

Merging PropBank, NomBank, TimeBank, Penn Discourse Treebank and Coreference

James Pustejovsky, Adam Meyers, Martha Palmer, Massimo Poesio

Abstract

Many recent annotation efforts for English have focused on pieces of the larger problem of semantic annotation, rather than initially producing a single unified representation. This paper discusses the issues involved in merging four of these efforts into a unified linguistic structure: PropBank, NomBank, the Discourse Treebank and Coreference Annotation undertaken at the University of Essex. We discuss resolving overlapping and conflicting annotation as well as how the various annotation schemes can reinforce each other to produce a representation that is greater than the sum of its parts.

1. Introduction

The creation of the Penn Treebank (Marcus et al, 1993) and the word sense-annotated SEMCOR (Fellbaum, 1997) have shown how even limited amounts of annotated data can result in major improvements in complex natural language understanding systems. These annotated corpora have led to high-level improvements for parsing and word sense disambiguation (WSD), on the same scale as previously occurred for Part of Speech tagging by the annotation of the Brown corpus and, more recently, the British National Corpus (BNC) (Burnard, 2000). However, the creation of semantically annotated corpora has lagged dramatically behind the creation of other linguistic resources: in part due to the perceived cost, in part due to an assumed lack of theoretical agreement on basic semantic judgments, in part, finally, due to the understandable unwillingness of research groups to get involved in such an undertaking. As a result, the need for such resources has become urgent.

Many recent annotation efforts for English have focused on pieces of the larger problem of semantic annotation, rather than producing a single unified representation like Head-driven Phrase Structure Grammar (Pollard and Sag 1994) or the Prague Dependency Tectogramatical Representation (Hajicova & Kucerova, 2002). PropBank (Palmer et al, 2005) annotates predicate argument structure anchored by verbs. NomBank (Meyers, et. al., 2004a) annotates predicate argument structure anchored by nouns. TimeBank (Pustejovsky et al, 2003)

annotates the temporal features of propositions and the temporal relations between propositions. The Penn Discourse Treebank (Miltsakaki et al 2004a/b) treats discourse connectives as predicates and the sentences being joined as arguments. Researchers at Essex were responsible for the coreference markup scheme developed in MATE (Poesio et al, 1999; Poesio, 2004a) and have annotated corpora using this scheme including a subset of the Penn Treebank (Poesio and Vieira, 1998), and the GNOME corpus (Poesio, 2004a). This paper discusses the issues involved in creating a *Unified Linguistic Annotation (ULA)* by merging annotation of examples using the schemata from these efforts. Crucially, all individual annotations can be kept separate in order to make it easy to produce alternative annotations of a specific type of semantic information without need to modify the annotation at the other levels. Embarking on separate annotation efforts has the advantage of allowing researchers to focus on the difficult issues in each area of semantic annotation and the disadvantage of inducing a certain amount of tunnel vision or task-centricity – annotators working on a narrow task tend to see all phenomena in light of the task they are working on, ignoring other factors. However, merging these annotation efforts allows these biases to be dealt with. The result, we believe, could be a more detailed semantic account than possible if the ULA had been the initial annotation effort rather than the result of merging.

There is a growing community consensus that general annotation, relying on linguistic cues, and in particular lexical cues, will produce an enduring resource that is useful, replicable and portable. We provide the beginnings of one such level derived from several distinct annotation efforts. This level could provide the foundation for a major advance in our ability to automatically extract salient relationships from text. This will in turn facilitate breakthroughs in message understanding, machine translation, fact retrieval, and information retrieval.

2. The Component Annotation Schemata

We describe below existing independent annotation efforts, each one of which is focused on a specific aspect of the semantic representation task: semantic role labeling,

coreference, discourse relations, temporal relations, etc. They have reached a level of maturity that warrants a concerted attempt to merge them into a single, unified representation, ULA. There are several technical and theoretical issues that will need to be resolved in order to bring these different layers together seamlessly. Most of these approaches have annotated the same type of data, Wall Street Journal text, so it is also important to demonstrate that the annotation can be extended to other genres such as spoken language. The demonstration of success for the extensions would be the training of accurate statistical semantic taggers.

PropBank: The Penn Proposition Bank focuses on the argument structure of verbs, and provides a corpus annotated with semantic roles, including participants traditionally viewed as arguments and adjuncts. An important goal is to provide consistent semantic role labels across different syntactic realizations of the same verb, as in *the window* in *[ARG0 John] broke [ARG1 the window]* and *[ARG1 The window] broke*. Arg0 and Arg1 are used rather than the more traditional Agent and Patient to keep the annotation as theory-neutral as possible, and to facilitate mapping to richer representations. The IM word Penn Treebank II Wall Street Journal corpus has been successfully annotated with semantic argument structures for verbs and is now available via the Penn Linguistic Data Consortium as PropBank I (Palmer, et. al., 2005). Coarse-grained sense tags, based on groupings of WordNet senses, are being added, as well as links from the argument labels in the Frames Files to FrameNet frame elements. There are close parallels to other semantic role labeling projects, such as FrameNet (Baker, et. al., 1998; Fillmore & Atkins, 1998; Fillmore & Baker, 2001), Salsa (Ellsworth, et.al, 2004), Prague Tectogramatics (Hajicova & Kucerova, 2002) and IAMTC, (Helmreich, et. al., 2004)

NomBank: The NYU NomBank project can be considered part of the larger PropBank effort and is designed to provide argument structure for instances of about 5000 common nouns in the Penn Treebank II corpus (Meyers, et. al., 2004a). PropBank argument types and related verb Frames Files are used to provide a commonality of annotation. This enables the development of systems that can recognize regularizations of lexically and syntactically related sentence structures, whether they occur as verb phrases or noun phrases. For example, given an IE system

tuned to a *hiring* scenario (MUC-6, 1995), NomBank and PropBank annotation facilitate generalization over patterns. PropBank and NomBank would both support a single IE pattern stating that the object (ARG1) of *appoint* is *John* and the subject (ARG0) is *IBM*, allowing a system to detect that *IBM hired John* from each of the following strings: *IBM appointed John*, *John was appointed by IBM*, *IBM's appointment of John*, *the appointment of John by IBM* and *John is the current IBM appointee*.

Coreference: Coreference involves the detection of subsequent mentions of invoked entities, as in *George Bush, ... he....* Researchers at Essex (UK) were responsible for the coreference markup scheme developed in MATE (Poesio et al, 1999; Poesio, 2004a), partially implemented in the annotation tool MMAX and now proposed as an ISO standard; and have been responsible for the creation of two small, but commonly used anaphorically annotated corpora – the Vieira / Poesio subset of the Penn Treebank (Poesio and Vieira, 1998), and the GNOME corpus (Poesio, 2004a). Parallel coreference annotation efforts funded by ACE have resulted in similar guidelines, exemplified by BBN's recent annotation of Named Entities, common nouns and pronouns. These two approaches provide a suitable springboard for an attempt at achieving a community consensus on coreference.

Discourse Treebank: The Penn Discourse Treebank (PDTB) (Miltsakaki et al 2004a/b) is based on the idea that discourse connectives are predicates with associated argument structure (for details see (Miltsakaki et al 2004a, Miltsakaki et al 2004b). The long-range goal is to develop a large scale and reliably annotated corpus that will encode coherence relations associated with discourse connectives, including their argument structure and anaphoric links, thus exposing a clearly defined level of discourse structure and supporting the extraction of a range of inferences associated with discourse connectives. This annotation references the Penn Treebank annotations as well as PropBank, and currently only considers Wall Street Journal text.

TimeBank: The Brandeis TimeBank corpus, funded by ARDA, focuses on the annotation of all major aspects in natural language text associated with temporal and event information (Day, et al, 2003, Pustejovsky, et al, 2004). Specifically, this involves three areas of the annotation: temporal expressions, event-denoting

expressions, and the links that express either an anchoring of an event to a time or an ordering of one event relative to another. Identifying events and their temporal anchorings is a critical aspect of reasoning, and without a robust ability to identify and extract events and their temporal anchoring from a text, the real aboutness of the article can be missed. The core of TimeBank is a set of 200 news reports documents, consisting of WSJ, DUC, and ACE articles, each annotated to TimeML 1.2 specification. It is currently being extended to AQUAINT articles. The corpus is available from the timeml.org website.

3. Unifying Linguistic Annotations

Since September, 2004, researchers representing several different sites and annotation projects have begun collaborating to produce a detailed semantic annotation of two difficult sentences. These researchers aim to produce a single unified representation with some consensus from the NLP community. This effort has given rise to both a listserv email list and this workshop: <http://nlp.cs.nyu.edu/meyers/pie-in-the-sky.html>, <http://nlp.cs.nyu.edu/meyers/frontiers/2005.html>. The merging operations discussed here would seem crucial to the furthering of this effort.

3.1 The Initial *Pie in the Sky* Example

The following two consecutive sentences have been annotated for Pie in the Sky.

Two Sentences From ACE Corpus File NBC20001019.1830.0181

- *but Yemen's president says the FBI has told him the explosive material could only have come from the U.S., Israel or two Arab countries.*
- *and to a former federal bomb investigator, that description suggests a powerful military-style plastic explosive c-4 that can be cut or molded into different shapes.*

Although the full *Pie-in-the-Sky* analysis includes information from many different annotation projects, the Dependency Structure in Figure 1 includes only those components that relate to PropBank, NomBank, Discourse annotation, coreference and TimeBank. Several parts of this representation require further explanation. Most of these are signified by the special arcs, arc labels, and nodes. Dashed lines represent transparent arcs, such as the transparent

dependency between the argument (ARG1) of modal *can* and the *or*. *Or* is transparent in that it allows this dependency to pass through it to *cut* and *mold*. There are two small arc loops -- *investigator* is its own ARG0 and *description* is its own ARG1. *Investigator* is a relational noun in NomBank. There is assumed to be an underlying relation between the *Investigator* (ARG0), the beneficiary or employer (the ARG2) and the item investigated (ARG1). Similarly, *description* acts as its own ARG1 (the thing described). There are four special coreference arc labels: ARG0-CF, ARG-ANAPH, EVENT-ANAPH and ARG1-SBJ-CF. At the target of these arcs are pointers referring to phrases from the previous sentence or previous discourse. The first three of these labels are on arcs with the noun *description* as their source. The ARG0-CF label indicates that the phrase *Yemen's president* (**1**) is the ARG0, the one who is doing the describing. The EVENT-ANAPH label points to a previous mention of the describing event, namely the clause: *The FBI told him the explosive material...* (**3**). However, as noted above, the NP headed by *description* represents the thing described in addition to the action. The ARG-ANAPH label points to the thing that the FBI told him *the explosive material can only come from ...* (**2**). The ARG1-SBJ-CF label links the NP from the discourse *what the bomb was made from* as the subject with the NP headed by *explosive* as its predicate, much the same as it would in a copular construction such as: *What the bomb was made from is the explosive C-4*. Similarly, the arc ARG1-APP marks *C-4 as an* appositive, also predicated to the NP headed by *explosive*. Finally, the thick arcs labeled SLINK-MOD represent TimeML SLINK relations between eventuality variables, i.e., the *cut* and *molded* events are modally subordinate to the *suggests* proposition. The merged representation aims to be compatible with the projects from which it derives, each of which analyzes a different aspect of linguistic analysis. Indeed most of the dependency labels are based on the annotation schemes of those projects.

We have also provided the individual PropBank, NomBank and TimeBank annotations below in textual form, in order to highlight potential points of interaction.

PropBank: and [_{ARG2} to a former federal bomb investigator], [_{ARG0} that description]
 [_{Rel_suggest.01} suggests] [_{ARG1} [_{ARG1} a powerful military-style plastic explosive c-4] that

[ArgM-MOD can] be [Rel_cut.01 cut] or [Rel_mold.01 molded] [ArgM-RESULT into different shapes]].

NomBank: and to a former [Arg2 federal] [Arg1 bomb] [Rel investigator], that description suggests a powerful [Arg2 military] - [Rel style] plastic [Arg1 explosive] c-4 that can be cut or molded into different shapes.

TimeML: and to a former federal bomb investigator, that description [Event = ei1 suggests] a powerful military-style plastic explosive c-4 that can be [Event = ei2 modal='can' cut] or [Event = ei3 modal='can' molded] into different shapes. <SLINK eventInstanceID = ei1 subordinatedEventID = ei2 relType = 'Modal' /> <SLINK eventInstanceID = ei1 subordinatedEventID = ei3 relType = 'Modal' />

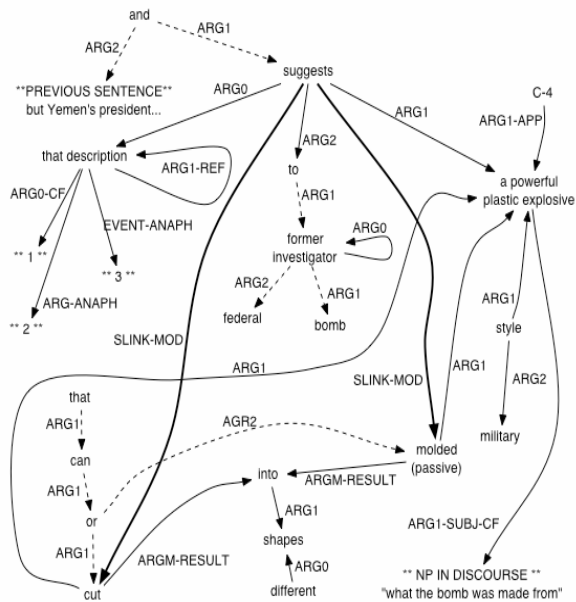


Figure 1. Dependency Analysis of Sentence 2

Note that the subordinating Events indicated by the TimeML SLINKS refer to the predicate argument structures labeled by PropBank, and that the ArgM-MODal also labeled by PropBank contains modality information also crucial to the SLINKS. While the grammatical modal on *cut* and *mold* is captured as an attribute value on the event tag, the governing event predicate *suggest* introduces a modal subordination to its internal argument, along with its relative clause. While this markup is possible in TimeML, it is difficult to standardize (or automate, algorithmically) since arguments are not marked up unless they are event denoting.

3.2 A More Complex Example

To better illustrate the interaction between annotation levels, and the importance of merging information resident in one level but not necessarily in another, consider the sentence below which has more complex temporal properties than the *Pie-in-the-Sky* sentences and its dependency analysis (Figure 2).

According to reports, sea trials for a patrol boat developed by Kazakhstan are being conducted and the formal launch is planned for the beginning of April this year.

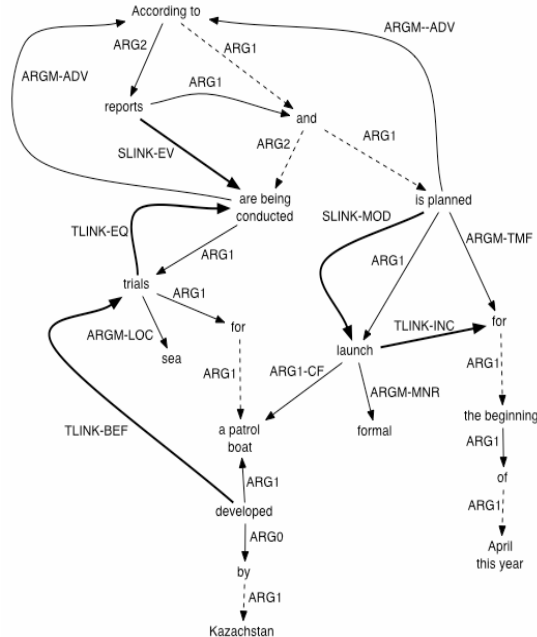


Figure 2. Dependency Analysis of a Sentence with Interesting Temporal Properties

The graph above incorporates these distinct annotations into a merged representation, much like the previous analysis. This sentence has more TimeML annotation than the previous sentence. Note the loops of arcs which show that *According to* plays two roles in the sentence: (1) it heads a constituent that is the ARGM-ADV of the verbs *conducted* and *planned*; (2) it indicates that the information in this entire sentence is attributed to *the reports*. This loop is problematic in some sense because the adverbial appears to modify a constituent that includes itself. In actuality, however, one would expect that the ARGM-ADV role modifies the sentence minus the adverbial, the constituent that you would get if you ignore the transparent arc from ARGM-

ADV to the rest of the sentence. Alternatively, a merging decision may elect to delete the ARGM-ADV arcs, once the more specific predicate argument structure of the sentence adverbial annotation is available.

The **PropBank** annotation for this sentence would label arguments for *develop*, *conduct* and *plan*, as given below.

[ArgM-ADV According to reports], [Arg1 sea trials for [Arg1 a patrol boat] [Rel_develop.02 developed] [Arg0 by Kazakhstan]] are being [Rel_conduct.01 conducted] and [Arg1 the formal launch] is [Rel_plan.01 planned] [ArgM-TMP for the beginning of April this year].

NomBank would add arguments for *report*, *trial*, *launch* and *beginning* as follows:

According to [Rel_report.01 reports], [Arg1 [ArgM-LOC sea [Rel_trial.01 trials] [Arg1 for [Arg1-CF_launch.01 a patrol boat] developed by Kazakhstan] are being conducted and the [ArgM-MNR formal] [Rel_launch.01 launch] is planned for the [[REL_beginning.01 beginning] [ARG1 of April this year]].

TimeML, however, focuses on the anchoring of events to explicit temporal expressions (or document creation dates) through TLINKs, as well as subordinating relations, such as those introduced by modals, intensional predicates, and other event-selecting predicates, through SLINKs. For discussion, only part of the complete annotation is shown below.

According to [Event = ei1 reports], sea [Event = ei3 trials] for a boat [Event = ei4 developed] by Kazakhstan are being [Event = ei5 conducted] and the formal [Event = ei6 launch] is [Event = ei7 planned] for the [Time3 = t1 beginning of April] [Time3 = t2 this year].
 <SLINK eventID="ei1" subordinatedEvent="ei5, ei7" relType=EVIDENTIAL/>
 <TLINK eventID="ei4" relatedToEvent="ei3" relType=BEFORE/>
 <TLINK eventID="ei6" relatedToTime="t1" relType=IS_INCLUDED />
 <SLINK eventID="ei7" subordinatedEvent="ei6" relType="MODAL"/>
 <TLINK eventID="ei5" relatedToEvent="ei3" relType=IDENTITY/>

Predicates such as *plan* and nominals such as *report* are lexically encoded to introduce SLINKs with a specific semantic relation, in this

case, a “MODAL” relType. This effectively introduces an intensional context over the subordinated events.

These examples illustrate the type of semantic representation we are trying to achieve. It is clear that our various layers already capture many of the intended relationships, but they do not do so in a unified, coherent fashion. Our goal is to develop both a framework and a process for annotation that allows the individual pieces to be automatically assembled into a coherent whole.

4.0 Merging Annotations

4.1 First Order Merging of Annotation

We begin by discussing issues that arise in defining a single format for a merged representation of PropBank, NomBank and Coreference, the core predicate argument structures and referents for the arguments. One possible representation format would be to convert each annotation into features and values to be added to a larger feature structure.¹ The resulting feature structure would combine stand alone and offset annotation – it would include actual words and features from the text as well as special features that point to the actual text (character offsets) and, perhaps, syntactic trees (offsets along the lines of PropBank/NomBank). Alternative global annotation schemes include annotation graphs (Cieri & Bird, 2001), and MATE (Carletta, et. al., 1999). There are many areas in which the boundaries between these annotations have not been clearly defined, such as the treatment of support constructions and light verbs, as discussed below. Determining the most suitable format for the merged representation should be a top priority.

4.2 Resolving Annotation Overlap

There are many possible interactions between different types of annotation: aspectual verbs have argument labels in PropBank, but are also important roles for temporal relations. Support

¹ The Feature Structure has many advantages as a target representation including: (1) it is easy to add lots of detailed features; and (2) the mathematical properties of Feature Structures are well understood, i.e., there are well-defined rule-writing languages, subsumption and unification relations, etc. defined for Feature Structures (Carpenter, 1992) The downside is that a very informative Feature Structure is difficult for a human to read.

constructions also have argument labels, and the question arises as to whether these should be associated with the support verb or the predicative nominal. Given the sentence *They gave the chefs a standing ovation*, a PropBank component will assign role labels to arguments of *give*; a NomBank component will assign argument structure to *ovation* that labels the same participants. If the representations are equivalent, the question arises as to which of them (or both) should be included in the merged representation. The following graph (Figure 3) is a combined PropBank and NomBank analysis of this sentence. "They" is the ARG0 of both "give" and "ovation"; "the chefs" is the ARG2 of "give", but the "ARG1" of ovation; "ovation" is the ARG1 of "give" and "give" is a support verb for "ovation". For this case, a reasonable choice might be to preserve the argument structure from both NomBank and PropBank, and to do the same for other predicative nominals that have *give* (or *receive*, *obtain*, *request*...) as a support verb, e.g., (*give a kiss/hug/squeeze*, *give a lecture/speech*, *give a promotion*, etc.). For other support constructions, such as *take a walk*, *have a headache* and *make a mistake*, the noun is really the main predicate and it is questionable whether the verbal argument structure carries

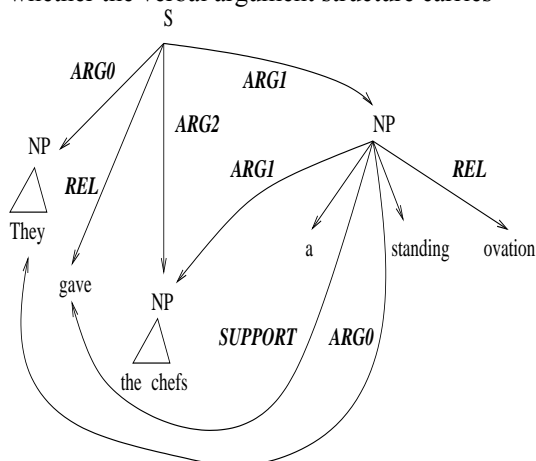


Figure 3. Merged PropBank/NomBank representation of *They gave the chefs a standing ovation*. much information, e.g., there are no selection restrictions between light verbs and their subject (ARG0) -- these are inherited from the noun. Thus *make a mistake* selects a different type of subject than *make a gain*, e.g., people and organizations make mistakes, but stock prices make gains. For these constructions, the merged representation might not need to include the (ARG0) relation between the subject of the sentence and *make*, and future propbanking efforts might do well to ignore the shared

arguments of such instances and leave them for NomBank. However, the merged representation would inherit PropBank's annotation of some other light verb features including: negation, e.g., *They did not take a walk*; modality, e.g., *They might take a walk*; and sentence adverbials, e.g., *They probably will take a walk*.

4.3 Resolving Annotation Conflicts

Interactions between linguistic phenomena can aid in quality control, and conflicts found during the deliberate merging of different annotations provides an opportunity to correct and fine-tune the original layers. For example, predicate argument structure (PropBank and NomBank) annotation sometimes assumes different constituent structure than the Penn Treebank. We have noticed some tendencies that help resolve these conflicts, e.g., prenominal noun constituents as in *Indianapolis 500*, which forms a single argument in NomBank, is correctly predicted to be a constituent, even though the Penn Treebank II assumes a flatter structure.

Similarly, idioms and multiword expressions often cause problems for both PropBank and NomBank. PropBank annotators tend to view argument structure in terms of verbs and NomBank annotators tend to view argument structure in terms of nouns. Thus many examples that, perhaps, should be viewed as idioms are viewed as special senses of either verbs or nouns. Having idioms detected and marked before propbanking and nombanking could greatly improve efficiency.

Annotation accuracy is often evaluated in terms of inter-annotation consistency. Task definitions may need to err on the side of being more inclusive in order to simplify the annotators task. For example, the NomBank project assumes the following definition of a support verb (Meyers, et.al., 2004b): "... a verb which takes at least two arguments NP_1 and XP_2 such that XP_2 is an argument of the head of NP_1 . For example, in *John took a walk*, a support verb (*took*) shares one of its arguments (*John*) with the head of its other argument (*walk*).” The easiest way to apply this definition is without exception, so it will include idiomatic expressions such as *keep tabs on*, *take place*, *pull strings*. Indeed, the dividing line between support constructions and idioms is difficult to draw (Meyers 2004b). PropBank annotators are also quite comfortable with associating general meanings to the main verbs of idiomatic expressions and labeling their

argument roles, as in cases like *bring home the bacon* and *mince words with*. Since idioms often have interpretations that are metaphorical extensions of their literal meaning, this is not necessarily incorrect. It may be helpful to have the literal dependencies and the idiomatic reading both represented. The fact that both types of meaning are available is evidenced by jokes, irony, and puns.

With respect to idioms and light verbs, TimeML can be viewed as a mediator between PropBank and NomBank. In TimeML, light verbs and the nominalizations accompanying them are marked with two separate EVENT tags. This guarantees an annotation independent of textual linearity and therefore ensures a parallel treatment for different textual configurations. In (a) the light verb construction "make an allusion" is constituted of a verb and an NP headed by an event-denoting noun, whereas in (b) the nominal precedes a VP, which in addition contains a second N:

(a) *Max [made an allusion] to the crime.*

(b) *Several anti-war [demonstrations have taken place] around the globe.*

Both verbal and nominal heads are tagged because they both contribute relevant information to characterizing the nature of the event. The nominal element plays a role in the more semantically based task of event classification. On the other hand, the information in the verbal component is important at two different levels: it provides the grammatical features typically associated with verbal morphology, such as tense and aspect, and at the same time it may help in disambiguating cases like *take/give a class*, *make/take a phone call*. The two tagged events are marked as identical by a TLINK introduced for that purpose. The TimeML annotation for the example in (a) is provided below.

Max [_{Event = ei1} made] an [_{Event = ei2} allusion] to the crime.

<TLINK eventID="ei1"relatedToEvent="ei2" relType=IDENTITY>

Some cases of support in NomBank could also be annotated as "bridging" anaphora. Consider the sentence: *The pieces make up the whole*. It is unclear whether *make up* is a support verb linking *whole* as the ARG1 of *pieces* or if *pieces* is linked to *whole* by bridging anaphora. There are also clearer cases. In *Nastase, a rival player defeated Jimmy Connors in the third round*, the word *rival* and *Jimmy Connors* are clearly linked by bridging. However, a wayward

NomBank annotator might construct a support chain (*player + defeated*) to link *rival* with its ARG1 *Jimmy Connors*. In such a case, a merging of annotation could reveal annotation errors. In contrast, a NomBank annotator would be correct in linking *John* as an argument of *walk* in *John took a series of walks* (the support chain *took + series* consists of a support verb and a transparent noun), but this may not be obvious to the non-NomBanker. Thus the merging of annotation may result in the more consistent specifications for all.

In our view, this process of annotating all layers of information and then merging them in a supervised manner, taking note of the conflicts, is a necessary prerequisite to defining more clearly the boundaries between the different types of annotation and determining how they should fit together. Other areas of annotation interaction include: (1) NomBank and Coreference, e.g. deriving that *John teaches Mary* from *John is Mary's teacher* involves: (a) recognizing that *teacher* is an argument nominalization such that the teacher is the ARG0 of *teach* (the one who teaches); and (b) marking *John* and *teacher* as being linked by predication (in this case, an instance of type coreference); and (2) Time and Modality - when a fact used to be true, there are two time components: one in which the fact is true and one in which it is false. Clearly more areas of interaction will emerge as more annotation becomes available and as the merging of annotation proceeds.

5. Summary

We proposed a way of taking advantage of the current practice of separating aspects of semantic analysis of text into small manageable pieces. We propose merging these pieces, initially in a careful, supervised way, and hypothesize that the result could be a more detailed semantic analysis than was previously available. This paper discusses some of the reasons that the merging process should be supervised. We primarily gave examples involving the interaction of PropBank, NomBank and TimeML. However, as the merging process continues, we anticipate other conflicts that will require resolution.

References

- C. F. Baker, F. Collin, C. J. Fillmore, and J. B. Lowe (1998), The Berkeley FrameNet project. In *Proc. of COLING/ACL-98*, 86--90

- O. Babko-Malaya, M. Palmer, X. Nianwen, S. Kulick, A. Joshi (2004), Propbank II, Delving Deeper, In *Proc. of HLT-NAACL Workshop: Frontiers in Corpus Annotation*.
- R. Carpenter (1992), *The Logic of Typed Feature Structures*. Cambridge Univ. Press.
- J. Carletta and A. Isard (1999), The MATE Annotation Workbench: User Requirements. In *Proc. of the ACL Workshop: Towards Standards and Tools for Discourse Tagging*. Univ. of Maryland, 11-17
- C. Cieri and S. Bird (2001), Annotation Graphs and Servers and Multi-Modal Resources: Infrastructure for Interdisciplinary Education, Research and Development *Proc. of the ACL Workshop on Sharing Tools and Resources for Research and Education*, 23-30
- D. Day, L. Ferro, R. Gaizauskas, P. Hanks, M. Lazo, J. Pustejovsky, R. Sauri, A. See, A. Setzer, and B. Sundheim (2003), The TIMEBANK Corpus. *Corpus Linguistics*.
- M. Ellsworth, K. Erk, P. Kingsbury and S. Pado (2004), PropBank, SALSA, and FrameNet: How Design Determines Product, in *Proc. of LREC 2004 Workshop: Building Lexical Resources from Semantically Annotated Corpora*.
- C. Fellbaum (1997), *WordNet: An Electronic Lexical Database*, MIT Press..
- C. J. Fillmore and B. T. S. Atkins (1998), FrameNet and lexicographic relevance. In the *Proc. of the First International Conference on Language Resources and Evaluation*.
- C. J. Fillmore and C. F. Baker (2001), Frame semantics for text understanding. In *Proc. of NAACL WordNet and Other Lexical Resources Workshop*.
- E. Hajivcova and I. Kuvcerov'a (2002). Argument/Valency Structure in PropBank, LCS Database and Prague Dependency Treebank: A Comparative Pilot Study. In the *Proc. of the Third International Conference on Language Resources and Evaluation (LREC 2002)*, 846--851.
- S. Helmreich, D. Farwell, B. Dorr, N. Habash, L. Levin, T. Mitamura, F. Reeder, K. Miller, E. Hovy, O. Rambow and A. Siddharthan,(2004), Interlingual Annotation of Multilingual Text Corpora, *Proc. of the HLT-EACL Workshop on Frontiers in Corpus Annotation*.
- A. Meyers, R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, and R. Grishman (2004a), The NomBank Project: An Interim Report, *Proc. of HLT-EACL Workshop: Frontiers in Corpus Annotation*.
- A. Meyers, R. Reeves, and C. Macleod (2004b), NP-External Arguments: A Study of Argument Sharing in English. In *The ACL 2004 Workshop on Multiword Expressions: Integrating Processing*.
- E. Miltsakaki, R. Prasad, A. Joshi and B. Webber. (2004a), The Penn Discourse Treebank. In *Proc. 4th International Conference on Language Resources and Evaluation (LREC 2004)*.
- E. Miltsakaki, R. Prasad, A. Joshi and B. Webber (2004b), Annotation of Discourse Connectives and their Arguments, in *Proc. of HLT-NAACL Workshop: Frontiers in Corpus Annotation*
- M. Marcus, B. Santorini, and M. Marcinkiewicz (1993), Building a large annotated corpus of english: The penn treebank. *Computational Linguistics*, 19:313--330.
- M. Palmer, D. Gildea, P. Kingsbury (2005), The Proposition Bank: A Corpus Annotated with Semantic Roles, *Computational Linguistics Journal*, 31:1.
- M. Poesio (2004a), The MATE/GNOME Scheme for Anaphoric Annotation, Revisited, *Proc. of SIGDIAL*
- M. Poesio (2004b), Discourse Annotation and Semantic Annotation in the GNOME Corpus, *Proc. of ACL Workshop on Discourse Annotation*.
- M. Poesio and M. Alexandrov-Kabadjov (2004), A general-purpose, off-the-shelf system for anaphora resol.. *Proc. of LREC*.
- M. Poesio, F. Bruneseaux, and L. Romary (1999), The MATE meta-scheme for coreference in dialogues in multiple language, *Proc. of the ACL Workshop on Standards for Discourse Tagging*.
- M. Poesio and R. Vieira (1998), A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2).
- C. Pollard and I. A. Sag (1994), *Head-driven phrase structure grammar*. Univ. of Chicago Press.
- J. Pustejovsky, R. Sauri, J. Castaño, D. R. Radev, R. Gaizauskas, A. Setzer, B. Sundheim and G. Katz (2004), Representing Temporal and Event Knowledge for QA Systems. In Mark T. Maybury (ed.), *New Directions in Question Answering*, MIT Press.
- J. Pustejovsky, B. Ingria, R. Sauri, J. Castaño, J. Littman, R. Gaizauskas, A. Setzer, G. Katz, and I. Mani (2003), The Specification Language TimeML. In I. Mani, J. Pustejovsky, and R. Gaizauskas, editors, *The Language of Time: A Reader*. Oxford Univ. Press.