



Marcello Frixione
Antonio Lieto

REPRESENTING CONCEPTS IN FORMAL ONTOLOGIES: COMPOSITIONALITY VS. TYPICALITY EFFECTS

Abstract. The problem of concept representation is relevant for many sub-fields of cognitive research, including psychology and philosophy, as well as artificial intelligence. In particular, in recent years it has received a great deal of attention within the field of knowledge representation, due to its relevance for both knowledge engineering as well as ontology-based technologies. However, the notion of a concept itself turns out to be highly disputed and problematic. In our opinion, one of the causes of this state of affairs is that the notion of a concept is, to some extent, heterogeneous, and encompasses different cognitive phenomena. This results in a strain between conflicting requirements, such as compositionality, on the one hand and the need to represent prototypical information on the other. In some ways artificial intelligence research shows traces of this situation. In this paper, we propose an analysis of this current state of affairs. Since it is our opinion that a mature methodology with which to approach knowledge representation and knowledge engineering should also take advantage of the empirical results of cognitive psychology concerning human abilities, we outline some proposals for concept representation in formal ontologies, which take into account suggestions from psychological research. Our basic assumption is that knowledge representation systems whose design takes into account evidence from experimental psychology (and which, therefore, are more similar to the human way of organizing and processing information) may therefore give better results in many applications (e.g. in the fields of information retrieval and semantic web).

Keywords: concept representation, formal ontologies, prototypical effects, compositionality.

1. Introduction

The computational representation of concepts is a central problem for the development of ontologies and knowledge engineering. Concept representation is a multidisciplinary topic of research that involves different disciplines such as Artificial Intelligence, Philosophy, Cognitive Psychology and Cognitive Science in general. However, the notion of concept itself turns out to be highly disputed and problematic. In our opinion, one of the causes of this current state of affairs is that the very notion of a concept is, to a certain extent, heterogeneous and encompasses different cognitive phenomena. This results in a strain between conflicting requirements such as compositionality, on the one hand, and the need to represent prototypical information on the other. This has several consequences for the practice of knowledge engineering as well as the technology of formal ontologies.

In this article, we propose an analysis of this situation. The rest of the article is organised as follows. In section 2, we point out some differences between the way concepts are conceived in philosophy and psychology. In section 3, we argue that AI research in some way shows traces of the contradictions identified in section 2. In particular, the requirement of compositional, logical style semantics conflicts with the need to represent concepts in terms of the typical traits that allow for exceptions. In section 4, we review several attempts to resolve this conflict in the field of knowledge representation, while paying particular attention to description logics. It is our opinion that a mature methodology with which to approach knowledge representation and knowledge engineering should take advantage of the empirical results of cognitive psychology that concern human abilities as well as philosophical analyses. With this in mind, in section 5, we identify several possible suggestions coming from different aspects of cognitive research: the distinction between two different types of reasoning processes, developed within the context of the so-called “dual process” accounts of reasoning; the proposal to keep prototypical effects separate from the compositional representation of concepts; and the possibility to develop hybrid, prototype and exemplar-based representations of concepts. In section 6, we describe the possible organisation of a system based on such hypotheses. Some conclusions follow in section 7.

2. Concepts in Philosophy and Psychology

Within the field of cognitive science, the notion of a concept is highly disputed and problematic. Artificial intelligence (from now on AI) and, in general, the computational approach to cognition reflect this current state of affairs. Conceptual representation seems to be constrained by conflicting requirements, such as compositionality, on the one hand and the need to represent prototypical information on the other.

A first problem (or, better, a first symptom indicating that a problem exists) lies in the fact that the use of the term “concept” in the philosophical tradition is not homogeneous with the use of the same term in empirical psychology (see e.g. Dell’Anna and Frixione 2010). Briefly¹, we can say that in cognitive psychology a concept is essentially intended as the mental representations of a category, and the emphasis is on processes such as categorisation, induction and learning. According to philosophers, concepts are above all the components of thoughts. Even if we leave aside the problem of specifying exactly what thoughts are, this requires a more demanding notion of concept. In other words, some phenomena that are classified as “conceptual” by psychologists turn out to be “nonconceptual” for philosophers. There are, thus, mental representations of categories that philosophers would not consider genuine concepts. For example, according to many philosophers, concept possession involves the ability to make high level inferences explicit and also sometimes the ability to justify them (Peacocke 1992; Brandom 1994). This clearly exceeds the possession of the mere mental representation of categories. Moreover, according to some philosophers, concepts can be attributed only to agents who can use natural language (i.e. only adult human beings). On the other hand, a position that can be considered in some sense representative of an “extremist” version of the psychological attitude towards concepts is expressed by Lawrence Barsalou in an article symptomatically entitled “Continuity of the conceptual system across species” (Barsalou 2005). He refers to knowledge of scream situations in macaques, which involve different modality-specific systems (auditory, visual, affective systems, etc.). Barsalou interprets these data in favour of the thesis of a continuity of conceptual representations in different an-

¹ Things are made more complex by the fact that also within the two fields considered separately this notion is used in a heterogeneous way, as we shall concisely see in the following. Consequently, the following characterisation of the philosophical and psychological points of view is highly schematic.

imal species, in particular between humans and non-human mammals: “this same basic architecture for representing knowledge is present in humans. [...] knowledge about a particular category is distributed across the modality-specific systems that process its properties” (p. 309). Therefore, according to Barsalou, a) we can also speak of a “conceptual system” in the case of non-human animals; b) low-level forms of categorisation which depend on some specific perceptual modality also belong to the conceptual system. Elizabeth Spelke’s experiments on infants (see e.g. Spelke 1994; Spelke and Kinzler 2007) are symptomatic of the difference in approach between psychologists and philosophers. These experiments demonstrate that some extremely general categories are very precocious and presumably innate. According to the author, they show that newborn babies already possess certain *concepts* (e.g., the physical object concept). However, some philosophers (Bermudez 1995, Bermudez and Cahen 2011) have interpreted these same data as a paradigmatic example of the existence of *nonconceptual* contents in agents (babies) who have yet to develop a conceptual system.

2.1. Compositionality

The fact that philosophers consider concepts mainly as the components of thoughts has given greater emphasis to *compositionality*, as well as to other related features, such as productivity and systematicity, which are often ignored by the psychological treatment of concepts. On the other hand, it is well known that compositionality is at odds with *prototypicality effects*, which are crucial in most psychological characterisations of concepts (we shall develop this point in greater detail in the next section).

Let us first consider the *compositionality* requirement. In a compositional system of representations, we can distinguish between a set of *primitive*, or *atomic*, *symbols* and a set of *complex symbols*. Complex symbols are generated from primitive symbols through the application of a set of suitable recursive syntactic rules (generally, a potentially infinite set of complex symbols can be generated from a finite set of primitive symbols). Natural languages are the paradigmatic example of compositional systems: primitive symbols correspond to the elements of the lexicon (or, better, to *lessemes*), and complex symbols include the (potentially infinite) set of all sentences.

In compositional systems, the meaning of a complex symbol s functionally depends on the syntactic structure of s as well as the meaning of primitive symbols in it. In other words, the meaning of complex symbols can be determined by means of recursive semantic rules that work in parallel with syntactic composition rules. This is the so-called *principle of compositionality of meaning*, which Gottlob Frege identified as one of the main features of human natural languages.

In classical cognitive science, it is often assumed that mental representations are compositional. One of the clearest and most explicit formulations of this assumption was proposed by Jerry Fodor and Zenon Pylyshyn (1988). They claim that the compositionality of mental representations is mandatory in order to explain some fundamental cognitive phenomena. In the first place, human cognition is *generative*: in spite of the fact that the human mind is presumably finite, we can conceive and understand an unlimited number of thoughts that we have never encountered before. Moreover, the *systematicity* of cognition also seems to depend on compositionality: the ability to conceive certain contents is systematically related to the ability to conceive other contents. For example, if somebody can understand the sentence *the cat chases a rat*, then (s)he is presumably also able to understand *a rat chases the cat*, by virtue of the fact that the forms of the two sentences are syntactically related. We can conclude that the ability to understand certain propositional contents systematically depends on the compositional structure of the contents themselves. This can easily be accounted for if we assume that mental representations have a structure similar to a compositional language.

2.2. Against “Classical” Concepts

Compositionality is less important for many psychologists. In the field of psychology, most research on concepts moves from the critiques to the so-called classical theory of concepts, i.e. the traditional point of view according to which concepts can be defined in terms of necessary and sufficient conditions. Empirical evidence favours those approaches to concepts that account for prototypical effects. The central claim of the classical theory of concepts (i.e.) is that every concept c is defined in terms of a set of features (or conditions) f_1, \dots, f_n that are individually necessary and jointly sufficient for the application of c . In other words, everything that satisfies features f_1, \dots, f_n is a c , and if anything is a c ,

then it must satisfy f_1, \dots, f_n . For example, the features that define the concept *bachelor* could be *human*, *male*, *adult* and *not married*; the conditions defining *square* could be *regular polygon* and *quadrilateral*. This point of view was unanimously and tacitly accepted by psychologists, philosophers and linguists until the middle of the 20th century.

The first critique of classical theory is due to a philosopher: in a well known section from the *Philosophical Investigations*, Ludwig Wittgenstein observes that it is impossible to identify a set of necessary and sufficient conditions to define a concept such as GAME (Wittgenstein, 1953, §66). Therefore, concepts exist which cannot be defined according to classical theory, i.e. in terms of necessary and sufficient conditions. Concepts such as GAME rest on a complex network of *family resemblances*. Wittgenstein introduces this notion in another passage in the *Investigations*: «I can think of no better expression to characterise these similarities than “family resemblances”; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament, etc. etc.» (*ibid.*, §67).

Wittgenstein’s considerations were corroborated by empirical psychological research: starting from the seminal work by Eleanor Rosch, with the psychological experiments that showed how common-sense concepts do not obey the requirement of the classical theory²: common-sense concepts cannot usually be defined in terms of necessary and sufficient conditions (and even if for some concepts such a definition is available, subjects do not use it in many cognitive tasks). Concepts exhibit *prototypicality effects*: some members of a category are considered better instances than others. For example, a robin is considered a better example of the category of birds than, say, a penguin or an ostrich. More central instances share certain typical features (e.g. the ability of flying for birds, having fur for mammals) that, in general, are neither necessary nor sufficient conditions.

Prototypical effects are a well established empirical phenomenon. However, the characterisation of concepts in prototypical terms is difficult to reconcile with the compositionality requirement. According to a well known argument by Jerry Fodor (1981), prototypes are not compositional (and, since concepts in Fodor’s opinion must be compositional, concepts cannot be prototypes). In brief, Fodor’s argument runs as fol-

² On the empirical inadequacy of the classical theory and the psychological theories of concepts see (Murphy 2002).

lows: consider a concept like PET FISH. It results from the composition of the concept PET as well as the concept FISH. However, the prototype of PET FISH cannot result from the composition of the prototypes of PET and FISH. For example, a typical PET is furry and warm, a typical FISH is greyish, but a typical PET FISH is neither furry and warm nor greyish.

Moreover, things are made more complex by the fact that, even within the two fields of philosophy and psychology considered separately, the situation is not very encouraging. In neither of the two disciplines does a clear, unambiguous and coherent notion of concept seem to emerge. Consider for example psychology. Different positions and theories on the nature of concepts are available (prototype view³, exemplar view, theory theory) that can hardly be integrated. From this point of view, the conclusions of Murphy (2002) are of great significance, since in many respects this book reflects the current status of empirical research on concepts. Murphy contrasts the approaches mentioned above in relation to different classes of problems, including learning, induction, lexical concepts as well as children's concepts. His conclusions are rather discouraging: the result of comparing the various approaches is that "there is no clear, dominant winner" (*ibid.*, p. 488) and that "[i]n short, concepts are a mess" (p. 492). This situation persuaded some scholars to doubt whether concepts constitute a homogeneous phenomenon from the point of view of a science of the mind (see e.g. Machery 2005 and 2009; Frixione 2007).

3. Concept Representation in Artificial Intelligence

The situation outlined in the section above is, to some extent, reflected by the state of the art in AI and, in general, in the field of computational modelling of cognition. This research area often seems to hesitate between different (and hardly compatible) points of view. In AI, the representation of concepts is faced mainly within the field of knowledge representation (KR). Symbolic KR systems (KRs) are formalisms whose

³ Note that the so-called prototype view does not coincide with the acknowledgement of prototypical effects: as stated before, prototypical effects are a well established phenomenon that all psychological theories of concepts are bound to explain; the prototype view is a particular attempt to explain empirical facts concerning concepts (including prototypical effects). On these aspects, see again Murphy 2002.

structure is, broadly speaking, language-like. This usually entails assuming that KRs are compositional.

In their early development (historically corresponding to the late 1960s and the 1970s), many KRs oriented to conceptual representations attempted to take into account suggestions from psychological research. Examples are early semantic networks and frame systems. Frame and semantic networks were originally proposed as alternatives to the use of logic in KR. The notion of frame was developed by Marvin Minsky (1975) as a solution to the problem of representing structured knowledge in AI systems⁴. Both frames and most semantic networks allowed for the possibility to characterise concepts in terms of prototypical information.

However, such early KRs were usually characterised in a rather rough and imprecise way. They lacked a clear formal definition, with the study of their meta-theoretical properties being almost impossible. When AI practitioners tried to provide a stronger formal foundation to concept oriented KRs, it turned out to be difficult to reconcile compositionality and prototypical representations. As a consequence, they often chose to sacrifice the latter.

In particular, this is the solution adopted in a class of concept-oriented KRs which were (and still are) widespread within AI, namely the class of formalisms that stem from the so-called structured inheritance networks and the KL-ONE system (Brachman and Schmolze 1985). Such systems were subsequently called terminological logics, and today are usually known as *description logics* (DLs) (Baader et al. 2010).

A standard inference mechanism for this kind of network is *inheritance*. The representation of prototypical information in semantic networks usually takes the form of allowing exceptions to inheritance. Networks in this tradition do not admit exceptions to inheritance, and therefore do not allow for the representation of prototypical information. In fact, representations of exceptions cannot be easily accommodated with other types of inference defined on these formalisms, first and foremost concept classification (Brachman 1985). Since the representation of prototypical information is not allowed, inferential mechanisms defined on these networks (e.g. inheritance) can be traced back to classical logical inferences.

⁴ Many of the original articles describing these early KRs can be found in (Brachman & Levesque 1985), a collection of classic papers of the field.

In more recent years, representation systems in this tradition have been directly formulated as logical formalisms (the above mentioned description logics, Baader et al., 2010), in which Tarskian, compositional semantics is directly associated to the syntax of the language. Logical formalisms are paradigmatic examples of compositional representation systems and, as a result, this kind of system fully satisfies the compositionality requirement. This has been achieved at the cost of not allowing exceptions to inheritance. However, in so doing, they have forsaken the possibility of representing concepts in prototypical terms. From this point of view, such formalisms can be seen as a revival of the classical theory of concepts, in spite of its empirical inadequacy in dealing with most common-sense concepts.

Nowadays, DLs are widely adopted within many application fields, in particular within that of the representation of ontologies. For example, the OWL (Web Ontology Language) system⁵ is a formalism in this tradition that has been endorsed by the World Wide Web Consortium for the development of the semantic web.

3.1. Artificial Systems: Why Prototypical Effects are Needed

Prototypical effects in categorisation and, in general, category representation are not only crucial for the empirical study of human cognition, but they are also of the greatest importance in representing concepts in artificial systems. Let us first consider human cognition. Under what conditions should we say that somebody *knows* the concept DOG (or, in other terms, that (s)he possesses an adequate mental representation of it)? It is not easy to say. However, if a person does not know that, for example, dogs usually bark, that they typically have four legs and that their body is covered with fur, that in most cases they have a tail and that they wag it when they are happy, then we probably should conclude that this person does not grasp the concept DOG. Nevertheless, all these pieces of information are neither necessary nor sufficient conditions for being a dog. In fact, they are traits that characterise dogs in typical (or prototypical) cases. The problem is exactly the same if we want to represent knowledge in an artificial system. Let us suppose that we want to provide a computer program with a satisfactory representation of DOG. Then we probably also want to represent the kind of information men-

⁵ <http://www.w3.org/TR/owl-features/>

tioned above: for many applications, a representation of DOG that does not include the information that dogs usually bark is a bad representation also from a technological point of view. Therefore, if a system does not allow information to be represented in typical/prototypical terms (as is the case of standard description logics), then it is not adequate in this respect. With standard DLs, the only way to tackle this problem should be the recourse to tricks or ad hoc solutions (as often happens in many applications).

The concept DOG is not exceptional from this point of view. The majority of everyday concepts behave in this way. For most concepts, a classical definition in terms of necessary and sufficient conditions is not available (or, even if it is available, it is unknown to the agent). On the other hand, it may be that we know the classical definition of a concept, but typical/prototypical knowledge still plays a central role in many cognitive tasks. Consider the following example: nowadays most people know necessary and sufficient conditions for being WATER: water is exactly the chemical substance whose formula is H_2O , i.e., the substance whose molecules are formed by one atom of oxygen and two atoms of hydrogen. However, in most cases of everyday life, when we categorise a sample of something as WATER, we do not take advantage of this piece of knowledge. We use such prototypical traits such as the fact that (liquid) water is usually a colourless, odourless and tasteless fluid. As a further example, consider the concept GRANDMOTHER. Everybody knows a classical definition for it: x is the grandmother of y if and only if x is the mother of a parent of y . However, in many cases we do not use this definition to categorise somebody as a grandmother. We resort to typical traits: grandmothers are old women who take care of children, who are tender and polite with them, and so on. Once more, the problem is not different in the case of artificial systems: generally a system that has to categorise something as WATER cannot perform chemical analyses, and it must trust in prototypical evidence.

Therefore, the use of prototypical knowledge in cognitive tasks such as categorisation is not a “fault” of the human mind, as it could be the fact that people are prone to fallacies and reasoning errors (leaving aside the problem of establishing whether recurrent errors in reasoning could have a deeper “rationality” within the general economy of cognition). It has to do with the constraints that concern every finite agent that has a limited access to the knowledge which is relevant for a given task. This is the case of both natural and artificial cognitive systems.

4. Non-classical Concepts in Computational Ontologies

Within symbolic, logic oriented KR, rigorous approaches exist that make it possible to represent exceptions, and that would therefore be, at least in principle, suitable for representing “non-classical” concepts. Examples are fuzzy logics and non-monotonic formalisms. Therefore, the adoption of logic oriented semantics is not necessarily incompatible with prototypical effects. Nevertheless, such approaches pose various theoretical and practical difficulties, with many problems remaining unsolved.

In this section, we review some recent proposals to extend concept-oriented KRs, and in particular DLs, with a view to representing non-classical concepts. Recently, different methods and techniques have been adopted to represent non-classical concepts within computational ontologies. These are based on extensions of DLs as well as standard ontology languages such as OWL. The different proposals that have been put forward can be grouped into three main classes: a) fuzzy approaches, b) probabilistic and Bayesian approaches, c) approaches based on non-monotonic formalisms.

a) Following this direction, as with the integration of *fuzzy logics* in DLs and ontology oriented formalisms (see for example Gao and Liu 2005, and Calegari and Ciucci 2007), Stoilos et al. (2005) propose a fuzzy extension of OWL, f-OWL, capable of capturing imprecise and vague knowledge, and a fuzzy reasoning engine that lets f-OWL reason about such knowledge. Bobillo and Straccia (2009) propose a fuzzy extension of OWL 2 in order to represent vague information in semantic web languages. However, it is well known (Osherson and Smith 1981) that approaches to prototypical effects based on fuzzy logic encounter difficulties with compositionality.

b) Current literature also offers several *probabilistic generalizations* of web ontology languages. Many of these approaches, as pointed out in Lukasiewicz and Straccia (2008), focus on combining the OWL language with probabilistic formalisms based on Bayesian networks.); Ding et al. (2006) propose a probabilistic generalization of OWL, called Bayes-OWL, which is based on standard Bayesian networks. Bayes-OWL provides a set of rules and procedures for the direct translation of an OWL ontology into a Bayesian network. A possible problem might lie in the “translation” from one form of “semantics” (OWL based) to another.

c) In the field of non-monotonic extensions of DLs, Baader and Hollunder (1995) propose an extension of the ALCF system based on Reiter’s

default logic.⁶ The same authors, however, point out both the semantic and computational difficulties of this integration and, for this reason, propose a restricted semantics for open default theories, in which the default rules are only applied to individuals explicitly represented in the knowledge base. Bonatti et al. (2006) propose an extension of DLs with circumscription. One of the reasons for applying circumscription is the possibility to express prototypical properties with exceptions, something which is done by introducing “abnormality” predicates whose extension is minimized. A different approach, investigated by Klinov and Parsia (2008), is based on the use of the OWL 2 annotation properties (APs) in order to represent vague or prototypical information. The limit of this approach is that APs are not taken into account by the reasoner, and therefore have no effect on the inferential behaviour of the system (Bobillo and Straccia 2009).

5. Some Suggestions from Cognitive Science

Even though a relevant field of research exists, in the scientific community there is no agreement on the use of non-monotonic and, in general, non-classical logics in ontologies. For practical applications, systems that are based on classical Tarskian semantics and that do not allow for exceptions (as it is the case of “traditional” DLs), are still preferred. Some researchers, such as Pat Hayes (2001), argue that non-monotonic logics (and, therefore, the non-monotonic “machine” reasoning for the semantic web) can be adopted for local uses only or for specific applications because it is “unsafe on the web”. Nevertheless, the question about which “logics” must be used in the semantic web (or, at least, to what degree and in which cases certain logics could be useful) is still open.

Empirical results from cognitive psychology show that most common-sense concepts cannot be characterised in terms of necessary/sufficient conditions. Classical, monotonic DLs seem to capture the compositional aspects of conceptual knowledge, but are inadequate in representing prototypical knowledge. However, a “non-classical” alternative, a general DL able to represent concepts in prototypical terms still does not exist.

⁶ The authors pointed out that “Reiter’s default rule approach seems to fit well into the philosophy of terminological systems because most of them already provide their users with a form of ‘monotonic’ rules. These rules can be considered as special default rules where the justifications — which make the behaviour of default rules non-monotonic — are absent”.

As a possible way out, we outline a tentative proposal based on several suggestions from cognitive science. Some recent trends in psychological research favour the hypothesis that reasoning is not a unitary cognitive phenomenon. At the same time, empirical data on concepts seem to suggest that prototypical effects could stem from different representation mechanisms. To this end, we identify some suggestions that, in our opinion, could be useful in developing artificial representation systems, namely: (i) the distinction between two different types of reasoning processes, which has been developed within the context of the so-called “dual process” accounts of reasoning (sect. 5.1 below); (ii) the proposal to keep prototypical effects separate from the compositional representation of concepts (sect. 5.2); and (iii) the possibility to develop hybrid, prototype and exemplar-based representations of concepts (sect. 5.3).

5.1. A “Dual Process” Approach

Cognitive research on concepts seems to suggest that concept representation does not constitute a unitary phenomenon from a cognitive point of view. In this perspective, a possible solution should be inspired by the experimental results of empirical psychology, in particular by the so-called dual process theories of reasoning and rationality (Stanovich and West 2000, Evan and Frankish 2008). In such theories, the existence of two different types of cognitive systems is assumed. The systems of the first type (type 1) are phylogenetically older, unconscious, automatic, associative, parallel and fast. The systems of type 2 are more recent, conscious, sequential and slow. They are also based on explicit rule following. In our opinion, there are good *prima facie* reasons to believe that, in human subjects, classification (a monotonic form of reasoning which is defined on semantic networks, and which is typical of DL systems) is a type 2 task (it is a difficult, slow, sequential task). On the contrary, exceptions play an important role in processes such as categorization and inheritance, which are more likely to be type 1 tasks: they are fast, automatic, usually do not require a particular conscious effort, and so on.

Therefore, a reasonable hypothesis is that a concept representation system should include different “modules”: a monotonic type 2 module, involved in classification and in similar “difficult” tasks, as well as a non-monotonic module involved in the management of exceptions. The latter module should be a “weak” non-monotonic system, able to perform only some simple forms of non-monotonic inferences (mainly related to

categorization as well as exceptions inheritance). This solution goes in the direction of a “dual” representation of concepts within the ontologies, with the realization of hybrid reasoning systems (monotonic and non-monotonic) on semantic network knowledge bases.

5.2. A “Pseudo-Fodorian” Proposal

As previously stated (section 2.2), Fodor claims that concepts cannot be prototypical representations, since concepts must be compositional, and prototypes do not compose. On the other hand, by virtue of the criticisms of “classical” theory, concepts cannot be definitions. Therefore, Fodor argues that (most) concepts are atomic, i.e. they are symbols with no internal structure. Their content is determined by their relation to the world, and not by their internal structure and/or by their relations with other concepts (Fodor 1987, 1998). Naturally, Fodor acknowledges the existence of prototypical effects. However, he claims that prototypical representations are not part of concepts. Prototypical representations make it possible to recognize the reference of concepts, but they must not be identified with concepts. Consider for example the concept DOG. Unquestionably, in our minds there is some prototypical representation associated to DOG (dogs usually have fur, they typically bark, and so on). However, this representation does not coincide with the concept DOG: DOG is an atomic, unstructured symbol.

We borrow from Fodor the suggestion that compositional representations and prototypical effects are requested by different components of the representational architecture. We assume that there is a compositional component of representations, which admits no exceptions and exhibits no prototypical effects, and which can be represented, for example, in the terms of some classical DL knowledge base. In addition, a prototypical representation of categories is responsible for processes such as categorisation, but it does not affect the inferential behaviour of the compositional component.

It must be noted that our present proposal is not entirely “Fodorian”, at least in the following three senses:

- i. We leave out the problem of the nature of the semantic content of conceptual representations. Fodor endorses a causal, informational theory of meaning, according to which the content of concepts is constituted by a nomic mind-world relation. We are in no way committed to such an account of semantic content. (In general, the philosophical

problem of the nature of the intentional content of representations is largely irrelevant to our present purposes).

ii. Fodor claims that concepts are compositional, and that prototypical representations, as they are not compositional, cannot be concepts. We do not take any position on which part of the system we propose must be considered as truly “conceptual”. In our opinion, the notion of concept is spurious from a cognitive point of view. Both the compositional and the prototypical components contribute to the “conceptual behaviour” of the system (i.e., they have a role in those abilities that we usually describe in terms of the possession of concepts).

iii. According to Fodor, the majority of concepts are atomic. In particular, he claims that almost all the concepts that correspond to lexical entries have no structure. We maintain that many lexical concepts, even though indefinable in classical theory terms, should exhibit some form of structure, and that such a structure can be represented, for example, by means of a DL taxonomy.

5.3. Prototypes and Exemplars

As previously anticipated (section 2.2), within the field of psychology, different positions and theories on the nature of concepts are available. They are generally grouped into three main classes, namely prototype views, exemplar views and theory-theory views (see e.g. Murphy 2002, Machery 2009). All of these succeed in accounting for (some aspects of) the prototypical effects in conceptualisation.

According to the prototype view, knowledge about categories is stored in terms of prototypes, i.e. in terms of some representation of the “best” instances of the category. For example, the concept CAT should coincide with a representation of a prototypical cat. In the simpler versions of this approach, prototypes are represented as (possibly weighted) lists of features.

According to the exemplar view, a given category is mentally represented as a set of representations of specific exemplars explicitly stored within memory: the mental representation of the concept CAT is the set of the representations of (some of) the cats we have encountered during our lifetime.

Theory-theories approaches adopt some form of holism about concepts. According to some versions of the theory-theories, concepts are analogous to theoretical terms in a scientific theory. For example, the

concept CAT is identified by the role it plays in our mental theory of zoology. In other versions of the approach, concepts themselves are identified with micro-theories of some sort. For example, the concept CAT should be identified with a mentally represented micro-theory about cats.

These approaches turn out to be not mutually exclusive. They seem to succeed in explaining different classes of cognitive phenomena, and many researchers hold that all of them are needed in order to explain psychological data. In this perspective, we propose integrating some of them in computational representations of concepts. More precisely, we propose combining prototypical and exemplar based representations in order to account for category representation as well as the prototypical effects (for a similar, hybrid prototypical and exemplar based proposal developed in the field of machine learning, see Gagliardi 2008). We do not take into consideration the theory-theory approach, since it is, to a certain extent, more vaguely defined when compared to both prototypes and exemplar based approaches. As a consequence, at present its computational treatment seems to be more problematic.

6. A Possible Architecture

In this section we outline the proposal of a possible architecture for concept representation, which takes advantage of the suggestions presented in the section above. It is based on a hybrid approach, and combines a compositional component for the representation of concepts in terms of necessary and/or sufficient conditions, with a second component implementing prototypical reasoning.

Concepts in the compositional component (which is based on a DL formalism) are represented as in Fig. 1. Every concept can be subsumed by a certain number of superconcepts, and it can be characterised by means of a number of attributes, which relate it to other concepts in the knowledge base. In the terms of traditional predicate logic, concepts correspond to one-argument predicates, and attributes to two-argument relations. Restrictions on the number of possible fillers can be associated to each attribute. Given a concept, its attributes and its concept/superconcept relations express necessary conditions for it. DL formalisms make it possible to specify which of these necessary conditions also count as sufficient conditions. Concepts can have any number



of individual instances, that are represented as individual concepts in the taxonomy.

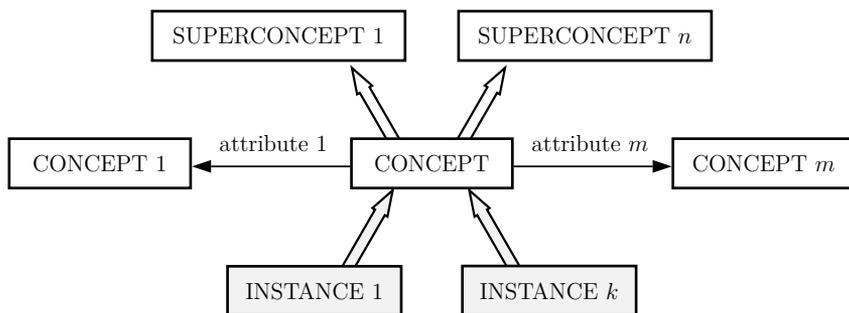


Figure 1.

As an example, consider the fragment of network shown in Fig. 2. The concept DOG is represented as a subconcept of MAMMAL. Since DL networks can express only necessary and/or sufficient conditions, some details of the representation are very loose. For example, according to Fig. 2, a DOG may or may not have a tail (this is expressed by the number restriction 0/1 imposed on the attribute *has_tail*), and has an unspecified number of limbs (since some dogs could have lost a limb or have more than four limbs). LASSIE and RIN TIN TIN are represented as individual instances of DOG (of course, concepts describing individual instances can be further described, specifying for example the values of the attributes inherited from parent concepts).

Prototypes describing typical instances of concepts are represented as data structures that are external to the DL knowledge base. Such structures could, for example, be lists of (possibly weighted) attribute/value pairs that are linked to the corresponding concept. Some attributes of the list should correspond to attributes of the DL concept, for which the value is further specified. For example, the prototypical dog does have a tail, and has exactly four limbs. Other attributes of the prototype could have no counterpart in the corresponding DL concept.

As far as the exemplar-based component of the representations is concerned, exemplars are directly represented in the DL knowledge base as instances of concepts. (It may also happen that some information concerning exemplars is represented outside the DL component, in the

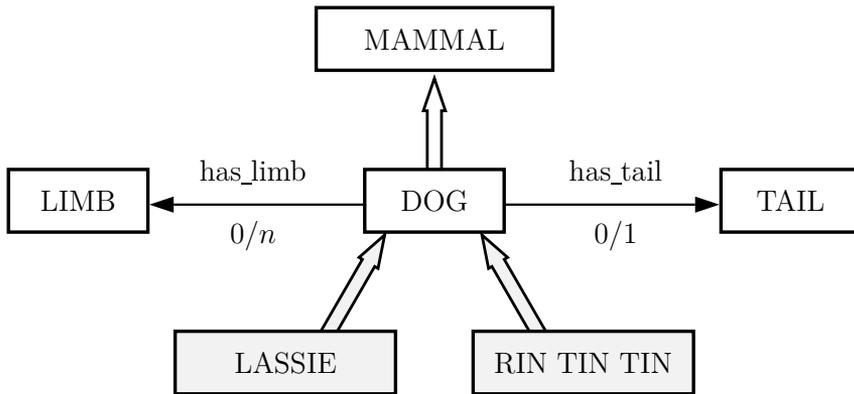


Figure 2.

form of Linked Data — see section 7 below. Typically, this could be the case of “non symbolic” information, such as images, sounds, etc.)

In section 5.3, we anticipated that prototype and exemplar based approaches to concept representation are not mutually exclusive, and that that they succeed in explaining different phenomena. Exemplar based representations can be useful in many situations. According to various experiments, it may happen that instances of a concept that are rather dissimilar to the prototype, but are very close to a known exemplar, are categorized quickly and with high confidence. For example, a penguin is rather dissimilar to the prototype of BIRD. However, if I already know an exemplar of penguin, and if I know that it is an instance of BIRD, it is easier for me to classify a new penguin as a BIRD. This is particularly relevant for concepts (such as FURNITURE, or VEHICLE) whose members differ significantly from one another.

Exemplar based representations are easier and faster to acquire, when compared to prototypes. In some situations, it may happen that there is not enough time to extract a prototype from the available information. Moreover, the exemplar based approach makes the acquisition of concepts that are not linearly separable easier (see Medin and Schwanenflugel 1981).

As far as the relation between the DL based, compositional component of the system on the one hand, and typicality based knowledge on the other, it must be noted that prototypical information about con-

cepts (either stored in the form of prototypes or extracted from the representation of exemplars) extends the information coded within the compositional formalism: the knowledge represented at the level of the DL formalism provides necessary and/or sufficient conditions for the application of concepts. As a consequence, such conditions hold for every instance of concepts, and cannot be violated by any specific instance. Therefore, what can be inferred on the basis of prototypical knowledge can extend, but can in no way conflict with what can be deduced from the DL based component.

According to our proposal, the categorisation of a new exemplar should follow the following steps:

a) Perform deductive reasoning on the DL knowledge base. Purely deductive, monotonic categorization requires sufficient conditions for concepts to be available. Since, in the majority of commonsense knowledge domains, sufficient conditions for defining concepts are scarce (or, when available, they cannot be used for many practical purposes — see section 3.1 above), we can hypothesize that in most cases this step will give poor results. However, if something can be deductively categorized, then this inference is definitive (in the sense that it is not defeasible), and steps b) and c) can be eschewed.

b) Compare the exemplar to be classified to the prototypes associated to the concepts, and evaluate their degree of similarity. If, for some concept, such a similarity value exceeds a given threshold, then the exemplar can be tentatively categorized as an instance of the corresponding concept.

c) Compare the exemplar to be classified to the exemplars of the concepts stored in the knowledge base, and, again, evaluate their degree of similarity.

Of course, categorization performed on the basis of steps b) and c) is always defeasible.

7. Some Conclusions

In recent years, one of the main application areas for concept representation has been the development of formal ontologies for the semantic web. In the field of web ontology languages, the developments proposed above could be achieved within the framework of the so-called Linked Data approach. In the semantic web research community, the Linked Data perspective is assuming a prominent position (Bizer, Heath and

Berners-Lee 2009). One of the main objectives of this approach is the integration of different data representations (often stored in different data sources) within a unique representational framework. This makes it possible to enlarge the answer-space of a query through the realization of “semantic bridges” between different pieces of data (and, often, between different data sources). Such integration is made possible through specific constructs provided by Semantic Web languages, such as OWL, SKOS etc.

The implementation of both the exemplar and prototype points of view (see section 5.3 above) can take advantage of the Linked Data approach. Let us consider prototypes. Concepts can be represented as classes in a formal ontology, based on a classical, compositional DL system. Prototypes can be associated to such representations; they can be implemented using the Open Knowledge-Base Connectivity (OKBC) protocol⁷. The knowledge model of the OKBC protocol is supported by Protegé Frames, an ontology editor that makes it possible to build frame representations (the so called Frame Ontologies). Since it is possible to export the Frame Ontologies built with Protegé in the OWL language, the connection between these two types of representation can be made using the standard formalisms provided by the semantic web community in the linked data perspective (e.g. using the owl:sameAs construct).

In this way, according to our hypothesis, different types of categorization processes can follow different paths: monotonic categorization involves only the DL ontology, while typicality-based categorization, which involves exemplars and prototypes, could also take advantage of Linked Data structures that are external to the compositional ontology.

The possibility of performing forms of non-monotonic reasoning (namely, non-monotonic categorization of instances) only outside the compositional component of the representation system is one of the main features of our proposal. Among other things, this solution makes it possible to avoid consistency problems in the compositional part⁸, introducing at the same time within the ontology (intended in a broad sense) the possibility to expand the allowed types of reasoning. Distributed reasoning is a feature of many KR systems but, in the field of formal ontologies, it has often been limited to performing the same type of rea-

⁷ <http://www.ai.sri.com/~okbc/>

⁸ This was one of the main problems both in frame based systems as well as hybrid knowledge representation approaches.

soning (e.g. classification using classical deductive logics) on multiple, modular, local knowledge bases.

Hybrid non-monotonic categorization, based on both prototypes and exemplars, should take advantage of suggestions from the field of machine learning, where the prototype-exemplar dichotomy in concept representation has been investigated. Consider for example the *PEL-C algorithm*, where *PEL-C* stands for *Prototype-Exemplar Learning Classifier* (Gagliardi 2010). The *PEL-C* is a hybrid instance-based algorithm used for machine learning tasks, which accounts for typicality effects in categorization using both prototypes and exemplars. It is based on a learning phase as well as a test phase, and it can also be adopted for both a semi-automatic ontology population as well as updating processes.

Naturally, the advantage of associating prototypical knowledge to concepts is not limited to categorization. Consider for example a task as *property checking*. Property checking consists of answering questions such as “does the class A have the property b?”. Let us suppose that a user runs an informational query⁹ on a knowledge base representing fruits, in order to know which kind of citrus fruit is yellow (i.e. (s)he asks the knowledge base the question: “does any citrus fruit have the property of being yellow?”). Intuitively, the expected answer that fits the information needs of the user is “lemon”. However, in the DL knowledge base, any form of citrus fruit that has the property of being yellow as a defining condition does not exist. Being yellow is not a necessary condition for being a lemon and, therefore, this property is not associated to the class LEMON of the DL ontology. However, from a cognitive point of view, the property of being yellow is relevant to characterize the concept LEMON. According to our hybrid approach, this can be represented in the prototypical information associated to LEMON (either in terms of the value of the attribute *colour* of the corresponding prototype, or stored in the knowledge concerning exemplars). In this way, it is possible to retrieve the desired information from the prototype and/or exemplar part of the representation. Thus, given a query such as:

```
SELECT the CONCEPT citrus WHERE {?CONCEPT citrus :has colour :
YELLOW}
```

⁹ According to the Information Retrieval literature, *informational queries* are different from *transactional* and *navigational* queries. In *informational queries*, the intention of the user is to obtain specific information concerning a given object (Jansen et al 2008).

the result returned from the DL representation will be null, while the “correct” answer (i.e. correct with respect to the intention of the user) will be generated from the prototypical component of the representation.

References

- Baader, F., D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, 2010, *The Description Logic Handbook: Theory, Implementations and Applications*, 2nd edition, Cambridge University Press.
- Baader, F., and B. Hollunder, 1995, “Embedding defaults into terminological knowledge representation formalisms”, *J. Autom. Reasoning* 14, 1: 149–180.
- Barsalou, L. W., 1985, “Continuity of the conceptual system across species”, *Trends in Cognitive Science* 9, 7: 305–311.
- Bizer, C., T. Heath, and T. Berners-Lee, 2009, “Linked Data – The Story So Far”, *International Journal on Semantic Web and Information System* 5, 3: 1–22.
- Bermudez, J. L., 1995, “Nonconceptual content: From perceptual experience to subpersonal computational states”, *Mind and Language* 10: 333–369.
- Bermudez, J. L., and A. Cahen, 2011, “Nonconceptual mental content”, *Stanford Encyclopaedia of Philosophy*, <http://plato.stanford.edu/>
- Bobillo, F., and U. Straccia, 2009, “An OWL Ontology for Fuzzy OWL 2”, *Proceedings of the 18th International Symposium on Methodologies for Intelligent Systems (ISMIS-09)*, Lecture Notes in Computer Science, Springer Verlag.
- Bonatti, P. A., C. Lutz, and F. Wolter, F., 2006, “Description logics with circumscription”, pages 400–410 in: *Proc. of KR*.
- Brachman, R., 1985, “I lied about the trees”, *The AI Magazine* 3, 6: 80–95.
- Brachman, R., and J. G. Schmolze, 1985, “An overview of the KL-ONE knowledge representation system”, *Cognitive Science* 9: 171–216.
- Brachman, R., and H. Levesque (eds.), 1985, *Readings in Knowledge Representation*, Morgan Kaufmann, Los Altos, CA.
- Brandom, R., 1994, *Making it Explicit*, Harvard University Press, Cambridge, MA.
- Calegari, S., and D. Ciucci, 2007, “Fuzzy Ontology, Fuzzy Description Logics and Fuzzy-OWL”, *Proc. WILF 2007*, LNCS, volume 4578.
- Dell’Anna, A., and M. Frixione, 2010, “On the advantage (if any) and disadvantage of the conceptual/nonconceptual distinction for cognitive science”, *Minds & Machines* 20: 29–45.
- Ding, Z., Y. Peng, and R. Pan, 2006, “BayesOWL: Uncertainty modeling in Semantic Web ontologies”, in: *Soft Computing in Ontologies and Semantic Web*, Z. Ma (ed.), Studies in Fuzziness and Soft Computing, volume 204, Springer.

- Evans, J. S. B. T., and K. Frankish (eds.), 2008, *In Two Minds: Dual Processes and Beyond*, Oxford UP, New York, NY.
- Fodor, J., 1981, “The present status of the innateness controversy”, in: J. Fodor *Representations*, The MIT Press, Cambridge, MA.
- Fodor, J., 1987, *Psychosemantics*, The MIT Press/A Bradford Book, Cambridge, MA.
- Fodor, J., 1998, *Concepts: Where Cognitive Science Went Wrong*, Oxford, UK: Oxford University Press.
- Fodor, J., and Z. Pylyshyn, 1988, “Connectionism and cognitive architecture: A critical analysis”, *Cognition* 28: 3–71.
- Frixione, M., 2007, “Do concepts exist? A naturalistic point of view”, in: *Explaining the Mental*, C. Penco, M. Beaney, M. Vignolo (eds.), Cambridge Scholars Publishing, Cambridge, UK.
- Gagliardi, F., 2008, “A Prototype-Exemplars Hybrid Cognitive Model of ‘Phenomenon of Typicality’ in Categorization: A Case Study in Biological Classification”, pages 1176–1181 in: *Proc. 30th Annual Conf. of the Cognitive Science Society*, Austin, TX.
- Gagliardi, F. 2010, “Cognitive Models of Typicality in Categorization with Instance-Based Machine Learning”, pages 115–130 in: *Practices of Cognition. Recent Researches in Cognitive Science*, University of Trento Press.
- Gao, M., and C. Liu, 2005, “Extending OWL by fuzzy Description Logic”, *Proc. 17th IEEE Int. Conf. on Tools with Artificial Intelligence (ICTAI 2005)*, IEEE Computer Society, Los Alamitos, pp. 562–567.
- Hayes, P., 2001, “Dialogue on rdf-logic. Why must the web be monotonic?. World Wide Web Consortium (W3C)”, <http://lists.w3.org/Archives/public/www-rdf-logic/2001Jul/0067.html>
- Klinov, P., and B. Parsia, 2008, “Optimization and evaluation of reasoning in probabilistic description logic: Towards a systematic approach”, pages 213–228 in: *ISWC 2008*, A. P. Sheth, S. Staab, M. Dean, M. Paolucci, D. Maynard, T. Finin, and K. Thirunarayan, (eds.), LNCS, volume 5318, Springer, Heidelberg.
- Lukasiewicz, Th., and U. Straccia, 2008, “Managing uncertainty and vagueness in description logics for the Semantic Web”, *Journal of Web Semantics* 6: 291–308.
- Machery, E., 2005, “Concepts are not a natural kind”, *Philosophy of Science* 72: 444–467.
- Machery, E., 2009, *Doing without Concepts* Oxford University Press, Oxford, UK.
- Medin, D. L., and P. J. Schwanenflugel, 1981, “Linear separability in classification learning”, *J. of Exp. Psyc.: Human Learning and Memory* 7: 355–368.

- Minsky, M., 1975, “A framework for representing knowledge”, in: *The Psychology of Computer Vision*, Patrick H. Winston (ed.), McGraw-Hill, New York (republished: Brachman & Levesque, New York, 1985).
- Murphy, G. L., 2002, *The Big Book of Concepts*, The MIT Press, Cambridge, MA.
- Osherson, D. N., and E. E. Smith, 1981, “On the adequacy of prototype theory as a theory of concepts”, *Cognition* 11: 237–262.
- Peacocke, C., 1992, *A Study of Concepts*, The MIT Press, Cambridge, MA.
- Rosch, E., 1975, “Cognitive representation of semantic categories”, *Journal of Experimental Psychology* 104: 573–605.
- Spelke, E. S., 1994, “Initial knowledge: six suggestions”, *Cognition* 50: 431–445.
- Spelke, E. S., and K. D. Kinzler, 2007, “Core knowledge”, *Developmental Science* 10, 1: 89–96.
- Stanovich, K.E., and R. West, 2000, “Individual Differences in Reasoning: Implications for the Rationality Debate?”, *The Behavioural and Brain Sciences* 23, 5: 645–665.
- Stoilos, G., G. Stamou, V. Tzouvaras, J. Z. Pan, I. Horrocks, 2005, “Fuzzy OWL: Uncertainty and the Semantic Web”, *Proc. Workshop on OWL: Experience and Directions (OWLED 2005)*, CEUR Workshop Proceedings, volume 188.
- Straccia, U., 1993, “Default inheritance reasoning in hybrid kl-one-style logics”, *Proc. IJCAI*, pp. 676–681.
- Wittgenstein, L., 1953, *Philosophische Untersuchungen*, Oxford, Blackwell.

MARCELLO FRIXIONE
DAFIST
University of Genova, Italy
marcello.frixione@gmail.com

ANTONIO LIETO
Department of Computer Science
University of Turin, Italy
lieto@di.unito.it