

# Modelling metaphor with attribute-based semantics

**Luana Bulat**

Computer Laboratory  
University of Cambridge  
ltf24@cam.ac.uk

**Stephen Clark**

Computer Laboratory  
University of Cambridge  
sc609@cam.ac.uk

**Ekaterina Shutova**

Computer Laboratory  
University of Cambridge  
es407@cam.ac.uk

## Abstract

One of the key problems in computational metaphor modelling is finding the optimal level of abstraction of semantic representations, such that these are able to capture and generalise metaphorical mechanisms. In this paper we present the first metaphor identification method that uses representations constructed from property norms. Such norms have been previously shown to provide a cognitively plausible representation of concepts in terms of semantic properties. Our results demonstrate that such property-based semantic representations provide a suitable model of cross-domain knowledge projection in metaphors, outperforming standard distributional models on a metaphor identification task.

## 1 Introduction

According to the Conceptual Metaphor Theory (Lakoff and Johnson, 1980), metaphors are not merely a linguistic, but also a cognitive phenomenon. They arise when one concept (or conceptual domain) can be understood in terms of the properties of another. For example, we interpret the metaphorical expression “He shot down my argument” by projecting our knowledge about *battles* (the source domain) onto our reasoning about *arguments* (the target domain).

Multiple studies have established the prevalence of metaphor in language (Cameron, 2003; Shutova and Teufel, 2010) and confirmed the key role that it plays in human reasoning (Thibodeau and Boroditsky, 2011). These findings make computational processing of metaphor essential for any NLP application that is focused on semantics, from machine translation (Shutova, 2011) to

recognising textual entailment (Agerri, 2008). Numerous approaches to metaphor processing have been proposed, modelling generalisations over source and target domains using hand-constructed lexical resources (e.g. WordNet) (Tsvetkov et al., 2014), distributional clustering (Shutova et al., 2010), LDA topic modelling (Heintz et al., 2013) and, more recently, multimodal word embeddings (Shutova et al., 2016). While these works have established that it is possible to generalise metaphorical mappings using the above techniques, one important question remains unanswered – that of the optimal level of abstraction of semantic representations needed to capture and generalise metaphorical mechanisms. On the one hand, such representations need to be sufficiently informative for the task, and on the other hand generalise well enough as to obtain a broad coverage of metaphorical language.

Much work in cognitive science suggests that human concept representation relies on salient *attributes* or *properties*<sup>1</sup> (Tyler et al., 2000; Randall et al., 2004). Property norm datasets (McRae et al., 2005; Devereux et al., 2013) are constructed by asking human participants to identify the most important attributes of a concept (see Table 1) and are widely used to test models of conceptual representation (McRae et al., 1997; Randall et al., 2004; Cree et al., 2006; Tyler et al., 2000; Grondin et al., 2009). Yet, to the best of our knowledge, such property norms have not been investigated in the context of metaphor processing.

Recent studies (Fagarasan et al., 2015; Bulat et al., 2016) have shown that wide-coverage property-norm based semantic representations can be automatically constructed using cross-modal maps and that these perform comparably to dense semantic representations (Mikolov et al., 2013)

<sup>1</sup>Throughout the paper we will be using the terms *properties* and *attributes* interchangeably.

SHOES	ANT	DISHWASHER
has_heels, 15	an_insect, 18	an_appliance, 19
has_laces, 13	is_small, 18	requires_soap, 15
worn_on_feet, 13	is_black 15	is_electrical, 14

Table 1: Examples of properties from McRae et al. (2005) together with their production frequencies

on standard word similarity tasks. In this paper we hypothesise that such attribute-based representations provide a suitable means for generalisation over the source and target domains in metaphorical language and test this hypothesis. Our results show that these property-based representations can perform better than dense context-predicting (Mikolov et al., 2013) and context-counting (Turney and Pantel, 2010) vectors in a metaphor classification task, thus providing a suitable model of cross-domain property projection in metaphorical language.

## 2 Related work

Much previous research on metaphor processing casts the problem as classification of linguistic expressions as metaphorical or literal. Gedigian et al. (2006) classified verbs using a maximum entropy classifier and the verbs’ nominal arguments and their semantic roles as features. Dunn (2013) used a logistic regression classifier and high-level properties of concepts extracted from the SUMO ontology, including domain types (ABSTRACT, PHYSICAL, SOCIAL, MENTAL) and event status (PROCESS, STATE, OBJECT). Tsvetkov et al. (2013) also used logistic regression and coarse semantic features, such as concreteness, animateness, named entity types and WordNet supersenses. They have shown that the model learned with such coarse semantic features is portable across languages. The work of Hovy et al. (2013) is notable as they focused on compositional features. They trained an SVM with dependency-tree kernels to capture compositional information, using lexical, part-of-speech tag and WordNet supersense representations of parse trees. Mohler et al. (2013) derived semantic signatures of texts as sets of highly-related and interlinked WordNet synsets. The semantic signatures served as features to train a set of classifiers (maximum entropy, decision trees, SVM, random forest) that map new metaphors to the semantic signatures of the known ones.

Turney et al. (2011) hypothesized that metaphor is commonly used to describe abstract con-

cepts in terms of more concrete or physical experiences. They developed a method to automatically measure concreteness of words and applied it to identify verbal and adjectival metaphors. Shutova et al. (2010) pointed out that the metaphorical uses of words constitute a large portion of the dependency features extracted for abstract concepts from corpora. As a result, distributional clustering of abstract nouns with such features identifies groups of diverse concepts metaphorically associated with the same source domain. Shutova et al. (2010) exploit this property of co-occurrence vectors to identify new metaphorical mappings starting from a set of examples. Shutova and Sun (2013) used hierarchical clustering to derive a network of concepts in which metaphorical associations are learned in an unsupervised way.

## 3 Method

### 3.1 Learning dense linguistic representations

We construct two types of linguistic representations: context-predicting – based on the skip-gram model of Mikolov et al. (2013) – and context-counting.

**EMBED** We employ 100-dimensional word embeddings constructed by Shutova et al. (2016) from Wikipedia using the standard log-linear skip-gram model with negative sampling of Mikolov et al. (2013). The embeddings were trained using a symmetric window of 5 words either side of the target word, 10 negative samples per word-context pair and number of epochs set to 3.

**SVD** We use Wikipedia to build count-based distributional vectors, using the top 10K most frequent lemmatised words (excluding stopwords) as contexts. Context windows are defined as sentence boundaries and counts are re-weighted using positive pointwise mutual information (PPMI). We obtain 100-dimensional dense semantic representations by applying singular value decomposition (SVD) (Deerwester et al., 1990) to the sparse 10K-dimensional PPMI weighted vectors.

### 3.2 Learning attribute-based vectors through cross-modal mapping

**Property norms** The property norm dataset collected by McRae et al. (2005) is one of the largest and most widely used attribute datasets in cognitive science. It contains a total of 541 concrete

	is_loud	has_keys	requires_air	is_long
ACCORDION	6	17	11	0
CLARINET	0	9	0	8
CROCODILE	0	0	0	6

Table 2: A subspace of the property-norm semantic space (PROPERTY)

concepts annotated with properties and production frequencies (i.e. the number of participants that produced that property). Examples of concepts and properties can be found in Table 1. Each concept was shown to 30 participants and only features listed by more than 5 annotators were recorded. The published dataset contains a total of 2526 properties, with a mean of 13.7 features per concept. The McRae et al. (2005) property-norm dataset can be used to obtain distributed representations of concepts over attributes (henceforth PROPERTY). We can view it as a bag of 2526 properties, with the standard co-occurrence counts being replaced by the production frequencies. Table 2 shows a subspace of such a property-norm semantic space.

**Cross-modal maps** Even though MCRAE only contains annotations for 541 concepts, cross-modal maps can be used to induce property-based representations for words outside of this dataset. Fagarasan et al. (2015) propose a method for obtaining such representations for any concept from its distributional behaviour and Bulat et al. (2016) show that these can be also inferred from images. Cross-modal maps represent a formalisation of the reference problem. For example, by inducing a cross-modal map between linguistic representations and property-based representations, we can learn to predict properties for new (unseen) concepts (Figure 1).

**Property-based vectors** Following Fagarasan et al. (2015), we obtain property-based vectors by using partial least squares regression<sup>2</sup> (PLSR) to learn a cross-modal mapping function between the dense linguistic representations (SVD and EMBED) and the property-norm semantic space (PROPERTY), using the 541 concepts in MCRAE as training data. We learn two different maps, hence two different attribute-based representations: one from SVD to PROPERTY (ATTR-SVD) and one from EMBED to PROPERTY (ATTR-EMBED).

<sup>2</sup>We set the number of latent variables in the cross-modal PLSR map to 100.

Metaphorical	Literal
black humor	black dress
filthy mind	filthy garment
young moon	young boy
ripe age	ripe banana
shallow argument	shallow grave
stormy applause	stormy sea

Table 3: Annotated adjective–noun pairs from TSV-TEST

### 3.3 Metaphor classification

We compare the performance of the aforementioned semantic representations (SVD, EMBED, ATTR-SVD and ATTR-EMBED) on a metaphor classification task, in order to test our hypothesis as to whether attribute-based semantic representations provide better concept generalisations for metaphor modelling than the widely-used dense linguistic representations. We use an SVM (Joachims, 1998) to perform the classification<sup>3</sup>.

## 4 Experiments

### 4.1 Experimental data

We evaluate our method using the dataset of adjective–noun pairs manually annotated for metaphoricality, created by Tsvetkov et al. (2014). This corpus was created by extracting the nouns that co-occur with a list of 1000 frequent adjectives in the TenTen Web Corpus<sup>4</sup> using SketchEngine and in collections of metaphor on the Web. The data is divided into a training set (TSV-TRAIN) and test set (TSV-TEST). TSV-TRAIN contains 884 literal and 884 metaphorical pairs annotated for metaphoricality. TSV-TEST contains 100 literal and 100 metaphorical pairs, annotated by 5 annotators with an inter-annotator agreement of  $\kappa = 0.76$ . Table 3 shows a portion of the test set. Metaphorical phrases that depend on wider context for their interpretation (e.g. *drowning students*) were removed.

This dataset is well-suited to our task since it includes examples of the same adjective used in both metaphorical and literal phrases (e.g. “hot topic” and “hot chocolate”). This is important since we want our model to be able to discriminate between different word senses, as opposed to selecting the most frequent class for any given word.

<sup>3</sup>Experiments were performed using the sklearn.svm toolkit.

<sup>4</sup><https://www.sketchengine.co.uk/xdocumentation/wiki/Corpora/enTenTen>

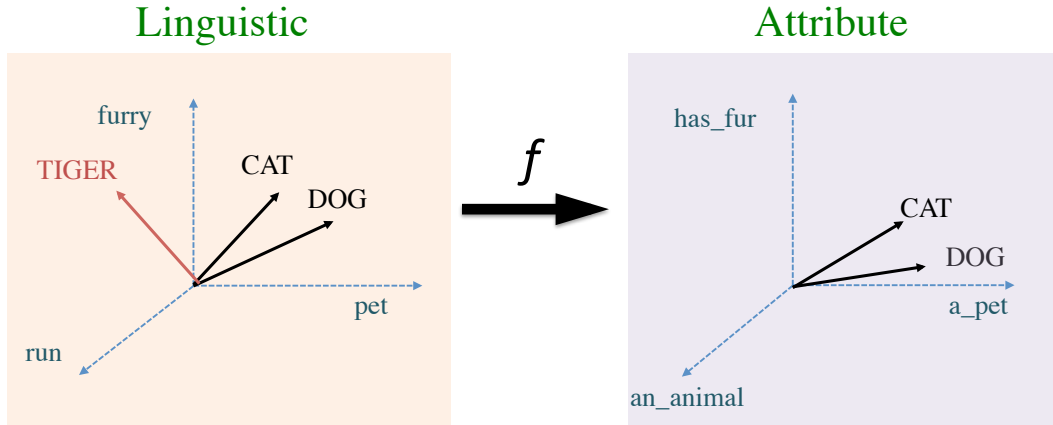


Figure 1: Example of cross-modal mapping: learn  $f$  using aligned representations (linguistic and attribute) for DOG and CAT, then predict attribute representation for TIGER as  $f(\text{TIGER}_{\text{linguistic}})$

## 4.2 Experimental setup and results

We obtain four types of semantic vectors (SVD, EMBED, ATTR-SVD, ATTR-EMBED) for all nouns and adjectives in Tsvetkov et al. (2014) as described in Section 3. It is important to note that up to now, attribute-based representations as those described in Section 3.2 have only been used for nouns. To our knowledge, this is also the first work that uses cross-modal maps learned on nouns to predict attribute-based representations for other parts of speech.

The input to our SVM classifier is the concatenation of the L2-normalised adjective and noun vectors. We use the phrases in TSV-TRAIN and TSV-TEST to train and test our system, respectively. We evaluated the performance of our classifier on TSV-TEST in terms of precision, recall and F-score; the results are presented in Table 4. Both types of attribute-based vectors outperform their dense counterparts, which lends support to our hypothesis that property norms offer a suitable level of generalisation of the source and target domains. The best performance is achieved when using the attribute-based representation learned from the embedding space (ATTR-EMBED), with an improvement of 4% in F1 score over EMBED.

## 5 Qualitative analysis and discussion

The results in Table 4 show that the systems are able to reliably distinguish between metaphorical and literal expressions both when using dense and attribute-based semantic representations. This is an effect of modelling word meanings as distributed representations over semantic primitives.

Vectors	P	R	F1
EMBED	0.84	0.65	0.73
ATTR-EMBED	0.85	0.71	0.77
SVD	0.86	0.64	0.73
ATTR-SVD	0.74	0.77	0.75

Table 4: System performance on Tsvetkov et al. test set (TSV-TEST) in terms of precision (P), recall (R) and F-score (F1)

Intuitively, one may expect the noun and the adjective in a metaphorical expression to share fewer properties than in the case of literal language, due to a semantic distinction between its source and target domains. And it is likely that all of our models capture this effect, by implicitly learning some notion of similarity between the semantic domains in the literal and metaphorical phrases. Our hypothesis is that attribute-based methods outperform the EMBED and SVD baselines because the attribute-based dimensions are cognitively-motivated and represent cognitively salient properties for concept distinctiveness. As such, they provide a more suitable means of generalisation in the metaphor identification task, as inferred from our results.

Another advantage of using attribute-based vectors (ATTR-EMBED, ATTR-SVD) in the metaphor identification task is that they are interpretable, i.e. every dimension in the space has a fixed interpretation (*is\_round*, *a\_bird* etc.) as opposed to the abstract dimensions of SVD and EMBED. We can thus identify the most salient attributes of a word by looking at the highest weighted dimensions in its attribute-based representation. This, in turn, can yield in-

sights into how the attributes of metaphorical expressions differ from those of the literal ones. For example, in the metaphorical expression “woolly liberal”, the highest weighted attributes for *woolly* (AN\_ANIMAL, A\_FRUIT, IS\_SMALL, A\_MAMMAL, IS\_BROWN, IS\_LONG) are ranked low for *liberal* and vice-versa. When we look at a literal expression using the same adjective, “woolly mammoth”, we observe many overlapping features among the top 200 highest-weighted ones, with 48% of these attributes being shared (e.g. AN\_ANIMAL, IS\_SMALL, IS\_BROWN, HAS\_4\_LEGS, A\_MAMMAL, IS\_LARGE). The same trend was observed for the majority of the AN pairs in TSV-TEST<sup>5</sup>, demonstrating that the components of literal expressions share many more features than the components of the metaphorical ones.

## 6 Conclusion

We presented the first method that uses large-scale attribute-based semantic representations for metaphor identification. Our results demonstrate that these provide a suitable level of generalisation for capturing metaphorical mechanisms. Our experiments also suggest interesting future research avenues in the investigation of the attribute-based representations of abstract concepts, more generally. For instance, we have observed that many of the highly-weighted attributes for abstract concepts are metaphorical in nature (e.g. A\_BIRD for “liberal”). This echoes previous research in cognitive science, which has shown that while concrete concepts are well represented through their internal properties and relation to similar concepts, abstract concepts tend to be represented through associations with many diverse concepts (Crutch and Warrington, 2005). We believe that our methods provide a framework for a data-driven investigation of this issue in the future.

## 7 Acknowledgments

LB is supported by an EPSRC Doctoral Training Grant. SC is supported by ERC Starting Grant DisCoTex (306920) and ERC Proof of Concept Grant GroundForce (693579). ES is supported by the Leverhulme Trust Early Career Fellowship. We are grateful to Jean Maillard for providing help

<sup>5</sup>ATTR-EMBED was used for this analysis as it performs best in the metaphor classification task.

with the embeddings and thank the anonymous reviewers for their helpful comments.

## References

- Rodrigo Agerri. 2008. Metaphor in textual entailment. In *COLING 2008: Companion volume: Posters*, pages 3–6, Manchester, UK, August.
- Luana Bulat, Douwe Kiela, and Stephen Clark. 2016. Vision and feature norms: Improving automatic feature norm learning through cross-modal maps. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 579–588, San Diego, California, June. Association for Computational Linguistics.
- Lynne Cameron. 2003. *Metaphor in Educational Discourse*. Continuum, London.
- George S. Cree, Chris McNorgan, and Ken McRae. 2006. Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4):643.
- Sebastian J. Crutch and Elizabeth K. Warrington. 2005. Abstract and concrete concepts have structurally different representational frameworks. *Brain*, 128(3):615–627.
- Scott Deerwester, Susan T Dumais, George W Furnas, Thomas K Landauer, and Richard Harshman. 1990. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391.
- Barry J. Devereux, Lorraine K. Tyler, Jeroen Geertzen, and Billi Randall. 2013. The centre for speech, language and the brain (CSLB) concept property norms. *Behavior Research Methods*, pages 1–9.
- Jonathan Dunn. 2013. Evaluating the premises and results of four metaphor identification systems. In *Proceedings of CICLing’13*, pages 471–486, Samos, Greece.
- Luana Fagarasan, Eva Maria Vecchi, and Stephen Clark. 2015. From distributional semantics to feature norms: grounding semantic models in human perceptual data. In *Proceedings of the 11th International Conference on Computational Semantics (IWCS’15)*, pages 52–57, London, UK, April. Association for Computational Linguistics.
- Matt Gedigian, John Bryant, Srinu Narayanan, and Branimir Cicic. 2006. Catching metaphors. In *In Proceedings of the 3rd Workshop on Scalable Natural Language Understanding*, pages 41–48, New York.
- Ray Grondin, Stephen J. Lupker, and Ken McRae. 2009. Shared features dominate semantic richness effects for concrete concepts. *Journal of Memory and Language*, 60(1):1–19.

- Iana Heintz, Ryan Gabbard, Mahesh Srivastava, Dave Barner, Donald Black, Majorie Friedman, and Ralph Weischedel. 2013. Automatic extraction of linguistic metaphors with lda topic modeling. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 58–66, Atlanta, Georgia.
- Dirk Hovy, Shashank Shrivastava, Sujay Kumar Jauhar, Mrinmaya Sachan, Kartik Goyal, Huying Li, Whitney Sanders, and Eduard Hovy. 2013. Identifying metaphorical word use with tree kernels. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 52–57, Atlanta, Georgia.
- T. Joachims. 1998. Text categorization with support vector machines: Learning with many relevant features. In *European Conference on Machine Learning (ECML)*, Berlin.
- George Lakoff and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press, Chicago.
- Ken McRae, Virginia R. de Sa, and Mark S. Seidenberg. 1997. On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126(2):99.
- Ken McRae, George S. Cree, Mark S. Seidenberg, and Chris McNorgan. 2005. Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, 37(4):547–559.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *ICLR workshop*.
- Michael Mohler, David Bracewell, Marc Tomlinson, and David Hinote. 2013. Semantic signatures for example-based linguistic metaphor detection. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 27–35, Atlanta, Georgia.
- Billi Randall, Helen E. Moss, Jennifer M. Rodd, Mike Greer, and Lorraine K. Tyler. 2004. Distinctiveness and correlation in conceptual structure: behavioral and computational studies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2):393.
- Ekaterina Shutova and Lin Sun. 2013. Unsupervised metaphor identification using hierarchical graph factorization clustering. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 978–988, Atlanta, Georgia, June. Association for Computational Linguistics.
- Ekaterina Shutova and Simone Teufel. 2010. Metaphor corpus annotated for source - target domain mappings. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA).
- Ekaterina Shutova, Lin Sun, and Anna Korhonen. 2010. Metaphor identification using verb and noun clustering. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 1002–1010, Beijing, China, August. Coling 2010 Organizing Committee.
- Ekaterina Shutova, Douwe Kiela, and Jean Maillard. 2016. Black holes and white rabbits: Metaphor identification with visual features. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 160–170, San Diego, California, June. Association for Computational Linguistics.
- Ekaterina Shutova. 2011. *Computational Approaches to Figurative Language*. Ph.D. thesis, University of Cambridge, UK.
- Paul H. Thibodeau and Lera Boroditsky. 2011. Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE*, 6(2):e16782, 02.
- Yulia Tsvetkov, Elena Mukomel, and Anatole Gershman. 2013. Cross-lingual metaphor detection using common semantic features. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 45–51, Atlanta, Georgia.
- Yulia Tsvetkov, Leonid Boytsov, Anatole Gershman, Eric Nyberg, and Chris Dyer. 2014. Metaphor detection with cross-lingual model transfer. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 248–258, Baltimore, Maryland, June. Association for Computational Linguistics.
- Peter D Turney and Patrick Pantel. 2010. From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37(1):141–188.
- Peter Turney, Yair Neuman, Dan Assaf, and Yohai Cohen. 2011. Literal and metaphorical sense identification through concrete and abstract context. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 680–690, Edinburgh, Scotland, UK., July. Association for Computational Linguistics.
- Lorraine K. Tyler, Helen E. Moss, MR Durrant-Peatfield, and JP Levy. 2000. Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain and language*, 75(2):195–231.