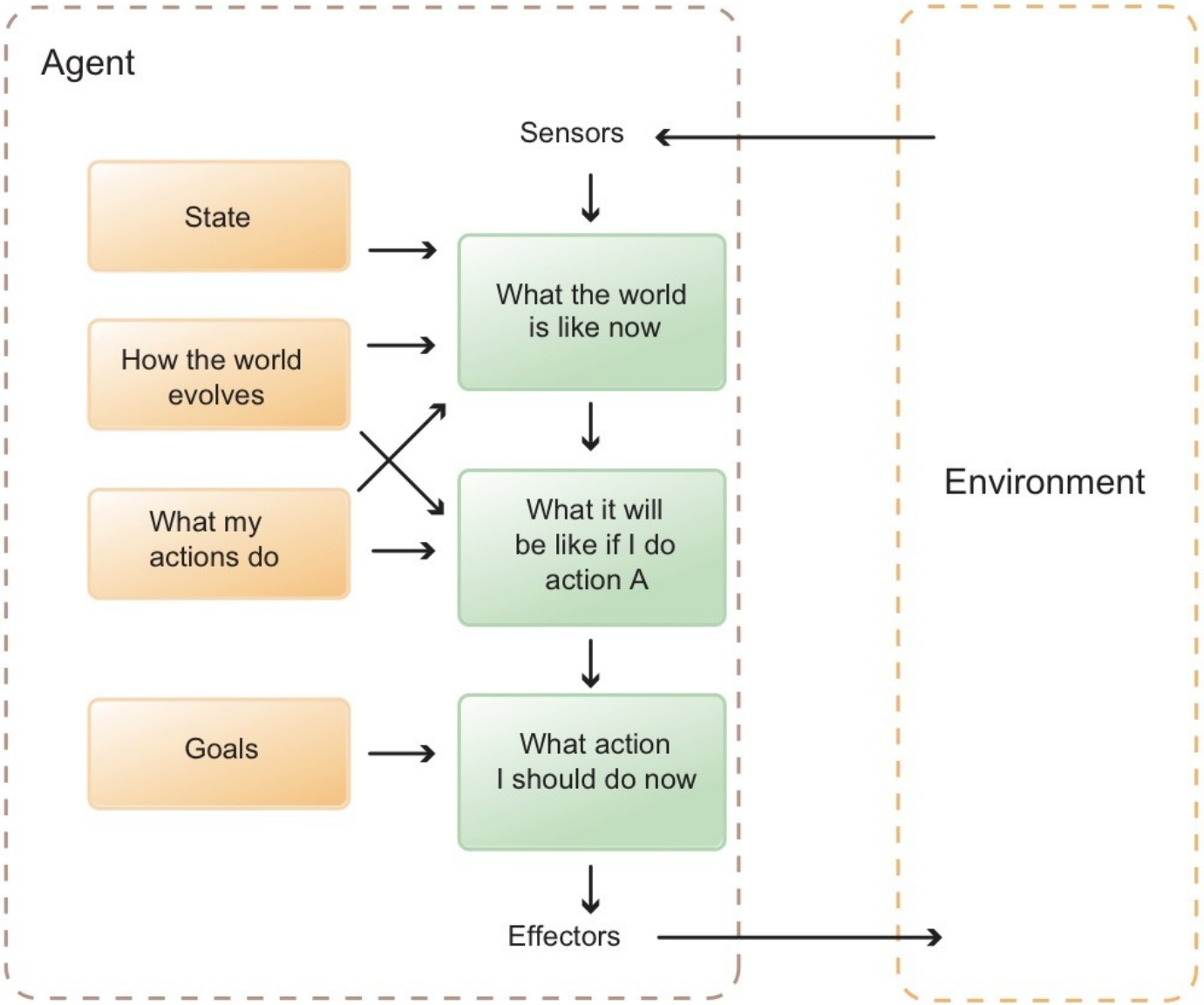

Figure 8.2 The architecture of a goal-based agent. There are information-processing systems intervening between input and output. (Adapted from Russell and Norvig 2009)

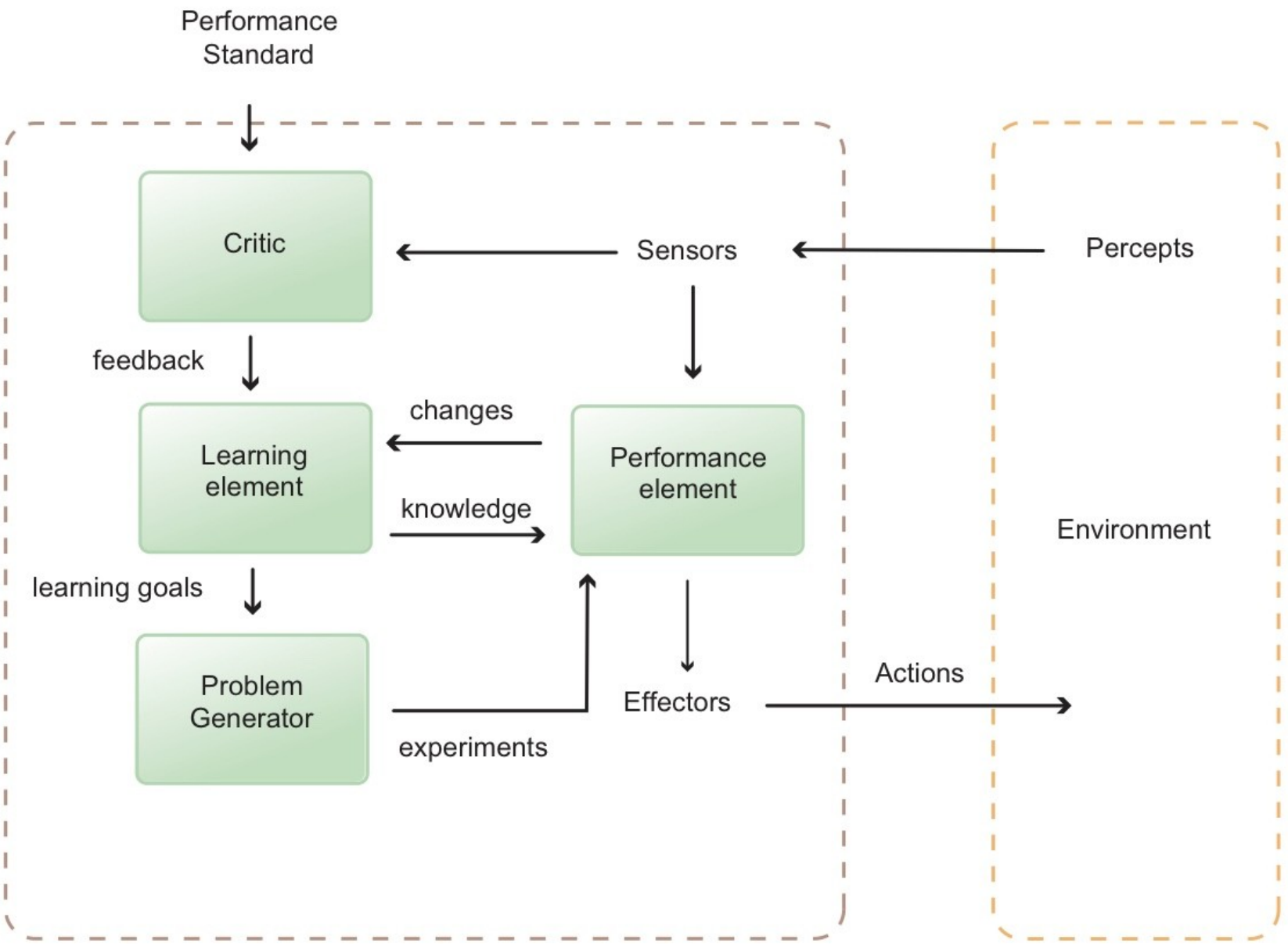

Figure 8.3 The architecture of a learning agent. Mismatches between sensory feedback and the performance standards are detected by the Critic subsystem. The Learning subsystem determines learning goals and allows the system to experiment with different ways of achieving its goals. (Adapted from Russell and Norvig 2003)

Learning agents

- Critic submodule detects mismatches between sensory feedback and the peformance standard
- These mismatches are taken into the Learning subsystem which determines learning goals and makes it possible for the system to experiment/explore
- Even from these simple efforts we get a number of important questions

Questions (solve them all for homework :)

- How can we identify and distiguish cognitive subsystems?
- Are there important diff between sensoryProc-motorBehaviour and Input-Output subsys?
- Do all subsys process information in the same way? Same representations?
- How „autonomous” are different subsys?

Modularity of mind

(Finally an interesting theory from Fodor)

- Modules! Monads?
- Nonmodular processes are high-level, open-ended, involve a large number of elements
  - Remember the compositionality assumption for processes (Process ontology paper)

Modules

- Domain-specific: this is not really precise, but it depends on the ontology behind it (detecting eyes or detecting staring)
- Informational encapsulation: modular processing is not affected by what is going on elsewhere in the mind
- Mandatory application: modules respond automatically to stimuli
- Speed
- Fixed neural architecture: fusiform gyrus is believed to do facial recognition
- Specific breakdown patterns: modular processing fails in highly determinate ways

Modules

- Perception and motor skills are most easily explained via modules
- Fodor also insists on something interesting:
  - There are psychological processes that cut across cognitive modules
- These processes form what Fodor calls „central processing”
  - He calls modular processing „peripheral”

Fodor's two feats of central proc

- Quinean:
  - An organism's belief system is like a scientific theory:
    - Empirical subdetermination:

      \Wedge_i T \to O
      \neg O

      ?
- Isotropic: not informationally encapsuled
- Influential movements in cogSci are skeptical on the importantance and even existence of central processing

Massive modularity hypothesis

- All information processing is modular, no central processing
- Darwinian, unlike Fodorean modules are not informationally encapsulated
- Darwinian modules:
  - Cheater detection
  - Kin detection
  - Mate selection
- But also:
  - Face recognition
  - Emotion detection
  - Gaze following
  - Folk psychology, Folk physics, Folk biology

Cheater detection module (Darwinian module)

- If a card has a vowel on one side then it has an even number on the other side
- If a person is drinking beer then that person has to be over 18

Evolution of cooperation

- Why should there be a cheater detection module?
  - Evolution favors free riders and exploiters above altruism
  - A popular way to an explanation is the prisoner's dillema (see Bermudez)
  - A simpler explanation is that humans are not self-sustainable, and need communities to survive
    - Cheaters damage communities (take more than the fair share), therefore it makes sense to weed them out