

A Resource-Driven Approach for Anchoring Linguistic Resources to Conceptual Spaces

Antonio Lieto, Enrico Mensa, Daniele P. Radicioni

Dipartimento di Informatica,
Università degli Studi di Torino
Turin, Italy
`{lieto, mensa, radicion}@di.unito.it`

Abstract. In this paper we introduce the TTCS system, so named after Terms To Conceptual Spaces, that exploits a resource-driven approach relying on BabelNet, NASARI and ConceptNet. TTCS takes in input a term and its context of usage and produces as output a specific type of vector-based semantic representation, where conceptual information is encoded through the Conceptual Spaces (a geometric framework for common-sense knowledge representation and reasoning). The system has been evaluated in a twofold experimentation. In the first case we assessed the quality of the extracted common-sense conceptual information with respect to human judgments with an online questionnaire. In the second one we compared the performances of a conceptual categorization system that was run twice, once fed with extracted annotations and once with hand-crafted annotations. In both cases the results are encouraging and provide precious insights to make substantial improvements.

Keywords: NLP, Lexical Semantics, Lexical Resources Integration

1 Introduction

The development of reliable knowledge sources to use in different scenarios (such as automatic reasoning, recognition, categorization, *etc.*) represents an active area of research in the AI community. In this paper we face the problem of automatically generating a Conceptual Space representation starting from text and passing through a pipeline involving the integrated use of different linguistic resources: BabelNet [1], NASARI [2] and ConceptNet [3]. The resulting representation enjoys the interesting property of being anchored to such resources, thus providing a uniform interface between the linguistic and the conceptual level.

Conceptual Spaces (CSs) can be thought of as a particular class of vector representations where knowledge is represented as a set of quality dimensions, and where a geometrical structure is associated to each quality dimension. For example, the concept *color* is characterized by 3 quality dimensions: brightness, saturation and hue. Brightness varies from white to black, so it can be represented as a linear dimension with two endpoints; saturation ranges from grey to full intensity and it is, therefore, isomorphic to an interval of the real line; hue can be arranged in a circle where complementary

colors (e.g. red-green) lie opposite one another. Then, a possible CS for representing colors is a three-dimensional space with a structure resembling the color spindle.

In this setting, concepts correspond to convex regions,¹ and regions with different geometrical properties correspond to different sorts of concepts [5]. Here common-sense conceptual representation and reasoning have a natural geometrical interpretation, since prototypes (the most relevant representatives of a category from a cognitive point of view, see [6]) correspond to the geometrical centre of a convex region (the centroid). Also exemplars-based representations can be mapped onto 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. Of course, exemplars can be used to calculate the centroid (i.e. the prototype) of each conceptual region.

It is widely accepted that knowledge acquisition is a severe and long-standing bottleneck in many applications [7]. However, while it is possible to ingest existing broad coverage resources such as the formal ontology OpenCyc or domain ontologies, unfortunately no broad coverage resources exist containing common-sense knowledge compliant to the CSs framework.² Also, wide-coverage semantic resources such as DBPedia and ConceptNet, in fact, mostly fail to represent the sort of common-sense information based on prototypical and default information which is usually required to perform forms of plausible reasoning.

Focus of this work is precisely how to extract common-sense information, suitable to be encoded through CSs.

This work has the following main strengths: the TTCS can be used to build a broad-coverage, basically domain-independent knowledge base implementing the geometrical representational tenets of CSs; additionally, the TTCS can be used to integrate different recent, state-of-the-art, lexical and semantic resources to operate them in a novel fashion (so to exploit both BabelNet concepts and ConceptNet relations). Finally, the TTCS builds on the relations harvested from ConceptNet to design a procedure to fill the appropriate dimensions in the CS representation thus producing Lexicalized Conceptual Spaces (i.e., CSs whose representations are fully endowed with BabelSynsetIds). On the whole, the TTCS is part of a broader effort aimed at collecting common-sense knowledge to overcome some of the limitations proper to most symbolic-oriented resources (like formal ontologies) in handling forms of non-monotonic reasoning [8].

The paper is structured as follows: we first survey related literature, elaborate on the existing approaches and about the differences with the current proposal (Section 2); we then illustrate in full detail the strategy implemented by the TTCS system and report

¹ Recently the convexity constraint of conceptual spaces has been argued as a plausible but not necessary condition for the characterisation of concepts within this framework in the case, for example, of the adoption of non-euclidean metrics (see [4]). In our case-study we considered such constraint as proposed in the original theory since we didn't consider non-euclidean metrics.

² In fact, we remark that differently from CSs, formal ontologies are not suited for representing defeasible, prototypical knowledge and for dealing with the corresponding typicality-based conceptual reasoning (e.g., non-monotonic inference). For example, for the concept *dog*, OpenCyc does not represent that "typically" dogs bark and woof because common-sense traits are not necessary/sufficient for defining this category.

some elements of the build of a resource where information is encoded based on the CSs framework (Section 3). The experimentation to assess the obtained results is described in Section 4, where we show how the obtained resource has been evaluated through an on-line questionnaire, and employed as the knowledge base used by a conceptual categorization system. The final remarks on future work conclude the paper (Section 5).

2 Related Work

Automatically extracting semantic information and annotating texts is an open problem in various fields of AI, and especially for the NLP community [9]. In the last few years many different methodologies and systems for the construction of unified lexical and semantic resources have been proposed.

Some of them are directly referred to the extraction of Conceptual Spaces representations. Existing approaches, for example, try to induce Conceptual Spaces based on distributional semantics by directly accessing huge amounts of textual documents to extract the multidimensional feature vectors that describe the Conceptual Spaces. In particular, [10] try to learn a different vector space representation for each semantic type (e.g. movies), given a textual description of the entities in that domain (e.g., movie reviews). Specifically, they use multi-dimensional scaling (MDS) to construct the space and identify directions corresponding to salient properties of the considered domain in a *post-hoc* analysis.

Other approaches that show some similarities with this proposal aim at learning word embeddings from text corpora. Word embeddings [11,12,13] represent the meaning of words as points in a high-dimensional Euclidean space, and are in this sense reminiscent of Conceptual Spaces. However, they differ from Conceptual Spaces in at least two crucial ways that limit their usefulness for applications in knowledge representation, e.g., in automatically dealing with inconsistencies. First, word embedding models are mainly aimed at modelling similarity (and notions such as analogy), and are not aimed at providing a geometric representation of the conceptual information (e.g., by representing concepts as convex regions where prototypical effects are naturally modelled). Moreover, the dimensions of a word embedding space are essentially meaningless, while quality dimensions in Conceptual Spaces directly reflect salient cognitive properties of the underlying domain.

Differently from such approaches that aim at extracting Conceptual Spaces (or similar multidimensional vector representations) from text or textual corpora, we use an approach that explicitly relies on existing linguistic resources. We assume that such resources already represent an intermediate step between the lexical level and the conceptual level of the representation, which we are targeting.

Existing resources, in general, can be arranged into two main classes: hand-crafted resources –created either by expert annotators, such as WordNet [14], FrameNet [15] and VerbNet [16], or through collaborative initiatives, such as ConceptNet [17]–; and resources built by automatically combining the above ones, like in the case of BabelNet [1]. Recently, great efforts have been invested to make such resources interoperable, such as the UBY platform [18,19] and the LEMON model [20]. Specifically, UBY is a lexical resource that combines a wide range of lexica (*WordNet*, *Wiktionary*, *Wikipedia*,

FrameNet, VerbNet and *OmegaWiki*),³ by converting them into lexica compliant to the ISO standard Lexical Markup Framework (LMF). Similar to the UBY-LMF project, also the LEMON project relies on the adoption of the LMF for standardisation purposes: LEMON builds on the LexInfo project [21], and it has the purpose of mapping lexical information onto symbolic ontologies. Ontologies, in turn, record the linguistic realizations for classes, properties and individuals. Our system does not directly compare with approaches based on formal ontologies (and standard logic-oriented symbolic representations in general), since the notion of meaning we are currently considering is complementary to ontological information that, on the other hand, is not explicitly committed to represent and reason on common-sense information. In our case, meaning is associated to terms that are mapped onto CSs via the identifiers provided by BabelNet, which is used as a reference framework for concept identifiers.

Some similarities can be drawn with works aiming at aligning WordNet (WN) and FrameNet (FN) [22,23]. The TTCS system shares some traits with the latter approaches, in that we provide a method to put together different linguistic resources. However, at the current stage of development, we do not align the exploited resources, but we rather provide a method for the intelligent multi-resource integration and exploitation, aimed at extracting relevant common-sense information that can be useful to fill Conceptual Spaces dimensions.

None of the mentioned proposals addresses the issue of integrating resources and extracting information to the ends of providing common-sense conceptual representations. The rationale underlying the TTCS is to extract the conceptual information hosted in BabelNet (and its vectorial counterpart, NASARI) and to exploit the relations in ConceptNet so to rearrange BabelNet concepts in a semantic network enriched with ConceptNet relations. Differently from the surveyed works, this is done by leveraging the lexical-semantic interface provided by such resources. In the next Section we illustrate our strategy.

3 TTCS: Terms To Conceptual Spaces

The TTCS system takes in input a pair $\langle t, ctx_t \rangle$ where t is a term and ctx_t is the context in which t occurs,⁴ and produces as output a set of attribute-value pairs:

$$\bigcup_{d \in \mathcal{D}} \{ \langle ID_d, \{v_1, \dots, v_n\} \rangle \} \quad (1)$$

where ID_d is the identifier of the d -th quality dimension, and $\{v_1, \dots, v_n\}$ is the set of values chosen for d . Such values will be used as fillers for Conceptual Space dimensions $d \in \mathcal{D}$. The output of the system is then a Conceptual Space representation of the input term.

The control strategy implemented by the TTCS is described in Algorithm 1, and it includes two main steps:

³ Resources marked with emphasized fonts are harmonized in UBY in both the English and the German version.

⁴ Typically, the context is composed by one or more sentences; without loss of generality, in the present setting the context has been retrieved by accessing the DBpedia page associated to t .

Algorithm 1 The control strategy of the TTCS system.

```
input : the pair  $\langle t, ctx_t \rangle$ 
output :  $\bigcup_{d \in \mathcal{D}} \{ \langle ID_d, \{v_1, \dots, v_n\} \rangle \}$ 
1: /* Associate  $t$  to a NASARI vector  $v_i$  */
2: /* and yield the lexicalized concept  $c^t$  */
3:  $c^t \leftarrow \arg \max_i (\text{WO}(ctx_t, v_i))$ 
4: for each edge  $e_i \in E$  in ConceptNet, such that  $t \xrightarrow{e_i} t'_i$  do
5:   retrieve the terms  $t'_i$  related to  $t$ 
6:   if  $t'_i$  is Relevant to  $t$  for the meaning conveyed by  $c^t$  then
7:      $C \leftarrow C \cup c_i^t$  /*  $C$  set of concepts related to  $c^t$  */
8:   end if
9: end for
10: for each pair  $c_i^t \in C, d \in \mathcal{D}$  do
11:   if  $d$  is filled by ConceptNet edges  $E^d$  and  $c^t \xrightarrow{e_i} c_i^t \mid e_i \in E^d$  then
12:     /* ConceptNet-driven mapping */
13:     fill the dimension  $d$  identified by  $ID_d$  with  $c_i^t$  as value
14:   else if  $d$  is filled by dictionary  $Dic$  and  $c_i^t \in Dic_d$  then
15:     /* Dictionary-driven mapping */
16:     fill the dimension  $d$  identified by  $ID_d$  with  $c_i^t$  as value
17:   end if
18: end for
19: return  $\bigcup_{d \in \mathcal{D}} \{ \langle ID_d, \{v_1, \dots, v_n\} \rangle \}$ 
```

- **semantic extraction** phase (lines 1-9): starting from the input term and its context, we access NASARI to provide that term with a BabelSynsetId (simply ID in the following) so to identify the correct concept. This step corresponds to a simple though effective form of word-sense disambiguation, which relies on NASARI vectors. Once the concept underlying t has been identified, we explore its ConceptNet connections and extract a bag-of-concepts semantically related to the seed term;
- **semantic matching** phase (lines 10-18): a new empty exemplar is created, corresponding to an empty vector in the CSs; we then use the bag-of-concepts extracted in the previous phase to identify the values suitable as fillers for the Conceptual Space quality dimensions.

3.1 Semantic Extraction

The semantic extraction creates a bag-of-concepts C containing a set of values for the conceptual representation of a given concept. We can distinguish two steps: the *concept identification* and the *extraction*.

Concept identification. As regards as the concept identification (Algorithm 1, line 3), given the pair $\langle t, ctx_t \rangle$, we employ NASARI to acquire an ID id_t for the input term. At the end of this step we obtain a lexicalized concept c^t , which is referred to as *seed concept*.

The problem of assigning an ID to the term t can be cast to the problem of associating a NASARI vector to t . A set of *candidate vectors* V is individuated; a NASARI

IsA	PartOf	MemberOf	HasA
CapableOf	AtLocation	HasProperty	Attribute
MadeOf	SymbolOf	UsedFor	InstanceOf

Table 1: The list of the considered ConceptNet relations.

vector v is a candidate for the meaning of t iff t is contained in the synset (the set of synonyms) associated to the head of the vector v .⁵ We distinguish three cases:

- V is empty: we cannot identify any concept for t , and the process stops;
- V contains exactly one element v_1 : t is identified by the ID of v_1 ;
- V has more than one element: for each $v_i \in V$ we compute the weighted overlap between ctx_t and all the synsets that appear as body of v_i . The vector v_k that has maximum weighted overlap [24] is then chosen as best candidate, so t is identified by the ID of v_k . If the weight of all the candidates is zero, we cannot identify any concept for t and the process stops.

Once the concept identification task has terminated, we have a lexicalized concept c^t that represents the semantics of t by means of a NASARI vector.

Extraction. In the extraction step (Algorithm 1, lines 4-9), we access the ConceptNet node associated with t , scan t 's outgoing edges $e_i \in E$, and retrieve the related terms $\{t'_1, \dots, t'_n\}$, such that $t \xrightarrow{e_1} t'_1, \dots, t \xrightarrow{e_n} t'_n$. The list of 12 relations that are presently considered –out of the 57 relations available in ConceptNet– is provided in Table 1.

Since ConceptNet does not provide any anchoring mechanism to associate its terms to meaning identifiers (the BabelSynsetIds), it is necessary to determine which edges are relevant for the concept associated to t , that is in the meaning conveyed by c^t . In particular, when we access the ConceptNet page for t , we find not only the edges regarding t intended as c^t , but also all the edges regarding t in any possible meaning. Ultimately, in this phase we look for the set of concepts related to c^t , that is the set $C = \{c_1^t, \dots, c_k^t\}$, with $k \leq n$.

To select only (and possibly all) the edges that concern c^t we introduce the notion of *relevance*. The devised algorithm is as follows:

1. Access the ConceptNet node regarding t and consider its set E of edges.
2. For each $e_i \in E$, we call t'_i the term linked to t via e_i , and verify that t'_i is relevant to t in the meaning intended by c^t . The term t'_i is relevant if either it appears within the first (highest weighted) α synsets⁶ in the NASARI vector of c^t , or if the set of nodes directly linked to the node t'_i in ConceptNet shares at least β terms⁷ with the NASARI vector of c^t .

⁵ NASARI *unified* vectors are composed by a *head* concept (represented by its ID in the first position) and a *body*, that is a list of synsets related to the head concept. Each synset ID is followed by a number that grasps its correlation with the head concept. It is worth noting that in order to reduce the number of required accesses to BabelNet we built an all-in-one resource that maps each ID referred in NASARI vectors onto its synset terms.

⁶ α is presently set to 100.

⁷ β is presently set to 3.

3. If t'_i is relevant, we then instantiate a concept c_i^t , that we identify through the ID of the first synset in the c^t NASARI vector. Finally, c_i^t is added to the result set C .⁸

For example, given in input the term ‘bank’ and a usage context such as ‘A bank is a financial institution that creates credit by lending money to a borrower’, we first disambiguate the term to identify the concept c^{bank} . Then, we inspect the edges of the ConceptNet node ‘bank’ and thanks to the relevance notion we get rid of sentences such as ‘bank isA flight maneuver’ (which has nothing to do with the sense bank-financial institution) since the term ‘flight maneuver’ is not present in the vector associated to concept c^{bank} ; conversely, we’ll accept sentences such as ‘bank HasA branch’, as related to bank-financial institution. Finally, ‘branch’ will be identified as a concept and then added to C .

3.2 Semantic Matching

The semantic matching phase consists in generating a new exemplar ex in the CS representation, and in filling it with the information previously extracted. An exemplar is a list of sets of IDs, where each set corresponds to a quality dimension; it is named and identified in accordance with the seed term t and its meaning c^t .

However, only in some cases C will be rich enough to completely fill the exemplar, which thus can be partially empty; interestingly, Conceptual Spaces are robust to lacking and/or noisy information.

Dimension anchoring. The process of assigning a certain value to a quality dimension is called *dimension anchoring*, and it is carried out for every pair $c_i^t \in C$ and $d \in \mathcal{D}$, where \mathcal{D} is the set of quality dimensions of the Conceptual Space. Quality dimensions can be either directly filled based on ConceptNet edges or checked through a dictionary (please also refer to Algorithm 1, lines 13 and 16).

In the former case (ConceptNet-driven approach) the process of extracting values to fill d leverages the set of edges E^d : if c_i^t is related to c^t by an edge included in E^d (the set of edges that are relevant to dimension d), then c_i^t is a valid value and it is added to ex as value for the quality dimension d , identified by ID_d like indicated in Equation 1. In the latter case (dictionary-driven approach) we exploit the dictionary associated with d : if c_i^t is included in the dictionary, then it is a valid value and it is added to ex as value for the quality dimension d , like indicated in Equation 1.

Additionally, every quality dimension can be metric or not (the whole picture is provided in Table 2). For metric quality dimensions we devised a set of translation maps (e.g., in the aforementioned case of *color*, we directly translate the *red* color into its L*a*b color space: $\langle 53, 80, 67 \rangle$).

Translation of metric quality dimensions. In the Conceptual Spaces theory metrical values are fundamental to carry out forms of common-sense reasoning by exploiting the distances between exemplars in the resulting geometrical framework. After the new exemplar ex is filled with the values extracted through the above mentioned procedure, we

⁸ We note that the presence of t'_i in the vector of c is guaranteed only if t'_i was detected as relevant through the first relevance condition. So, if t'_i does not appear in the vector of c^t , the identification process fails, and the term will not be added to the result set C .

Name	BabelSynsetId	Metric	DD	CND
<i>class</i>	00016733n	no	-	IsA
<i>family</i>	00032896n	no	✓	-
<i>shape</i>	00021751n	no	✓	-
<i>color</i>	00020726n	yes	✓	-
<i>locationEnv</i>	00057017n	yes	✓	-
<i>atLocation</i>	00051760n	no	-	AtLocation
<i>feeding</i>	00029546n	yes	✓	-
<i>hasPart</i>	00021395n	no	-	HasA
<i>partOf</i>	00021394n	no	-	PartOf
<i>locomotion</i>	00051798n	yes	✓	-
<i>symbol</i>	00075653n	no	-	SymbolOf
<i>function</i>	00036822n	no	-	UsedFor

Table 2: List of the considered quality dimensions; the last two columns indicate respectively whether each dimension is filled in a dictionary-driven (DD column) or in a ConceptNet-driven (CND column) way.

translate the values of the *metric* quality dimensions by exploiting the related translation maps.

Translation maps have been devised to map the extracted values onto the corresponding set of metric values in the Conceptual Space. For example, the *locomotion* 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). In the Conceptual Space representation the above mentioned values are translated into a numerically ordered scale such that the distance between indexes of values mirrors the semantic distance between the different types of locomotion: e.g., in this setting ‘dig’ and ‘crawl’ are assumed to be closer than ‘swim’ and ‘fly’ [25].

3.3 Building the CSs Representation

In order to build an actual Conceptual Space, the TTCS took in input a set of 593 cross-domain pairs term-context; such contexts have been obtained by crawling DBPedia, and by extracting the abstracts therein. To briefly account for the computational effort required in the concept identification step, the TTCS handled over 2.8M NASARI vectors (restricting to consider the first 100 features). By and large, 592 terms out of the initial 593 were associated to a NASARI vector: the failure was due to the fact that no NASARI vector was found containing the given term in the synset associated to its head.

In the extraction step, the TTCS accessed around 10M of ConceptNet assertions, linking about 3M concepts. In this step 28K ConceptNet nodes (on average 47.6 per concept) were extracted; 2.3K out of 28K concepts were selected, by finally retaining the relevant and correctly identified ones.⁹

⁹ Correctly identified concepts are those for which the whole procedure produces an output.

The semantic extraction phase ended up with 516 success cases (the 76 failures were caused by the lack of the ConceptNet node, rather than by the extraction of irrelevant concepts); where the bag-of-concepts C contains at least one extracted concept. For 30 lexicalized concepts the resulting bag-of-concept did not contain any suitable value to fill the exemplars. This led to a total of 486 correctly extracted exemplars, and producing overall 2,388 dimension values (on average 4.9 per exemplar).

Whether and to what extent the information extracted by the TTCS approaches human common-sense knowledge is the object of the next Section.

4 Evaluation

We devised a twofold experimentation aimed at assessing *i*) the quality of the extracted common-sense conceptual information via human assessment; and *ii*) the usefulness of the obtained representations in the context of a specific conceptual categorization task.

4.1 Human Evaluation of the Extracted Conceptual Information

The first evaluation regards the assessment of the quality of the extracted conceptual space representations through the TTCS, based on human common-sense judgement. More precisely, the evaluation is intended to assess both correctness (does TTCS output reasonable information for the considered concept?) and completeness (does TTCS output all relevant information for the considered concept?) of the extracted information.¹⁰

Twenty-seven volunteers, 17-52 years of age (average 37) were recruited for this experiment, mostly from the Department of Computer Science from the authors' University (11 females and 16 males), all naïve to the experiment. The human subjects have been provided with a concept (e.g., *dog*) and some related common-sense statements obtained from the representations extracted by the TTCS (e.g., "Dogs have fur", "Dogs are animals", *etc.*). In this setting, participants had to assess each statement by indicating *i*) whether it was appropriate or not for the concept at hand; and *ii*) any further statement reputed essential in order to complete the common-sense description of the considered concept. Participants were randomly split into 2 groups (respectively composed by 12 and 15 participants); subjects had to provide their assessment through an on-line questionnaire containing statements about 15 concepts randomly extracted from the obtained Conceptual Space resource.¹¹ Each piece of information available in the conceptual space has been used to automatically generate a statement, in such a way that all information collected by the TTCS has undergone this experimentation. Overall 173 statements have been proposed to human subjects for evaluation.

According to the experimental design, for each concept we recorded the number of statements that were found inappropriate (*deletions*), and the number of added statements (*insertions*). In particular, we recorded two distinct metrics: one considering *all* answers, and one examining *relevant* answers. As '*all* answers' metrics we recorded all

¹⁰ Questionnaires are available at: <http://goo.gl/am0S2f>.

¹¹ We acknowledge that compared to similar experiments (such as [26,27]) such data is rather small, and defer to future work an extensive evaluation.

Table 3: The accuracy results on the deletions (Table 3-a) and insertions (Table 3-b).

a. Analysis of the deletions suggested by human subjects.

	accuracy agreement	
all deletions	83.12%	83%
relevant deletions	68.21%	57%

b. Analysis of the insertions suggested by human subjects.

	# insertions agreement	
all insertions	149	3%
relevant insertions	7	58%

of the responses, whilst the second metrics has been designed to tame the sparsity of human answers. As regards as the ‘*relevant* answers’ metrics, we defined as *relevant* a deletion (or an insertion) that occurs when a statement about a given concept is not accepted (or felt to be lacking, but necessary) by at least 3 participants. For example, given the concept *mouse* and the statements ‘Mouse lives in desert’ and ‘Mouse is snake food’, we recorded a relevant deletion for the former statement which was refused by 3 participants, but not for the latter one, which was refused by only 1 participant.

Results. The final figures of our results are reported in Table 3. Let us start by considering the result on deletions (Table 3-a). The participants produced on the whole 2,340 judgements; the *all answers* raw datum is that in 1,945 cases they accepted the considered statement (thereby determining a 83.12% of correct results). As regards as *relevant deletions*, 55 statements were refused by at least 3 participants, thus leading to 32% *relevant deletions* (and to its complement, 68% of accuracy). The relevant deletions occurred mainly for statements rather obscure or incorrect, such as ‘A bullet is spherical’, ‘Banana is a dessert’, or ‘Singer is at location show’. Besides, in order to evaluate how reliable are the collected judgements, we also measured the agreement in participants’ answers. Agreement values were measured about the *all deletions* metrics.¹² The average agreement concerning the *deletions* amounts to 83%; this datum grasps that there is a neat consensus on which statements are acceptable and which ones are not (be them counter-intuitive, or explicitly incorrect, and irrespective of individual preferences and experiences).

On the other hand, as regards as insertions (please refer to Table 3-b), we registered a clear data sparsity which resulted in a low agreement (3%). The few cases where the participants provided *relevant insertions* (that is, proposing at least 3 times the same statement for insertion), point out an information lacking from the set of statements

¹² As the ratio between the number of deletions expressed for a given statement and the number of assessments obtained by that statement: e.g., the statement ‘Soap has function of scent’ has been questioned by 2 participants out of 12. The agreement on such deletion was computed as $2/12 = 16.7\%$.

Table 4: Results in a conceptual categorization task, where the output of the system fed with information extracted by the TTCS system is compared against the results obtained by the categorization system fed with information annotated by hand.

CSs annotation	number of input descriptions	correct categorization
manual	60	52 (87.7%)
TTCS	60	41 (68.3%)

extracted for that concept: e.g., the TTCS missed to extract that ‘Airplanes fly’, that ‘Camels have two humps’ and that ‘Chameleons change color’. Despite there is a consistent difference in the agreement between the *all insertions* and *all deletions* metrics, there is on the other hand a similar agreement rate on relevant deletions and insertions. Or, equivalently, there is a prominent analogous accord on both mistaken and missing information.

A general insight emerging from the collected data is that the output obtained by the TTCS system mostly corresponds to the characterization of the conceptual information in terms of prototypical, common-sense knowledge. This element is crucial, since most of the available KBs are not equipped with this sort of distilled but salient information. On the other hand, common sense knowledge represents exactly the type of knowledge crucially used by humans for efficient heuristic reasoning, and it could be adopted by artificial systems aiming at providing forms of plausible automatic reasoning. In next Section we show a concrete application requiring this sort of knowledge.

4.2 A Case Study in Conceptual Reasoning and Categorization

The obtained Conceptual Space resource produced by the TTCS has been additionally evaluated in a practical case study involving a basic conceptual categorization task. In particular, a target concept illustrated by a simple common-sense linguistic description had to be identified; for this experiment we exploited an existing categorization system that relies on a hybrid knowledge base coupling annotated Conceptual Spaces encoding common-sense information, and an external ontological component, represented by the OpenCyc ontology [28,29,30]. In this evaluation we compared the output provided by this system in two different executions: in the former case the categorization system made use of manually annotated Conceptual Spaces, whilst in the latter one it was fed with the Conceptual Spaces extracted by the TTCS system. A set of 60 common-sense textual descriptions has been given in input to the categorization system in both conditions (with manual or automatically obtained CSs); these stimuli have been built by a multidisciplinary team composed of neuro-psychologists, linguists and philosophers in the frame of a project aimed at investigating the brain activation of visual areas in tasks of lexical processing.

The whole categorization pipeline works as follows. The input to the system is a simple sentence, like ‘The feline with mane and big jaws’, and the correct output is the

category evoked by the description (the category *lion* in this case).¹³ Correctly identified categories represent a gold standard which has been individuated based on the results of an experiment involving human subjects [28]; both outputs have thus been compared to human answers.

Results. The obtained results are reported in Table 4, that provides a comparison between the accuracy of the categorization system adopting the automatically obtained conceptual representations and the accuracy of the same system endowed with manually annotated representations. Although the obtained accuracy is lower than that obtained when using hand-crafted knowledge, the performance of the system using data extracted by the TTCS is still acceptable, especially if we consider the increase in the coverage. In fact, the extracted KB includes around 500 conceptual representations, while only 300 conceptual representations were present in the manually annotated one. Also, we are now able to extract salient information to fill CSs representations with virtually no domain restriction, thus attaining a much broader resource representing common-sense knowledge in terms of Lexicalized Conceptual Spaces (i.e. CSs whose representations are fully endowed with BabelSynsetIds).

5 Conclusions

In this paper we have presented TTCS, a system that takes a textual input and returns the corresponding common-sense conceptual representation encoded in terms of Lexicalized Conceptual Spaces. This representation is obtained through a novel method leveraging different linguistic resources such as BabelNet, NASARI and ConceptNet. The results obtained through the human assessment are encouraging for what concerns the acceptability of the extracted representations; also, the TTCS output has been fed as input to a broader cognitively-inspired categorization system, with an interesting outcome.

The ongoing and future work is represented by the attempt to build a wider, general, common-sense resource in term of Lexicalized Conceptual Space. Such knowledge resource is complementary, and easily integrable onto existing encyclopedic knowledge resources such as BabelNet, since the interface between the lexical and conceptual level would be grounded on the BabelNet synsets. All of the TTCS IDs are actually anchored to BabelNet synset IDs, so the output of TTCS is, *de facto*, already connected to the semantic network of BabelNet. In addition, such a resource would benefit from the geometrical features proper to Conceptual Spaces representations, that are especially helpful in applications that mix different types of reasoning strategies for tasks such as conceptual categorization, question answering, *etc.*. The obtained resource will also enable us to extend the present evaluation towards a larger coverage and more quantitative scenario which will furnish further insights for iteratively refining the TTCS.

¹³ In essence, the employed system executes a two-steps categorization process: it first computes a result based on Conceptual Spaces, and it then checks the validity of the obtained result against an ontological knowledge base.

References

1. Navigli, R., Ponzetto, S.P.: BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence* **193** (2012) 217–250
2. Camacho-Collados, J., Pilehvar, M.T., Navigli, R.: NASARI: a novel approach to a semantically-aware representation of items. In: *Proceedings of NAACL*. (2015) 567–577
3. Speer, R., Havasi, C.: Representing General Relational Knowledge in ConceptNet 5. In: *LREC*. (2012) 3679–3686
4. Hernández-Conde, J.V.: A case against convexity in conceptual spaces. *Synthese* (2016) 1–27
5. Gärdenfors, P.: *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press (2014)
6. Rosch, E.: Cognitive Representations of Semantic Categories. *Journal of experimental psychology: General* **104** (1975) 192–233
7. Lenat, D.B., Prakash, M., Shepherd, M.: CYC: Using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks. *AI magazine* **6** (1985) 65
8. Frixione, M., Lieto, A.: Representing concepts in formal ontologies: Compositionality vs. typicality effects. *Logic and Logical Philosophy* **21** (2012) 391–414
9. Reeve, L., Han, H.: Survey of semantic annotation platforms. In: *Proceedings of the 2005 ACM symposium on Applied computing*, ACM (2005) 1634–1638
10. Derrac, J., Schockaert, S.: Inducing semantic relations from conceptual spaces: a data-driven approach to plausible reasoning. *Artificial Intelligence* **228** (2015) 66–94
11. Pennington, J., Socher, R., Manning, C.D.: Glove: Global Vectors for Word Representation. In: *EMNLP*. Volume 14. (2014) 1532–1543
12. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*. (2013) 3111–3119
13. Turney, P.D., Pantel, P., et al.: From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research* **37** (2010) 141–188
14. Miller, G.A.: WordNet: a lexical database for English. *Communications of the ACM* **38** (1995) 39–41
15. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The Berkeley Framenet Project. In: *Proceedings of the 17th international conference on Computational linguistics-Volume 1*, Association for Computational Linguistics (1998) 86–90
16. Levin, B.: *English verb classes and alternations: A preliminary investigation*. University of Chicago press (1993)
17. Havasi, C., Speer, R., Alonso, J.: ConceptNet: A lexical resource for common sense knowledge. *Recent advances in natural language processing V: selected papers from RANLP* **309** (2007) 269
18. Eckle-Kohler, J., Gurevych, I., Hartmann, S., Matuschek, M., Meyer, C.M.: UBY-LMF-A Uniform Model for Standardizing Heterogeneous Lexical-Semantic Resources in ISO-LMF. In: *LREC*. (2012) 275–282
19. Gurevych, I., Eckle-Kohler, J., Hartmann, S., Matuschek, M., Meyer, C.M., Wirth, C.: Uby: A large-scale unified lexical-semantic resource based on LMF. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics* (2012) 580–590
20. McCrae, J., Montiel-Ponsoda, E., Cimiano, P.: Integrating WordNet and Wiktionary with lemon. In: *Linked Data in Linguistics*. Springer (2012) 25–34

21. Buitelaar, P., Cimiano, P., Haase, P., Sintek, M.: Towards linguistically grounded ontologies. In: *The semantic web: research and applications*. Springer (2009) 111–125
22. Tonelli, S., Pianta, E.: A novel approach to mapping framenet lexical units to Wordnet synsets. In: *Proceedings of the Eighth International Conference on Computational Semantics*, Association for Computational Linguistics (2009) 342–345
23. Ferrández, O., Ellsworth, M., Munoz, R., Baker, C.F.: Aligning FrameNet and WordNet based on Semantic Neighborhoods. In: *LREC*. Volume 10. (2010) 310–314
24. Pilehvar, M.T., Jurgens, D., Navigli, R.: Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarity. In: *ACL* (1). (2013) 1341–1351
25. Bejan, A., Marden, J.H.: Constructing Animal Locomotion from New Thermodynamics Theory. *American Scientist* **94** (2006) 342–349
26. Tonelli, S., Pighin, D.: New features for framenet: Wordnet mapping. In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*, Association for Computational Linguistics (2009) 219–227
27. De Cao, D., Croce, D., Basili, R.: Extensive evaluation of a framenet-wordnet mapping resource. In: *LREC*. (2010)
28. Lieto, A., Minieri, A., Piana, A., Radicioni, D.P.: A knowledge-based system for prototypical reasoning. *Connection Science* **27** (2015) 137–152
29. Lieto, A., Radicioni, D.P., Rho, V.: A Common-Sense Conceptual Categorization System Integrating Heterogeneous Proxytypes and the Dual Process of Reasoning. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Buenos Aires, AAAI Press (2015) 875–881
30. Lieto, A., Radicioni, D.P., Rho, V.: Dual PECCS: a cognitive system for conceptual representation and categorization. *Journal of Experimental & Theoretical Artificial Intelligence* (2016) 1–20