View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Institutional Research Information System University of Turin

Integrating a Cognitive Framework for Knowledge Representation and Categorization in Diverse Cognitive Architectures

Antonio Lieto, Daniele P. Radicioni and Valentina Rho
{lieto,radicion}@di.unito.it, rho.valentina@gmail.com

Abstract

This paper describes the rationale followed for the integration of Dual-PECCS, a cognitively-inspired knowledge representation and reasoning system, into two rather different cognitive architectures, such as ACT-R and CLARION. The provided integration shows how the representational and reasoning mechanisms implemented by our framework may be plausibly applied to computational models of cognition based on different assumptions.

1 Introduction

In this work we illustrate how the knowledge representation and reasoning system Dual-PECCS,[1] aimed at performing the conceptual categorization tasks, was integrated into two different cognitive architectures: ACT-R [1] and CLARION [2]. Our system represents a unifying suite, where different sorts of cognitively-inspired common-sense reasoning (*prototypical* reasoning and *exemplars-based* reasoning) and standard monotonic categorization procedures are integrated and autonomously executed according to the stimulus being categorized. Although potentially conflicting, these different types of reasoning are harmonized according to the theoretical tenets coming from the dual process theories [3]. On the other hand, from a representational perspective, the system relies on the hypothesis that concepts are "heterogeneous proxytypes" [4]. This work is organized as follows: in Section 2 we sketch the main elements inspiring our system and its theoretical bases as well as its overall architecture, in Section 3 we show how our hybrid system for conceptual categorization was integrated into ACT-R and CLARION, and finally we elaborate on the future works.

2 Heterogeneous Proxytypes and Dual Process Model of Categorization

As mentioned, the two main theoretical cornerstones inspiring our system are the heterogeneous proxyxtypes approach and the dual process theory of reason-

[1]So named after 'Dual Prototypes and Exemplars-based Conceptual Categorization System'.

ing and rationality. According to the 'heterogeneous proxytypes' approach (for a detailed account, please refer to [4]), conceptual structures in cognitive systems and architectures are assumed to be formed by heterogeneous representations (or bodies of knowledge) referring to the same conceptual entity. That is, the different bodies of knowledge act as 'semantic pointers' in the sense intended by [5] towards the same reference concept. Each body of knowledge provides specific types of information and specific access and reasoning procedures to the concept they refer to. Such heterogeneous representations are 'proxytypes' [6], in the sense that they can be contextually activated by external stimuli, coming from the environment, and 'go proxy', in working memory, for their reference category. The proxyfication may be then the result of activities such as concept identification, recognition, retrieval, and so forth. This approach also allows to tackle the problem of the 'contextual activation' of knowledge: i.e., given a specific perceived stimulus to be categorized, only a specific portion of the available conceptual knowledge is activated (specifically, only what is "contextually relevant" w.r.t. the stimulus at hand). In other terms, according to this perspective we only proxyfy (i.e., activate in our working memory) the type of representation which is closer to the percept (see [4] for further details).

The different types of conceptual representations hypothesized to co-exist in the heterogeneous proxytypes approach are typicality-based representations of a given concept, as well as representations in terms of necessary and/or sufficient conditions. The typicality-based representations included in this approach regard not only prototypes but also exemplars-based representations of a given category.[2] In this respect, Dual-PECCS is then equipped with a hybrid knowledge base composed of heterogeneous representations for the same conceptual entities: that is, for a given concept the hybrid knowledge base includes prototypes, exemplars and classical representations (representations in terms of necessary and sufficient conditions). For example: the heterogeneous representation of the concept *tiger* includes prototypical and exemplar-based representations semantically pointing to the same conceptual entity, as well as a representation encoding necessary information. Namely, the prototypical representation grasps information such as that tigers are wild animals, their fur in most cases has yellow and black stripes, etc.; the exemplar-based representations grasp information on individuals, such as a given individual of white-tiger, which is a particular tiger with white fur. 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*.

From a reasoning perspective 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.[3] On the other hand,

[2] According to the exemplars perspective, a given category is mentally represented as set of specific exemplars explicitly stored within memory: e.g., the mental representation of the concept *cat* is the set of the representations of (some of) the cats we encountered during our past experience. For a detailed review regarding the differing theories about concepts, prototypes and exemplars, please refer to [7, 8].

[3] 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

the retrieval of 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 [3]. 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 turn by many sub-systems and processes. Following the hypotheses in [9, 10], the heterogeneous conceptual representation of our system includes, then, two main sorts of components that are based on different types of representations and that share these two sorts of processes: *Type 1* processes have been designed to deal with prototypes- and exemplar-based retrieval and categorization, while *Type 2* processes have been designed to deal with deductive inference.

The two sorts of system processes interact, 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 is associated to the Conceptual Spaces framework [11], where the reasoning functions are implemented as similarity calculations in a metric space. 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 Description Logics based ontological representation (in our case the OpenCyc ontology containing more than 230K concepts was used). The details of such integrated framework as well as the results of different experiments can be found in [12, 13]. In the next section we briefly describe the categorization pipeline of the system by presenting the dynamics of the interaction between $\mathcal{S}1$ and $\mathcal{S}2$ processes; a fuller account of this process is documented in [14].

2.1 Categorization Pipeline of the DUAL-PECCS

The whole categorization pipeline of the systems works as follows. The current 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 (the category *monkey* in this case).

The system answers rely on the the output of $\mathcal{S}1$ and $\mathcal{S}2$, respectively. The categorization provided by $\mathcal{S}1$ is based on approximate, defeasible, inference and is error prone. It runs on the conceptual spaces framework and implements both forms of typicality based reasoning: prototype and exemplar based categorization. In particular, according to the linguistic stimulus being categorized Dual-PECCS chooses, based on a similarity calculation between the stimulus and the typical representations available in $\mathcal{S}1$ knowledge base, whether to select an exemplar or a prototype (we refer to this process as $\mathcal{S}1$ categorization). By following a preference that has been experimentally observed in human cognition [15], our algorithm favors the results of the exemplars-based categorization

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 [8]. This type of common sense categorization is known in literature as *exemplars-based categorization.*

if the knowledge-base stores any exemplars similar to the input being categorized. Once the result of $\mathcal{S}1$ is selected (i.e., either a prototype or an exemplar is *proxyfied* in working memory), such approximate categorization result is then checked with the ontological knowledge base of $\mathcal{S}2$ (we refer to this process as $\mathcal{S}1$-$\mathcal{S}2$ categorization). This check is implemented by type 2 processes, and it is therefore based on deductive inference. 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 $\mathcal{S}2$ is returned along with the top scoring class returned by $\mathcal{S}1$. Otherwise, the system evaluates a fixed amount of $\mathcal{S}1$ candidates, meantime keeping track of the inconsistent elements: in case all such 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 returned along with the top scoring class initially returned by $\mathcal{S}1$. The control strategy implements a tradeoff between ontological inference and the output of $\mathcal{S}1$, which is more informative but also formally less reliable.

3 Integrating Dual-PECCS into ACT-R and CLARION

The proposed system has been integrated into two of the most widely known cognitive architectures: ACT-R [1] and CLARION [2]. The underlying rationale behind such integration efforts is to investigate whether our approach is compatible with architectures implementing different cognitive theories of mind; in this case, it can be considered a candidate general framework for representing and reasoning on conceptual information, and eventually tested with even further architectures.

One main difference between the two architectures is that CLARION natively assumes the perspective of the dual process theory; ACT-R, on the other hand, is not natively dual process based. Therefore, in the latter architecture, the dual mechanisms of reasoning needed to be explicitly designed and instantiated within an already existing general framework. In particular, in ACT-R cognitive mechanisms emerge from the interaction of two types of knowledge: the *declarative* knowledge, that encodes explicit facts that the system knows, and the *procedural* knowledge, that encodes rules for processing declarative knowledge. The declarative module is used to store and retrieve pieces of information called *chunks*, that are featured by a *type* and a set of attribute-value pairs, similar to frame slots. Finally, the central production system connects these modules by using a set of IF-THEN production rules.

Differently, in CLARION, cognitive processes are mainly subject to the activity of two sub-systems, called Action Centered Sub-system (ACS) and the Non-Action Centered Sub-system (NACS). Both sub-systems store information using a two-layered architecture, i.e., they both include an *explicit* and an *implicit* level of representation. The working memory, acting as temporary storage for decision making, is a part of the ACS, which also maintains the active behavior strategies. To hold general knowledge, the NACS provides a semantic memory consisting of both a rule-based layer that encodes explicit, symbolic knowledge, and of an underlying distributed layer with implicit, sub-symbolic representations. For both architectures we mainly focused on the Declarative

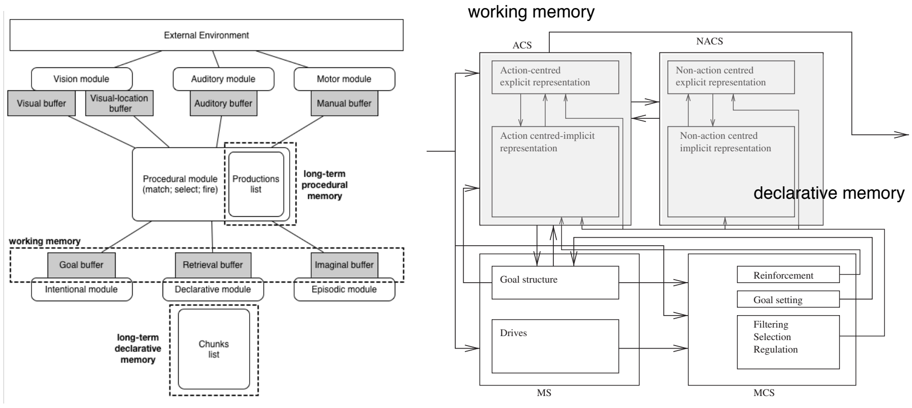

Figure 1: ACT-R Architecture with the used modules in dotted frames (left side), adapted from [1]; CLARION cognitive architecture with the working memory and the declarative memory emphasized through shaded frames (right side), adapted from [16].

Memory and Working Memory buffers, and on the corresponding retrieval mechanisms.

Besides, the dual process strategies of concept categorization have been integrated into the ACT-R and CLARION processes and connected to the retrieval request executed in the Working Memory. In the *Extended* Declarative Memory (equivalent to its counterpart, NACS, in CLARION) every concept is represented as an empty chunk (that is, a chunk having no associated information, except for its *WordNet synset ID* and a human readable *name*), referred to by the external bodies of knowledge (prototypes and exemplars) acting like semantic pointers. The novel dual process-based categorization mechanism triggers both the $\mathcal{S}1$ categorization and the $\mathcal{S}1$-$\mathcal{S}2$ categorization procedures. In this setting, when the categorization result of $\mathcal{S}1$ is returned, the representation activated in the Extended Declarative Long Term Memory is *proxyfied* (i.e., recalled to the working memory) in order to perform the $\mathcal{S}2$ consistency check, in the dual process perspective.

As regards as the ACT-R implementation, we have integrated our hybrid knowledge base directly into the declarative memory, differently from other approaches that have extended the knowledge capabilities of ACT-R based on the introduction of a new, *ad-hoc*, external module of declarative memory [17, 18]. We designed a novel retrieval request implementing the $\mathcal{S}1$-$\mathcal{S}2$ categorization mechanism by extending the repertoire of the retrieval buffer through a new action (symbolized by the operator `$`). Such action allows a direct access to the heterogeneous information represented by the $\mathcal{S}1$-$\mathcal{S}2$ external bodies of knowledge. We designed two types of `$` requests that are executed according to the specific type of request received from the retrieval buffer, the *approximate categorization request* and the *consistency request*. The *approximate categorization request* is activated when the retrieval request is generic (the request chunk does not contain a filler for the *concept_id* slot), and does not include information about the concept type to be retrieved (this task is similar to the *open request* that is possible to execute in ACT-R [19]).

This kind of request triggers the $\mathcal{S}1$ retrieval system, and its output of a classification request is a chunk-like translation of the exemplar or the prototype resulting from the execution of the $\mathcal{S}1$ retrieval on the typicality-based knowledge. We introduced the *conceptual finsts*, by building on the notion of *declarative finsts* [20] delivered in ACT-R, to keep track of the representations that have been recently retrieved by the system $\mathcal{S}1$. In our implementation, conceptual finsts allow $\mathcal{S}1$ to exclude the elements already inspected and found inconsistent by $\mathcal{S}2$. On the other hand, if the $ request specifies the concept to which it refers, then we are dealing with a *consistency check* request, to be sent to the $\mathcal{S}2$ system (i.e., we want to know whether the category assigned by $\mathcal{S}1$ is compliant with a general ontological model): in this case, if the $ request chunk contains a filler for the *concept_id* slot, we convert the request and redirect it to the $\mathcal{S}2$ system that checks whether the features of the chunk are compatible with the proposed classification. The output of this request is a chunk where the slot *concept_id* is filled with the conceptual representation resulting from $\mathcal{S}2$. The integration at the representational and reasoning level in CLARION followed the same rationale indicated in ACT-R, but it has been adapted to the specific requirements of the architecture. In particular, we adopted both implicit and explicit representational layers provided by the NACS in order to create a direct mapping with our hybrid architecture: $\mathcal{S}1$ (and its typicality-based information represented with conceptual spaces) has been mapped onto the implicit layer, while $\mathcal{S}2$ (the classical, ontology-based representation, has been mapped onto the explicit one). The mapping between the sub-symbolic module of CLARION and the dimension-based representations of the conceptual spaces has been favored, since such architecture also synthesizes the implicit information in terms of dimensions-values pairs. The dual process based categorization mechanisms have been implemented based on the following procedure: every request is encoded in working memory as a particular type of instance (instance chunk). The dimensions and values of every instance chunk are filled through an update of the implicit module with the information extracted from the external stimulus (in the present case a linguistic description). Such process is executed in the ACS module, and it is arranged as a series of *rounds*, each producing a query to the implicit $\mathcal{S}1$ component and to the explicit $\mathcal{S}2$ module. The ACS module initializes the input layer of the $\mathcal{S}1$ module, based on the instance chunk being considered. This initialization requires to handle external stimuli (the *world*) along with internal information, at disposal of CLARION agents. Let us start from the *sensory input* space: this space represents the agent's percepts, and is encoded as a set of pairs $\langle$*dimension,value*$\rangle$. Populating the sensory input space (and therefore building the instance chunk for the request to be sent to the NACS declarative memory) involves adding the appropriate set of $\langle$*dimension,value*$\rangle$ pairs (when such information, extracted from the input *stimuli*, is available). Filling a value of a dimension in CLARION is based on the sub-symbolic activation of that dimension when the external input is processed. In our implementation, the available dimensions that a chunk can assume is based on the set of dimensions defined in [13] for encoding Conceptual Spaces (therefore the internal information that CLARION can process is fixed). It is worth noting that the activated chunk can lack of some information (i.e., a *dimension* not filled with its corresponding *value*), since by definition percepts include noisy or partially missing information. After building the chunk request, a retrieval request is executed on the $\mathcal{S}1$ knowledge base, with the aim at retrieving an exemplar or a prototype-based representation. The obtained $\mathcal{S}1$ result is then *proxyfied* and temporarily stored in working memory, and checked, as previously illustrated, with the external $\mathcal{S}2$ knowledge base, the Cyc ontology.

4 Conclusions

We have illustrated the integration between the representational and reasoning assumptions presented in the DUAL-PECCS with the ACT-R and CLARION representational and reasoning modules. Such integration has shown a good level of compatibility with two general cognitive systems making different theoretical assumptions about the architecture of human cognition. As a future work, we plan to integrate the proposed representational and reasoning framework into further general cognitive architectures (e.g., SOAR, Micro PSI, OpenCog). Such set of integrations, should it prove to be feasible, will allow us to simulate brain disorders related to the activation and retrieval of conceptual information. Disorders such as the Semantic Dementia can be thought, in fact, as involving access to conceptual structures that inhibit, in different ways, the "proxyfication" process of the different representational elements of our hybrid framework.

References

[1] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, Y. Qin, An integrated theory of the mind, Psychological Review 111 (4) (2004) 1036.

[2] R. Sun, The CLARION Cognitive Architecture: Extending Cognitive Modeling to Social Simulation, Cognition and multi-agent interaction (2006) 79–99.

[3] J. S. B. Evans, K. E. Frankish, In two Minds: Dual Processes and Beyond, Oxford University Press, 2009.

[4] A. Lieto, A Computational Framework for Concept Representation in Cognitive Systems and Architectures: Concepts as Heterogeneous Proxytypes, Procedia Computer Science, Proceedings of BICA 2014 41 (2014) 6–14.

[5] C. Eliasmith, T. C. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, D. Rasmussen, A large-scale model of the functioning brain, Science 338 (6111) (2012) 1202–1205.

[6] J. J. Prinz, Furnishing the Mind: Concepts and Their Perceptual Basis, MIT press, 2002.

[7] E. Machery, Doing without concepts, OUP, 2009.

[8] M. Frixione, A. Lieto, Representing non classical concepts in formal ontologies: Prototypes and exemplars, in: New Challenges in Distributed Information Filtering and Retrieval, Springer, 2013, pp. 171–182.

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

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

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

[12] A. Lieto, A. Minieri, A. Piana, D. P. Radicioni, A knowledge-based system for prototypical reasoning, Connection Science 27 (2) (2015) 137–152. doi:10.1080/09540091.2014.956292.

[13] A. Lieto, D. P. Radicioni, M. Frixione, V. 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, 2015, pp. 1–16. doi:10.1007/978-3-319-25840-920.

[14] A. Lieto, D. P. Radicioni, V. Rho, A Common-Sense Conceptual Categorization System Integrating Heterogeneous Proxytypes and the Dual Process of Reasoning, in: In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Buenos Aires, AAAI Press, 2015, pp. 875–881.

[15] D. L. Medin, M. M. Schaffer, Context theory of classification learning, Psychological Review 85 (3) (1978) 207.

[16] R. Sun, The importance of cognitive architectures: An analysis based on clarion, Journal of Experimental & Theoretical Artificial Intelligence 19 (2) (2007) 159–193.

[17] A. Oltramari, C. Lebiere, Pursuing artificial general intelligence by leveraging the knowledge capabilities of ACT-R, in: Artificial General Intelligence, Springer, 2012, pp. 199–208.

[18] D. D. Salvucci, Endowing a cognitive architecture with world knowledge, in: Procs. of the 36th Annual Meeting of the Cognitive Science Society, 2014, pp. 1353–1358.

[19] R. Thomson, C. Lebiere, J. R. Anderson, J. Staszewski, A general instance-based learning framework for studying intuitive decision-making in a cognitive architecture, Journal of Applied Research in Memory and Cognition.

[20] Z. Pylyshyn, The Role of Location Indexes in Spatial Perception: a Sketch of the FINST Spatial-index Model, Cognition 32 (1) (1989) 65–97.