

Annotating Noun Argument Structure for NomBank

Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekely,
Veronika Zielinska, Brian Young, Ralph Grishman

New York University
719 Broadway, 7th Floor
New York, New York 10003
U.S.A.

meyers/reevesr/macleod/szekely/zielinsk/byoung/grishman@cs.nyu.edu

Abstract

When complete, NomBank will provide annotation of noun arguments in Penn Treebank II (PTB). In PropBank, University of Pennsylvania annotators provide similar information for verbs. Given nominalization/verb mappings, the combination of NomBank and PropBank allows for generalization of arguments across parts of speech. This paper describes our annotation task including factors which make assigning role labels to noun arguments a challenging task.

1. Introduction

When complete, NomBank will provide annotation of noun arguments in the Penn Treebank II (PTB), including arguments of both nominalizations (*his resignation*) and other nouns (*his height*). Each triple {noun, sense-label, argument-label} defines a relation between a noun and one of its arguments. For example, the CAPITALIZED items are ARG1s of one sense of *destruction: the destruction of SOME 100,000 WEAPONS* and *minor PROPERTY destruction*. Assuming statistical procedures trained on NomBank can identify such ARG1s, a few argument-based patterns can replace many string based patterns for applications like Information Extraction, Machine Translation and Question Answering. In PropBank(Kingsbury et al., 2002; Kingsbury and Palmer, 2002; University of Pennsylvania, 2002), University of Pennsylvania annotators provide similar information for verbs. Given nominalization/verb mappings, the combination of NomBank and PropBank affords even greater generalization.

This paper describes our annotation task including factors which make assigning role labels to noun arguments a challenging task.

2. Overview of the Task

There are approximately 200,000 instances of markable common nouns (nouns that can take arguments) in the PTB. We estimate that 5000 instances are nominalizations of adjectives and that the remaining items are evenly split between nominalizations of verbs (including argument nominalizations) and various semantic noun classes. Our task is for annotators to look at these 200,000 instances and mark those instances that actually take arguments.

We estimate that we will finish this task at the end of 2004. We began writing the specifications in January, 2003. We hired our first annotator to “debug” the specs and the system in February 2003. The second annotator was hired in September, 2003 and two more were hired in January 2004. We started slowly and expanded as we understood the task better. We believe that what we learned in the first year will help us to finish building NomBank more efficiently and accurately with a larger staff.

During training, we have achieved inter-annotator consistency scores ranging from 82% to 90%. Now that the annotators are trained, we will perform additional consistency testing on a monthly basis.

3. Lexical Entries

Annotators are provided with initial lexical entries for each of the approximately 6500 argument-taking common nouns in the Penn Treebank. Each entry lists one role-set (set of role labels) for each noun sense. Consequently, sense distinctions are closely aligned with differences in inventories of possible arguments. For our purposes, sense distinctions may be coarser than, for example, WordNet.

When necessary, an annotator will alter a lexical entry to fit the data. Annotators then apply the resulting lexical entry to instances of that noun in the PTB. Each roleset lists the possible argument labels for that noun sense from the set {ARG0, ARG1, ARG2, ARG3, ARG4}. These labels allow annotators to differentiate among arguments of each sense of each noun in a consistent manner. For example, *destruction* has exactly one sense, the ARG0 (subject) of *destruction*’s roleset corresponds to the agent of destruction and the ARG1 corresponds to the patient, e.g., *Richard* is the ARG0 and *the secret tapes* is the ARG1 in *Richard’s destruction of the secret tapes*.

3.1. The Initial Lexical Entries

In the initial stages of NomBank, each common noun occurring in the corpus was classified by a combination of automatic and manual means. The idea was to generate first approximations of lexical entries as quickly as possible. This subsection describes a set of defaults which the annotators can override.

For nominalizations of verbs that have been annotated in PropBank, the default is to choose the set of argument labels used in PropBank for the related verb, e.g., the ARG1 of sense 1 of *destroy* should correspond to the ARG1 of sense 1 of *destruction* (*Frodo destroyed THE RING* vs. *Frodo’s destruction OF THE RING*). This applies to both verbal nominalizations like *destruction* and argument nominalizations like *maker*. In a similar vein, we attempt to insure consistency across predicates. Typical subjects (e.g.,

agents), objects (patients, themes, etc.) and indirect objects (recipients, goals, etc.) tend to be marked as ARG0, ARG1 and ARG2 respectively, much the same as assumed under the Universal Alignment Hypothesis of Relational Grammar (Perlmutter and Postal, 1984; Rosen, 1984) and similar ideas in other frameworks. Some nominalization-like nouns are not morphologically related to commonly occurring verbs. We have nevertheless found it convenient to relate these “cousins” of nominalizations to verbs that take similar arguments, e.g., we assume that *aggression* takes the same set of arguments as the verb *attack*.

For adjective nominalizations (*his writing ability*), we use a combination of a quick classification scheme, COMLEX Syntax and heuristics to provide automatically generated approximations of their lexical entries (which are then hand-corrected). Adjectival nominalizations typically take 1 or 2 arguments and vary regarding whether one of their arguments is a typical ARG0 (e.g., *his* in *his anger*). Our initial assumption is that adjectives take one of the following rolesets: {ARG0}, {ARG1}, {ARG0, ARG1}, {ARG1, ARG2} depending on our classification scheme and the presence or absence of complements in the entry for the corresponding adjective in COMLEX Syntax.

In addition to nominalizations, we allow 16 additional classes of argument taking nouns. The five most frequent ones are:

1. PARTITIVE (*a SET of meetings, a wide VARIETY of crops*)
2. RELATIONAL (*Mindy Hymowitz’s MOTHER, PRESIDENT of the garden club*)
3. ENVIRONMENT (*a PERIOD of industry consolidation, a recent MORNING of working at home*)
4. ATTRIBUTE (*the BREADTH of inquiries, the HUMOR of his uncombed appearance*)
5. ABILITY (*the absolute RIGHT of everyone to disseminate materials, the ART of selling*)

For each class, we create one or more rolesets and then assume that all members of the class share these rolesets in their lexical entries, e.g., one roleset for ABILITY nouns includes one ARG0 (AGENT) and one ARG1 (ACTION).

3.2. Fine-tuning the Lexical Entries

When an annotator is given a word to annotate, they must look at every instance of that word in the corpus to annotate. During annotation, they may realize that the given lexical entry is not adequate for the task. In this case, they edit it.

Figure 1 is the current lexical entry for the noun *complaint*. The automatically produced lexical entry assumed that *complaint* was merely a nominalization of the verb *complain* – this only generated sense 2 corresponding to the Propbank Roleset “complain.01”. However, in actual text *complaint* is often used in its legal sense with the set of arguments listed in sense 1. The annotator, therefore, had to add this sense before creating the entry. Note that the examples in figure 1 mark items with the features REL (the

1. Legal Sense

Roles: ARG0 = CLAIMANT, ARG1 = REASON, ARG2 = ADJUDICATOR, ARG3 = DEFENDANT

Noun Example: *They filed a police complaint against the pasta maker for conspiracy*
REL = complaint, ARG0 = they, SUPPORT = filed, ARG1 = for conspiracy, ARG2 = police, ARG3 = the pasta maker

2. Sense based on verbal sense *complain.01*

Roles: ARG0 = AGENT, ARG1 = TOPIC, ARG2 = RECIPIENT

Noun Example: *There have been no customer complaints about that issue.*
REL = complaints, ARG0 = customer, ARG1 = about that issue, ARGM-NEG = no

Verb Example: *They complained about that issue*
REL = complain, ARG0 = they, ARG1 = about that issue

Figure 1: Two Senses of *complaint*

main predicate), ARGX (an argument with a number X, between 0 and 5), ARGM-XYZ (an adjunct, where XYZ is some adjunct class, e.g. NEG for negation) and SUPPORT (indicating a predicate that links the noun with an argument outside the NP, in this case the ARG0. Each of these items will be discussed further in later sections. Each set of feature value pairs that make up one of these examples is called a *proposition*. Each NomBank proposition must include a REL (or predicate) and one or more arguments or ARGMs. Only some propositions include SUPPORT.

4. The Annotation Task

Given all the instances of a markable noun, the annotator must mark all arguments and certain ARGMs that co-occur with these nouns. Nouns that do not co-occur with either arguments or ARGMs are left unmarked. Thus *the destruction* would not be marked, but *the destruction of wild America* would be marked because *of wild America* would be tagged as the ARG1. There are a few factors that can make this particularly challenging. First of all, SUPPORT and other phenomena make it possible for an argument of a noun to occur outside the NP headed by that noun. Secondly, not all adjuncts of the noun are legitimate ARGMs. In the interest of limiting the scope of this study, we only mark those classes of adjuncts that roughly correspond to verbal and sentential adjuncts. While this excludes noun-specific adjuncts out of hand (e.g., adjectives like *red*, and most relative clauses), it is not always easy to figure out how to mark or whether to mark some nominal constituents, particularly adjectives. A third issue has to do with hyphenation. Tokenization rules in the Penn Treebank provide that strings containing hyphens are considered single words, e.g., *the profit-maker*. Unfortunately, hyphens sometimes separate nouns from their arguments or separate

two arguments of a noun. This has led to some special annotation tags and annotation procedures to accommodate this. These issues are discussed in the following subsections.

4.1. Support and Related Phenomena

While most arguments of nouns occur locally inside the noun phrase, a substantial number of arguments violate this assumption. In *the judge made demands on his staff*, the judge is the ARG0 of *demands*. This is a SUPPORT verb construction in which the verb *make* links *demands* with an argument (ARG0) outside the NP it heads. The relation between these verbs and nominalizations is much the same as that between raising/control verbs and subordinate predicates. By far the most difficult part of annotating SUPPORT constructions is drawing the line between inference and SUPPORT. As with raising and control verbs, SUPPORT verbs may change certain aspects of meaning. Compare *John attacked*, *John planned to attack* and *John planned the attack*. SUPPORT constructions may reduce an agent-like role to the level of a participant, a helper or a planner (*John aided the attack*, *John planned the attack*). In order to insure consistent annotation, we found that a liberal interpretation of SUPPORT was necessary. We nevertheless require that SUPPORT be lexically licensed by intervening predicates, thus ruling out cases where coreference and inference imply relations. For example, in *Proponents say this view is correct*, the word *proponents* implies an argument relation with *view* by something akin to coreference (compare *his view* where *his* = ARG0). However, *say* does not license this connection. Therefore, this is not an instance of SUPPORT. We know that *say* does not license this connection because subjects other than proponents do not appear to be arguments of *view*, e.g., *John says this view is correct*. We know that *proponents* implies this relation because this inference can extend across sentences, e.g., *This view is very controversial. However, proponents say that soon everybody will agree with them*.

Other instances where arguments of a noun occur outside its NP include: (1) PP parenthetical constructions, e.g., *John left* is the ARG1 of *request* in *At Mary's request, John left* (2) victim noun constructions, e.g., *victim* is an ARG1 of *assassination* in *the victim of an assassination*; and (3) partitive constructions, e.g., *we* is the ARG0 of *debate* in *we had lots of internal debate*. In the last case, the SUPPORT verb *had* interacts with the partitive noun *lots* to link the ARG0 to *debate*. In general, support, partitive and victim noun constructions do interact to form what we call SUPPORT chains. Chains are successive head/complement links that connect the topmost support item (*had*) in the tree to the noun predicate (*debate*), where some argument (ARG0 = *We*) is also an argument of the topmost support item.

4.2. Prenominal Modifiers

Given a classification task, there is a strong tendency for humans to classify each interesting or important item, even if that item does not fit smoothly into one of the potential classes. However, we have asked our annotators to fight this instinct in the case of adjuncts. Specifically, we ask

our annotators to only mark adjuncts that fit into one of the adjunct classes allowed in PropBank. We needed to narrow our scope to make the task tractable and this was the minimum set that we needed to make sure that NomBank and PropBank were compatible.

Consider the phrase: *the legal battle*. Clearly, *legal* is an important item in the interpretation of this phrase. However, it does not fall neatly into one of the classes of allowable adjuncts: DIRectional, LOCative, manner (MNR), temporal (TMP), EXTent, purpose (PRP), CAUse, (sentence) ADverbial, and NEGative.¹ It therefore should be left unmarked. However, it is easy to see how an annotator might provide any of the following analyses: (1) “this is a battle about legal matters”, i.e., ARG1 = *legal*; (2) “they battled in a legal manner”, i.e., ARGM-MNR = *legal*; or (3) “they battled in the legal arena” which by a metaphoric stretch could entail that ARGM-LOC = *legal*. Unfortunately, none of these adjunct classes are really good fits. First of all, not one of these interpretations is clearly better than the other two. Secondly, clear cases of ARG1, ARGM-MNR and ARGM-LOC can co-occur with *legal*. Thus *legal* would be left out of a NomBank analysis of the following phrase *the petty Boston legal battle about parking*. The NomBank proposition would be: REL = battle, ARG1 = about parking, ARGM-MNR = petty, ARGM-LOC = Boston

It is precisely adjectives in prenominal modifier position which most commonly pose this sort of problem. To battle this issue, we have created a dictionary of adjectives (ADJADV)(Meyers et al., 2004) in which adjectives are marked with adverbial classes. While the initial dictionary was constructed by semi-automatic means, we expect to update it during annotation. We have observed that many adjectives are classified (or not) the same way each time. All annotation of adjectives is compared with the entry in ADJADV. Incompatible annotations are rechecked. In some cases, this results in an update of our dictionary and in others it results in a change of annotation.

4.3. Hyphenated Modifiers

As illustrated in figure 2, we have developed a way of dividing hyphenated words by numbering the segments before and after the hyphens. The first segment is labeled H0, the segment after the first hyphen is labeled H1 and the segment after the Nth hyphen is labeled HN. In this way, we are able to, for example, indicate that *self* is the ARG1 and *criticism* is the REL in the word *self-criticism* in spite of the limitations imposed by the tokenization used in the Penn Treebank. The only limitation that we have come up against so far is that noun compounding sometimes does not involve hyphenation. Thus instances of *bondholder* and *bond-holder* will be treated differently. In the former case, the same word would be assigned the ARG0, ARG1 and REL slots; in the latter case, *bond* would be assigned the ARG1 slot via the tag ARG1-H0 and *holder* would be assigned the ARG0 and REL slots via the tags ARG0-H1 and

¹Connoisseurs of PropBank will note some PropBank function tags are not listed here. There are two reasons for this: (1) ARGMDIS adjuncts do not occur with nouns; (2) Other tags occur primarily with argument (e.g. -PRD) rather than adjunct phrases.

1. *This is a time of self-criticism*
REL-H1 = self-criticism, ARG1-H0 = self-criticism
2. *a second daily Chicago-Paris flight*
REL = flight, ARG4-H0 = Chicago-Paris, ARG3-H1 = Chicago-Paris, ARGM-TMP = daily
3. *an active bell-ringer*
REL-H1 = bell-ringer, ARG0-H1 = bell-ringer, ARG1-H0 = bell-ringer, ARGM-MNR = active

Figure 2: Annotation of Hyphenated Segments

REL-H1.

5. Error Detection

We have implemented a simple error detection system. After an annotator has finished all the instances of a particular noun, they can send their annotation to an error detection program. The program detects various possible errors including:

1. The propositions are ill-formed. For example, an error is flagged if there are multiple instances of the same argument role (e.g., two ARG1s).
2. There is a conflict with one of our dictionaries. Items are flagged if prenominal modification is marked differently than would be predicted by the ADJADV dictionary or if the annotation is not compatible with the nominalization type. For example, an error would be detected if the noun was *destruction* and the noun itself was given an ARG0 tag (as if it were a subject nominalization like *destroyer*).
3. There is a probable error, e.g., first words are rarely correctly tagged as arguments unless the nominal predicate is close to the beginning of the sentence.

6. Summary and Future Work

We have provided an overview of the NomBank annotation task, highlighting the difficult parts of the task. Lexical entries are initialized based on PropBank lexical entries and a combination of other resources and procedures. Annotators take these lexical entries and fine tune them for each of the 200,000 instances of the 6500 argument-taking common nouns in the Penn Treebank. Annotation is basically the task of assigning role labels to arguments and adjuncts of these nouns. However, this task is complicated by various factors including: (a) some noun arguments are outside the NP, typically linked by SUPPORT verbs; and (b) prenominal adjectives can be difficult to fit into adjunct classes originally designed for adverbials. After annotation is complete, we run our automatic error detection procedures.

The NomBank project started small and has been slowly gaining momentum. This way we could work on specifications and procedures before doing most of the annotation. We are currently implementing a number of speed and quality measures to make it likely that we will complete the

project by the end of 2004. These include a preprocessor to produce automatic annotation (for checking by annotators); improved dictionaries for initial lexical entries and error checking; and more feedback mechanisms so that the annotators can help improve the preprocessing programs.

Acknowledgments

Nombank is supported under Grant N66001-001-1-8917 from the Space and Naval Warfare Systems Center San Diego. This paper does not necessarily reflect the position or the policy of the U.S. Government.

7. References

- Kingsbury, P. and M. Palmer, 2002. From treebank to propbank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*. Las Palmas, Spain.
- Kingsbury, P., M. Palmer, and Mitch Marcus, 2002. Adding semantic annotation to the penn treebank. In *Proceedings of the Human Language Technology Conference*. San Diego, California.
- Meyers, A., R. Reeves, Catherine Macleod, Rachel Szekeley, Veronkia Zielinska, Brian Young, and R. Grishman, 2004. The Cross-Breeding of Dictionaries. In *Proceedings of LREC-2004*. Lisbon, Portugal.
- Perlmutter, D. M. and P. M. Postal, 1984. The 1-Advancement Exclusiveness Law. In D. M. Perlmutter and C. G. Rosen (eds.), *Studies in Relational Grammar 2*. Chicago: The University of Chicago Press.
- Rosen, C. G., 1984. The Interface between Semantic Roles and Initial Grammatical Relations. In D. M. Perlmutter and C. G. Rosen (eds.), *Studies in Relational Grammar 2*. Chicago: The University of Chicago Press.
- University of Pennsylvania, 2002. Annotation guidelines for PropBank. <http://www.cis.upenn.edu/~ace/propbank-guidelines-feb02.pdf>.