A Dual Process Architecture for Ontology-Based Systems

This is the author's manuscript

Original Citation:
A Dual Process Architecture for Ontology-Based Systems / Antonio Lieto; Andrea Minieri; Alberto Piana; Daniele P. Radicioni; Marcello Frixione. - (2014), pp. 48-55. ((Intervento presentato al convegno 6th International Conference on Knowledge Engineering and Ontology Development, KEOD 2014 tenutosi a Rome, Italy nel 21-24 October 2014.

Availability:
This version is available http://hdl.handle.net/2318/157130 since

Publisher:
INSTICC Press

Terms of use:
Open Access
Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)
This is an author version of the contribution published on:

Antonio Lieto, Andrea Minieri, Alberto Piana, Daniele P. Radicioni, Marcello Frixione
A Dual Process Architecture for Ontology-Based Systems
Editor: INSTICC Press
2014
ISBN: 9789897580499

in

KEOD 2014 - Proceedings of the International Conference on Knowledge Engineering and Ontology Development
48 - 55
6th International Conference on Knowledge Engineering and Ontology Development, KEOD 2014
Rome, Italy
21-24 October 2014

The definitive version is available at:
http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6982079
A Dual Process Architecture for Ontology-Based Systems

Antonio Lieto\textsuperscript{1,3}, Andrea Minieri\textsuperscript{1}, Alberto Piana\textsuperscript{1},
Daniele P. Radicioni\textsuperscript{1} and Marcello Frixione\textsuperscript{2}
\textsuperscript{1}Dipartimento di Informatica, Universit\`a di Torino, Torino, Italy
\textsuperscript{2}DAFIST, Universit\`a di Genova, Genova, Italy
\textsuperscript{3}ICAR-CNR, Palermo, Italy
\{antonio.lieto,daniele.radicioni\}@unito.it,
friz@dist.unige.it

Keywords:
Knowledge Representation, Formal Ontologies, Conceptual Spaces, Common Sense Reasoning,
Dual Process Theory, Prototypical Reasoning.

Abstract:
In this work we present an ontology-based system equipped with a hybrid architecture for the
representation of conceptual information. The proposed system aims at extending the representa-
tional and reasoning capabilities of classical ontology-based systems towards more realistic and
cognitively grounded scenarios, such as those envisioned by the prototype theory. The resulting
system attempts to reconcile the heterogeneous approach to the concepts in Cognitive Science and
the \emph{dual process theories} of reasoning and rationality. The system has been experimentally assessed
in a conceptual categorization task where common sense linguistic descriptions were given in input,
and the corresponding target concepts had to be identified. The results show that the proposed
solution substantially improves on the representational and reasoning “conceptual” capabilities of
standard ontology-based systems.

1 INTRODUCTION

One of the main open problems in the field
of ontology engineering is that formal ontolo-
gies do not allow \textemdash for technical convenience\textemdash neither the representation of concepts in prototypical terms nor forms of approximate, non mono-
tonic, conceptual reasoning. Conversely, in Cog-
nitive Science evidences exist in favor of pro-
totypical concepts, and typicality-based concep-
tual reasoning has been widely investigated in the
fields of the studies regarding the human cogni-
tion. In this field the early work of Rosch (Rosch,
1975) showed that ordinary concepts do not obey
the classical theory (stating that concepts can be
deﬁned in terms of sets of necessary and sufﬁcient
conditions). Rather, they exhibit \emph{prototypical}
traits: e.g., some members of a category are
considered \emph{better instances} than other ones; more
central instances share certain typical features \textemdash
such as the ability of ﬂying for birds\textemdash that, in
general, cannot be thought of as necessary nor
sufﬁcient conditions. These results inﬂuenced pi-
oneering KR research, where some efforts were in-
vested in trying to take into account the sugges-
tions coming from Cognitive Psychology: arti-
fﬁcial systems were designed \textemdash e.g., frames (Minsky,
1975) and semantic networks \textemdash to represent and to
conduct reasoning on concepts in “non classical”,
prototypical terms.

However, these systems were later sacrificed in
favor of a class of formalisms stemmed from struc-
tured inheritance semantic networks and based in
a more rigorous semantics: the ﬁrst system in this
line of research was the KL-ONE system (Brach-
mann and Schmolze, 1985). These formalisms are
known today as description logics (DLs) (Nardi
and Brachman, 2003). In this setting, the repre-
sentation of prototypical information (and there-
fore the possibility of performing non monotonic
reasoning) is not allowed,\textsuperscript{1} since the formalisms
in this class are primarily intended for deduc-

\textsuperscript{1}This is the case, for example, of \emph{exceptions} to the
inheritance mechanism.
tive, logical inference. However, under a historical perspective, the choice of preferring classical systems based on a well defined ‘Tarskian-like’ semantics left unsolved the problem of representing concepts in prototypical terms. Although in the field of logic oriented KR various fuzzy and non-monotonic extensions of DL formalisms have been designed to deal with some aspects of “non-classical” concepts (Straccia, 2011; Giordano et al., 2013), nonetheless various theoretical and practical problems remain unsolved (Frixione and Lieto, 2010).

In this paper a conceptual architecture is presented that, embedded in a larger knowledge-based system, aims at extending the representational and reasoning capabilities available to traditional ontology-based frameworks.

The paper is structured as follows: in Section 2 we illustrate the general architecture and the main features of the knowledge-based system. In Section 3 we provide the results of a twofold experimentation to assess the accuracy of the system in a categorization task. Finally, we conclude by presenting the related work (Section 4) and outlining future work (Section 5).

2 THE SYSTEM

Two cornerstones inspiring the current proposal are the dual process theory and the heterogeneous approach to concepts in Cognitive Science. The theoretical framework known as dual process theory 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 each system type can be composed by many sub-systems and processes; according to the hypotheses in (Frixione and Lieto, 2012; Frixione and Lieto, 2014), the conceptual representation of our system includes two main sorts of components, based on two sorts of processes. Type 1 processes are used to perform fast and approximate categorization, and benefit from prototypical information associated to concepts. Type 2 processes, used in classical inference tasks, and not considering the contribution of prototypical knowledge. The two sorts of system processes are assumed to interact, since type 1 processes are executed first and their results are then refined by type 2 processes. We also based our work on the heterogeneous approach to the concepts in Cognitive Science (Machery, 2009). According to this perspective, concepts do not constitute a unitary element; rather, artificial concepts can be composed by several bodies of knowledge, each one conveying a specific type of information.

The whole system embedding the proposed conceptual architecture works as follows. The input to the system is a simple description, like ‘The animal that eats bananas’: and the expected output is the category monkey. We devised an Information Extraction (IE) step in order to map the linguistic input to an internal representation. In turn, the internal representation is fed into the proposed architecture, which is concerned with the categorization task. Due to space restrictions, we are presently concerned with the architecture, deferring the details about the IE task to future contributions.

2.1 Knowledge Base Architecture

We designed a hybrid conceptual architecture that builds on a classical ontological component, and on a typical one. Each component represents a specific conceptual body of knowledge together with the related reasoning procedures as in the dual process perspective. A classical representation grounded on a DL formalism is the base of the ontological component. It permits listing necessary and/or sufficient conditions to define concepts. For example, if we consider the concept water, the classical representation contains the information that water is a natural substance, whose chemical formula is \( H_2O \). On the other hand, the prototypical traits include information about the fact that water usually occurs in liquid state, and it is mostly a tasteless, odorless and colorless fluid.

According to the “dual process” approach, in the implemented system the representational and reasoning functions are assigned to the system 1 (hereafter S1), that executes processes of type 1, and are associated to the Conceptual Spaces framework (Gärdenfors, 2000). On the other hand, the reasoning and representational functions are assigned to the system 2 (hereafter S2), to execute processes of type 2, and are associated to a classical DL-based ontological representation.
2.1.1 The $S_1$ and $S_2$ components

Conceptual spaces (CS) are a geometrical representational framework where knowledge is represented in terms of a set of quality dimensions (Gärdenfors, 2000). A geometrical structure is associated to each quality dimension. In this framework instances are represented as points in a multidimensional space, and their similarity can be computed as the intervening distance, based on some suitable distance metrics. In this setting, concepts correspond to regions and regions with different geometrical properties correspond to different sorts of concepts.

Conceptual spaces are suitable to represent concepts in “typical” terms, since the regions representing concepts can have soft boundaries. Prototypes have a natural geometrical interpretation, in that they correspond to the geometrical centre of the convex region; conversely, given a convex region we can provide each point with a certain centrality degree, that can be interpreted as a measure of its typicality.

The conceptual space defines a metric space that can be used to compute the proximity between any two entities, and between entities and prototypes. To compute the distance between two points $p_1, p_2$ we apply a distance metrics based on the combination of the Euclidean distance and the angular distance intervening between the points. Namely, we use Euclidean metrics to compute within-domain distance, while for dimensions from different domains we use the Manhattan distance metrics, as suggested in (Gärdenfors, 2000; Adams and Raubal, 2009). Weights assigned to domain dimensions are affected by the context, too, so the resulting weighted Euclidean distance $\text{dist}_E$ is computed as follows

$$\text{dist}_E(p_1, p_2, k) = \sqrt{\sum_{i=1}^{n} w_i (p_{1,i} - p_{2,i})^2},$$

where $i$ varies over the $n$ domain dimensions, $k$ is the context that indicates the set of weights associated to each domain, and $w_i$ are dimension weights.

We represent points as vectors (with as many dimensions as required by the considered domain), whose components correspond to the point coordinates, so that a natural metrics to compute the similarity between them is cosine similarity. In the metric space being defined, the distance between an individual and prototypes is computed with the Manhattan distance, enriched with information about context $k$. Also, the distance between any two concepts can be computed as the distance between two regions in a given domain.

Also, we can compute the distance between any two region prototypes, or the minimal distance between their individuals, or we can apply more sophisticated algorithms. Further details about technical issues can be found in (Ghignone et al., 2013).

Inference in conceptual spaces is mostly performed on incomplete and/or noisy information: that is, it is frequent the case that only partial information is available to categorize a given input individual, and some points are not defined for one or more dimensions. Conceptual spaces are robust to this sort of lack of information, which is conversely a decisive factor in the context of formal ontologies. In these cases we restrict to considering domains that contain points in the input individual: if the description for a given individual does not contain points of some domains, the distance for those domains is set to a default value.

The basic representational structure processed by the system is named $\text{genericDescription}$; it encodes the salient aspects of the entities being considered. A $\text{genericDescription}$ is a super-domain that hosts information about physical and non physical features arranged into nine domains: size, shape, color, location, feeding, locomotion, hasPart, partOf, manRelationship. The size of entities is expressed through the three Euclidean dimensions; the shape allows expressing that an object has circular, square, spherical, cubic, etc., shape. The color space maps object’s features onto the $L^*a^*b^*$ color space. $L^*$ ($0 \leq L \leq 100$) is the correlate of lightness, $a^*$ ($-128 \leq a \leq 127$) is the chromaticity axis ranging from green to red, and $b^*$ ($-128 \leq b \leq 127$) is the chromaticity axis ranging from blue to yellow.

The location space indicates the place where the object being modeled can be typically found. It actually results from the combination of five dimensions, and namely: humidity, indicated as a percentage; temperature, ranging in $[-40^\circ, 50^\circ]$; altitude, ranging in $[-11000, 8848]$; vegetation, ranging in $[0, 100]$; time. In turn, time contains a partitioning of the hours of the day into sunrise (4–6 AM), morning (6–12 AM), afternoon (12–5 PM), evening (5–10 PM) and night (10 PM–4 AM).

The domain feeding is currently specific to animals, and it allows mapping an element over two dimensions, typeOfFood and amountOfFood. The typeOfFood is associated to an integer indicat-
ing 1: herbivore, 2: lectivore, 3: detritivore, 4: necrophage, 5: carnivore. The underlying rationale is that close elements (e.g., necrophage and carnivore, that are one step apart in the proposed scale) are represented as close in this space due to their proximity under an ethological viewpoint, whilst different categories (e.g., herbivore and carnivore) are featured by larger distances in the considered scale (Getz, 2011).

Similar to the previous one, also the locomotion domain combines two dimensions: the former dimension is used to account for the type of movement (1: swim, 2: dig, 3: crawl, 4: walk, 5: run, 6: roll, 7: jump, 8: fly), and the latter one is used to account for the speed, expressed in km/h (Bejan and Marden, 2006). Finally, the manRelationship space is used to grasp entities as related to man by function (both a train and a horse can be used as ‘transport’), product (chicken produce ‘eggs’, and ‘chicken’ per se are a food product), symbol (‘lion’ can be used as a symbol for ‘strength’ and ‘royalty’). A simplified example of the lion prototype information is reported below.

<object name="lion">
  <genericPhysicalDescription>
    <feeding>
      <foodType>5</foodType>
      <foodQuantity>100</foodQuantity>
    </feeding>
  </genericPhysicalDescription>
  <manRelationship>
    <symbol id="08153437">royalty</symbol>
    <symbol id="05036872">strength</symbol>
  </manRelationship>
</object>

On the other hand, the representation of the classical component S2 is implemented through a formal ontology. As already pointed out, the standard ontological formalisms leave unsolved the problem of representing prototypical information. Furthermore, it is not possible to execute non monotonic inference, since classical ontology-based reasoning mechanisms contemplate deductive processes.

In this setting we cannot represent even simple prototypical information, such as ‘A typical rose is red’. This is due to the fact that being red is neither a necessary nor a sufficient condition for being a rose, and therefore it is not possible neither to represent and to automatically identify a prototypical rose (let us assume #roseF) nor to describe (and to learn from new cases) the typical features of the class of prototypical roses. Such aspect has, on the other hand, a natural interpretation by using the conceptual spaces framework.

2.2 Inference in the hybrid system

Categorization (i.e., to classify a given data instance into a predefined set of categories) is one of the classical processes automatically performed both by symbolic and sub-symbolic artificial systems. In our system categorization is based on a two-step process involving both the typical and the classical component of the conceptual representation. These components account for different types of categorization: approximate or non monotonic (performed on the conceptual spaces), and classical or monotonic (performed on the ontology). Different from classical ontological inference, in fact, categorization in conceptual spaces proceeds from prototypical values. In turn, prototypical values need not be specified for all class individuals, that vice versa can overwrite them: one typical example is the case of birds that (by default) fly, except for special birds, like penguins, that do not fly.

The whole categorization process can be summarized as follows. The system takes in input a textual description d and produces in output a pair of categories (c0, cc), the output of S1 and S2, respectively (see Algorithm 1). If the S2 system classifies it as consistent with the ontology, then the classification succeed and the category provided by S2 (cc) is returned along with c0, the top scoring class returned by S1 (Algorithm 1: line 8). If cc –the class computed by S2– is a subclass of one of those identified by S1 (c1), both cc and c0 are returned (Algorithm 1: line 11). Thus, if S2 provides more specific output, we follow a specificity heuristics; otherwise, the output of S2 is returned, following the rationale that it is safer.2 A pair of results is always returned, including both the output of S1 and the output of S2, thereby providing typically valid answers (through S1) that are checked against a logically valid reasoning conducted on the ontological knowledge base (through S2). In so doing, we follow the rationale that despite the S1 output can contain errors, it furnishes approximate answers that cannot be obtained by resorting only

2The output of S2 cannot be wrong on a purely logical perspective, in that it is the result of a deductive process. The control strategy implements a tradeoff between ontological inference and the output of S1, which is more informative but also less reliable from a formal point of view. However, in next future we plan to explore different conciliation mechanisms to ground the overall control strategy.
Cetaceans are in the order of whale, which is a fish that is a mammal because mammal and fish are mutually disjoint classes. The disjointness of the superclass mammal with the other classes in the ontology is not contextually relevant in this phase. Therefore we designed and implemented a software layer that runs on top of Jena explanation utilities to extract a laconic explanation from the longer one: main focus of the laconic explanation is to make apparent the cause of the inconsistency. In so doing, we adopted a simple heuristic according to which the only motivation reported is that focused on the tuples of classes which are under investigation.

3 EXPERIMENTATION

We have designed a twofold experimentation on a categorization task. In the first experiment we tested the system over a restricted domain (the animal kingdom domain) to compare it with the results obtained in a previous experimentation with S2 using a handcrafted ontology. In the second experiment we tested the system in a broader context, and we were interested in assessing its robustness and the discriminative features of the S1 component in a multi-domain setting. Two datasets, one composed of 39 “common-sense” linguistic descriptions, and one composed of 40 linguistic descriptions were used. Each stimulus st = ⟨d, T⟩ is a pair of description and target, such as “The big carnivore with yellow and black stripes”, “tiger”. The target T is the “prototypically correct” category, and in the following it is referred to as the expected result. The set of stimuli was devised by a team of neuropsychologists and philosophers in the frame of a broader project aimed at investigating the role of visual load in concepts involved in inferential and referential tasks. The expected prototypical target category represents a gold standard, since it corresponds to the results provided within a psychological experimentation. In this experimentation 30 subjects were requested to provide the corresponding target concept for each description. The full list of the stimuli is available at the URL: http://www.di.unito.it/~radicion/datasets/keod_2014/stimuli.txt.

---

Algorithm 1: Inference in the hybrid system.

```
input : textual description d
output : a class assignment, as computed by S1 and S2
1: C ← S1(d) /* conceptual spaces output */
2: for each ci ∈ C do
3:   cc ← S2(⟨d, ci⟩) /* ontology based output */
4:   if cc == NULL then
5:     return ⟨c0, cc⟩ /* inconsistency detected */
6:   end if
7:   if cc equals c0, then
8:     return ⟨c0, cc⟩
9:   else
10:  if cc is subclass of c0 then
11:   return ⟨c0, cc⟩
12:  end if
13:  end if
14: end for
15: cc ← S2(⟨d, Thing⟩)
16: return ⟨c0, cc⟩
```

to classical ontological inference.

If all results in C are inconsistent with those computed by S2, a pair of classes is returned including c0 and the output of S2 having for actual parameters d and Thing, the meta-class of all the classes in the ontological formalism.

An important function provided by S2 regards the explanation of the detected inconsistencies. This function is obtained by recurring to standard DL reasoners. One main problem encountered in the explanation of inconsistencies regards the fact that reasoners’ output is usually quite verbose, since it provides the whole chain of all the possible reasons explaining why a given model is not consistent w.r.t. the represented assertions. For example, let us suppose that the ontological KB is provided with an assertion about the fact that whale isA fish. Whales are not fishes: they are in the order of cetacea. If an inconsistency is detected, the initial results obtained by the reasoner report the following clauses: since whale isA mammal, and since mammal is disjoint with other animal classes, it must also be disjoint with the classes fish, reptile, bird and so on, as well as with all respective subclasses of fish, reptile, bird. Although factually correct and complete,

\[3\]

To actually access the KBs we used the Jena framework, https://jena.apache.org.
Such input was then used for querying our system as in a typicality based question-answering task. In Information Retrieval such queries (which is by far the most common and complex to interpret) are known to belong to the class of “informational queries”, i.e., queries where the user intends to obtain information regarding a specific information need (Jansen et al., 2008).

Experiment 1

The first experiment was made to compare the results obtained in the previous study (Ghignone et al., 2013) —where S2 made use of a handcrafted ontological knowledge base— to those obtained with S2 using publicly available common sense domain ontologies. With respect to the previous experiment a larger number of new descriptions were considered, all in the animal domain.

Two ontologies have been selected for the comparison: the Animal in Context Ontology (ACO), developed by the Veterinary Medical Informatics Laboratory at the Virginia-Maryland Regional College and the BBC Wildlife Ontology. They were both retrieved by using a mixed search over Sindice and Swoogle, and were selected as guaranteeing a granularity of information similar to that of the handcrafted ontology of our previous study.

The results concerning both experimental setups (ACO and BBC ontologies within S2 KB) are reported in Table 1.

By considering the ACO ontology, in most cases (37/39) S1 and S2 return the same category. Differently from the previous experiment, however, this is given by the fact that such ontology, in certain cases, is not precise enough to provide a different correct/better specified category w.r.t. S1. In particular, in the case of the description “The big fish eating plankton”, the result whale obtained by S1 is not detected as inconsistent in such ontology, since the concept representation along with its restrictions are not present in the ontology itself. Therefore, checking the consistency of the class (which was added to the ontology without the conceptual constraints associated to the concept) does not reveal inconsistencies with the original model of the ontology. The only differences w.r.t. the S1 and S2 output is obtained by considering the two descriptions: “A honey producing insect. (Bee)” and “A bird kept on a farm for its meat and eggs. (Chicken)”.

In both cases, more specific and detailed results are provided by S2 w.r.t. the already correct categorization of S1.

Analogous results, in terms of accuracy, have been obtained by plugging the BBC Wildlife ontology into S2. The testing of such ontology has been possible thanks to a manual import of the missing classes/instances through the browsing of the BBC directory containing the full classification considered by the BBC archives. Despite the two experiments produced similar figures, they also revealed some differences between the two ontologies. The answer whale for the above mentioned description is correctly detected as inconsistent with the BBC ontology, and the ontologically correct result whale-shark is returned.

Summing up, the results of this experiment show that the conceptual categorization obtained by S2 using external ontological knowledge bases are directly comparable to those obtained in a previous study where a handcrafted domain ontology was considered. A major problem encountered in using external knowledge bases stems from axiomatization, that is not always sufficient to solve the task of detecting and correcting errors produced by S1. In particular, in the case of ACO, a finer-grained result is obtained in two cases after a correct categorization already provided by S1. This finer-grained result is not obtained by considering the BBC Wildlife ontology that, however, is able to detect the expected inconsistencies.

Experiment 2

In order to assess the accuracy of the system in a more demanding experimental setting, we devised a second experimental task, where S2 used the knowledge base OpenCyc. OpenCyc is one of the largest ontologies publicly available, in that it is an enormous attempt at integrating many diverse semantic resources (such as, e.g., WordNet, DBpedia, Wikicompany, etc.). Its coverage and depth was therefore its most attractive feature (it contains about 230,000 concepts, 2,000,000 triples and 22,000 predicates). Additionally, we adopted OpenCyc to use a knowledge base independent of our own representational commitments in order to more effectively assess the flexibility of the proposed system when using general-purpose, well known, existing resources.
A second dataset of 40 “common-sense” linguistic descriptions was collected with the same rationale considered for the first experiment.\(^8\)

The obtained results are reported in Table 2.

Despite being less accurate with respect to the previous experiment, the hybrid knowledge based S1-S2 system was able to categorize and retrieve most of the new typicality-based stimuli provided as input and still showed a better performance w.r.t. the general purpose search engines Google and Bing used in question-answering mode.

The major problems encountered in this experiment derive from the difficulty of mapping the linguistic structure of stimuli containing very abstract meaning in the representational framework of S1 as they are actually encoded according to the conceptual space. For example, it was impossible to map the information contained in the description “the place where kings, princes and princesses live in fairy tales” onto the features used to characterize the prototypical representation of the concept Castle. Similarly, the information extracted from the description “Giving something away for free to someone” could not be mapped onto the features associated to the concept Gift. On the other hand, the system shows good performances when dealing with less abstract descriptions based on perceptual features such as shape, color, size, and with some typical information such as function, symbol, product.

In this experiment, differently from the previous one (e.g., that one in the case of whale for the BBC ontology), S2 did not reveal any inconsistency, in that S1 mostly provided an output coherent with the model in OpenCyc. However, by analyzing in detail the different answers, we notice that at least one inconsistency should have been detected by S2. This is the case of the description “An intelligent grey fish” associated to the target concept Dolphin. In this case, the S1 system returned the expected target, but S2 did not raise the inconsistency since OpenCyc erroneously represents Dolphin as a subclass of Mammal, rather than a subclass of Fish.

### 4 RELATED WORK

In the context of a different field of application, a solution similar to the one adopted here has been proposed in (Chella et al., 1997). The main difference with their proposal concerns the underlying assumption on which the integration between symbolic and sub-symbolic system is based. In our system the conceptual spaces and the classical component are integrated at the level of the representation of concepts, and such components are assumed to convey different—though complementary—conceptual information. On the other hand, the previous proposal is mainly used to interpret and ground raw data coming from sensors in a high level symbolic system through the mediation of conceptual spaces.

In other respects, our system is also akin to that ones developed in the field of the computational approach to the above mentioned dual process theories. A first example of such “dual based systems” is the mReasoner model (Khemlani and Johnson-Laird, 2013), developed with the aim of providing a computational architecture of reasoning based on the mental models theory proposed by Philip Johnson-Laird (Johnson-Laird, 1980). The mReasoner architecture is based on three components: a system 0, a system 1 and a system 2. The last two systems correspond to those hypothesized by the dual process approach. System 0 operates at the level of linguistic pre-processing. System 1 uses this intensional representation to build an extensional model, and uses heuristics to provide rapid reasoning conclusions; finally, system 2 carries out more demanding processes to search for alternative models, if the initial con-

---

\(^8\)An extended version of the present experiment is illustrated in (Lieto et al., 2014).

---

### Table 1: The results of the first experiment, with S2 using external ontologies.

| Cases where S1 and S2 returned the same category | ACO (37/39) | BBC (36/39) |
| Cases where S1 returned the expected category | 94.8% | 92.3% |
| Cases where S2 returned the expected category | 94.8% | 92.3% |

### Table 2: The results of the second experiment.

<table>
<thead>
<tr>
<th>Test cases categorized</th>
<th>40</th>
<th>100.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-S2</td>
<td>31</td>
<td>77.5%</td>
</tr>
<tr>
<td>Google</td>
<td>26</td>
<td>65.0%</td>
</tr>
<tr>
<td>Bing</td>
<td>23</td>
<td>57.5%</td>
</tr>
</tbody>
</table>
clusion does not hold or if it is not satisfactory.

A second system that is close to our present work has been proposed by (Pilato et al., 2012). The authors do not explicitly mention the dual process approach; however, they build a system for conversational agents (chatbots) where agents’ background knowledge is represented using both a symbolic and a sub-symbolic approach. They also associate different sorts of representation to different types of reasoning. Namely, deterministic reasoning is associated to symbolic (system 2) representations, and associative reasoning is accounted for by the sub-symbolic (system 1) component. Differently from our system, however, the authors do not make any claim about the sequence of activation and the conciliation strategy of the two representational and reasoning processes. It is worth noting that other examples of this type of systems can be considered that are in some sense analogous to the dual process proposal: for example, many hybrid, symbolic-connectionist systems—including cognitive architectures such as, for example, CLARION9—and in which the connectionist component is used to model fast, associative processes, while the symbolic component is responsible for explicit, declarative computations. However, at the best of our knowledge, our system is the only one that considers this hybridization with a granularity at the level of individual conceptual representations.

5 CONCLUSIONS

In this work we have presented a knowledge-based system relying upon a cognitively inspired architecture for the representation of conceptual knowledge. The system is grounded on a hybrid framework coupling classical and prototypical representation and reasoning, and it aims at extending the representational and reasoning capabilities of classical ontological-based systems towards more realistic and cognitively grounded scenarios, such as those envisioned by the prototype theory. The results obtained in a twofold experimentation show that in the restricted domain of the animal kingdom the proposed architecture is comparable with the results obtained with a handcrafted domain ontology. The overall results corroborate the hypothesis that matching in conceptual spaces can be fruitfully coupled to ontological inference. However, the ontological inference would not suffice to categorize the presented common sense stimuli presented in both experiments, with only one exception.

In next future we plan to test the proposed approach in the area of biomedical domain to assess disease diagnosis tasks by grounding $S_2$ on SNOWMED, and $S_1$ on conceptual spaces representing the typical symptoms of a given disease.

REFERENCES


9http://www.cogsci.rpi.edu/~rsun/clarion.html


