DELTA: A Unifying Categorization Algorithm Integrating Prototypes, Exemplars and Theory-Theory Representations and Mechanisms

Antonio Lieto
Università di Torino, Dipartimento di Informatica e ICAR-CNR (Palermo)

This paper details how to reconcile, under a heterogeneous representational perspective, different theories of typicality about conceptual representation and reasoning that have been traditionally seen as incompatible. In particular, it provides a novel theoretical hypothesis - as well as a novel categorization algorithm called **DELTA** - (i.e. unifie**D** Cat**E**gorization a**L**gorithm for he**T**erogeneous represent**A**tions) able integrate, in a cognitive artificial agent, the representational and reasoning assumptions of the theory-theory of concepts with the those ascribed to the prototype and exemplars-based theories[1].
The DELTA algorithm (detailed below) has the goal of selecting, given a certain stimulus d perceived from the environment, the most appropriate **typicality-based** representation available in the declarative memory of a cognitive agent (i.e. a prototype, an exemplar or a theory-like structure).

Data: Stimulus d; list of candidate representations: $closed^{\mathcal{S}1}$.
Result: A typicality based representation of a category.

$closed^{\mathcal{S}1} = \{\emptyset\}$
$\mathcal{S}1_{\text{EX}} \leftarrow categorizeExemplars(d)$;
if $firstOf(\mathcal{S}1_{\text{EX}}, closed^{\mathcal{S}1}).distance(d) < similarityThreshold$ **then**
  return $firstOf(\mathcal{S}1_{\text{EX}}, closed^{\mathcal{S}1})$;
else
  $\mathcal{S}1_{\text{PR}} \leftarrow categorizePrototypes(d)$ **return** $firstOf(\mathcal{S}1_{\text{PR}}, closed^{\mathcal{S}1})$;
end
if $firstOf(\mathcal{S}1_{\text{PR}}, closed^{\mathcal{S}1}).distance(d) > ConceptualCoherenceThreshold$ **then**
  return $firstOf(\mathcal{S}1_{\text{PR}}, closed^{\mathcal{S}1})$;
else
  $\mathcal{S}1_{\text{T}} \leftarrow categorizeTheory(d)$;
  return $firstOf(TheoryBasedCategorization, closed^{\mathcal{S}1})$;
end

Algorithm 1: A Unified categorization algorithm for prototypes, exemplars and theory-like representations.

Following a preference that has been experimentally observed in human cognition (Medin and Schaffer, 1978), DELTA assumes that exemplars and prototypes-based categorization are executed first (and in the following order: first exemplars-based categorization is attempted and then a prototype-based one) but also include a theory-like mechanism able to eventually discard the categorization result based on prototypes in favor of more coherent theory-like representations (as originally shown in the experiments by (Keil, 1989). The heuristics used for the choice between prototype and theory-like representation is grounded on the notion of Conceptual Coherence (Thagard, and Verbeurgt, 1998). In particular: the prototypical answer is maintained in case the considered stimulus results to be "coherent enough" with respect to the corresponding micro-theory related to the selected prototype, otherwise it is overridden by the theory-like representation which is closer to the stimulus.

Representational assumptions for DELTA

The DELTA algorithms relies on the representational hypothesis according to which conceptual structures, in natural and artificial systems, are *heterogeneous proxytypes* (see Lieto, 2014 for details). In this view, a concept

[1] Unification strategies for knowledge level processing are a crucial aspect in the current research on computational cognitive architectures (Lieto, Lebiere, Oltramari 2018).

is composed by heterogeneous bodies of knowledge containing different types of information associated to the the same conceptual entity. Each body of knowledge is a proxytype (i.e. an element of a representational network stored in the agent's long term memory corresponding to a particular category and that can tokenized in working memory to "go proxy" for that category). Furthermore, each body of conceptual knowledge is assumed to be featured by specific processes in which such representations are involved (e.g., in cognitive tasks like recognition, learning, categorization, etc.). Such heterogeneous perspective has been explicitly taken into account in the DUAL-PECCS system (http://www.dualpeccs.di.unito.it). In particular, the current version of DUAL-PECCS exhibits both prototype and exemplars-based categorization.

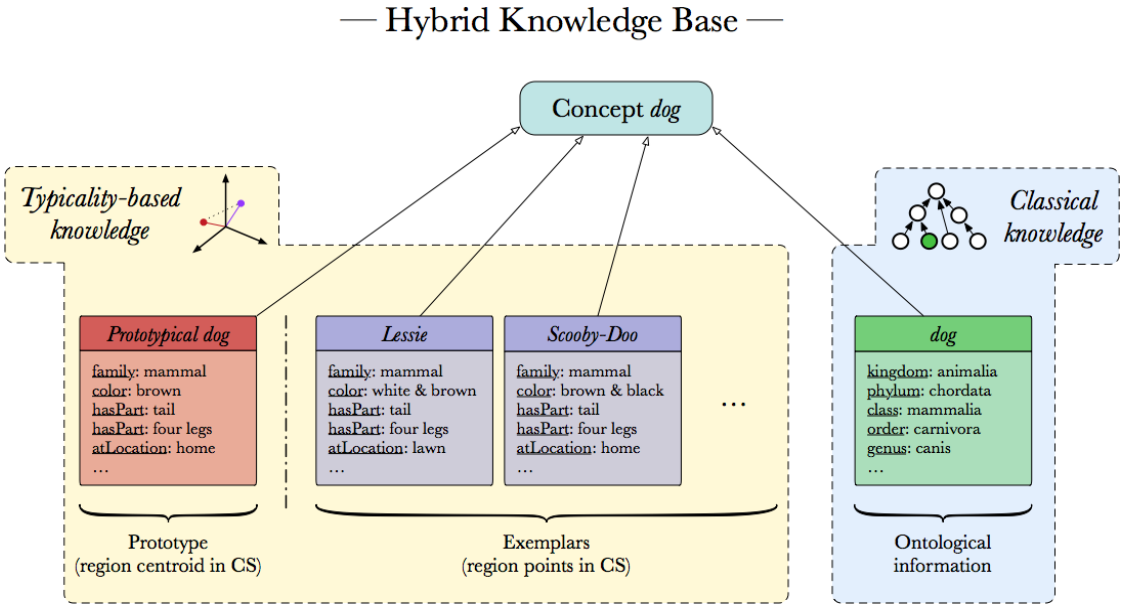

Fig. 1. An example of heterogeneous representation of the concept DOG in DUAL-PECCS

A missing part of its the current heterogeneous conceptual architecture concerns the representation of the default knowledge in terms of theory-like representational structures (while it already integrates classical, prototypical and exemplars based knowledge representation and processing mechanisms). To overcome these limits, I propose: i) to use graphical models (Danks, 2004) in order to represent the type of common-sense knowledge assumed in the theory-theory hypothesis; ii) to adopt the proposed unifying algorithm to explicitly integrate different types of typicality-based categorization mechanisms. A preliminary test on the task of common-sense linguistic categorization (involving only prototypes and exemplars based representations and reasoning procedures) has obtained promising results when compared with human performances (with an overlapping of the 89% of the responses, see (Lieto, Radicioni, and Rho, 2017)).

References

Danks, D. (2004). Psychological theories of categorizations as probabilistic models, CMU Technical report.

Keil, F. (1989). Concepts, kinds, and cognitive development. Cambridge, MA. . MIT press.

Lieto, A. (2014). A computational framework for concept representation in cognitive systems and architectures: Concepts as heterogeneous proxytypes. *Procedia Computer Science*, *41*, 6-14.

Lieto, A., Radicioni, D.P., Rho, V. (2017). Dual PECCS: A Cognitive System for Conceptual Representation and Categorization. *Journal of Experimental & Theoretical Artificial Intelligence* 29(2), 433–452 http://dx.doi.org/10.1080/0952813X.2016.1198934

Lieto, A., Lebiere, C., & Oltramari, A. (2018). The knowledge level in cognitive architectures: Current limitations and possible developments. *Cognitive Systems Research*, *48*, 39-55.

Lieto, A. (2019). Heterogeneous Proxytypes Extended: Integrating Theory-like Representations and Mechanisms with Prototypes and Exemplars, in *Proceedings of BICA 2018, 9th Internat.Conference on Biologically Inspired Cognitive Architecture*s, Springer, Advances in Intelligent Systems and Computing.

Medin, D.L., Schaffer, M.M. (1978). Context theory of classification learning. *Psychological review* 85(3), 207.

Thagard, P., Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cognitive Science* 22(1), 1–24.