Building a class-based verb lexicon using TAGs

Karin Kipper, Hoa Trang Dang, William Schuler, and Martha Palmer

Department of Computer and Information Sciences
University of Pennsylvania
200 South 33rd Street
Philadelphia, PA 19104, USA
{kipper,htd,schuler,mpalmer}@linc.cis.upenn.edu

Abstract

We present a class-based approach to building a verb lexicon that makes explicit the close relation between syntax and semantics for Levin classes. We have used a Lexicalized Tree Adjoining Grammar to capture the syntax associated with each verb class and have added semantic predicates to each tree, which allow for a compositional interpretation.

1. Introduction

We describe a computational verb lexicon called VerbNet which utilizes Levin verb classes (Levin, 1993) to systematically construct lexical entries. We have used Lexicalized Tree Adjoining Grammar (LTAG) (Joshi, 1985; Schabes, 1990) to capture the syntax associated with each verb class, and have added semantic predicates. We also show how regular extensions of verb meaning can be achieved through the adjunction of particular syntactic phrases. We base these regular extensions on intersective Levin classes, a fine-grained variation on Levin classes, as a source of semantic components associated with specific adjuncts (Dang *et al.*, 1998). Whereas previous research on tying semantics to Levin classes (Dorr, 1997) has not explicitly implemented the close relation between syntax and semantics hypothesized by Levin, our lexical resource combines traditional lexical semantic information, such as thematic roles and semantic predicates, with syntactic frames and selectional restrictions. In order to increase the utility of VerbNet, we also include links to entries in WordNet, which is one of the most widely used online lexical databases in Natural Language Processing applications.

2. Levin Classes and WordNet

Two current approaches to English verb classifications are WordNet and Levin classes. WordNet is an on-line lexical database of English that currently contains approximately 120,000 sets of noun, verb, adjective, and adverb synonyms, each representing a lexicalized concept. A synset (synonym set) contains, besides all the word forms that can refer to a given concept, a definitional gloss and – in most cases – an example sentence. Words and synsets are interrelated by means of lexical and semantic-conceptual links, respectively. Antonymy or semantic opposition links individual words, while the super-/subordinate relation links entire synsets. WordNet was designed principally as a semantic network, and contains little syntactic information. Even as a semantic resource, however, it is missing some of the information that has traditionally been required by NLP applications, including explicit predicate-argument structures. WordNet senses are often too fine-grained as well, lacking an underlying notion of semantic components and a systematic extension of basic senses to produce these fine-grained senses.

The Levin verb classification, on the other hand, does explicitly state the syntax for each class, but still falls short of assigning semantic components to each class. The classes are based on the ability or inability of a verb to occur in pairs of syntactic frames that are in some sense meaning preserving (diathesis alternations) (Levin, 1993). The sets of syntactic frames associated with a particular Levin class are supposed to reflect underlying semantic components that constrain allowable arguments and adjuncts. For example, *break* verbs and *cut* verbs are similar in that they can all participate in the transitive and middle constructions. However, only *break* verbs can also occur in the simple intransitive, and *cut* verbs can occur in the conative, where *break* verbs cannot. The explanation given is that *cut* describes a series of actions directed at achieving the goal of separating some object into pieces. It is possible for these actions to be performed without the end result being achieved, but where the *cutting* manner can still be recognized (i.e., "John cut at the loaf"). For *break*, the only thing specified is the resulting change of state where the object becomes separated into pieces. If the result is not achieved, no attempted *breaking* action can be recognized.

1. Transitive construction
   (a) John broke the window.
   (b) John cut the bread.
2. Middle construction
   (a) Glass breaks easily.
   (b) This loaf cuts easily.
3. Intransitive construction
   (a) The window broke.
   (b) *The bread cut.
4. Conative construction
   (a) *John broke at the window.
   (b) John valiantly cut/hacked at the frozen loaf, but his knife was too dull to make a dent in it.

The fundamental assumption is that the syntactic frames are a direct reflection of the underlying semantics. However, Levin classes exhibit inconsistencies that have hampered researchers' ability to reference them directly in applications. Many verbs are listed in multiple classes, some of which have conflicting sets of syntactic frames. For instance, *carry* verbs are described as not taking the conative (*"The mother carried at the baby"), and yet many of the verbs in the *carry* class *(push, pull, tug, shove, kick)* are also listed in the *push/pull* class, which does take the conative. Dang et al. (1998) showed that multiple listings could in some cases be interpreted as regular sense extensions, and defined intersective Levin classes, which are a more syntactically and semantically coherent refinement of basic Levin classes. We implement these verb classes and their regular sense extensions in the Lexicalized Tree Adjoining Grammar formalism.

3. Verb lexicon

VerbNet can be viewed in both a static and a dynamic way. The static aspect refers to the verb entries and how they are organized, providing the characteristic descriptions of a verb sense or a verb class. The dynamic aspect of the lexicon constrains the entries to allow a compositional interpretation in LTAG derivation trees, capturing extended verb meanings by incorporating adjuncts.

3.1. *Static description*

Each verb entry refers to a set of classes, corresponding to the different senses of the verb. For example, the manner of motion sense of "run" is a member of the *Manner of Motion* class, whereas "run" as in "the street runs through the district" is a member of the *Meander* class. For each verb sense there is a verb class as well as specific selectional restrictions (e.g., an instrument of "kick" must be of type $foot$) and semantic characteristics (e.g., a particular manner of directed motion) that may not be captured by the class membership. In order to provide a mapping to other dictionaries, we also include links to WordNet synsets. Because WordNet has more fine-grained sense distinctions than Levin, each verb sense in VerbNet references the set of WordNet synsets (if any) that captures the meaning appropriate to the class.

Verb classes allow us to capture generalizations about verb behavior. This reduces not only the effort needed to construct the lexicon, but also the likelihood that errors are introduced when adding a new verb entry. Each verb class lists the thematic roles that the predicate-argument structure of its members allows, and provides descriptions of the syntactic frames corresponding to licensed constructions, with selectional restrictions defined for each argument in each frame. Each frame also includes semantic predicates describing the participants at various stages of the event described by the frame.

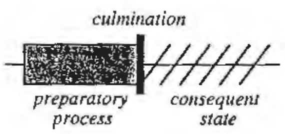


Figure 1: Moens and Steedman's tripartite structure of events

We decompose each event E into a tripartite structure in a manner similar to Moens and Steedman (1988), introducing a time function for each predicate to specify whether the predicate is true in the preparatory ($during(E)$), culmination ($end(E)$), or consequent ($result(E)$) stage of an event. The tripartite event structure (Figure 1) allows us to express the semantics of classes of verbs like change of state verbs whose adequate description requires reference to a complex event structure. In the case of a verb such as "break", it is important to make a distinction between the state of the object before the end of the action ($during(E)$), and the new state that results afterwards (*result(E)*).

Verb classes are hierarchically organized, ensuring that each class is coherent – that is, all its members have common semantic elements and share a common set of thematic roles and basic syntactic frames. This requires some manual restructuring of the original Levin classes, which is facilitated by using intersective Levin classes. In addition, a particular verb may add more semantic information to the basic semantics of its class.

Figure 2 shows a partial entry for the *Hit* class. This class allows for three thematic roles: Agent, Patient and Instrument, with constraints that the Agent is generally animate; the Patient concrete; and the Instrument concrete and inanimate.[1] These selectional restrictions refer to

[1]These constraints are more like preferences that generate a preferred reading of a sentence. They may be

HIT class

⟨⟨MEMBERS⟩⟩	$[\langle hit, 1\rangle, \langle kick, 1\rangle, \langle slap, 1\rangle, \langle tap, 1\rangle, \ldots]$
⟨⟨THEMATIC ROLES⟩⟩	Agent(A), Patient(P), Instrument(I)
⟨⟨SELECT RESTRICTIONS⟩⟩	Agent[+animate], Patient[+concrete], Instrument[+concrete,-animate]

⟨⟨FRAMES and PREDICATES⟩⟩

Basic Transitive	A V P	manner(during(E),directedmotion,A) ∧ manner(end(E),forceful,A) ∧ contact(end(E),A,P)
Transitive with Instrument	A V P with I	manner(during(E),directedmotion,I) ∧ manner(end(E),forceful,I) ∧ contact(end(E),I,P)
Conative	A V at P	manner(during(E),directedmotion,A)
With/against alternation	A V I against/on P	manner(during(E),directedmotion,I) ∧ manner(end(E),forceful,I) ∧ contact(end(E),I,P)
Transitive	I V P	manner(during(E),directedmotion,I) ∧ manner(end(E),forceful,I) ∧ contact(end(E),I,P)

Figure 2: Partial entry for the *Hit* class

a feature hierarchy where *animate* subsumes *animal* and *human*, *concrete* subsumes both *animate* and *inanimate*, and so forth. This representation does not suffer from some drawbacks of theta role analysis because our roles are not global primitives, but are only used to describe relationships within a class.

The strength of our representation comes from the explicit relationship between syntax and semantics captured in each entry. Figure 2 shows some of the syntactic frames allowed for the *Hit* class and the semantic predicates for each frame. Thematic roles are used as descriptors which are mapped into arguments of semantic predicates as well as the argument positions in a TAG elementary tree.

The tripartite event structure also handles the conative construction, in which there is an intention of a goal during the event which is not achieved at the end of the event. The example shown in Figure 2 for the conative construction has the predicate

manner(during(E),directedmotion,A)

but because the intended contact by sudden impact is not satisfied, the semantics does not include the predicates

manner(end(E),forceful,A) ∧ *contact(end(E),A,P)*.

3.2. *Compositional Semantics*

We use TAG elementary trees to describe syntactic frames and associate semantic predicates and selectional restrictions with each tree. Elementary trees capture the basic semantics of the verbs in each class. Each frame in the static aspect of the lexicon maps onto a TAG elementary tree, in which the thematic roles correspond to substitution sites. Some auxiliary trees are class-based because they interact with the verbs in the class in peculiar ways and add seman-

relaxed depending on the domain of a particular application.

$$manner(during(E), directedmotion, X_{arg0}) \wedge$$
$$manner(end(E), forceful, X_{arg0}) \wedge$$
$$contact(end(E), X_{arg0}, X_{arg1})$$

$$meets(E_{arg0}, E) \wedge$$
$$motion(during(E), X_{arg0.arg1}) \wedge$$
$$via(during(E), X_{arg0.arg1}, X_{arg1})$$

Figure 3: Initial transitive tree for "hit" and auxiliary tree for "across"

tic content specific to the class. Others, such as temporal adjuncts, bring the same semantic predicate independent of the verb. We use a flat semantic representation like that of Joshi and Vijay-Shanker (1999) in which the semantics of a sentence is the conjunction of the semantic predicates of the trees used to derive the sentence.

We ensure that all the semantic arguments of basic predicates are local to the syntactic initial tree. For example, the basic transitive frame in Figure 2 shows that the Agent is in direct motion and contacts the Patient in a forceful manner. If an instrument is specified, it replaces the Agent in these predicates. Since the instrument can be an argument in the basic predicates of the *Hit* class, it must appear in the elementary trees whenever it is specified, even if it is in a prepositional phrase.

The ability of certain verbs to take on extended senses based on their adjuncts is captured in a natural way by the TAG operation of adjunction and our conjunction of semantic predicates. Figure 3 shows an initial transitive tree anchored by "hit" and the semantic predicates associated with this syntactic frame. The original *Hit* verb class does not include movement of the direct object as part of the meaning of "hit" – only one event of contact by sudden impact is described. This event is subdivided into three predicates: the first,

$manner(during(E),directedmotion,X_{arg0})$

specifies that during the event E, X_{arg0} is in directed motion; the second,

$manner(end(E),forceful,X_{arg0})$

refers to the forceful contact of X_{arg0} at the end of E; and the third,

$contact(end(E),X_{arg0},X_{arg1})$

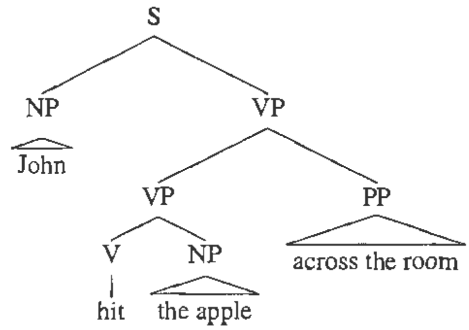
establishes that at the end of event E, contact between X_{arg0} and X_{arg1} has been achieved.

By adjoining a path PP such as "across NP", we get an extended meaning, and a change in Levin class membership to the *Throw* class. Figure 3 shows the auxiliary tree anchored by the preposition "across" together with its semantic predicates. The class-specific path PP adds the predicates

$meets(E_{arg0},E) \wedge motion(during(E),X_{arg0.arg1}) \wedge via(during(E),X_{arg0.arg1},X_{arg1})$,

introducing a motion event that immediately follows (*meets*) the contact event, which is the basic sense of the *Hit* class.

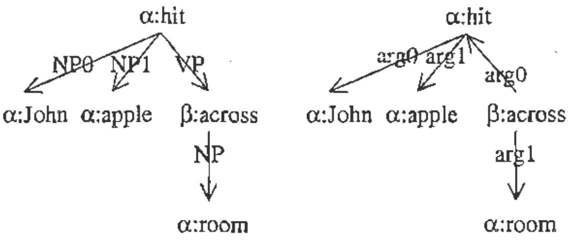
In Figure 4, we show the derived tree for the sentence "John hit the apple across the room" with all the predicates instantiated. The arguments are recovered from the derivation tree, following Candito and Kahane (1998). When an initial tree, such as α_{John}, is substituted into another



$$manner(during(e1), directedmotion, john) \wedge manner(end(e1), forceful, john) \wedge$$
$$contact(end(e1), john, apple) \wedge meets(e1, e2) \wedge motion(during(e2), apple) \wedge$$
$$via(during(e2), apple, room)$$

Figure 4: Sense extension of "hit" through adjunction of a path PP

tree α_{hit}, the dependency mirrors the derivation structure, so the variables associated with the substituting tree can be referenced as arguments in the host tree's predicates (see Figure 5). When an auxiliary tree β_{across} is adjoined, the dependency for the adjunction is reversed, so that variables associated with the host tree can be referenced as arguments in the adjoining tree's predicates. With this dependency from β_{across} to α_{hit} (labeled $arg0$), it is now possible for the semantic predicates associated with β_{across} to predicate over variables in the dependent tree α_{hit}, including the variable $X_{arg0.arg1}$ instantiated as *apple*, resulting in the predicates *motion(during(e2),apple)* ∧ *via(during(e2),apple,room)*.



Derivation structure   Dependency structure

Figure 5: Derivation and dependencies

Verbs in the intersective class formed by the *Push/Pull* verbs and the *Carry* verbs behave in a similar manner. The core meaning of this verb class is exertion of force. Adjunction of a path PP implying motion modifies membership of these verbs to the *Carry* class. *Push/Pull* verbs can appear in the conative construction, which emphasizes their forceful semantic component and ability to express an *attempted* action where any result that might be associated with the verb is not necessarily achieved; *Carry* verbs (used with a goal or directional phrase) cannot take the conative alternation because this would conflict with the causation of motion which is the intrinsic meaning of the class (Dang *et al.*, 1998).

Palmer et al. (1999) and Bleam et al. (1998) also defined compositional semantics for classes of verbs implemented in FB-LTAG, but they represented general semantic components (e.g., motion, manner) as features on the nodes of the trees. Our use of separate logical forms gives a more detailed semantics for the sentence, so that for an event involving motion, it is possible to know not only that the event has a *motion* semantic component, but also which entity is actually in motion.

4. Conclusion

We have presented a class-based approach to building a verb lexicon that makes explicit the close association between syntax and semantics, as postulated by Levin. By using verb classes we capture generalizations about verb behavior and reduce not only the effort needed to construct the lexicon, but also the likelihood that errors are introduced when adding new verbs. Another important contribution of this work is that by dividing each event into a tripartite structure, we permit a more precise definition of the associated semantics, which is necessary for applications such as animation of natural language instructions (Bindiganavale *et al.*, 2000). The power of the lexicon comes from its dynamic aspect which is based on the LTAG formalism. The operation of adjunction in TAGs provides a principled approach to representing the type of regular polysemy that has been a major obstacle in building verb lexicons.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments. This research was partially supported by NSF grants IIS 9800658 and IIS-9900297, Office of Naval Research AASERTs N00014-97-1-0603, and CAPES grant 0914-95.

References

BINDIGANAVALE R., SCHULER W., ALLBECK J. M., BADLER N. I., JOSHI A. K. & PALMER M. (2000). Dynamically Altering Agent Behaviors Using Natural Language Instructions. *Fourth International Conference on Autonomous Agents.*

BLEAM T., PALMER M. & VIJAY-SHANKER K. (1998). Motion Verbs and Semantic Features in TAG. In *Proceedings of the Fourth TAG+ Workshop*, Philadelphia, PA.

CANDITO M.-H. & KAHANE S. (1998). Can the TAG Derivation Tree Represent a Semantic Graph? An Answer in the Light of Meaning-Text Theory. In *Proceedings of the Fourth TAG+ Workshop*, p. 21–24, Philadelphia, PA.

DANG H. T., KIPPER K., PALMER M. & ROSENZWEIG J. (1998). Investigating Regular Sense Extensions Based on Intersective Levin classes. In *Proceedings of COLING-ACL98*, Montreal, CA.

DORR B. J. (1997). Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation. *Machine Translation*, **12**, 1–55.

JOSHI A. K. (1985). How Much Context Sensitivity Is Necessary for Characterizing Structural Descriptions: Tree Adjoining Grammars. In L. K. D. DOWTY & A. ZWICKY, Eds., *Natural language parsing: Psychological, computational and theoretical perspectives*, p. 206–250. Cambridge, U.K.: Cambridge University Press.

JOSHI A. K. & VIJAY-SHANKER K. (1999). Compositional Semantics with Lexicalized Tree-Adjoining Grammar: How Much Under-Specification Is Necessary? In *Proceedings of the Third International Workshop on Computational Semantics (IWCS-3)*, p. 131–145, Tilburg, The Netherlands.

LEVIN B. (1993). *English Verb Classes and Alternation: A Preliminary Investigation*. The University of Chicago Press.

MOENS M. & STEEDMAN M. (1988). Temporal Ontology and Temporal Reference. *Computational Linguistics*, 14, 15–38.

PALMER M., ROSENZWEIG J. & SCHULER W. (1999). Capturing Motion Verb Generalizations in Synchronous Tree-Adjoining Grammar. In P. SAINT-DIZIER, Ed., *Predicative Forms in Natural Language and in Lexical Knowledge Bases*. Kluwer Press.

SCHABES Y. (1990). *Mathematical and Computational Aspects of Lexicalized Grammars*. PhD thesis, Computer Science Department, University of Pennsylvania.