A Composite Model of Concept Representation

Paul Davidsson
Department of Computer Science, Lund University, Box 118, S–221 00 Lund, Sweden
paul.davidsson@dna.lth.se

Many different approaches to representing concepts has been suggested, e.g., rule- and exemplar-based descriptions and neural networks. This diversity has led to disagreement about which type of representation is the most appropriate. However, the general opinion seems to be that a single one of these is sufficient to capture most relevant aspects of a concept. This state of affairs might be satisfying if we only wanted to use and learn concepts in restricted domains. However, it is not sufficient when dealing with autonomous agents, natural or artificial, acting in the real world, since they need concepts to serve multiple functions. But which are these functions and how do they influence the representation of concepts?

The Functions of Concepts

It is possible to distinguish several functions of concepts, some of them are: (i) *Stability functions*, concepts give our world stability in the sense that we can compare the present situation with similar past experiences. Actually, there are two types of stability functions, *intrapersonal* and *interpersonal*. Intrapersonal stability is the basis for comparisons of cognitive states within an agent, whereas interpersonal stability is the basis for comparisons of cognitive states between agents. (ii) *Cognitive economical functions*, by partitioning the set of objects in the world into categories, in contrast to always treating each individual entity separately, we decrease the amount of information we must perceive, learn, remember, communicate and reason about. (iii) *Linguistic functions*, provides semantics for linguistic entities, so that they can be translated and synonymy relations be revealed. (iv) *Metaphysical functions*, are those that determine what makes an entity an instance of a particular category. (v) *Epistemological functions*, are those that determine how we decide whether the entity is an instance of a particular category (cf. perceptual categorization). (vi) *Inferential functions*, concepts allow us to infer non-perceptual information from the perceptual information we get from perceiving an entity.

A Novel Framework for Composite Concepts

Most present models of concepts concern representations to be used in some categorization task. Thus, they can be said to serve an epistemological function. What we need is a richer composite representation that, in some way, is structured according to the functions of the concept to be represented. My candidate for such a representation contains five components: the *internal* and the *external designator*, the *epistemological*, the *metaphysical*, and the *inferential components*. I argue that this composite structure enables concepts to serve all the functions listed earlier. The last three support the functions indicated by their names, the internal designator supports the intrapersonal stability, and the external designator supports both the interpersonal stability and the linguistic function.[1]

Let us illustrate the composite representation by the category "chair". For the external designator it is natural to choose "chair" (given that the agent communicates in English). The choice of the internal designator, on the other hand, is entirely up to the agent, it should be as convenient and effective as possible for the agent. The epistemological component could for instance be a 3-D model of a prototypical chair, but any representation that can be used by the perceptual system to successfully identify members of the category would be adequate. Since perceptual classification often is dependent of the situation, it might be convenient to have more than one epistemological representation. Similarly, it is useful to have several external designators in domains where many communication languages are used. The metaphysical component defines that something is a chair if it could seat one person. Finally, the inferential component represents the facts that a chair usually has four legs, is often made of wood and so on. There are no sharp distinctions between what types of information is included in these components. They may even contain redundant information, e.g., besides being a part of the epistemological component, the fact that chairs have legs is a natural part of the inferential component. However, the fact is not represented in the same way in these components. For instance, it may be implicitly represented in a 3-D model for the epistemological component and explicitly represented in a logic-based notation for the inferential component.

Depending on the situation, the composite category representation is activated in different ways. External stimuli in the form of direct perception of objects activates the concept via the epistemological component. If the external stimulus is on the linguistic level, as when communicating with other agents, the concept is accessed via the external designator. Finally, if the stimulus is internal, i.e., when the agent is reasoning, the concept is accessed via the internal designator.

To sum up, concepts should not only be used for some limited classification task, they should provide the basis for most of an agent's cognitive tasks.

[1] A more detailed description of this model is provided in my PhD thesis "Autonomous Agents and the Concept of Concepts" (1996).