Automatic Metaphor Detection using Large-Scale Lexical Resources and Conventional Metaphor Extraction

Yorick Wilks, Lucian Galescu, James Allen, Adam Dalton
Florida Institute for Human and Machine Cognition
15, SE Osceola Ave
Ocala, FL, 34471, USA
{ywilks,lgalescu,jallen,adalton}@ihmc.us

Abstract

The paper presents an experimental algorithm to detect conventionalized metaphors implicit in the lexical data in a resource like WordNet, where metaphors are coded into the senses and so would never be detected by any algorithm based on the violation of preferences, since there would always be a constraint satisfied by such senses. We report an implementation of this algorithm, which was implemented first the preference constraints in VerbNet. We then derived in a systematic way a far more extensive set of constraints based on WordNet glosses, and with this data we reimplemented the detection algorithm and got a substantial improvement in recall. We suggest that this technique could contribute to improve the performance of existing metaphor detection strategies that do not attempt to detect conventionalized metaphors. The new WordNet-derived data is of wider significance because it also contains adjective constraints, unlike any existing lexical resource, and can be applied to any language with a semantic parser (and WN) for it.

1 Introduction

Metaphor is ubiquitous in standard language; it is not a fringe or add-on phenomenon. The work described concerns detecting and interpreting metaphor on a large scale in corpora. If metaphor is ubiquitous, then locating and interpreting it must be central to any NLP project that aims to understand general language. This paper focuses on the initial phase of detection: the identification in text of conceptual combinations that might be deemed metaphoric by a pre-theoretic observer, e.g., "*Brazil has economic muscle*", "*Tom is a brick*", or "*The unions have built a fortress round their pensions*". There is a long cultural tradition of describing and interpreting such phenomena but our goal here is computational: to provide criteria for automatically detecting such cases as candidates for further analysis and interpretation.

The key fact is that metaphors are sometimes new and fresh but can be immediately understood: producing them is often the role of poets, creative journalists and writers of all kinds. But many are simply part of the history of the language, and are novel only to those who do not happen to know them already: for example "*Tom is a brick*" – taken to mean that he is a reliable man, but which cannot be literally true – is actually encoded as a sense of *brick* in WordNet (WN) (Miller, 1995) even though it is more familiar to UK than US English speakers.

This means that lexical resources already contain conventionalized metaphors. We propose a simple method for locating and extracting these into the metaphor candidate pool, even when they are not indicated as such in resources like WN (which marks figurative senses very infrequently, unlike some traditional dictionaries). However, we believe these implicit metaphors in WN – a resource we intend to use as a semantic/lexical database, though transformed as we shall show below – can be extracted by a simple algorithm, and without any need for *a priori* distinction of literal versus metaphorical. That distinction, as we noted, depends to a large degree on the temporal snapshot of a language; e.g., no one now would think "*taking a decision*" was metaphor, even though decisions are not literally taken anywhere.

In this paper, we shall present an algorithm for conventionalized metaphor detection, and show results over a standard corpus of examples that

Proceedings of the First Workshop on Metaphor in NLP, pages 36–44,
Atlanta, Georgia, 13 June 2013.

demonstrate a possible useful gain in recall of metaphors, our original aim. The algorithm is described in two implementations (or pipelines) corresponding, respectively, to the use of WN and VerbNet (Kipper et al., 2000; Kipper et al., 2008) as semantic knowledge-bases, and to their replacement by our automatically recomputed form of WN, which enables predictions about the preference behavior (see below) of English verbs and adjectives to be better founded than in VerbNet (VN) and on a much larger scale.

2 Background on Metaphor Detection using Preference Violation as Cue

In early work on metaphor detection, long preceding access to large-scale or annotated corpora, it was suggested as sufficient a criterion for being a metaphor that a "semantic preference" of a verb or adjective was violated (Wilks, 1978). So, for example, one might say that the verb *drink* had a preference for animate agents and liquid objects, in which case "*My car drinks gasoline*" violates its subject preference, which might then be a cue to look for metaphor at that point. Similarly, in the "*economic muscle*" case mentioned earlier one might say that *economic* has a preference for abstract entities as objects, as in "*economic value*", and *muscle* is not an abstract entity.

There was discussion in those early days of syntactic-semantic interface cases like "*John ran a mile*" where *a mile* might be said to violate the preference of the (intransitive) verb for a zero object and so again trigger a metaphor. The preference notion was not initially intended to detect metaphor but to semantically disambiguate candidates at those sites by preferring those conceptual entities that did **not** violate such restrictions. In early work, preferences were largely derived by intuition and sometimes ordered by salience. Later (e.g. Resnik, 1997) there was a range of work on deriving such preferences from corpora; however, in VN the semantic preferences of verbs were again largely intuitive in origin.

Early work linking preference violation to metaphor detection (summarised in Fass and Wilks, 1983, also Martin 1990) worked with handcrafted resources, but by 1995 Dolan had noted (Dolan, 1995) that large-scale lexical resources would have implications for metaphor detection, and WN was used in conjunction with corpora, by (Peters and Wilks, 2003) using symbolic methods and by Mason (2004) and Krishnakumaran and Zhu (2007) using a combination of WN and statistical methods. Mason also acquires preferences automatically from corpora, and the latter two papers treat metaphor as a form of anomaly based on rare combinations of surface words and of WN-derived hypernyms, a notion that appears in (Guthrie et al., 2007) but based only on corpus sparsity and not WN codings. Other work on the automatic acquisition of preferences (McCarthy and Carrol, 2003) for WSD has also its considered extension to the detection of classes of metaphor. More recently, work by Shutova (Shutova et al., 2010) has shown that the original preference violation insight can be combined with large-scale investigations, using notions of machine learning and large-scale resources like WN. Our approach is smaller scale and does not involve machine learning: it simply seeks access to implicit metaphors built into the structure of WN by its creators, and which a preference-violation detection criterion cannot, by definition, access. Thus, we view our contribution as complementary to larger efforts on metaphor and interpretation detection, rather than a competing approach. We have not made comparisons here with the work of (Li and Sporleder, 2010), which is explicitly concerned with idioms, nor with (Markert and Nissim, 2009) which is focused on metonymy.

3 The Conventional Metaphor Detection Hypotheses

Where WN codes conventionalized metaphors as senses, as in the initial cases described, then the senses expressing these will NOT violate preferences and so will not be detected by any metaphor-as-violation hypothesis. For example, in "*Jane married a brick*" this will not be a preference violation against WN senses because WN explicitly codes *brick* as a reliable person, though we would almost certainly want to say this sentence contains a metaphor to be detected.

The hypothesis we propose is simply this: if we have a word whose main (usually first) sense in WN fails the main preference for the sentence slot it fills, but has a lower, less frequent, sense that satisfies that preference, then we declare that lower sense a metaphorical one. In the case of *brick*, whose main sense is a PHYSICAL OBJECT, one which clearly fails the equivalence to *Tom* in the example "*Tom is a brick*". Yet the less frequent listed sense for a reliable person does satisfy the same preference. The work at this stage is not concerned with the metaphor-metonymy distinction and this criterion may well capture both, their distinction being, as is well known (e.g. in Fass and Wilks, 1983) hard to establish in the limit. Ours is a purely empirical hypothesis and will work or not, and we argue that it does to a reasonable degree. It does not rest on any assumption of strict ordering of WN senses, only on a tendency (from literal to metaphorical) which is plainly there for any observer.

4 Metaphor Detection Experiments

We have implemented two versions of conventional metaphor detection, using two different lexical resources. We were thus able to divide the hypothesis into two parts, essentially one making use of VN and one within WN only. In this first pipeline, we use WN together with the verb preferences provided by VN even though those give only patchy coverage of common verbs. At the outset this was the only lexical resource for verb preferences available. VN includes classes of verbs that map members to specific WN senses. VN also provides a hierarchy of verb object/subject inclusions, which we use for assessing whether one sentence object/subject type appears below another in this simple inclusion hierarchy, and so can be said to be semantically included in it. The selectional restrictions, however, are not linked to any lexicons so a mapping was constructed in order to allow for automated detection of preference violations.

Our first experiment utilizes WN, VN, and the Stanford Parser (de Marneffe et al., 2006) and Named Entity Recognizer (Finkel et al., 2005). The Stanford Parser identifies the verbs, as well as their corresponding subjects and direct objects. The Stanford Named Entity Recognizer was used to replace sequences of text representing names with WN senses whose hypernyms exist in the selectional restriction hierarchy.

The first step in determining whether a sentence contains a metaphor is to extract all verbs along with the subject and direct object arguments for each verb. The Stanford Parser dependencies used to describe the relationships between verbs and their arguments include *agent, nsubj,* and *xsubj* for subjects and *dobj* and *nsubjpass* for direct objects. The parser also handles copular and prepositional verbs but additional steps are required to link these verbs to their arguments.

Once verbs have been extracted and parameterized from the sentence, each is checked for preference violations. A preference is violated if a selectional restriction on one of the thematic roles of a VN class is not satisfied for all VN classes the verb is a member of. In order for a VN class's preferences to be satisfied, there must be a WN sense for the argument of a verb such that either itself or its hypernym matches the WN senses allowed by the selectional restriction in VN class, where the terms in the VN hierarchy have been hand-matched to WN senses. If a sentence contains a verb that does not exist in VN then we must assume that it is not violated.

5 Conventionalized Metaphor Detection

Closer inspection of false negatives revealed that many of the verbs and the arguments that satisfied their selectional restrictions were unannotated conventionalized metaphors.

5.1 Conventionalized Verbs

In our approach, a conventionalized verb occurs when two VN Classes have the same member, but one maps to a lower WN sense (in the WN ordering, which can be taken roughly to mean less frequent) than the other. If the VN Class mapped to the lower sense is satisfied in a sentence, but the other VN Class is not, we say that the verb is used in a conventionalized sense. The verb *pour* is a member of four VN classes. Three of those classes, **Pour-9.5**, **Preparing-26.3-2**, and **Substance_Emission-43.4** all map to first sense of the word which means to *cause to run*. The fourth VN class of *pour,* **Weather-57,** maps to the sixth WN sense of the verb, which means *to rain heavily.* If we take the example sentence "*Bisciotti has poured money into the team*", we determine that all VN classes that map to the primary WN sense of *pour* are violated in some way. According to our semantic role labeling heuristic, **Pour-9.5** expects *money* to be a substance, **Preparing-26.3-2** expects *the team* to be an animate, and **Substance_Emission-43.4** is violated because Bisciotti is animate. The only Verb Class that is satisfied is

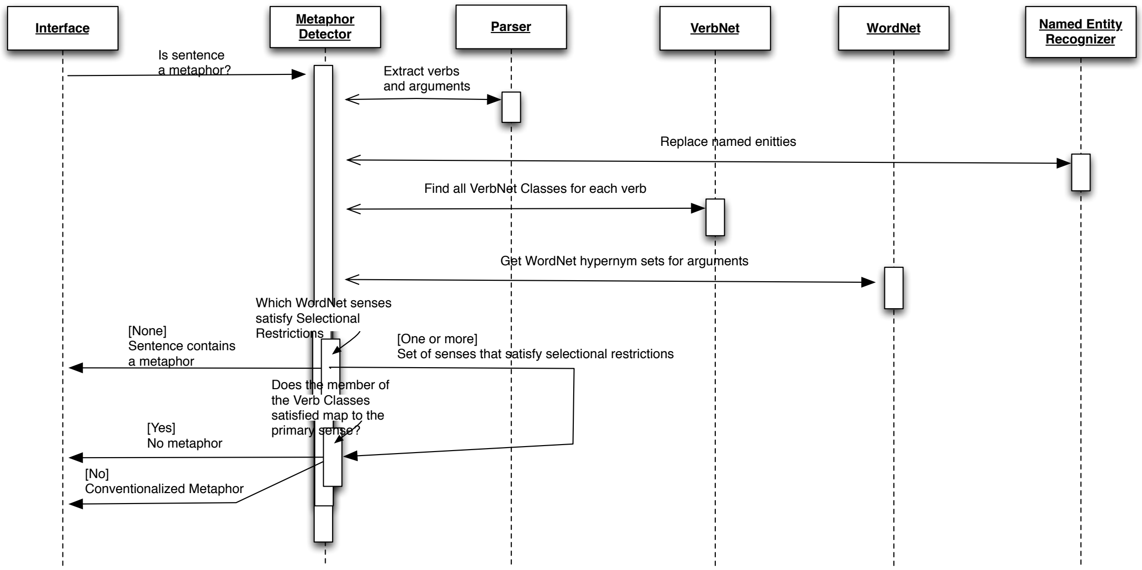

Figure 1. Conventionalized verb metaphor detection using WordNet senses and VerbNet selectional restrictions

Weather-57, and that class maps to the sixth sense of *pour*. Interestingly, there is no VN class member that maps to the fifth WN sense (*supply in large amounts or quantities*).

The pseudocode for detecting conventional metaphors used as verbs is as follows:

```
• for each VN Class
  • for each member of that class
    • for each WN sense of that member
    with Verb POS
      • get the sense number of the WN
      sense
      • associate the sense number to the
      verb member and selectional re-
      strictions for the Verb Class
• given a verb in a sentence, decide
that the verb is conventionalized if:
  • it satisfies the selectional re-
  strictions of one Verb Class V1 but…
  • it violates the selectional restric-
  tions of another Verb Class V2 and…
  • the sense number of the verb member
  in V2 is above the sense number of the
  verb member in V1
```

5.2 Conventionalized Nouns

Let us look again at the example of *brick,* where the primary sense of the noun is the building material most are familiar with and the secondary sense refers to a reliable person. For this reason, the noun *brick* will satisfy any VN class that requires a human or animate. Without the ability to detect conventional metaphors in noun arguments, *She married a brick* would pass through without detection by preference violation. Here are the WN entries for the two senses:

- **brick#1 (brick%1:06:00::)** (rectangular block of clay baked by the sun or in a kiln; used as a building or paving material)
- **brick#2 (brick%1:18:00::)** (a good fellow; helpful and trustworthy)

Less obvious are more abstract words such as *zone*:

- **zone#1** (**zone%1:15:00::**) (a locally circumscribed place characterized by some distinctive features)
- **zone#2 (zone%1:15:02::)**, **geographical zone#1** (**geographical_zone%1:15:00::**) (any of the regions of the surface of the Earth loosely divided according to latitude or longitude)
- **zone#3 (zone%1:15:01::)** (an area or region distinguished from adjacent parts by a distinctive feature or characteristic)
- **zone#4 (zone%1:08:00::)**, **zona#1** (**zona%1:08:00::**) ((anatomy) any encircling or beltlike structure)

Zone's primary sense, again, is the anticipated concept of circumscribed space. However, the fourth sense deals with anatomy, and therefore is a hyponym of *body part*. *Body part* is capable of satisfying any thematic role restricted to *animate* arguments.

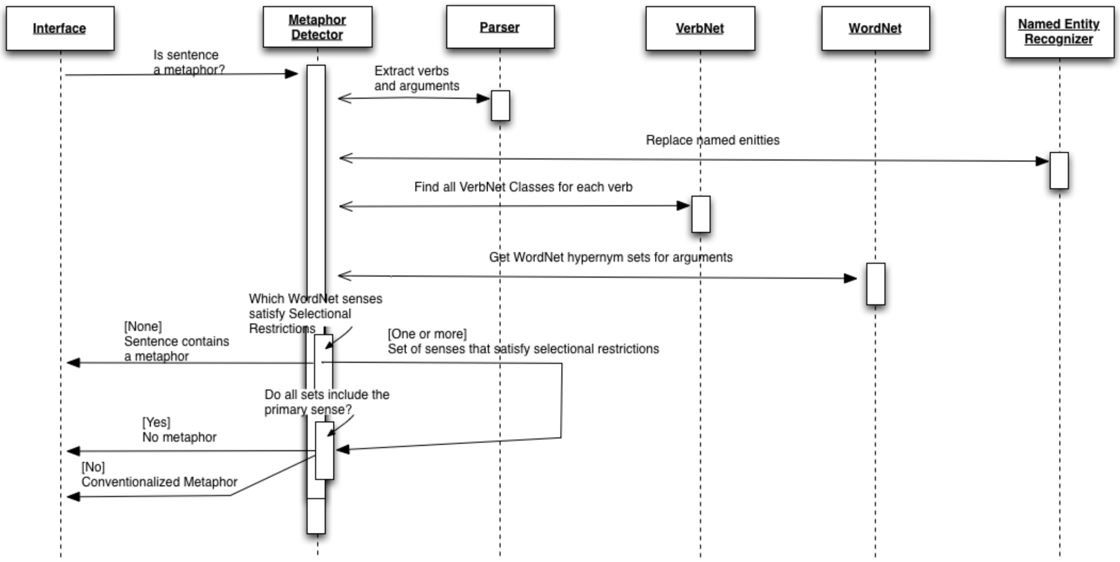

Figure 2. Conventionalized noun metaphor detection using WordNet senses and VerbNet selectional restrictions

The pseudocode for detecting conventional metaphors used as nouns is as follows:

```
• determine if verbs' subjects and di-
  rect objects satisfy the restriction
  • if not, it is a Preference Violation
    metaphor
  • if they do:
    • determine if the sense of the sat-
      isfying word is the primary sense
      in WN
      • if not, it is a conventional
        metaphor
      • otherwise, it is not a metaphor
```

Thus, our overall hypothesis is intended to locate in the very broad WN sense sets those that are actually conventionalized metaphors: we determine that only the first sense, hopefully literal, should be able to satisfy any restriction. If a lower sense satisfies a verb, but the primary sense does not, we classify the satisfaction as being conventionalized, but a metaphor nonetheless.

6 Deriving Preferences and an Ontology from WordNet

To date, VerbNet is the most extensive resource for verb roles and restrictions. It provides a rich semantic role taxonomy with some selectional restrictions. Still, VN has entries for less than 4000 verbs. PropBank (Palmer et al., 2005) has additional coverage, but uses a more surface oriented role set with no selectional restrictions. On the other hand, WordNet has many more verb entries but they lack semantic role information. However, we believe it is possible to extract automatically a comprehensive lexicon of verbs with semantic roles and selectional restrictions from WN by processing definitions in WN using deep understanding techniques. Specifically, each verb in WN comes with a gloss that defines the verb sense, and there we can find clues about the semantic roles and their selectional restrictions. Thus, we are testing the hypothesis that the semantic roles of the verb being defined are inherited from the roles in its definition, though roles in the latter may be elided or fully specified. For example, consider this entry from WN for one of the senses of the verb *kill*:

> S: (v) **kill** (cause to die; put to death, usually intentionally or knowingly) "*This man killed several people when he tried to rob the bank*"; "*the farmer killed a pig for the holidays*"

Let us assume we already know that the verb *cause* takes three roles, say, a CAUSER, an AFFECTED and an EFFECT role; this leads us to hypothesize that *kill* would take the same roles. However, the EFFECT role from cause is not inherited by *kill* as it is fully specified in the definition. The proof of

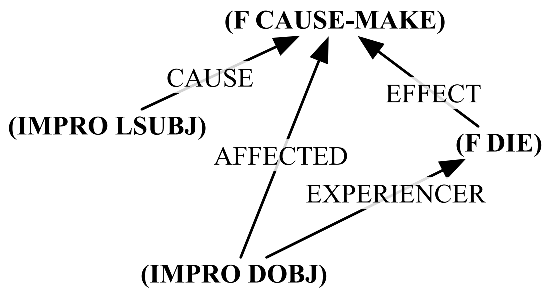

Figure 3: Abstracted Logical Form for "cause to die"

this hypothesis is ultimately in how well it predicts the role set. But intuitively, any role in the definition verb (i.e., *cause*) that is fully filled in the definition has no "space" for a new argument for that role. Therefore, we conclude that *kill* takes two roles, filling the CAUSER and AFFECTED roles in the definition.

We can now derive selectional restrictions for *kill* by looking at inherited restrictions from the definition, as well as those that can be derived from the examples. From the definition, the verb *cause* puts little to no restriction on what the CAUSER role might be. For instance, an animal may cause something, but natural forces cause things as well. Likewise, *cause* puts little constraint on what the PATIENT role might be, as one can cause the temperature to rise, or an idea to fade. The restriction from the verb *die* in the complement, however, suggests a restriction of some living object (if we can derive this constraint from *die*). We also look at the examples to find more informative restrictions. In the definition of *kill*, we have two examples of a CAUSER, namely a man and a farmer. Given the hypernym hierarchy of nouns in WordNet, we could look for the most specific subsuming concept in the hierarchy for the concepts MAN and FARMER, finding it to be **person%1:03:00**. The fillers for the AFFECTED role in the examples are PEOPLE and PIG, with the most specific WN node being **organism%1:03:00**). Putting all this together, we produce an entry for *kill* as follows:

kill: ACTOR/**person%1:03:00**
PATIENT/**organism%1:03:00**

To implement this idea we need a number of capabilities. First, semantic roles do not appear out of the ether, so we need an initial seed of semantic role information. In addition, to process the glosses we need a parser that can build a semantic representation, including the handling of elided arguments. As a start, we use the TRIPS parser (Allen et al., 2008). The TRIPS lexicon provides information on semantic roles, and the parser can construct the required semantic structures. TRIPS has been shown to be successful at parsing WN glosses in order to build commonsense knowledge bases (Allen et al., 2011). With around 3000 types, TRIPS offers a reasonable upper-level ontology to serve as the seed for semantic roles. We also use the TRIPS selectional restrictions to bootstrap the process of determining the restrictions for new words.

To attain broad lexical coverage, the TRIPS parser uses input from a variety of external resources. This includes a subsystem, Wordfinder, for unknown word lookup that accesses WN when an unknown word is encountered. The WN senses have mappings to semantic types in the TRIPS ontology, although sometimes at a fairly abstract level. When faced with an unknown word, the parser looks up the possible senses in WordNet, maps these to the TRIPS ontology and then uses the verb entries in the TRIPS lexicon associated with these types to suggest possible subcatgorization frames with mappings to roles. Thus, Wordfinder uses the combined information from WN and the TRIPS lexicon and ontology to dynamically build lexical entries with approximate semantic and syntactic structures for words not in the core lexicon. This process may produce a range of different possibilities based on the different senses and possible subcategorization frames for the verbs that share the same TRIPS type. We feed all of these to the parser and let it determine the entries that best match the definition and examples. While WordNet may have multiple fine-grained senses for a given word, we set a parameter that has the system use only the most frequent sense(s) of the word (cf. McCarthy et al. 2004).

We use TRIPS to parse the definitions and glosses into a logical form. Figure 3 shows the logical form produced for the definition *cause to die*. We then search the logical form for structures that signal a potential argument that would fill a role. Besides looking for gaps, we found some other devices that serve the same purpose and occur frequently in WordNet:

- elided arguments (an IMPRO in the logical form);
- indefinite pronouns (e.g., *something*, *someone*);
- prepositional/adverbial forms containing an IMPRO or an indefinite pronoun (e.g., *give a benediction to*);
- a noun phrase in parentheses (e.g., *to remove (people) from a building*).

The final condition is probably a WN specific device, and was discovered when working on a 10-verb development set, and occurred twice in that set.

Once these arguments are identified, we have a candidate set of roles for the verb. We identify candidate selectional restrictions as described above. Here are a few examples of verbs and their automatically derived roles and restrictions, as computed by our system (here we indicate WordNet entries by their sense index rather than their sense key, since the index is used in the conventional metaphor detection strategy – see below):

bend.v.06:	AGENT/being.n.02
	PATIENT/physical_entity.n.01
collect.v.03:	AGENT /person.n.01
	PATIENT/object.n.01
drive.v.01:	AGENT/person.n.01
	PATIENT/motor_vehicle.n.01
play.v.13:	CAUSE/instrumentality.n.03
	EFFECT/music.n.01
walk.v.08:	AGENT/being.n.02
	GOAL/location.n.01

The techniques described in this section have been used to provide a set of roles with selectional restrictions for the second IHMC pipeline, described below. The current system takes a list of verbs from a corpus and returns the role names and selectional restrictions for every sense of those words in WordNet.

The transformations described here all equally able to produce preferences for adjectives, as would be needed to detect "*economic muscle*" as a metaphor, which is a form of lexical information not present in any existing database, and the whole process can be applied to any language that possesses a WordNet type lexical resource, and for which we have a capable semantic parser. Hence, these techniques are amenable to being used for detecting metaphorical usage in constructions other than just verb-subject and verb-object, as we do here.

	Pipeline 1 (VerbNet SRs)	Pipeline 2 (WordNet SRs)
TP	24	50
FP	23	37
TN	48	24
FN	37	11
Precision	0.649	0.575
Recall	0.393	0.82
F1	0.49	0.676

Figure 4. Performance comparison between the first pipeline using VerbNet selectional restrictions (SRs) and the second pipeline using WordNet-derived selectional restrictions

7 Conventional Metaphor Detection based on WordNet-Derived Preferences

The preferences and ontology derived from WN definitions greatly improve the mapping between selectional restrictions and WN sense keys. This allows us to replace VN with a new lexical resource that both improves performance, and reduces the complexity of discovering preference violations. In the new pipeline, we can reuse the capabilities developed to extract verbs and their parameters from a sentence. We also reuse the tie-ins to WN that allow us to determine if one WN sense exists within another's hypernym set. It is the selectional restriction lookup that is greatly simplified in the new lexicon, where verbs are mapped directly to WN senses. The conventional metaphor detection is also simplified because the WN senses are included in the responses to the looked up verbs, allowing us to quickly determine if a satisfied verb is conventionalized or is satisfied with conventionalized arguments.

8 Results and Conclusion

Figure 4 shows the results obtained in a metaphor detection task over a small corpus of 122 sentences. Half of these sentences have metaphors and half do not. Of the half that do, approximately half are metaphors about Governance and half are other metaphors. This is not any sort of principled corpus but a seed set chosen to give an initial leverage and in a domain chosen by the sponsor (Governance); the selection and implicit annotation were done by consensus by a large group of twenty or so collaborators. The notion of baseline is irrelevant here, since the choice for every sentence is simply whether it contains a metaphor or not, and could thus be said to be 50% on random assignment of those categories.

From the figures above, it can be seen that the second pipeline does give significant improvement of recall over the first implementation above, even though there is some loss of precision, probably because of the loss of the information in VN. One possibility for integrating a conventional metaphor extraction pipeline like ours with a general metaphor detection pipeline (including, for example, pattern-based methods and top-down recognition from stored Conceptual Metaphors) would be to OR these two pipelines together and to hope to gain the benefits of both, taking anything as a metaphor that was deemed one by either.

However, that is not our aim here: our purpose is only to test the hypothesis that using knowledge derived from existing lexical resources, in combination with some form of the conventionalized metaphor hypothesis, we can achieve good recall performance. On this point we think we have shown the value of the technique.

Acknowledgements

This work was supported in part by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Defense US Army Research Laboratory contract number W911NF-12-C-0020, and NSF grant IIS 1012205. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoD/ARL, or the U.S. Government.

References

James Allen, William de Beaumont, Nate Blaylock, George Ferguson, Jansen Orfan, and Mary Swift. 2011. Acquiring commonsense knowledge for a cognitive agent. In *Proceedings of the AAAI Fall Symposium on Advances in Cognitive Systems* (*ACS 2011*), Arlington, Virginia.

James Allen, Mary Swift, and Will de Beaumont. 2008. Deep semantic analysis of text. In *Proceedings of the 2008 Conference on Semantics in Text Processing* (*STEP '08*), Venice, Italy. pp. 343-354.

Marie-Catherine de Marneffe, Bill MacCartney and Christopher D. Manning. 2006. Generating Typed Dependency Parses from Phrase Structure Parses. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, pp. 449-454.

William B. Dolan. 1995. Metaphor as an emergent property of machine-readable dictionaries. In *Proceedings of the AAAI 1995 Spring Symposium Series: Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity and Generativity*, pp. 27–32.

Dan Fass and Yorick Wilks. 1983. Preference semantics, ill-formedness, and metaphor. *American Journal of Computational Linguistics*, 9(3):178–187.

Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. In *Proceedings of the 43nd Annual Meeting of the Association for Computational Linguistics (ACL 2005),* pp. 363-370.

David Guthrie, Louise Guthrie, Ben Allison and Yorick Wilks. 2007. Unsupervised Anomaly Detection. In *Proceedings of the 20th international joint conference on Artifical intelligence (IJCAI'07), San Francisco, CA*, pp. 1624-1628.

Karin Kipper, Hoa Trang Dang, and Martha Palmer. 2000. Class-based construction of a verb lexicon. In *Proceedings of the 17th National Conference on Artificial Intelligence*, Austin, Texas. pp. 691-696.

Karin Kipper, Anna Korhonen, Neville Ryant, and Martha Palmer. 2008. A large-scale classification of English verbs. *Language Resources and Evaluation* 42(1):21-40.

Saisuresh Krishnakumaran and Xiaojin Zhu, 2007. Hunting Elusive Metaphors Using Lexical Resources, *Proceedings of the Workshop on Computational Approaches to Figurative Language*, pp. 13-20.

Linlin Li and Caroline Sporleder. 2010. Linguistic Cues for Distinguishing Literal and Non-Literal Usage. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, China, pp. 683-691.

Katia Markert and Nissim Malvina. 2009. Data and Models for Metonymy Resolution. In *Language Resources and Evaluation*, 43(2):123-138.

James Martin. 1990. *A Computational Model of Metaphor Interpretation*. Academic Press.

Zachary J. Mason. 2004. Cormet: A computational, corpus-based conventional metaphor extraction system. *Computational Linguistics*, 30(1):23–44.

Diana McCarthy and John Carrol. 2003. Disambiguating nouns, verbs and adjectives using automatically acquired selectional preferences. *Computational Linguistics.* 29(4): 639-654.

Diana McCarthy, Rob Koeling, Julie Weeds, and John Carroll. 2004. Finding predominant word senses in untagged text. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics* (*ACL '04*), Barcelona, Spain. pp. 280-287.

George Miller. 1995. Wordnet: A lexical database for English. *Communications of the ACM*, 38(11):39-41.

Martha Palmer, Dan Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71-106.

Wim Peters and Yorick Wilks. 2003. Data-Driven Detection of Figurative Language Use in Electronic Language Resources, *Metaphor and Symbol*, 18(3): 161-174.

Philip Resnik, 1997. Selectional preference and sense disambiguation. In *Proceedings of the ACL SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What and How?*, Washington, DC, pp. 52-57.

Ekaterina Shutova, Li-ping Sun and Anna Korhonen. 2010. Metaphor Identification Using Verb and Noun Clustering. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, Beijing, China, pp. 1002-1010.

Yorick Wilks, 1978. Making Preferences More Active. *Artificial Intelligence*, 11(3):197-223.