Dipartimento di elettronica e informazione
Politecnico di Milano

Artificial Intelligence 2010-11

3. Representations

3.1 What is a representation?

As we have seen, an agent's decision module operates on representations: it receives in input perceptual representations of the current state of the environment, it may exploit representations of unperceivable aspects of the environment (models) and representations of goals, and it produces in output representations of actions to be executed on the environment. But, *what is a representation*?

Human beings use many different kinds of representations, either internal (in their minds) or external (as reminders or for communication). Internal representations are implemented in the nervous system (we still do not know how). External representations may take many different forms: sounds (e.g., speech), gestures (e.g., waving one's hand to say hallo), written language, drawings, technical diagrams, and so on.

Content

Independently of their forms, there is something that we can say about representations in general. First, representations are always *representations of something*: that is, they have a *content*, in that they represent certain *states of affairs* (objects, facts, events, etc.). For example, the sentence "it will be sunny in Milan next Sunday" represents a state of affairs that every speaker of English can easily figure out.

Mode

Second, every representation has a *mode*. For example, the sentence "it will rain in Milan next Sunday" may express an expectation or, maybe, a hope. Different representations may have the same content, but different modes. For example, all the following sentences have the same content (that I meet Barbara tomorrow), but different modes:

- "I will meet Barbara tomorrow": the mode is *expectation*, that is, a belief about the future;
- "I'd like to meet Barbara tomorrow": the mode is *desire*, that is, the representation represents a state of affairs that the agent would like to achieve;
- "I hope to meet Barbara tomorrow": the mode is *hope*, that is, an expectation in conjunction with a desire;
- "I'm afraid I'll meet Barbara tomorrow": a *fear*, that is, an expectation in conjunction with a negative desire;
- "I want to meet Barbara tomorrow": a *goal*, that is, an state of affairs that the agent will actively try to bring about.

Aspect

The same state of affairs may be represented under different *aspects*. For example, a situation in which thirty students are attending a lecture may be described as:

- "thirty students are attending a lecture";
- "thirty people are listening to a guy who goes on talking for about one and a half hour";

月 Moon 

a b c

Figure 1. Three representations of the Moon.

- “thirty animals (of the species *Homo sapiens*) are sitting in a room, their hearts beating at a pace of 60 to 90 beats per minute”;
- “thirty physical systems are producing heat in a closed space of about 350 m^3”;

Given that a representation is used for some purpose, it has to represent states of affairs under aspects that are relevant to such a purpose.

Code

A representation represents something by encoding information about its content and its mode according to some *code*. For example, the three representations of Figure 1 all represent the Moon. Representation 1.a is the Chinese ideogram that directly encodes the *concept* of the Moon: it is used in both Chinese (where it is read differently depending on the local language) and in Japanese (where it is read differently whether it is used as a single word or as a part of a compound word). Representation 1.b encodes the English word ‘Moon’ using Latin letters to encode sounds or, more precisely, *phonemes* (/mu:n/). Representation 1.c is an *icon*, which encodes the concept of the Moon through one of its possible shapes, as perceived from the Earth.

Representations 1.a and 1.b are *symbolic*, in the sense that:

- they are conventional;
- there is no structural relationship between the signs (“月” and “Moon”) and the represented object.

On the contrary, representation 1.c is *iconic* or *analogic*, in the sense that there is a structural relationship between the sign and (our visual perception of) the moon. As we shall see, symbolic and iconic representations have different properties, that deeply affect how they can be processed by computer programs. A difference between the two types of representations should be noted from the very start: iconic representations necessarily encode certain aspects of their content which can be ignored in symbolic representations. For example, Figure 1.c represents a *crescent* moon, not a *full* or a *new* moon, and there is no way to avoid representing this aspect if you use an iconic representation. On the contrary, representations 1.a and 1.b represent the Moon without committing to its being crescent, full, or new.

Procedural vs. declarative representations

Another distinction related to coding is the one between *procedural* and *declarative representations*. A representation is procedural if it is embodied in a procedure that can be executed under suitable conditions; for example, a person’s procedural representation of swimming is embodied in a procedure (somehow encoded in the person’s neural system) that, if executed in the appropriate conditions, causes the person to swim. On the contrary, a declarative representation is embodied in static information structures that can be accessed for various purposes; for example, a person’s declarative representation of swimming is embodied in a network of concepts that encode what the person knows *about swimming*. The two types of representation are typically used for different purposes; for example, a procedural representation of swimming will be used to swim, and a declarative representation of swimming will be used to answer questions about swimming, to explain to somebody why he swims so badly, to judge whether a certain place is fit for swimming, and so on.

In general, both types of representations are useful and important. Many artificial systems, however, use only (or predominantly) one type of representations. As we shall see, for example, knowledge of actions is based on procedural representations in state space search methods, and on declarative representations in planning methods.

3.2 Representations and rational agents

Remember that our goal is to design the decision module of an artificial rational agent. Intuitively, a rational agent is a system that acts in the attempt to reach some goals, on the basis of its general models of the world and of its current perception of its environment. Such an agent will have to exploit representations ...

- ... whose contents concern the agent's environment ...
- ... represented under suitable aspects, ...
- ... in some mode, ...
- ... and suitably encoded.

Consider for example a chess-playing program. As a whole, the program will implement several functions, including reading inputs from a keyboard and displaying graphics on a monitor. Here we are only interested in its decision module, that is, the function of choosing the move to be performed next. Such a module will have to process different kinds of data structures, representing at least:

- the current state of the environment (the disposition of the pieces on the chessboard);
- the current goal or goals (like, for example, attacking the adversary's Queen);
- general knowledge about the rules of the game, winning strategies, and so on;
- the move to be performed next.

The first three kinds of representations are in input to the decision module, and the fourth is in output. Let us concentrate for example on the representation of the current state of the environment. A typical choice is to represent the 64 places of the chessboard as an 8×8 array, each cell of which contains a symbol, that is, either "empty" or the type and colour (white or black) of a piece. Using the concepts introduced so far we can describe this representation as follows:

- *content*: the current disposition of pieces on the chessboard;
- *mode*: belief (i.e., the agent assumes that the representation faithfully describes the current situation);
- *aspects*: only those aspects that are relevant for the agent's rational decision are represented, that is: the type, colour, and position on the chessboard of all pieces; all irrelevant aspects (like for example the size of the chessboard, the weight of the pieces, and so on) are neglected;
- *code*: the convention of representing the chessboard as an 8×8 array of suitable symbols.

As far as encoding is concerned, the representation is a mixture of iconic and symbolic codes: the chessboard is represented iconically by the 8×8 array, but the pieces are coded symbolically. Indeed, pure iconic representation are never used in AI: all representations are either totally symbolic, or partly iconic and partly symbolic. In other words, only some aspects of the world (if any) are represented iconically. In any case, for the sake of simplicity we shall call "iconic" all representations that include an iconic component, even if they also include a symbolic component.

3.3 Rational action and problem solving

Problem solving is an important subtype of rational action. In problem solving, an agent has:

- knowledge of the current state of the environments
- knowledge of the actions that the agent may perform to change the state of the environments
- a final goal, that is, a representation of certain desired states of the environment (typically different from the current state).

Solving the problem means finding a sequence of actions that, if executed, will change the state of the environment so that the goal is reached.

We shall start our journey in the field of AI by learning well-established methods to design artificial systems that can solve problems in the sense clarified above. Such methods can be roughly classified in two classes: *state space* methods, and *planning* methods. As we shall see, the two types of methods have many similarities, but also crucial differences. It is important to note from the start that such differences are justified by the two different approaches taken to representation: while state space methods are based on

- iconic representations of environmental states, and
- procedural representations of actions,

planning methods are based on

- symbolic representations of environmental states, and
- declarative representations of actions.