III. Intelligent Agents

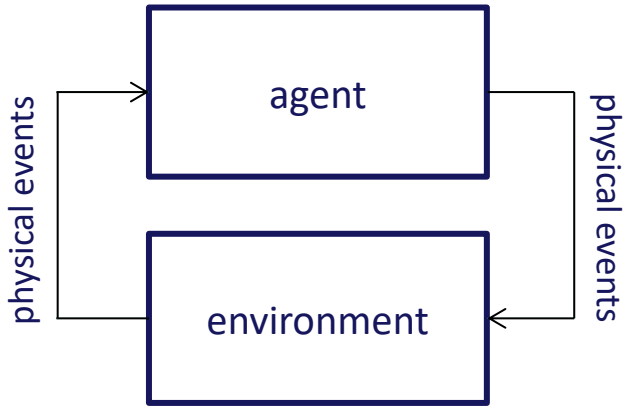
Intelligent agents

- In Part II we have seen that **intelligent** is a qualification that we are going to apply to **agents**; but then we have to answer a number of questions, in particular
  - what is an agent? are there different types of agents?
  - are there different types of agents?
- Intuitively, an agent is **a system that is able to act**
- What kinds of systems do we have in mind?
  - **biological organisms** endowed with a natural mind, in particular human beings (whom, quite immodestly, we consider as the ideal model of an intelligent agent)
  - **software systems**, e.g., pure “artificial minds” acting in the Web
  - **robotic systems**, endowed with both an artificial mind and an artificial body, e.g., autonomous vehicles
- Acting on what?
  - in general, an agent performs an action on its **environment**: therefore, what we call an action is actually an **agent-environment interaction**
  - (there are some exceptions, like performing a calculation mentally, which we can consider as an action that is not performed on an evironment)

Agents: first level of analysis

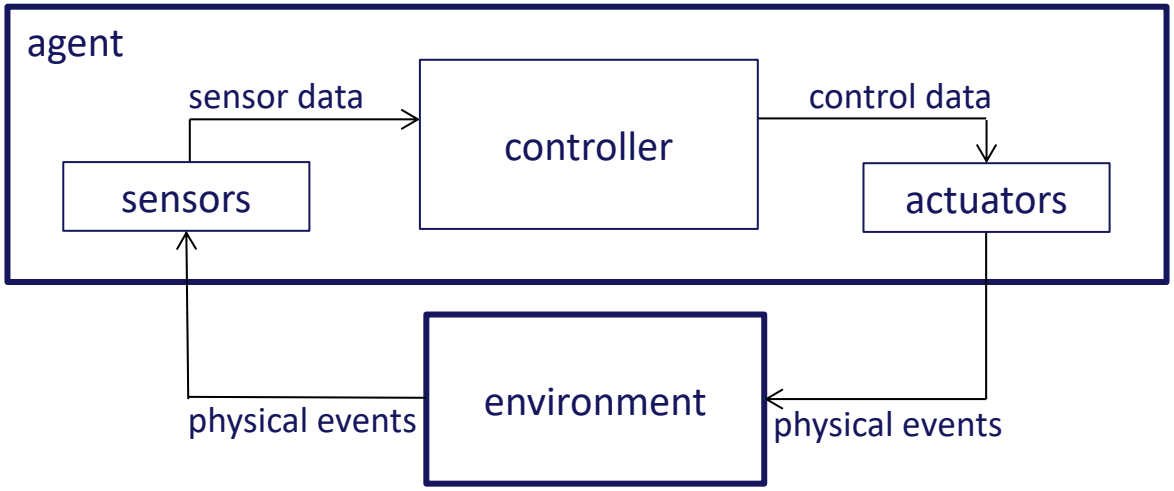
- To give a more general idea of what an agent is, we shall first consider agents that have both a “mind” and a “body” (biological organisms, robots); a pure software agent can then be viewed as a special case, in which there is no “body” properly so called
- The interaction between an agent and its environment consists of two streams of **physical events** (one in input to the agent, the other one in output; examples:
  - input: the photons reaching the agent’s visual sensors
  - output: the acts of physically grasping an object present in the agent’s environment
- By definition, an agent’s environment is that part of the world with which the agent can interact, in that:
  - it can influence the agent’s actions
  - it can be influenced by the agent’s actions

  Note that an agent’s environment may include elements that are internal to the agent’s “body” (e.g., a robot’s battery)



Agents: second level of analysis

- The function of an agent is to **control its environment**, in order to **achieve certain goals**
- At the interface with its environment, an agent has to:
  - transform the input physical events into input data suitable for internal processing performed by what we can call the agent's **controller**
  - tranform the output data of internal processing into the output physical events
- These operations are performed by two types of devices, called **transducers**:
  - input transducers, called **sensors** (e.g., a camera or a microphone), and
  - output transducers, called **actuators** or effectors (e.g., a robotic gripper)



Agents: third level of analysis

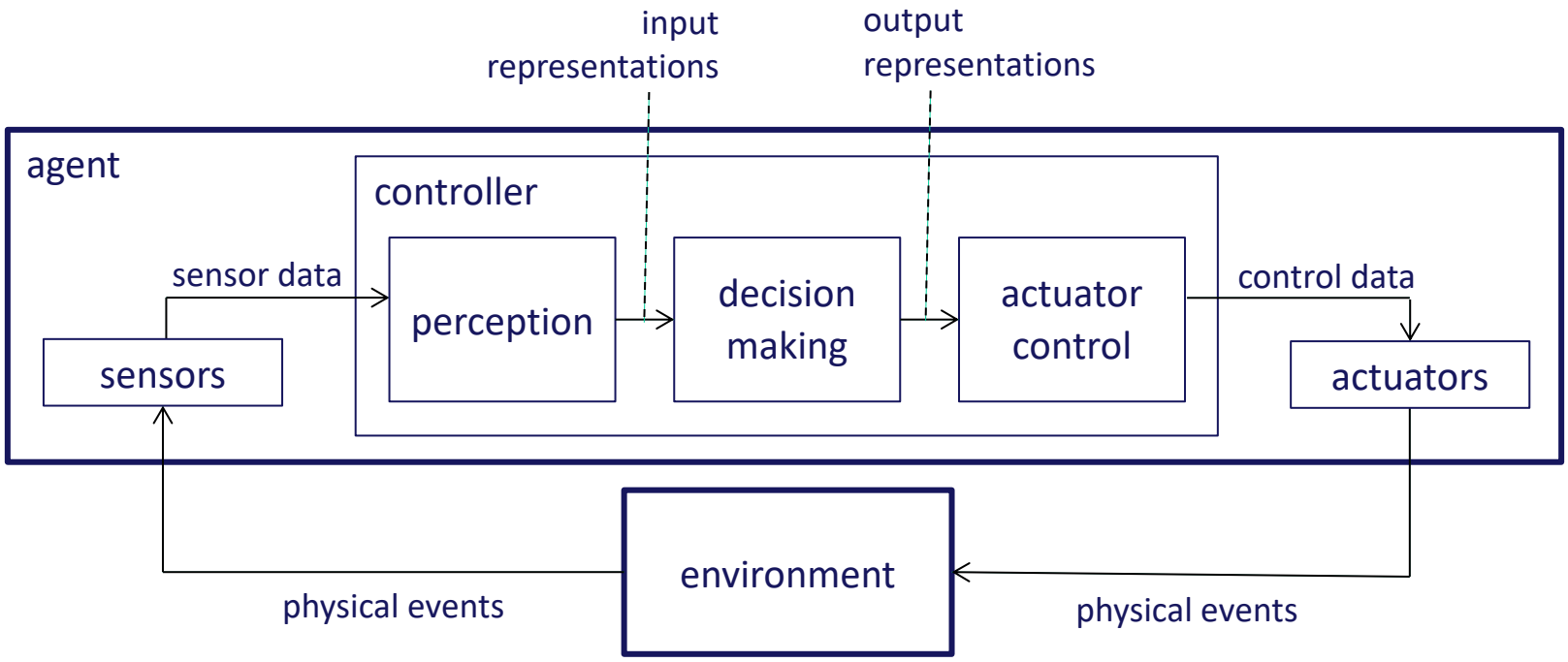
- The main function of the agent's controller is to **decide what to do** in order to fulfill the agent's goals, given the current state of the environment (and possibly other types of information)
- Often, such decisions cannot be taken directly from an inspection of the sensor data; rather, the sensor data need first to be interpreted

  For example, it may be relevant to an autonomous vehicle's decision whether the "thing" that suddenly appeared in front of it is or is not a human being

  But the sensor data generated by the vehicle's front camera is just an **array of pixels**: to understand whether this array is the **visual representation of a human being** requires a process of interpretation, that in this case is called visual perception
- In general, **perception** is a process that transforms sensor data into higher-level representations of the current state of the environment, which are suitable inputs for the agent's decision module

Agents: third level of analysis (2)

- The agent's decision module will output high level representation of what to do, like "brake" or "turn right"
- Symmetrically to the case of perception, such representations will have to be translated into control data to drive the actuators



What the analysis actually means

- The previous diagrams are intended to present the main **functions** that an agent has to implement in order to interact with its environment
- Such diagrams, on the contrary, are not meant to represent the software architecture of an artificial agent: in particular, they do not imply that an artificial agent has to be implemented as a pipeline of perception/decision/actuator-control processes
- In the rest of this course, we shall almost exclusively concentrate on an agent's decision-making function

  Other functions are analysed in details in other AI-related courses, like Machine Learning, Robotics, and Image Analysis and Computer Vision

Types of environments

- As we have already said, an agent's environment is the part of the world with which the agent can interact

  We should regard the environment not as "objectively given", but as "subjectively viewed" through the eyes of the agent

- Agent-environment interactions take place in **time**, whose flow is usually conceived to be continuous; however, an artificial agent will typically consider time to be discrete sequence of time instants t_0, t_1, t_2, and so on

- An environment (again, as viewed through an agent's eyes) may be
  - static or dynamic
  - deterministic or stochastic
  - fully or partially observable

  Moreover, an agent's environment may or may not contain other agents

Simple reflex agents

- In view of our goals it is important to classify agents according to the complexity of their decision-making function

  Different terminologies are used in the scientific literature; here we adopt the terminology of RN

- **Simple reflex agents** are agents whose decisions at time *t* are based on the agent's perceptions at time *t*

  Examples:

  - a typical domestic robot vacuum, whose decision-making module is a pre-programmed stateless controller
  - agents whose decision-making module is a stateless controller resulting from a process of Machine Learning

Model-based reflex agents

- Simple reflex agents are stateless and thereore have no memory of past perceptions and actions
- Many types of behaviours cannot be implemented by a simple reflex agent, because relevant information about the environment may not be currently accessible through the agent's sensors
- Model-based reflex agents maintain a model that represents aspect of the environment that are not currently accessible through the sensors
- Examples:
  - a domestic robot vacuum that has a map of the environment that must be kept clean and remembers which places have already been cleaned recently
- In any case, an agent of this type is still reflex, in the sense that its next action will be completely determined by the current perceptions and the current model ot the environment

Goal-based agents

- Reflex agents (either simple or model-based) are realised in such a way that they pursue some fixed goal (e.g., to keep an apartment clean)
- A goal-based agent is an agent that can, within limits, select what goals to pursue and how to pursue them; for example, an intelligent autonomous vehicle may be assigned a very general goal (e.g., “bring me to a nice restaurant”), while being free to choose how to realise such a goal
- Goal-based agents are the strong point of Classical AI, and for this reason we shall devote most of our effort to them

The textbook

- From our textbook RN, read Chapter 2 “Intelligent Agents”, but do not pay too much attention to the schematic diagrams of different types of agents (Figures 2.1, 2.9-2.15), which to the lecturers of this course appear to be at least partially misleading