Inducing Frame Semantic Verb Classes from WordNet and LDOCE

Abstract
This paper presents SemFrame, a system that induces frame semantic verb classes from WordNet and LDOCE. Semantic frames are thought to have significant potential in resolving the paraphrase problem challenging many language-based applications.

When compared to the handcrafted FrameNet, SemFrame achieves its best recall-precision balance with 83.2% recall (based on SemFrame’s coverage of FrameNet frames) and 73.8% precision (based on SemFrame verbs’ semantic relatedness to frame-evoking verbs). The next best performing semantic verb classes achieve 56.9% recall and 55.0% precision.

1 Introduction
Semantic content can almost always be expressed in a variety of ways. Lexical synonymy (She esteemed him highly vs. She respected him greatly), syntactic variation (John paid the bill vs. The bill was paid by John), overlapping meanings (Anna turned right at Elm vs. Anna rounded the corner at Elm), and various other phenomena interact to produce a broad range of choices for most language generation tasks (Hirst, 2003; Rinaldi et al., 2003; Kozlowski et al., 2003). At the same time, natural language understanding must recognize what remains constant across paraphrases.

The paraphrase phenomenon affects virtually all computational linguistic applications, including information retrieval, information extraction, question-answering, and machine translation. For example, documents that express the same content using different linguistic means should typically be found relevant to and retrieved for the same queries. Information sought after to answer a question needs to be recognized no matter how it is expressed.

Machine translation (MT) often addresses the paraphrase problem by adopting an interlingual approach, based on a language-neutral conceptual representation (Dorr, 1993). However, MT lacks as yet an interlingual representation of broad scope that is conceptually, rather than lexically, based. In the most successful phrase-based approaches for statistical MT (Och and Ney, to appear), failure to equate equivalent phrases is a contributor to data sparseness.

Semantic frames (Fillmore 1982; Fillmore and Atkins 1992) address the paraphrase problem by providing slot-and-filler templates to represent frequently occurring, structured experiences. Semantic frame (type) is of an intermediate granularity have the potential to fulfill the interlingua role within a solution to the paraphrase problem.

Until now, the generation of semantic frames has been based solely on native speaker intuition; the FrameNet project (Johnson et al., 2002) now couples this generation with empirical validation. Only recently has this project begun to achieve relative breadth in its inventory of semantic frames. To have a comprehensive inventory of semantic frames, however, we need the capacity to generate semantic frames semi-automatically (the need for manual post-editing is assumed).

To address these challenges, we have developed SemFrame, a system that induces semantic frames automatically. Overall, the system performs two primary functions: (1) identification of sets of verb senses that evoke a common semantic frame; and (2) identification of the conceptual structure of semantic frames. This paper explores the first task of identifying frame semantic verb classes.

2 Previous Work
The EAGLES (1998) report on semantic encoding differentiates between two approaches to the development of semantic verb classes: those based on syntactic behavior and those based on semantic criteria.

Levin (1993) groups verbs based on an analysis of their syntactic properties, especially their ability to be expressed in diathesis alternations; her

\[1\text{http://www.icsi.berkeley.edu/~framenet/}\]
approach reflects the assumption that the syntactic behavior of a verb is determined in large part by its meaning. Verb classes at the bottom of Levin’s shallow network group together (quasi-) synonyms, hierarchically related verbs, antonyms, alongside verbs with looser semantic relationships.

The verb categories based on Pantel and Lin (2002) and Lin and Pantel (2001) are induced automatically from a large corpus, using an unsupervised clustering algorithm, based on syntactic dependency features. The resulting clusters contain synonyms, hierarchically related verbs, and antonyms, as well as verbs more loosely related from the perspective of paraphrase.

The handcrafted WordNet (Fellbaum, 1998a) uses the hyperonymy/hyponymy relationship to structure the English verb lexicon into a semantic network. Each collection of a top-level node supplemented by its descendants may be seen as a semantic verb class.

In all fairness, resolution of the paraphrase problem is not the explicit goal of most efforts to build semantic verb classes. However, they can process some paraphrases through lexical synonymy, hierarchically related terms, and antonymy.

3 Resources Used in SemFrame

We adopt an approach that relies heavily on pre-existing lexical resources. Such resources have several advantages over corpus data in identifying semantic frames. First, both definitions and example sentences often mention their participants using semantic-type-like nouns, thus mapping easily to the corresponding frame element. Corpus data, however, are more likely to include instantiated participants, which may not generalize to the frame element. Second, lexical resources provide a consistent amount of data for word senses, while the amount of data in a corpus for word senses is likely to vary widely. Third, lexical resources provide their data in a more systematic fashion than do corpora.

Most centrally, the syntactic arguments of the verbs used in a definition often correspond to the semantic arguments of the verb being defined. For example, Table 1 gives the definitions of several verb senses in the Longman Dictionary of Contemporary English (LDOCE; Procter, 1978) that evoke the COMMERCIAL TRANSACTION frame, which includes as its semantic arguments a Buyer, a Seller, some Merchandise, and Money. Words corresponding to the Money (money, value), the Merchandise (property, goods), and the Buyer (buyer, buyers) are present in, and to some extent shared across, the definitions; however, no words corresponding to the Seller are present.

<table>
<thead>
<tr>
<th>Verb sense</th>
<th>LDOCE Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>buy 1</td>
<td>to obtain (something) by giving money (or something else of value)</td>
</tr>
<tr>
<td>buy 2</td>
<td>to obtain in exchange for something, often something of great value</td>
</tr>
<tr>
<td>buy 3</td>
<td>to be exchangeable for</td>
</tr>
<tr>
<td>purchase 1</td>
<td>to gain (something) at the cost of effort, suffering, or loss of something of value</td>
</tr>
<tr>
<td>sell 1</td>
<td>to give up (property or goods) to another for money or other value</td>
</tr>
<tr>
<td>sell 2</td>
<td>to offer (goods) for sale</td>
</tr>
<tr>
<td>sell 3</td>
<td>to be bought; get a buyer or buyers; gain a sale</td>
</tr>
</tbody>
</table>

Table 1. LDOCE Definitions for Verbs Evoking the COMMERCIAL TRANSACTION Frame

Of available machine-readable dictionaries, LDOCE appears especially useful for this research. It uses a restricted vocabulary of about 2000 words in its definitions and example sentences, thus increasing the likelihood that words with closely related meanings will use the same words in their definitions and thus support the pattern of discovery envisioned. LDOCE also includes subject field codes, which accomplish some of the same type of grouping as semantic frames.

WordNet is a machine-readable lexico-semantic database whose primary organizational structure is the synset—a set of synonymous word senses, or alternatively, a concept that has been lexicalized. A limited number of relationship types (e.g., antonymy, hyponymy, meronymy, troponymy, entailment) are also used to relate synsets within a part of speech.

Fellbaum (1998b) suggests that relationships in WordNet “reflect some of the structure of frame semantics. For example, WordNet relates verbs like
buy and sell, which are part of a common frame. In fact, both frame semantics and the relational semantics in WordNet share a great deal with semantic field analysis in that they all naturally relate words and concepts from a common semantic domain” (p. 5). Through the relational structure of WordNet, buy, purchase, sell, and pay are related together: buy and purchase comprise one synset; they entail paying and are opposed to sell.

The relationship of buy, purchase, sell, and pay to other COMMERCIAL TRANSACTION verbs—for example, cost, price, charge—is not made explicit in WordNet, however. Further, as Roger Chaffin has noted, the specialized vocabulary of, for example, tennis (e.g. racket, court, lob) is not collocated, but is dispersed across different branches of the noun network (Miller, 1998, p. 34).

4 SemFrame Approach

SemFrame gathers evidence about frame semantic relatedness between verb senses by analyzing LDOCE and WordNet data from a variety of perspectives. The overall approach used is shown in Figure 1. The first stage of processing extracts pairs of verb senses that potentially evoke favoring recall; subsequent stages improve the precision of the resulting data.

Figures 2 and 3 give details of the algorithms for extracting verb pairs based on different types of evidence. These include: clustering LDOCE verb senses/WordNet synsets on the basis of words in their definitions and example sentences (fig. 2); relating LDOCE verb senses defined in terms of the same verb (fig. 3a); relating LDOCE verb senses that share a common stem (fig. 3b); extracting explicit sense-linking relationships in LDOCE (fig. 3c); relating verb senses that share general or specific subject field codes in LDOCE (fig. 3d); extracting (direct or extended) semantic relationships in WordNet (fig. 3e); and relating LDOCE verb senses that map to the same WordNet synset (fig. 3f).

In the second stage, mapping between WordNet verb synsets and LDOCE verb senses relies on finding matches between the data available for the verb senses in each resource (e.g., other words in the synset; words in definitions and example sentences; words closely related to these words; and stems of these words). The similarity measure used is the average of the proportion of words on each side of the comparison that are matched in the other. This mapping is used both to relate LDOCE verb senses, as indicated above, and to translate WordNet verb synsets paired above into LDOCE verb sense pairs.

In the third stage, the resulting verb sense pairs are merged into a single data set, retaining only those pairs whose cumulative support exceeds thresholds for either the number of supporting data sources or strength of support, thus achieving higher precision in the merged data set than in the input data sets. Then, the graph formed by the verb sense pairs in the merged data set is analyzed to find the fully connected components.

Finally, these groups of verb senses become input to a clustering operation (Voorhees, 1986). Those groups whose similarity (due to overlap in membership) exceed a threshold are merged together, thus reducing the number of verb sense groups. The verb senses within each resulting group are hypothesized to evoke the same semantic frame and constitute a frameset.

Figure 1. Approach for Building Frame Semantic Verb Classes

the same frame, some from LDOCE and some from WordNet. By exploiting many different clues to semantic relatedness, we overgenerate these pairs,
5 Results

We explored a range of thresholds in the final stage of the algorithm. In general, the lower the threshold, the looser the verb grouping and the fewer the number of clusters produced. The number of verb senses retained (out of 12,663 non-phrasal verb senses in LDOCE) and the verb sense groups produced by using these thresholds are recorded in Table 2.

6 Evaluation

Our goal is to produce frameset-like verb sets capable of extending FrameNet’s coverage while requiring reasonably little post-editing to ensure precision. Therefore FrameNet, which is hand-crafted and of reliably high precision, provides a gold standard against which we evaluate SemFrame’s output. Specifically, we measure the degree of correspondence between SemFrame frames and FrameNet frames and, for frames identified by both systems, the degree to which the verbs identified by SemFrame can be shown to evoke those frames.

FrameNet includes frames of varying levels of generality, with some semantic areas being covered by a general frame, some by a combination of specific frames, and some by a mix of general and specific frames. Thus, we determined the degree to which SemFrame and FrameNet overlap by automatically finding and comparing corresponding frames instead of fully equivalent frames. Frames correspond if the semantic scope of one frame is included within the semantic scope of the other frame or if the semantic scopes of the two frames have significant overlap. (Since FrameNet lists evoking words, without specification of word sense, the comparison was done on the word level rather than on the word sense level, as if LDOCE verb senses were not specified, with multiple senses of an LDOCE verb counting only once.)

Table 2. Results of Frame Clustering Process

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Num verb senses</th>
<th>Num groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>6461</td>
<td>1338</td>
</tr>
<tr>
<td>1.0</td>
<td>6414</td>
<td>1759</td>
</tr>
<tr>
<td>1.5</td>
<td>5607</td>
<td>1421</td>
</tr>
<tr>
<td>2.0</td>
<td>5604</td>
<td>1563</td>
</tr>
</tbody>
</table>

2 Certain constraints imposed by FrameNet’s development strategy constitute drawbacks to its use as a gold standard, however. (1) As of summer 2003, only 382 frames had been identified within the FrameNet project. (2) Low recall affects not only the set of semantic frames identified by FrameNet, but also the individual sets of frame-evoking units listed for each frame. No verbs are listed for 38.5% of FrameNet’s frames, while another 13.1% of them list only 1 or 2 verbs. (3) Many of FrameNet’s frames are more syntactically than semantically motivated (e.g., EXPERIENCER-OBJECT, EXPERIENCER-SUBJECT).
Figure 3. Algorithms for Generating Non-clustering-based Verb Pairs

Correspondence between FrameNet and SemFrame frames is established, for purposes of evaluation, in two ways. In the first, a SemFrame frame is deemed to correspond to a FrameNet frame
if the two frames meet both a minimal-overlap criterion and a frame-name-relatedness criterion. The minimal-overlap criterion is met in either of two ways: (1) If the FrameNet frame lists four or fewer verbs (true of over one-third of the FrameNet frames that list associated verbs), minimal overlap occurs when any one verb associated with the FrameNet frame matches a verb associated with a SemFrame frame. (2) If the FrameNet frame lists five or more verbs, minimal overlap occurs when two or more verbs in the FrameNet frame are matched by verbs in the SemFrame frame.

Establishing frame-name relatedness involves identification of individual components of each frame name and augmentation of this set with morphological variants from CatVar (Habash and Dorr 2003). The resulting set for each FrameNet and SemFrame frame name is then used to search both the noun and verb WordNet networks to find all the synsets that might correspond to the frame name. To this set are added all synsets directly related to the synsets that correspond to the frame names. If the set of synsets gathered for a FrameNet frame name intersects with the set of synsets gathered for a SemFrame frame name, the two frame names are deemed to be semantically related.

For example, the FrameNet ADORNING frame contains 17 verbs: adorn, blanket, cloak, coat, cover, deck, decorate, dot, encircle, envelop, festoon, fill, film, line, pave, stud, and wreath. The SemFrame ORNAMENTATION frame contains 12 verbs: adorn, caparison, decorate, embellish, embroider, garland, garnish, gild, grace, hang, incrust, and ornament. Two of the verbs—adorn and decorate—are shared. In addition, the frame names are semantically related through a WordNet synset consisting of decorate, adorn (which CatVar relates to ADORNING), grace, ornament (which CatVar relates to ORNAMENTATION), embellish, and beautify. The two frames are therefore designated as corresponding frames by meeting both the minimal-overlap and the frame-name relatedness criteria.

In the second correspondence scenario, a SemFrame frame is deemed to correspond to a FrameNet frame if they meet one of two more stringent verb overlap criteria, the majority-match criterion or the majority-related criterion.

The majority-match criterion is met if the set of verbs shared by FrameNet and SemFrame framesets account for half or more of the verbs on either side of the comparison. For example, the APPLY HEAT frame in FrameNet includes 22 verbs: bake, blanch, boil, braise, broil, brown, char, coddle, cook, fry, grill, microwave, parboil, poach, roast, saute, scald, simmer, steam, steep, stew, and toast, while the BOILING frame in SemFrame includes 7 verbs: boil, coddle, jug, parboil, poach, seethe, and simmer. Five of these verbs—boil, coddle, parboil, poach, and simmer—are shared across the two frames and constitute over half of the SemFrame frameset. Therefore the two frames are deemed to correspond by meeting the majority-match criterion.

The majority-related criterion is met if half or more of the verbs from the SemFrame frame are semantically related to verbs from the FrameNet frame (that is, if the precision of the SemFrame verb set is at least 0.5). To evaluate this criterion, each FrameNet and SemFrame verb is associated with the WordNet verb synsets it occurs in, augmented by the synsets to which the initial sets of synsets are directly related. If the sets of synsets corresponding to two verbs share one or more synsets, the two verbs are deemed to be semantically related. This process is extended one further link, such that a SemFrame verb that is found by this process to be semantically related to a FrameNet verb, already found to be semantically related to a FrameNet verb, will also be established as a frame-evoking verb. If half or more of the verbs listed for a SemFrame frame are established as evoking the same frame as the list of WordNet verbs, then the FrameNet and SemFrame frames are hypothesized to correspond through the majority-related criterion.

For example, the FrameNet ABUNDANCE frame includes 4 verbs: crawl, swarm, teem, and throng. The SemFrame FLOW frame likewise includes 4 verbs: pour, teem, stream, and pullulate. Only one verb—teem—is shared, so the majority-match criterion is not met, nor is the related-frame-name criterion met, as the frame names are not
semantically related. The majority-related criterion, however, is met through a WordNet verb synset that includes *pour, swarm, stream, teem, and pullulate*.

Of the 197 FrameNet frames that include at least one LDOCE verb, 175 were found to have a corresponding SemFrame frame. But this 88.8% recall level should be balanced against the precision ratio of SemFrame verb framesets. After all, we could get 100% recall by listing all verbs in every SemFrame frame.

The majority-related function computes the precision ratio of the SemFrame frame for each pair of FrameNet and SemFrame frames being compared. By modifying the minimum precision threshold the balance between recall and precision, as measured using F-measure, can be investigated. The best balance for the SemFrame version is based on a clustering threshold of 2.0 and a minimum precision threshold of 0.4, which yields a recall of 83.2% and overall precision of 73.8%.

To interpret these results meaningfully, one would like to know if SemFrame achieves more FrameNet-like results than do other available verb category data, more specifically the 357 semantic verb classes of WordNet 1.7.1, the 258 verb classes from Levin, or the 272 verb clusters of Lin and Pantel, as described in Section 2.

For purposes of comparison with FrameNet, the name of a WordNet-based frame is taken from the words for the root-level synset; Levin’s verb class names have been hand-edited to isolate the word that best captures the semantic sense of the class; and the name of each Lin and Pantel cluster is taken to be the first verb in the cluster.4

Evaluation results for the best balance between recall and precision of the four comparisons are summarized in Table 3. FrameNet itself constitutes the upper bound on the task, i.e., 100% recall and 100% precision. The Lin & Pantel results may be treated as a lower bound for automatically induced semantic verb classes. The ability of SemFrame to develop verb classes that correspond to semantic frames is readily apparent.

<table>
<thead>
<tr>
<th>Semantic verb classes</th>
<th>Precision threshold</th>
<th>Recall</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SemFrame</td>
<td>0.40</td>
<td>0.832</td>
<td>0.738</td>
</tr>
<tr>
<td>WordNet</td>
<td>0.15</td>
<td>0.528</td>
<td>0.466</td>
</tr>
<tr>
<td>Levin</td>
<td>0.20</td>
<td>0.569</td>
<td>0.550</td>
</tr>
<tr>
<td>Lin &amp; Pantel</td>
<td>0.15</td>
<td>0.472</td>
<td>0.407</td>
</tr>
</tbody>
</table>

Table 3. Best Recall-Precision Balance When Compared with FrameNet

7 Conclusions and Future Work

We have demonstrated that sets of verbs evoking a common semantic frame can be induced from existing lexical tools. In a head-to-head comparison with frames in FrameNet, the frame semantic verb classes developed by the SemFrame approach achieve a recall of 83.2% and the verbs listed for frames achieve a precision of 73.8%; these results far outpace those of other semantic verb classes. On a practical level, a large number of frame semantic verb classes have been identified. Associated with clustering threshold 1.5 are 1421 verb classes, averaging 14.1 WordNet verb synsets. Associated with clustering threshold 1.5 are 1563 verb classes, averaging 6.6 WordNet verb synsets.

Despite these promising results, we are limited by the scope of our input data set. While LDOCE and WordNet data are generally of high quality, the relative sparseness of these resources has an adverse impact on recall. In addition, the mapping technique used for picking out corresponding word senses in WordNet and LDOCE is shallow, thus constraining the recall and precision of SemFrame outputs. Finally, the multi-step process of merging smaller verb groups into verb groups that are intended to correspond to frames sometimes fails to achieve an appropriate degree of correspondence (all the verb classes discovered are not distinct).

In our future work, we will experiment with the more recent release of WordNet (2.0). This version provides derivational morphology links between nouns and verbs, which will promote far greater precision in the linking of verb senses based on morphology than was possible in our initial implementation. Another significant addition to WordNet 2.0 is the inclusion of category domains,

---

4 Lin and Pantel have taken a similar approach, “naming” their verb clusters by the first three verbs listed for a cluster, i.e., the three most similar verbs.
which co-locate words pertaining to a subject (e.g., the verb overbid pertains to the auction, auction.sale, vendue domain) and perform the same function as LDOCE’s subject field codes. With the addition of derivational morphology links and category domain links, WordNet 2.0 is becoming a single, integrated network rather than a set of separate networks. This integration brings with it implied relationships, which can be further mined. For example, the set of verb senses with derivational morphology links to the same set of nouns are themselves both morphologically and semantically related.5

References


5 A near-term enhancement of WordNet is the sense-tagging of content words in WordNet glosses in Senseval-3 (Litkowski, 2004), making available another rich source of semantic data about relationships between word senses.