CROC: a Representational Ontology for Concepts

Aris van Dijk[*]
Universiteit Utrecht
aris.vandijk@phil.uu.nl

Huib Aldewereld
Universiteit Utrecht
Department of Information and Computing Sciences
huib@cs.uu.nl

Virginia Dignum
Universiteit Utrecht
Department of Information and Computing Sciences
virginia@cs.uu.nl

ABSTRACT

Reidentification has been recognised as the most central job of cognition [2]. In this paper, we motivate that *concepts* as *abilities to reidentify* [2], rather than *classifications*, should be the basis of an agent's conceptuology. Most concepts are not classes; class definitions are artificial, often context-dependent, and don't use inductive knowledge. We will present the basic concepts of CROC, a Representational Ontology for Concepts.

Artificial agents can have concepts through language representations alone. Language-like representations, based on lexical concepts, plus reasoning, will be able to solve the interoperability problem to a large extent. By using these concepts, agents can interoperate without need for shared ontologies and with freedom for own conceptions.

Categories and Subject Descriptors

I.2.0 [**General**]: Philosophical foundations; I.2.11 [**Distributed Artificial Intelligence**]: Coherence and coordination—*Semantic Web*; I.2.6 [**Learning**]: Concept learning; I.2.11 [**Distributed Artificial Intelligence**]: Languages and structures; I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems

General Terms

Theory, Standardization, Languages, Design

Keywords

Semantic Web, concepts, ontology, agents, representation

1. CONCEPTS

Without concepts, there would be no agents. With agents, one has to rely on the concepts of the other to communicate. CROC is an attempt to create a framework for concepts for agents.

1.1 Abilities to reidentify

In the perspective of [2], concepts are *abilities to reidentify* for a purpose. E.g., when I see Oscar again, I reidentify Oscar by his hair style, or for example by his voice, or by his name— although my ability may fail for someone who looks, speaks, or is named like him.

[*]The work presented is from my Master's thesis in Artificial Intelligence at Universiteit Utrecht.

Current Semantic Web technology provides agents with *classifications* (using 'ontologies'), but not with *concepts*. Semantic Web ontologies are specifying a classification of the world, not just the world of reference, but also the world of sense (see Frege, 1892). That is, e.g., an ontology could contain different terms for 'morning star' and 'evening star' (though both refer to Venus), and a term for 'Odysseus' (though he might not have existed).

Having a classification doesn't automatically give us an ability to reidentify, however. Semantic Web ontologies use *property restrictions* to define members of the class.

But most (natural) kinds, which we certainly have concepts of, are not classes (see [2]). A kind like the species dog is not a class (nor a fuzzy class), because it isn't true that each instance resembles (or has same properties with) *all* other instances. "There are no properties that *every dog* has in common with *every other dog*" [2]. Indeed, even the most basic properties may fail: there are dogs with three legs, deaf dogs. But at the first place think of the differences between a German Mastiff or a Maltese... Instead, what holds the group together is that its instances are *causally related*. This causal relation may have the *result* that certain properties are probable to be shared among all instances. "There *is* a *good explanation* of why *one* is likely to be *like the next*" [2]. — Though most things do not have a definition by common properties, they may have inductive properties. E.g., a swan is likely to be white.

1.2 Interoperability

A different problem for Semantic Web ontologies is that they cause an interoperability problem: agents that have different classifications cannot communicate using classifications. Here again the process of identification is the basis: before one can classify, one has to identify.[1]

We do not need to have *the same ability* to be able to identify the same. What we identify can be the same (our concepts can be the same), but how we identify is likely to be different (we are likely to have different *conceptions*) (see [2]). We use *representations* to communicate. The word 'fish' can give a fish-reidentification for both of us. We don't need, e.g., visual representations of fish to have a concept of a fish.

[1]One could map, align or merge classes, but each of these needs identification as well.

2. A REPRESENTATIONAL ONTOLOGY FOR CONCEPTS

Following these considerations, the aim of CROC is to use lexical representations of concepts to provide agents with abilities to reidentify. Using lexical expressions, descriptions of concept extension or intension can be given (including, e.g., statements about is-a relations or inductive properties). Due to space limitations, we cannot provide a comprehensive example here. Our poster will try to depict more extensively.

2.1 Step one: having concepts for every building block of representations

There are different kinds of concepts. [2] describes substance concepts ('things'), and distinguishes kinds, individuals (what [1] calls a type-token distinction) and stuffs (like silver). For CROC (where we use OWL for creating an ontology of concepts) we have substance concepts as a subclass of subject concepts.[2]

To make statements about subjects, we use happenings. Concepts for happenings are similar to concepts for substances; but happenings may involve subjects and predicates. We have happening concepts (kinds, like writing; as well as individuals, like the coronation of the queen) fall under subject concepts as well.

To subjects we can apply predicates. To support atom predicates (like poor, eager) and relation predicates (like by car, of mine) we add the concept classes atom predicate, relation, and (subclass of subject) property (like colour).

2.2 Step two: abilities to gather, store and query representational information for reidentification

Subjects can often be described by a single name. Using these names we can make representations, and state related knowledge for concepts (using happening representations). These representations may further include quantification and determination, and reference to kinds (see Carlson, 1980). Reasoning about those statements can be supported using semantic tableaux.

Artificial agents can have concepts through language representation alone. "It is common [to] have a substance concept *entirely* through the medium of language. It is possible to have it, that is, while lacking any ability to recognize the substance in the flesh. For most of us, that is how we have a concept of Aristotle, of molybdenum and, say, of African dormice. — There, I just handed you a concept of African dormice, in case you had none before. Now you can think of them nights if you like, wondering what they are like – on the assumption, of course, that you gathered from their name what sorts of questions you might reasonably ask about them (animal questions, not vegetable or mineral or social artifact questions). . . ." ([2], Chapter 6)

[2]There are many more specific classes of subject concepts possible, like concepts for places, for time moments, for units (like degrees Celsius), for numbers, and so on. One can think of these as involving different abilities, as distinct subclasses (using specific properties), for that reason. For the time being, we treat them simply as subject kinds.

The sorting of inductive properties is done by *subject templates*: these specify what are the interesting questions to ask about subjects [2]. E.g., a chemical element is identified well by its mass or atomic number.

Of course, there is more to concepts than lexical representations alone. Other conceptual abilities may perfectly extend the mechanism of language. Our aim is however to provide artificial agents with the necessary, that what they can use for communication. Language-like representations, based on lexical concepts, plus reasoning, will be able to solve the interoperability problem to a large extent.

2.3 Step three: communication with representing, and using concepts

Agents can carry over representations to other agents. Basic language functions (see also Millikan, 2005) are, e.g., interrogatives, descriptives, directives. The receiving agent uses its concepts to reidentify the representation, uses its reasoning capabilities to verify truth- and other kinds of satisfaction conditions.

3. WHERE AGENTS MEET

As stated, our goal is not primarily knowledge representation, but agent communication and understanding. Agents should use words for *concepts* — most concepts are not classes, because class definitions are artificial, often context-dependent, and don't use inductive knowledge.

Each agent can manage it on its own: there is no need for division of linguistic labour (where only experts 'own' the concepts), private concepts and conceptions are welcome. What agents share is the way of representing using words. Agents can easily learn new lexical concepts (see also [3]). That is, they can easily add new lexical representations and relate to knowledge they already have. But services and agents should have intelligent enough descriptions for their concepts — else they have to learn them by trial and error.

4. ACKNOWLEDGEMENTS

This work has been performed in the framework of the FP7 project ALIVE IST-215890, which is funded by the European Community. The authors would like to acknowledge the contributions of their colleagues from ALIVE Consortium (http://www.ist-alive.eu).

5. REFERENCES

[1] R. Jackendoff. What is a concept, that a person may grasp it? *Mind and Language*, 4(1 and 2), 1989.

[2] R. G. Millikan. *On Clear and Confused Ideas: An Essay about Substance Concepts.* Cambridge University Press, New York, 2000.

[3] L. Steels. Language learning and language contact. In *Proceedings of the workshop on Empirical Approaches to Language Acquisition*, 1997.