CROC: a Representational Ontology for Concepts

Aris van Dijk

Introduction

Classes are semantic, but only for the system that uses them (like a library book code).

For sharing information we need a different mechanism instead: identification. Identification takes place *before* the classification.

We define concepts as *abilities to reidentify for a purpose* [Millikan, 2000]. **CROC** provides agents with an 'ontology' for concepts; we call this a 'conceptuology'.

Main research question: is a 'conceptuology' realizable for artificial agents?

References

Aristotle. Categoriae. c. 350 B.C.

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

S. Laurence and E. Margolis. *Concepts — Core Readings*, chapter Introduction. The MIT Press, 1999.

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

L. Steels. Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. *Evolution of Human Language*, 1997.

Materials and methods

The 'conceptuology' is *an ontology of concepts*: we distinguish several basic kinds of concepts, such as substance, happening, and property concepts.

The 'conceptuology' is *representational* using lexical representations to yield abilities for identification. Names give a principal ability to reidentify. Where names fail, e.g., for an unknown or ambiguous name, reasoning with representations gives abilities to learn/match concepts.

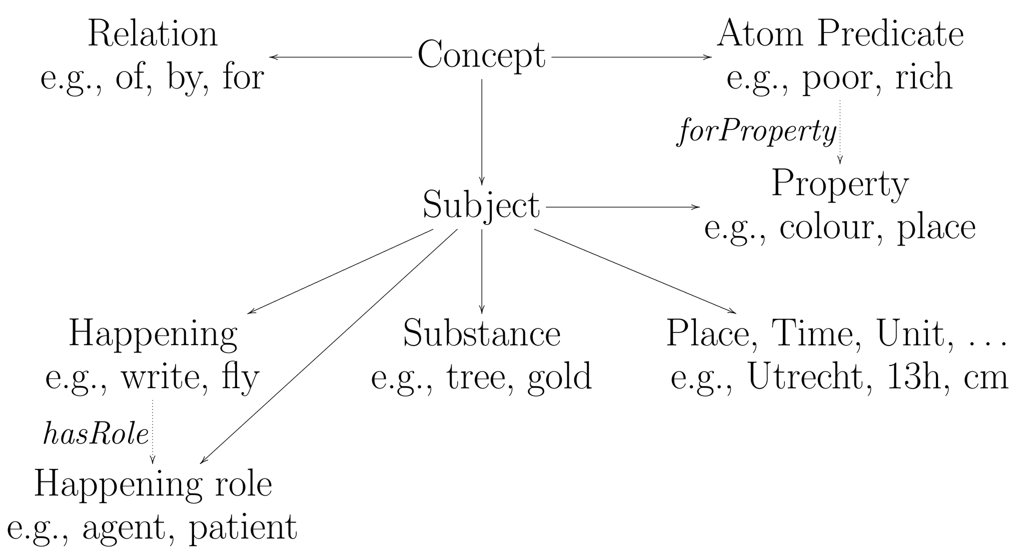
Concept ontology

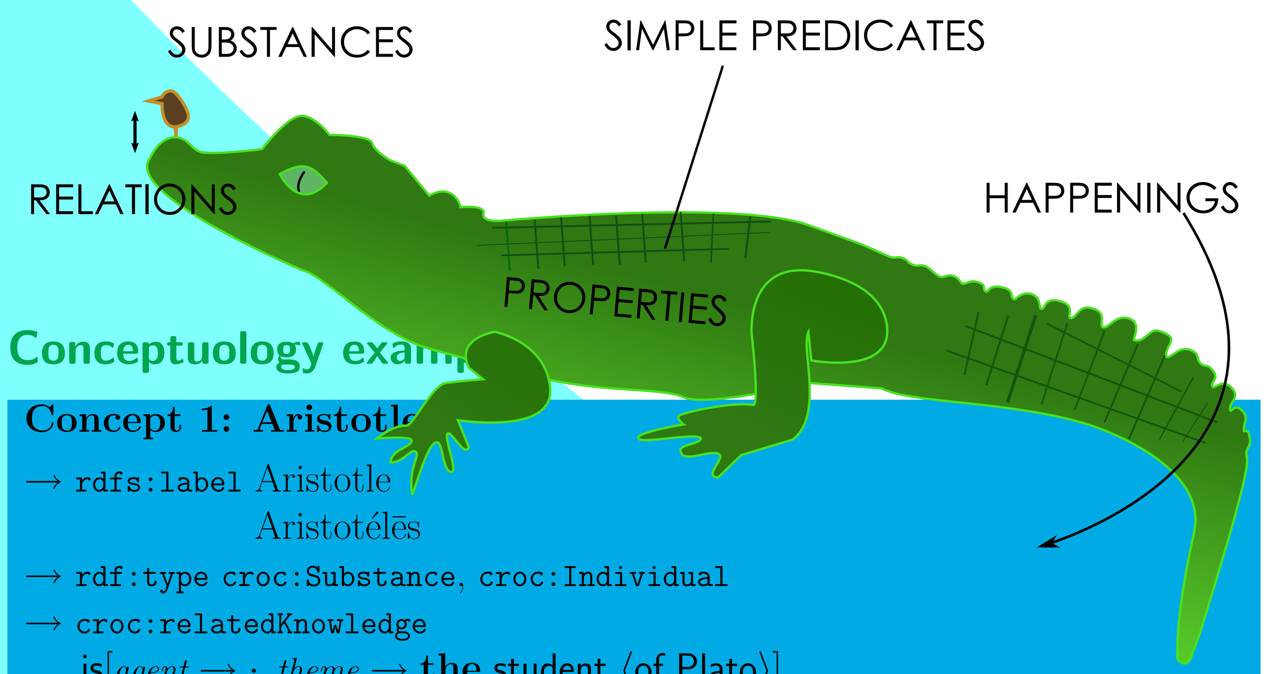
In the literature there have been many other sketches of 'categories' of concepts, starting with Aristotle. A recent example is Jackendoff [1989]. Our purpose is to place the categories we mention in the framework of Millikan [2000], which defines concepts as *abilities to reidentify for a purpose*.

Millikan [2000] describes substance concepts: concepts for 'things', and distinguishes individuals, stuffs, and kinds (such as MAMA, MILK, and A MOUSE).

Substances are one kind of subject concept; there are also more abstract subject concepts, as A CAR BRAND or A SPECIES. Concepts for happenings also resemble concepts for substances, but have one extra dimension: time.

Atom predicates (like *cold*, *healthy*) are a different kind of concepts: they are qualifications, sometimes on specific *properties* (like *temperature*). Other predicates make use of *relations*.





Conceptuology exam

Concept 1: Aristotle

→ `rdfs:label` Aristotle
  Aristotélēs
→ `rdf:type croc:Substance, croc:Individual`
→ `croc:relatedKnowledge`
  is[*agent* → ·, *theme* → **the** student ⟨of Plato⟩]
  is[*agent* → ·, *theme* → '**a** philosopher']
  was born[*patient* → ·, *place* → Stageira, *time* → 384 BC]

Concept 2: vehicles

→ `rdfs:label` vehicle
→ `rdf:type croc:Substance, croc:Kind`
→ `croc:inductionSupportingQuestion`
  has[*agent* → ··· , *theme* → **?** wheel]
  is[*agent* → ··· , fast]

Concept 3: cold

→ `rdfs:label` cold
→ `rdf:type croc:AtomPredicate`
→ `croc:forProperty Temperature`
→ `croc:relatedKnowledge` is[*agent* → ice, ·]

Subject matching

For **CROC** we implemented subject matching using the lexical representations and the corresponding identification and reasoning mechanisms. We did so on basis of subject *templates*: the inductive properties of a subject kind that are stable over time [Millikan, 2000].

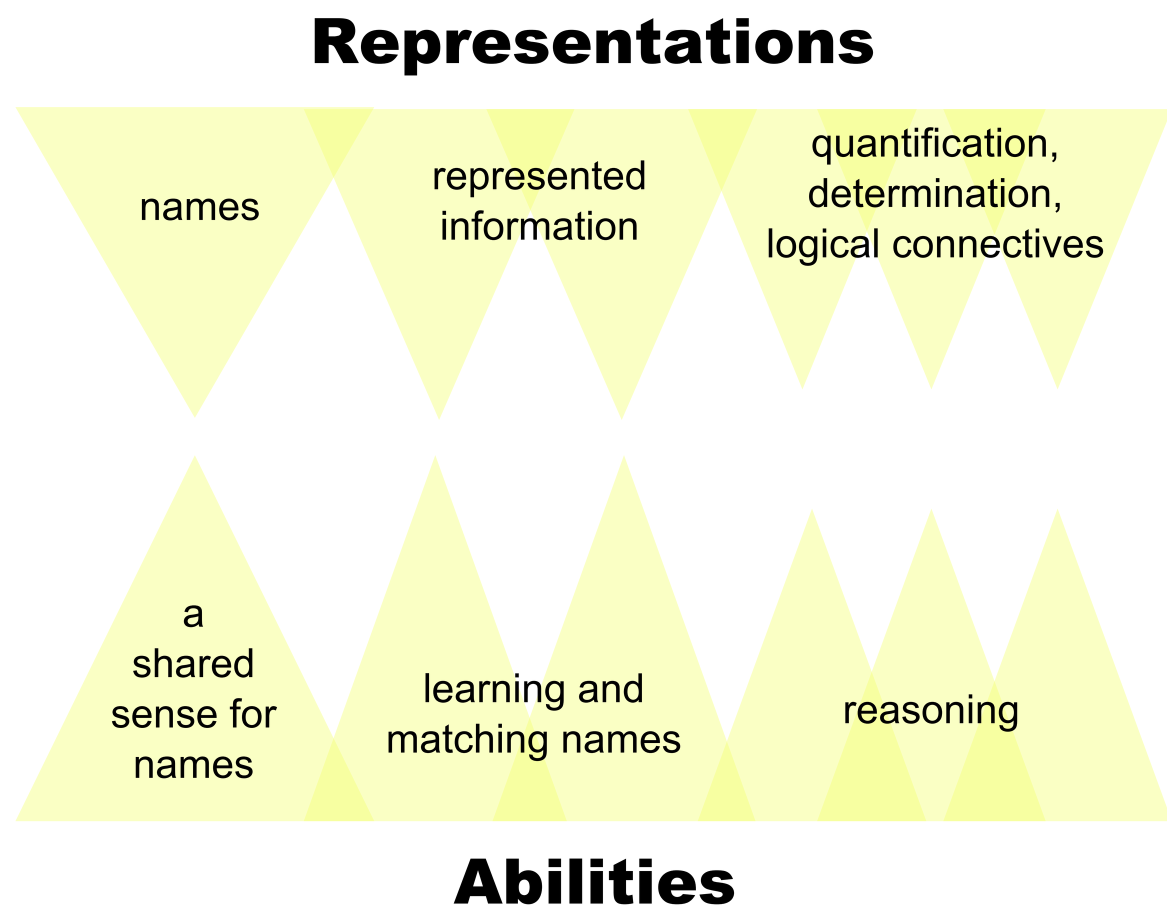
Representations and abilities

Having these concepts gives us concepts for every building block of lexical representations.

Statements or complete sentences are happening representations. For happening representations, we use *happening roles* for representing involved subjects ('thematic roles', see Jackendoff [1989]).

Happening representations may be combined with logical connectives. Handling representations therefore depends strongly on reasoning mechanisms.

For reasoning about subject intension, we use quantification and determination representations. Furthermore we use various predicate representations: atom predicates, relation, and happening predicates.



Dialogue 1: Explicating a concept.

Alice: I'm looking for a radiator. ‖ is[*agent* →?, *theme* → '**a** radiator']?
Bob: What is a radiator? ‖ explain[*patient* → radiator]!
Alice: A service that is able to keep the temperature of a building above some temperature. ‖ is[*agent* → radiator, *theme* → 'service'] ∧ is able to[*agent* → radiator, *theme* → keep[*patient* →?, ⟨warm⟩]].
Bob: OK. I am a radiator. ‖ is[*agent* → me, *theme* → '**a** radiator'].

Dialogue 2: Further completing knowledge about the subject by the subject template.

Alice: A car is a vehicle. ‖ is[*agent* → car, *theme* → 'vehicle'].
Bob: Thanks. (Ah. I know a vehicle by its number of wheels. Because I am gaining knowledge about something being a vehicle, and I don't know how many wheels this vehicle has, I will ask about it:) How many wheels does a car have? ‖ has[*agent* → car, *patient* → **?** wheel]?
Alice: Four wheels. ‖ has[*agent* → car, *patient* → **4** wheel].
Bob: Thanks. ‖ ···

Dialogue 3: Deriving equality by inductive properties.

Alice: A car is a vehicle with four wheels. ‖ ···
Bob: Thanks. (Ah. I have another concept for a subject that is a vehicle and has four wheels; perhaps they are equal:) Does a car equal an automobile? ‖ is[*agent* → car, *patient* → automobile]?
Alice: Yes (I have one concept for them; internalising yields identical entities). ‖ ···
Bob: Thanks. ‖ ···

These mechanisms yield abilities to learn and match subject concepts. In this first implementational phase, we have not yet implemented further specific concept matching abilities, such as matching predicates relative to a property (see also Steels [1997]), and matching happenings using temporal logic. The system may be usefully extended with these capabilities.

Results

Grounding concepts in lexical representations

A conceptuology for artificial agents can be grounded in lexical representations alone.

Language is a representation in which concepts can be grounded, just like our concepts can be grounded in picture representations (see [Millikan, 2000, §6.1]).

For purposes of artificial agents, 'common sense' knowledge is not needed for having a concept.

Some philosophical arguments

A concept is an *ability*, not a prototype or (fuzzy) definition from a set of representations.

Not a definition: they are often partial, context sensitive; if there may be complete definitions, it certainly is not efficient to use for identification.

Not a prototype: what is a typical DOG, comparing a German Mastiff and a Maltese? What are prototypical properties of CAR BRANDs?

Although we ground concepts in representations (including properties, descriptions), this does not mean these representations completely constitute the concept: they provide fallible ways of identifying.

Conclusions

Abilities to reidentify (concepts) can be grounded using lexical representations.

Representing in a universal way, artificial agents can use, learn and match concepts, for their purposes.

Other conceptual abilities may perfectly extend the mechanism of language. Our aim is however to provide artificial agents with the necessary for communication.

In the Semantic Web, agents or services have to rely on their concepts when they encounter new agents or available services and have to organize, share content or communicate.

While there may be asymmetries between background knowledge of the different agents involved — organizing agents have extensive conceptuologies, services may have limited conceptuologies — no agent needs to rely on the concepts of others.

Acknowledgements

Special thanks to Huib Aldewereld and Virginia Dignum for supervision and support of my master's project.

alive | Universiteit Utrecht

Further information

The project is available at `http://sourceforge.net/projects/croc`.

For a more extensive essay on this topic, please read my thesis which is downloadable from the project website.

ALIVE project: `http://ist-alive.eu/`.

Contact me: `aris.vandijk@phil.uu.nl`.