Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

A Common-Sense Conceptual Categorization System Integrating Heterogeneous Proxytypes and the Dual Process of Reasoning

Antonio Lieto°⋆, Daniele P. Radicioni° & Valentina Rho°
° Dipartimento di Informatica — Università di Torino — Italy
⋆ ICAR-CNR, Palermo — Italy
{lieto,radicion}@di.unito.it;rho.valentina@gmail.com

Abstract

In this article we present DUAL-PECCS, an integrated Knowledge Representation system aimed at extending artificial capabilities in tasks such as conceptual categorization. It relies on two different sorts of cognitively inspired common-sense reasoning: *prototypical* reasoning and *exemplars-based* reasoning. Furthermore, it is grounded on the theoretical tenets coming from the dual process theory of the mind, and on the hypothesis of *heterogeneous proxytypes*, developed in the area of the biologically inspired cognitive architectures (BICA). The system has been integrated into the ACT-R cognitive architecture, and experimentally assessed in a conceptual categorization task, where a target concept illustrated by a simple common-sense linguistic description had to be identified by resorting to a mix of categorization strategies. Compared to human-level categorization, the obtained results suggest that our proposal can be helpful in extending the representational and reasoning conceptual capabilities of standard cognitive artificial systems.

1 Introduction

In this work we present an integrated knowledge representation system aimed at performing conceptual categorization tasks. It is named DUAL-PECCS (Prototypes and Exemplars-based Conceptual Categorization System), since it relies on two different sorts of cognitively-inspired common-sense reasoning: *prototypical* reasoning and *exemplars-based* reasoning. Furthermore, it is grounded on the theoretical tenets coming from the dual process theory of mind, and on the hypothesis of "heterogeneous proxytypes" developed in the area of the biologically inspired cognitive architectures (BICA).

The system aims at providing a unified framework for the conceptual categorization simulating some of the common sense heuristic strategies exploited by humans in categorization tasks. More specifically, it integrates strategies based on prototypes and exemplars-based reasoning, as suggested by the psychological results coming form the area of experimental Cognitive Science. Furthermore, DUAL-PECCS has been also integrated and tested in the ACT-R cognitive architecture to investigate its compatibility with the model of mind herein implemented [Anderson *et al.*, 2004; Langley *et al.*, 2009].

While existing systems and architectures allow to perform either prototype or exemplar-based categorization rather than autonomously adapting their strategy to the input being categorized [Anderson and Betz, 2001], conversely, DUAL-PECCS addresses this issue. In addition to the deployment of such common sense categorization strategies, DUAL-PECCS also integrates such types of non monotonic reasoning with the classical categorizations based on standard, deductive, processes. The flow and the interaction of such diverse reasoning mechanisms has been devised based on the tenets coming from the dual process theory of reasoning.

This paper is organized as follows: in Section 2 we introduce the theoretical background inspiring our system; in Sections 3 and 4 we illustrate the system, the implemented categorization strategies and its integration into ACT-R; Section 5 describes the evaluation and discusses the obtained results, and Section 6 provides an outlook on future work.

2 Prototypes, Exemplars and Proxytypes

In Cognitive Science different theories about the nature of concepts have been proposed. According to the traditional view, known as "classical" or Aristotelian theory, concepts can be simply defined in terms of sets of necessary and sufficient conditions. Such theory was dominant until the mid '70s of the last Century, when Rosch's experimental results demonstrated its inadequacy for ordinary –or *common sense*– concepts [Rosch, 1975]. Such results showed, on the other hand, that ordinary concepts are characterized and organized in our mind in terms of *prototypes*. Since then, different theories of concepts have been proposed to explain different representational and reasoning aspects concerning the problem of typicality: we focus here on the prototype theory and on the exemplars theory.[1] According to the *prototype* view, knowledge about categories is stored in terms of prototypes, i.e., in terms of some representation of the "best" instance

[1]Due to space restrictions, we briefly survey these works; full-detailed reviews can be found in [Machery, 2009].

of the category. In this view, the concept *cat* should coincide with the representation of a typical 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 set of specific exemplars explicitly stored in memory: the mental representation of the concept *cat* is a set containing the representations of (some of) the cats we encountered during our past experience.

Although these approaches have been largely considered as competing ones (since they propose different models and predictions about how we organize and reason on conceptual information), they turned out to be not mutually exclusive [Malt, 1989]. Rather, they seem to succeed in explaining different classes of cognitive *phenomena*, such as the fact that human subjects use different representations to categorize concepts: some use exemplars, a few rely on prototypes, and often both exemplars and prototypes are employed [Smith and Minda, 1998]. This distinction also has neural plausibility, as witnessed by empirical research [Squire and Knowlton, 1995]. Such experimental evidences led to the development of the so called "heterogeneous hypothesis" about the nature of concepts: this approach assumes that concepts do not constitute a unitary phenomenon, and hypothesizes that different types of conceptual representations exist: prototypes, exemplars, classical representations, and so on [Machery, 2009]. All such representations, in this view, constitute different *bodies of knowledge* and contain different types of information associated to the the same conceptual entity (i.e., these different bodies of knowledge act like semantic pointers towards the same conceptual entity [Eliasmith *et al.*, 2012; Thagard, 2012]). Furthermore, each body of conceptual knowledge is featured by specific processes in which such representations are involved (e.g., in cognitive tasks like recognition, learning, categorization, *etc.*). In particular, prototypes and exemplars-based representation are associated with the possibility of dealing with non monotonic strategies of reasoning and categorization, while the classical representations are associated with standard deductive mechanisms of reasoning.[2]

In recent years an alternative theory of concepts has been proposed: the *proxytype theory*; it postulates a biological localization and interaction between different brain areas for dealing with conceptual structures, that have a direct counterpart in the distinction between *long term* and *working memory* [Prinz, 2002]. In this setting, concepts are seen as *proxytypes*.

Definition 1 (Proxytypes). Proxytype *is any element of a complex representational network* stored in long-term memory *corresponding to a particular* category *that can be tokenized in working memory to 'go proxy' for that category* [Prinz, 2002].

In other terms, proxytype theory considers concepts as *temporary constructs* of a given category, activated (tokenized) in working memory as a result of conceptual processing activities, such as concept identification, recognition and retrieval.

In its original formulation, however, proxytypes are depicted as monolithic conceptual structures, primarily intended as prototypes [Prinz, 2002]. A revised view of this approach has been recently proposed in the area of BICA [Lieto, 2014], hypothesizing the availability of a wider range of representation types than just prototypes corresponding to the kinds of representations postulated by the heterogeneous approach to concepts. In this sense, proxytypes are heterogeneous in nature.

Definition 2 (Heterogeneous Proxytypes). *Heterogeneous representations (such as* prototypes*,* exemplars*, etc.) for each conceptual category are* stored in long-term memory. *They can be activated and accessed by resorting to different categorization strategies. In this view, each representation has its associated access procedures.*

In the design of our system we followed the approach based on heterogeneous proxytypes for both the representational level (that is, we devised a hybrid knowledge base composed of heterogeneous representations, each endowed with specific reasoning mechanisms) and for the 'proxyfication' (that is, the set of procedures implementing the tokenization of the different representations in working memory). In the following section we illustrate how the heterogeneous conceptual representations can be implemented through the classical AI paradigms and representational frameworks.

2.1 Representational Paradigms in AI

In Artificial Intelligence, and in particular in the area of Knowledge Representation, several proposals have been carried out to handle different aspects of conceptual information (e.g., in processes such as learning and reasoning). A classical distinction was drawn between symbolic and sub-symbolic models. While sub-symbolic (connectionist) models were used for embedding knowledge structures by taking inspiration from human-like organizations and processes, on the other hand many forms of logic-based systems were developed mostly targeted on providing a clear formal semantics, thus enabling forms of logically-valid automatic reasoning. Examples of these systems are the KL-ONE semantic networks [Brachman and Schmolze, 1985]. In more recent years, a tripartition of representational levels has been proposed, where a further level is considered, the *conceptual level* [Gärdenfors, 2000; 2014]. This level of representation is considered as intermediate between the symbolic level and the sub-symbolic level, and is featured by a representation in terms of conceptual spaces, i.e., geometrical representations of knowledge (introduced in Section 3).

Such a tripartition provides several computational frameworks to encode the different bodies of conceptual knowledge. For example, the *prototypical body of knowledge* for a

[2] Let us assume that we have to categorize a *stimulus* with the following features: "it has fur, woofs and wags its tail". In this case, the result of a *prototype-based categorization* would be *dog*, since these cues are associated to the prototype of *dog*. Prototype-based reasoning, however, is not the only type of reasoning based on typicality. In fact, if an exemplar corresponding to the *stimulus* being categorized is available, too, it is acknowledged that humans use to classify it by evaluating its similarity w.r.t. the exemplar, rather than w.r.t. the prototype associated to the underlying concepts [Frixione and Lieto, 2013]. This type of common sense categorization is known in literature as *exemplars-based categorization*.

given concept can be represented as follows: *i)* from a symbolic perspective, in terms of frames [Minsky, 1975] or semantic networks [Quillian, 1968]; *ii)* from a conceptual space perspective, prototypes can be geometrically represented as centroids of a convex region (more on this aspect later); *iii)* from a sub-symbolic perspective, the prototypical knowledge concerning a concept can, on the other hand, be represented as reinforced patterns of connections in Artificial Neural Networks (ANNs). Similarly, for the *exemplars-based body of knowledge*, both symbolic and conceptual space representations can be used, as well as the sub-symbolic paradigm. In particular, exemplars can be represented as instances of a concept in symbolic systems, as points in a geometrical conceptual space, or as a particular (local) pattern of activation in a ANN. Finally, also for the classical body of knowledge it is –at least in principle–, possible to use the same frameworks. However, this seems to be a case where symbolic and conceptual levels are more appropriate w.r.t. the sub-symbolic one.

Summing up, all the different types of conceptual representations can be implemented in cognitive artificial systems and architectures. In addition, different computational mechanisms for "proxyfying" conceptual representations can be applied. In the next Section we illustrate and discuss the representational levels and the associated computational frameworks we adopted for each type of body of knowledge.

3 The DUAL-PECCS System

As mentioned, the DUAL-PECCS relies on the heterogeneous proxytypes approach and on the dual process theory. It is equipped with a hybrid knowledge base composed of heterogeneous representations of the same conceptual entities: that is, the hybrid knowledge base includes prototypes, exemplars and classical representations for the same concept. Both prototypes and exemplars are represented at the conceptual level (see Section 3.1), while classical information is represented through standard symbolic formalisms (i.e., by means of a formal ontology).

The retrieval of such representations is driven by different process types. In particular, prototype and exemplar-based retrieval is based on a fast and approximate kind of categorization, and benefits from common-sense information associated to concepts. On the other hand, the retrieval of the classical representation of concepts is featured by explicit rule following, and makes no use of common-sense information. These two differing categorization strategies have been widely studied in psychology of reasoning in the frame of the dual process theory, that postulates the co-existence of two different types of cognitive systems [Evans and Frankish, 2009; Kahneman, 2011]. The systems of the first type (*type 1*) are phylogenetically older, unconscious, automatic, associative, parallel and fast. The systems of the second type (*type 2*) are more recent, conscious, sequential and slow, and featured by explicit rule following. We assume that both systems can be composed in their turn by many sub-systems and processes. According to the hypotheses in [Frixione and Lieto, 2012; 2014], the conceptual representation of our system includes two main sorts of components, based on these two sorts of

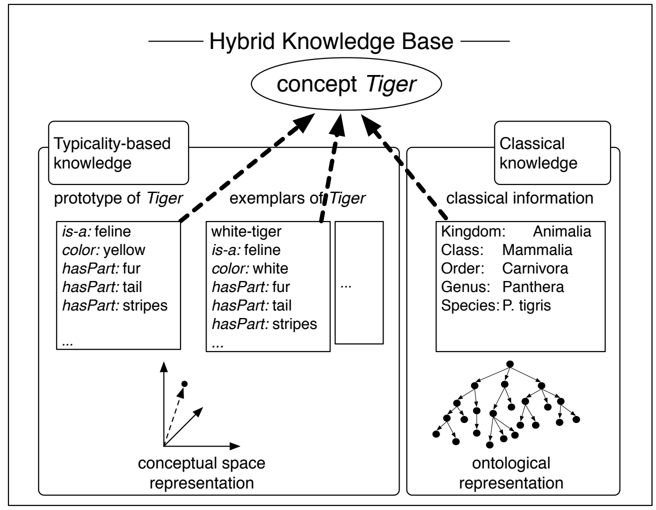

Figure 1: Heterogeneous representation of the *tiger* concept

processes. *Type 1* processes have been designed to deal with prototypes- and exemplar-based retrieval, while *Type 2* processes have been designed to deal with deductive inference.

The two sorts of system processes interact (Algorithm 1), since Type 1 processes are executed first and their results are then refined by Type 2 processes. In the implemented system the typical representational and reasoning functions are assigned to the System 1 (hereafter $\mathcal{S}1$), which executes processes of Type 1, and are associated to the Conceptual Spaces framework [Gärdenfors, 2000; 2014]. On the other hand, the classical representational and reasoning functions are assigned to the System 2 (hereafter $\mathcal{S}2$) to execute processes of Type 2, and are associated to a standard ontological representation.

Figure 1 shows the heterogeneous representation for the concept *tiger*, with prototypical and exemplar-based representations semantically pointing to the same conceptual entity. In this example, the exemplar and prototype-based representations make use of non classical information. Namely, the prototypical representation grasps information such as that tigers are wild animals, their fur has yellow and black stripes, *etc.*; the exemplar-based representations grasp information on individuals (such as *white-tiger*, which is a particular tiger with white fur). Both sorts of representations activate Type 1 processes. On the other hand, the classical body of knowledge is filled with necessary and sufficient information to characterize the concept (representing, for example, the taxonomic information that a tiger is a *mammal* and a *carnivore*), and activates Type 2 processes.

In the following we introduce the two representational and reasoning frameworks used in our system, by focusing *i)* on how typicality information (including both prototypes and exemplars) and their corresponding non monotonic reasoning procedures can be encoded through conceptual spaces; and *ii)* on how classical information can be naturally encoded in terms of formal ontologies.

3.1 $\mathcal{S}1$-$\mathcal{S}2$: Conceptual Spaces and Ontologies

Conceptual spaces (CS) are a representational framework where knowledge is represented as a set of *quality dimensions*, and where a geometrical structure is associated to each quality dimension. Instances can be represented as points in a multidimensional space, and their similarity can be computed as the intervening distance between each two points, based on some suitable metrics (such as Euclidean and Manhattan distance, or standard cosine similarity).[3] In this setting, concepts correspond to convex regions, and regions with different geometrical properties correspond to different sorts of concepts [Gärdenfors, 2000].

Prototypes have a natural geometrical interpretation in conceptual spaces, in that they correspond to the geometrical centre of a convex region. This can be thought of as a centroid, that is the mean position of all the points in all dimensions. This representation also allows us, given a convex region, to associate each point to a certain centrality degree, that can be interpreted as a measure of its typicality. This framework has been extended to consider the exemplars, that are also represented as points in the multidimensional space. Conceptual spaces can be also used to compute the proximity between any two entities, and between entities and prototypes. Concepts, in this framework, are characterized in terms of *domains*; a domain is "a set of integral dimensions that are separable from all other dimensions" [Gärdenfors, 2014]. Typical domain examples are color, size, shape, texture. In turn, domain information can be specified along some dimensions: e.g., in the case of the *color* domain, relevant dimensions are hue, chromaticity, and brightness. Inference in conceptual spaces is robust to incomplete and/or noisy information.

On the other hand, the representation of the classical information regarding a given concept is demanded to classical ontological formalizations. In this setting, ontologies provide the characterization of concepts in terms of necessary and sufficient conditions (if these conditions exists: as mentioned, most common sense concepts cannot be characterized in these terms). Additionally, the ontological representations are used by the $\mathcal{S}2$ component (in our implementation it is grounded on the OpenCyc ontology, one of the widest ontological resources currently available). In the next sections we present the details regarding the categorization strategies adopted in the DUAL-PECCS.

3.2 Categorization Pipeline of the DUAL-PECCS

The whole categorization pipeline works as follows. The input to the system is a simple linguistic description, like 'The animal that eats bananas', and the expected output is a given category evoked by the description (e.g., the category monkey in this case). After an Information Extraction (IE) step, the input information is encoded into an internal format devised to store conceptual spaces information, which is then used as input in the categorization task by adopting the strategies that will be described below.

[3]A full account of the semantic similarity calculated in the conceptual spaces is out of the scope of this contribution; in the present setting, distances are computed in a multi-dimensional space that can be thought of as a vectorial model [Lieto *et al.*, In Pressb].

Data: Linguistic d
Result: A class assignment, as computed by $\mathcal{S}1$ and $\mathcal{S}2$

trialCounter $\leftarrow 0$;
closed$^{\mathcal{S}1} = \{\emptyset\}$
while *trialCounter* $<$ *maxTrials* **do**
    `// conceptual spaces output`
    $\mathsf{c} \leftarrow \mathcal{S}1(d, closed^{\mathcal{S}1})$;
    if *trialCounter* $== 0$ **then** $\mathsf{c}^* \leftarrow \mathsf{c}$;
    `// ontology based consistency check`
    $\mathsf{cc} \leftarrow \mathcal{S}2(d, conceptPointedBy(\mathsf{c}))$;
    if cc *equals(conceptPointedBy*(c)) **then**
        return $\langle \mathsf{c}^*, \mathsf{cc} \rangle$;
    else
        closed$^{\mathcal{S}1}$ *add(conceptPointedBy*(c))
    end
    ++*trialCounter* ;
end
$\mathsf{cc} \leftarrow \mathcal{S}2(\langle d, \mathsf{Thing} \rangle)$;
return $\langle \mathsf{c}^*, \mathsf{cc} \rangle$;

Algorithm 1: The $\mathcal{S}1$-$\mathcal{S}2$ categorization process.

A shallow IE approach has been devised, where the morphological information computed from input sentences has been used to devise a simple finite-state automaton describing the input sentences' structure (more on the input descriptions in Section 5). This approach would not scale to handle more complex sentences; its limitations are due to using morphological information, and in general are inherent in finite-state machines (which prevents us from dealing with parenthetical clauses, like relative clauses). We defer to future work the adoption of richer language models. Despite these limitations, however, it allowed us to complete the automatization of the software pipeline going all throughout from the simple linguistic input description used for the evaluation (that will be described later) to its final conceptual categorization.

3.3 Dual Process Prototypes and Exemplars-based categorizations

The overall control strategy implemented by DUAL-PECCS regulates the flow of interaction between the $\mathcal{S}1$ and $\mathcal{S}2$ systems (it is thus referred to as $\mathcal{S}1$-$\mathcal{S}2$ categorization). Its underlying rationale is to assess the approximate categorization results obtained by Type 1 processes in $\mathcal{S}1$ with the ontological information and the deliberative processes of reasoning implemented by $\mathcal{S}2$. The $\mathcal{S}1$-$\mathcal{S}2$ categorization process can be summarized as follows (Algorithm 1). The system takes in input a textual description *d* and produces in output a pair $\langle \mathsf{c}^*, \mathsf{cc} \rangle$, the output of $\mathcal{S}1$ and $\mathcal{S}2$, respectively. If the categorization result provided by $\mathcal{S}1$ (based on the similarity calculation between the input and $\mathcal{S}1$ representations) is consistent with the ontology, then the categorization succeeded and the category provided by $\mathcal{S}2$ (cc) is returned along with c^*, the top scoring class returned by $\mathcal{S}1$; otherwise the system evaluates a fixed amount (*maxTrials*) of $\mathcal{S}1$ candidates. In case all the $\mathcal{S}1$ candidates are inconsistent w.r.t. the ontology in $\mathcal{S}2$, the output of $\mathcal{S}2$, computed independently of $\mathcal{S}1$, is provided along with c^*. The control strategy implements a tradeoff be-

Data: Linguistic description: d; list of inconsistent concepts: *closed*$^{\mathcal{S}1}$.
Result: A typicality based representation of a category.

$\mathcal{S}1_{\text{EX}} \leftarrow$ *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);
  `// in case of equal distance prefer exemplars`
  typicalityCategorization $\leftarrow$ *sortResults*($\mathcal{S}1_{\text{EX}}$, $\mathcal{S}1_{\text{PR}}$);
  return *firstOf*(*typicalityCategorization*, *closed*$^{\mathcal{S}1}$);
end

Algorithm 2: $\mathcal{S}1$ categorization with prototypes and exemplars implementing the instruction in Algorithm 1: line 4.

tween ontological inference and the output of $\mathcal{S}1$, which is more informative but also formally less reliable.

The second categorization algorithm governing the reasoning mechanisms is executed within the $\mathcal{S}1$ component (Algorithm 2), and extends the previous model of categorization by determining which kind of $\mathcal{S}1$ output must be selected and then checked against the deliberative $\mathcal{S}2$ module. In particular, the algorithm is designed to activate either the prototypical-based or the exemplar-based representation, based on the actual input description. The implemented procedure works as follows: when the input *stimulus* –in our case a simple linguistic description– is similar enough to an exemplar representation (a threshold has been fixed to these ends), the corresponding exemplar of a given category is retrieved. Otherwise, the prototypical representations are also scanned and the representation (prototype or exemplar) that is closest to the input is returned. By following a preference that has been experimentally observed in human cognition [Medin and Schaffer, 1978], this algorithm favors the results of the exemplars-based categorization if the knowledge-base stores any exemplars similar to the input being categorized.

4 Integrating Dual-PECCS into ACT-R

The proposed system has been integrated into the cognitive architecture ACT-R [Anderson *et al.*, 2004]. In ACT-R, cognitive mechanisms emerge from the interaction of two types of knowledge: *declarative* knowledge, that encodes explicit facts that the system knows, and *procedural* knowledge, that encodes rules for processing declarative knowledge. In particular, the declarative module is used to store and retrieve pieces of information (called *chunks*, featured by a *type* and a set of attribute-value pairs, similar to frame slots) in the declarative memory. Finally, the central production system connects these modules by using a set of IF-THEN production rules. We focused on the two following elements in ACT-R: the Declarative Memory, the Working Memory buffers, and the corresponding mechanisms of retrieval.

We have integrated the DUAL-PECCS heterogeneous representational model in ACT-R by extending the notion of chunk-based representation supported by such architecture. In particular, differently from other efforts made to extend the knowledge capabilities of ACT-R [Oltramari and Lebiere, 2012; Salvucci, 2014] based on the introduction of a new, *ad-hoc*, external module of declarative memory, we have directly integrated our hybrid knowledge base into the ACT-R declarative memory. Besides, the dual process strategies of concept categorization have been integrated into the ACT-R processes and connected to the retrieval request of the Working Memory. More in detail: in the *Extended* Declarative Memory every concept is represented as an empty chunk (that is, a chunk having no information except for its *WordNet ID* and a human readable *name*), referred to by different kinds of external bodies of knowledge (like prototypes and exemplars) acting like semantic pointers. The novel categorization mechanism — implemented by inserting a new retrieval request— activates both the $\mathcal{S}1$ and $\mathcal{S}2$ categorization procedures. In particular, in this setting, once the categorization result of $\mathcal{S}1$ is provided, the activated representation in the Extended Declarative long-term memory is proxyfied (i.e., stored in working memory) in order to perform the $\mathcal{S}2$ categorization check, in the dual process perspective.

5 Evaluation

A dataset composed of 90 descriptions (corresponding to very simple riddles) was collected and given in input to the implemented system: namely, we selected 45 descriptions for which an exemplar-based representation was expected to be retrieved, and 45 descriptions for which a prototype-based representation was expected to be retrieved. These *stimuli* have been built by a multidisciplinary team composed of neuropsychologists, linguists and philosophers in the frame of a project aimed at investigating the brain activation of visual areas in tasks of lexical processing even for very abstract concepts. An example of such descriptions is "The big carnivore with yellow fur and black stripes", where the expected category to be retrieved was *tiger*, and in particular its representation corresponding to the "prototype of *tiger*"; conversely, a description such as "The big carnivore with *white* fur and black stripes" was expected to lead as answer to "exemplar of *white_tiger*".[4]

The *expected* categorical targets represent a gold standard, since they correspond to the results provided by human subjects within a preliminary psychological experimentation. In the first experiment 30 subjects were requested to provide the corresponding target concept for each description (i.e., in this case no description referred to exemplar-based representations was provided) [Lieto *et al.*, In Pressa]; in the second experiment 10 subjects were requested to indicate the target (prototype or exemplar) corresponding to descriptions evoking both prototype- and exemplar-based representations. Such experimental design was borrowed from the third experiment described in [Malt, 1989].

We have considered a twofold experimental setting. In the first case we tested the whole pipeline, where the salient information is extracted by starting from the linguistic descrip-

[4]A prototype of the system is available at the URL http://goo.gl/sk9tJ1; the full list of descriptions is available at the URL http://goo.gl/PDLo4J.

Table 1: The accuracy results (Table 1-a) and the analysis of the proxyfication errors (Table 1-b).

a. Accuracy rates obtained for the CC-ACC and P-ACC metrics.

test	CC-ACC	P-ACC
with no IE	95.6% (86/90)	86.0% (74/86)
with IE	85.6% (77/90)	75.3% (58/77)

b. Analysis of the errors in the proxyfication (P-ACC metrics).

test	Proxyfication error		
	Ex-Proto	Proto-Ex	Ex-Ex
with no IE	12.8% (11/86)	1.1% (1/86)	0.0% (0/86)
with IE	18.2% (14/77)	0.0% (0/77)	6.5% (5/77)

tion, the corresponding representation is retrieved, proxyfied and loaded in working memory according to the dual process approach. The information extraction of the linguistic input is not implemented in ACT-R, but it relies on the CoreNLP Stanford Parser [Manning *et al.*, 2014], which is used to convert the textual description into a chunk request. This measure is intended to assess the robustness of the overall approach, from the input parsing to the final categorization. In the second case we tested the heterogeneous proxytypes approach by directly starting with a manual encoding of the linguistic *stimuli* in terms of chunk request: this measure is intended to assess the accuracy in the categorization task of the hybrid system, featured by dual process approach, heterogeneous representation and reasoning, proxyfication, integration in the ACT-R architecture, but no information extraction.

In both cases (all linguistic pipeline *vs.* clean input) we recorded two distinct metrics:
i) Concept-categorization accuracy (CC-ACC metrics) This metrics was designed to evaluate the final categorization, that is the accuracy in retrieving the expected concept (in this case, the wrong proxyfication did not count as error).
ii) Proxyfication accuracy (P-ACC metrics) This metrics was designed to evaluate whether given in input a description evoking a given concept, the expected proxyfied representation was retrieved. In this case the confusion between prototype and exemplar (or between exemplars) is scored as error even if the expected category is returned.

5.1 Results and Discussion

The system obtained an accuracy of 95.6% in conceptual retrieval (that reduces to 85.6% when performing the IE step), and 86.0% in the proxyfication task (75.3% in the setting with the IE). These figures are reported in Table 1.

The results in the conceptual categorization are in line with those previously reported in [Ghignone *et al.*, 2013; Lieto *et al.*, In Pressa], although the dataset was presently more diverse by including exemplars, as well. The results of the whole pipeline (Table 1-a, second row) provide a baseline for future implementations of the IE step; we observe that, although producing 10% error either in the POS tagging or in the IE step proper, this result is also in line with those reported in [Lieto *et al.*, In Pressb]. This fact shows that the approach –devised to match the simple linguistic structures in the considered *stimuli*, and which was expected not to generalize to handle further linguistic descriptions– maintains its performance when dealing with a broader dataset. The IE step significantly affects also the P-ACC metrics: that is, if we restrict to considering cases where the concept was categorized as expected, the proxyfication step is performed correctly in 86.0% of descriptions with 'clean' input, and only in 75.3% of cases with the Information Extraction step.

Table 1-b reports the detailed errors committed in the proxyfication phase; here we distinguish three cases. Provided that proxyfication errors occur only when the concept has been correctly categorized, three kinds of proxyfication errors were recorded: an exemplar returned in place of an expected prototype (column Ex-Proto); a prototype returned in place of an expected exemplar (column Proto-Ex), or by retrieving a wrong exemplar (e.g., siberian_tiger in place of malaysian_tiger, column Ex-Ex). Notably, the vast majority of errors are due to confusion between exemplars and prototypes: in particular, in the 12.8% of the considered *stimuli* an exemplar-proxyfied representation has been returned by the system in spite of the expected prototype. This sort of error raises to 18.2% in the implementation including the IE task. This error was caused by the fact that in case exemplar based representations in the KB are equipped with the typical information matching the linguistic description being categorized, then such representations are always favored w.r.t. their prototypical counterpart (see Section 3.3, Algorithm 2). The rest of errors are less relevant, except for the type Ex-Ex, where we observe a 6.5% error rate, which is mostly due to the noisy features extracted from the linguistic descriptions.

6 Conclusions

This paper has illustrated two main advancements, in that DUAL-PECCS provides the heterogeneous proxyfication approach governed by the dual process theory with a working implementation meantime integrating autonomous prototypical and exemplar-based categorization. Additionally, it has been integrated into ACT-R, thus showing a good level of compatibility with a general cognitive architecture. Although there is room both for refining the theory and tuning the implemented system, the obtained experimental results are encouraging. Further representational levels can be foreseen that fit the needs of cognitive architectures, such as sensory-motor representations, visual imageries, *etc.*. However, while it would be possible to use the proposed framework to integrate additional levels into the DUAL-PECCS (for example, *affordances* can be thought of as prototypical sensory-motor representations associated to a given concept), a complete account of these aspects is out of the scope of the current paper and deserves further investigation.

Acknowledgments

This work has been partly supported by the Project *the Role of the Visual Imagery in Lexical Processing*, grant TO-call03-2012-0046 funded by Università degli Studi di Torino and Compagnia di San Paolo.

References

[Anderson and Betz, 2001] John R Anderson and Jonathan Betz. A hybrid model of categorization. *Psychonomic Bulletin & Review*, 8(4):629–647, 2001.

[Anderson *et al.*, 2004] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. An integrated theory of the mind. *Psychological review*, 111(4):1036, 2004.

[Brachman and Schmolze, 1985] Ronald J Brachman and James G Schmolze. An Overview of the KL-ONE Knowledge Representation System. *Cognitive science*, 9(2):171–216, 1985.

[Eliasmith *et al.*, 2012] Chris Eliasmith, Terrence C Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. A large-scale model of the functioning brain. *science*, 338(6111):1202–1205, 2012.

[Evans and Frankish, 2009] Jonathan St BT Evans and Keith Ed Frankish. *In two minds: Dual processes and beyond.* Oxford University Press, 2009.

[Frixione and Lieto, 2012] Marcello Frixione and Antonio Lieto. Representing concepts in formal ontologies: Compositionality vs. typicality effects. *Logic and Logical Philosophy*, 21(4):391–414, 2012.

[Frixione and Lieto, 2013] Marcello Frixione and Antonio Lieto. Representing non classical concepts in formal ontologies: Prototypes and exemplars. In *New Challenges in Distributed Information Filtering and Retrieval*, pages 171–182. Springer, 2013.

[Frixione and Lieto, 2014] Marcello Frixione and Antonio Lieto. Towards an Extended Model of Conceptual Representations in Formal Ontologies: A Typicality-Based Proposal. *Journal of Universal Computer Science*, 20(3):257–276, March 2014.

[Gärdenfors, 2000] Peter Gärdenfors. *Conceptual spaces: The geometry of thought*. MIT press, 2000.

[Gärdenfors, 2014] Peter Gärdenfors. *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. MIT Press, 2014.

[Ghignone *et al.*, 2013] Leo Ghignone, Antonio Lieto, and Daniele P. Radicioni. Typicality-Based Inference by Plugging Conceptual Spaces Into Ontologies. In Antonio Lieto and Marco Cruciani, editors, *Procs. of the Int. Workshop on Artificial Intelligence and Cognition*. CEUR, 2013.

[Kahneman, 2011] Daniel Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

[Langley *et al.*, 2009] Pat Langley, John E Laird, and Seth Rogers. Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, 10(2):141–160, 2009.

[Lieto *et al.*, In Pressa] Antonio Lieto, Andrea Minieri, Alberto Piana, and Daniele P. Radicioni. A knowledge-based system for prototypical reasoning. *Connection Science*, In Press.

[Lieto *et al.*, In Pressb] Antonio Lieto, Daniele P. Radicioni, Marcello Frixione, and Valentina Rho. Extending Ontological Categorization through a Dual Process Conceptual Architecture. In *Selected Papers from the KEOD 2014 Proceedings*, LNCS – Communications in Computer and Information Science. Springer-Verlag, In Press.

[Lieto, 2014] Antonio Lieto. A computational framework for concept representation in cognitive systems and architectures: Concepts as heterogeneous proxytypes. *Procedia Computer Science*, 41:6–14, 2014.

[Machery, 2009] Edouard Machery. *Doing without concepts*. OUP, 2009.

[Malt, 1989] Barbara C Malt. An on-line investigation of prototype and exemplar strategies in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4):539, 1989.

[Manning *et al.*, 2014] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, 2014.

[Medin and Schaffer, 1978] Douglas L Medin and Marguerite M Schaffer. Context theory of classification learning. *Psychological review*, 85(3):207, 1978.

[Minsky, 1975] Marvin Minsky. A framework for representing knowledge. In P. Winston, editor, *The Psychology of Computer Vision*, pages 211–277. McGraw-Hill, New York, 1975.

[Oltramari and Lebiere, 2012] Alessandro Oltramari and Christian Lebiere. Pursuing artificial general intelligence by leveraging the knowledge capabilities of act-r. In *Artificial General Intelligence*, pages 199–208. Springer, 2012.

[Prinz, 2002] Jesse J Prinz. *Furnishing the mind: Concepts and their perceptual basis.* MIT press, 2002.

[Quillian, 1968] Ross Quillian. *Semantic Memory*, pages 216–270. MIT Press, 1968.

[Rosch, 1975] Eleanor Rosch. Cognitive representations of semantic categories. *J. Exp. Psychol. Gen.*, 104(3):192–233, 1975.

[Salvucci, 2014] Dario D Salvucci. Endowing a cognitive architecture with world knowledge. In *Procs. of the 36th Annual Meeting of the Cognitive Science Society*, 2014.

[Smith and Minda, 1998] J David Smith and John Paul Minda. Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(6):1411, 1998.

[Squire and Knowlton, 1995] Larry R Squire and Barbara J Knowlton. Learning about categories in the absence of memory. *Proceedings of the National Academy of Sciences*, 92(26):12470–12474, 1995.

[Thagard, 2012] Paul Thagard. *The cognitive science of science: Explanation, discovery, and conceptual change*. Mit Press, 2012.