

Bridging Sentential and Discourse-level Semantics through Clausal Adjuncts

Rashmi Prasad¹, Bonnie Webber², Alan Lee³, Sameer Pradhan⁴, Aravind Joshi³

¹Department of Health Informatics and Administration, University of Wisconsin-Milwaukee
prasad@uwm.edu

²School of Informatics, University of Edinburgh
Bonnie.Webber@ed.ac.uk

³Institute for Research in Cognitive Science, University of Pennsylvania
{aleewk, joshi}@seas.upenn.edu

⁴Boulder Language Technologies
pradhan@bltek.com

Abstract

It is in PropBank’s ARGM annotation of clausal adjuncts that sentential semantics meets discourse relation annotation in the Penn Discourse TreeBank. This paper discusses complementarities between the two annotation systems: How PropBank ARGM annotation can be used to seed annotation of additional discourse relations in the PDTB, and how PDTB annotation can be used to refine or enrich PropBank ARGM annotation.

1 Introduction

Discourse relations between *abstract objects*, such as facts, events, propositions, etc. (Asher, 1993), can hold either across sentences (i.e., *inter-sententially*), or within a single sentence (i.e., *intra-sententially*), as in Ex. 1–4. (Italics and bold-face highlight the two related abstract objects, respectively, and relation signals, when present, are underlined.)

- (1) *The federal government suspended sales of U.S. savings bonds* because **Congress hasn’t lifted the ceiling on government debt.**
- (2) *The House has voted to raise the ceiling to \$3.1 trillion,* but **the Senate isn’t expected to act until next week at the earliest.**
- (3) *Now, we regard this as a largely phony issue,* **but the “long term” is nonetheless a big salon topic all around the Beltway.**
- (4) *The U.S. wants the removal of . . . barriers to investment;* **Japan denies there are real barriers.**

Researchers working on discourse parsing have commented that intra-sentential (intra-S) discourse relations are, in general, easier to recognize than ones whose arguments are found in separate sentences (Joty et al., 2012; Lin et al.,

2012; Feng, 2014). They are also quite useful in Language Technology applications that exploit sentence-level relations. Thus, there is particular value in improving the quality of recognizers capable of determining what, if any, discourse relations hold between intra-S units.

Taking abstract objects to be expressed (arguably) typically as clauses headed by verbs or other predicates, the Penn Discourse Treebank (PDTB) (Prasad et al., 2008) includes annotations of intra-S discourse relations but, as noted by Prasad et al. (2014), they are significantly under-annotated in the corpus. At the same time, Prasad et al. (2014) point to possible overlaps between intra-S discourse relations in the PDTB and a subset of verb-argument annotations in PropBank (Palmer et al., 2005). The PropBank annotations of particular interest here are those in which the arguments are clausal adjuncts, labeled ARGM, and further assigned a semantic role. For example, the PropBank annotation of the verb *suspend* in Ex. 1 is shown in (5), with the adjunct clause annotated as ARGM and assigned the role CAU (causal). The PDTB annotation for the same example, shown in (6), marks *because* as the connective, ‘Contingency.Cause.Reason’ as the sense, the adjunct clause as Arg2 (defined as the argument attached to the connective), and the matrix clause as Arg1 (defined as the non-Arg2 argument).

- (5) **PropBank:** Verb = *suspend*
Arg0 = *The federal government*
Arg1 = *sales of U.S. savings bonds*
ARGM-CAU = *because Congress hasn’t lifted the ceiling on government debt*
- (6) **PDTB:** Connective = *because*
Arg1 = *The federal government suspended sales of U.S. savings bonds*
Arg2 = *Congress hasn’t lifted the ceiling on government debt*
Sense = Contingency.Cause.Reason

	TEMPORAL	CONTINGENCY	COMPARISON	EXPANSION	TOTAL
ARGM-ADV (2235)	222	1067	907	157	2353
ARGM-CAU (657)	14	650	0	0	664
ARGM-TMP (2503)	2258	523	73	23	2877
ARGM-PNC (66)	0	65	1	0	66
ARGM-MNR (13)	0	5	1	7	13
TOTAL (5475)	2494	2310	982	187	5973

Table 1: Correspondences between PropBank ARGM- roles and PDTB senses

Given possible overlaps between the PDTB and PropBank, this paper addresses the following questions: (1) To what extent can the PropBank clausal ARGM annotations be taken as conveying information relevant for intra-S discourse relations (Section 2), and can they be useful for increasing the number of intra-S discourse relations annotated in the PDTB (Section 3)?; and (2) Can PDTB annotations be useful for enriching PropBank in any way (Section 4)? Section 5 concludes the paper.

2 PropBank ARGM Roles and Discourse Relations

The PDTB 2.0 (the current version of the corpus) lacks extensive annotation of intra-S relations. Annotations of intra-S relations are provided primarily for relations that are signaled by explicit connectives (subordinating conjunctions (Ex. 1), coordinating conjunctions (Ex. 2), and adverbials (Ex. 3)). The only implicit relations currently annotated are those between clauses connected by a punctuation such as the semi-colon or colon (Ex. 4). Among the relations that are missing are implicit relations linking adjunct clauses that are not subordinated by any explicit form, as in Ex. 7, and adjunct clauses introduced by prepositional subordinators like *by*, *for*, *with*, *without*, *to*, as in Ex. 8-9.

- (7) *Second, they channel monthly mortgage payments into semiannual payments, **reducing the administrative burden on investors.***
- (8) ***To avoid this deficit,** Mr. Lawson inflated the pound in order to prevent its rise.*
- (9) Critics say *South Carolina is paying a price by **stress-ing improved test scores so much.***

These types of unannotated relations involving adjunct clauses in the PDTB have, on the other hand, been annotated in PropBank, as described in Section 1. Hence, a natural question to ask is whether the semantic roles of such adjunct clauses in PropBank can be used to fill in

the gap when annotations of intra-S discourse relations are not present in the PDTB for these clauses, thus avoiding duplicate annotation efforts. To explore this possibility, we considered a parallel case: the annotation of adjunct clauses introduced by *explicit* connectives in the PDTB, such as those in Ex. 1, which have been annotated in the corpus. We investigated the extent of the overlap between PropBank and PDTB annotations in such cases. Using the underlying syntactic annotations of the Penn Treebank (PTB) (Marcus et al., 1993), 11534 clausal adjuncts with either of the following six roles were extracted from PropBank:¹ ARGM-ADV (adverbial), ARGM-CAU (causal), ARGM-MNR (manner), ARGM-PNC (purpose), ARGM-PRD (secondary predication), and ARGM-TMP (temporal). Other roles (ARGM-MOD/DIR/EXT/DIS/LOC) were excluded because we did not see them as representing discourse relations. The 11534 ARGMs were then aligned with the PDTB and 48% (5475) were found to contain an explicit subordinating form annotated as a discourse connective in the PDTB. Except for ARGM-PRD, all the discourse-relevant ARGM roles were observed in this set (Table 1). We then looked at the correspondence between the roles assigned to these ARGMs in PropBank and the senses annotated for the connectives in the PDTB. Because the PDTB sense classification is hierarchical and contains many fine-grained relations, we simplified the comparison by considering only the four top level classes of the PDTB – Temporal (TEMP), Contingency (CONT), Comparison (COMP), and Expansion (EXP). The correspondences are shown in Table 1. The numbers in parentheses in the first column represent the total number of ARGM instances annotated with the role shown.

There are several observations to make from Table 1. First, based on the definitions of the ARGM roles in PropBank and those of the senses

¹For this part of the work, we used PropBank-I.

in the PDTB, we would expect PropBank ARGM-CAU to align with PDTB Contingency, and PropBank ARGM-TMP to align with PDTB Temporal. This is largely borne out for ARGM-CAU, with 99% of the instances labeled as Contingency in PDTB. ARGM-TMP, while showing a greater association with non-Temporal PDTB senses, nevertheless corresponds to PDTB Temporal 90% of the time. While the current PDTB sense classification does not include Purpose and Manner relations, PDTB guidelines followed a convention to label Purpose connectives such as *so that* as ‘Contingency.Cause.Result’, so we would expect to see ARGM-PNC aligned with Contingency. This is largely borne out as well, with only one instance labeled otherwise. Manner relations, on the other hand, are not addressed in the existing PDTB guidelines at all, which may explain the variable sense annotation of the ARGM-MNR cases in the PDTB. In contrast to the previous four ARGM roles, however, the ARGM-ADV role, which constitutes 41% of all the roles, fails to provide semantically meaningful alignment with the PDTB senses. According to the PropBank guidelines (Bonial et al., 2010), ARGM-ADV is used for syntactic elements which clearly modify the event structure of the verb in question, but which cannot be classified as any of the other roles. As the table shows, this role is ambiguous among all four sense classes in the PDTB, although we see a much higher proportion of Contingency and Comparison than Temporal and Expansion.

The second observation from Table 1 is that the total number of PDTB senses associated with an ARGM role (last column) is in some cases more than the total number of instances for that ARGM role (first column). Altogether, the table shows a total of 5973 PDTB senses associated with a total of 5475 PropBank ARGM roles. This is due to the fact that the PDTB allows multiple relations to be inferred between abstract objects, whereas PropBank only allows a single role to be assigned to any given ARGM. Notably, however, multiple PDTB senses do not appear at all for ARGM-PNC and ARGM-MNR, and appear in only seven instances for ARGM-CAU. ARGM-TMP had the most instances (374) with multiple senses, while ARGM-ADV had 118.

What these observations suggest is that while the correspondence between PropBank clausal ARGM roles and PDTB senses is not exact, they

can still be leveraged to some extent. On the one hand, relations with the ARGM-ADV role would need to be manually annotated for the PDTB sense. But on the other hand, the high degree of correspondences seen for other roles suggest the possibility of their straightforward mapping from PropBank to PDTB. The possibility of multiple senses in the PDTB would require further annotation, but this may be needed for only the Temporal sense. And here too, there may be less effort required since the annotator would not need to reason about the Temporal sense but only consider the possibility of inferring a second sense.

We must note that the semantics of the relation is not the only kind of correspondence to consider between PropBank and the PDTB. Mismatches in alignment can also arise between PDTB arguments and PropBanks semantic role structure, in large part because the PropBank annotation is tied directly to the syntactic trees in the PTB. Ex. 10 shows a sentence containing a *when*-clause annotated as ARGM-TMP in PropBank and with a Temporal sense in the PDTB. But the relation between the *when*-clause and its other argument is different between the two corpora. In the PDTB, where annotation is done over the raw text spans, the Arg1 of the connective excludes *he says*, and the temporal relation is annotated between the *winning* and *awarding* events. In contrast, in PropBank, the *when*-clause is taken to modify the verb *say*. Hence, we cannot use PropBank’s projected clause to automatically annotate the Arg1 of the corresponding PDTB relation as this would be inconsistent with the PDTB guidelines.

- (10) **When Mr. Green won a \$240,000 verdict in a land condemnation case against the state in June 1983,** he says Judge O’Kicki unexpectedly *awarded* him an additional \$100,000.

Given the difference in annotation practice, the extent of such mismatches between the PDTB and PropBank is expected to be the same as that between the PDTB and the PTB (Dinesh et al., 2005). Nevertheless, since the majority of semantic conflicts are due to attribution verbs, one can reduce the annotation effort by automatically highlighting instances with attribution verbs, in contexts that may lead to inconsistent semantics.

3 Using PropBank to Seed New PDTB Annotations

Despite the partial correspondence described above, PropBank is richly annotated with clausal

adjunct tokens that in PDTB would be Arg2 of a discourse relation. Therefore, in preparing the next version of the PDTB, we have used these PropBank tokens to seed the corpus with new annotations of intra-S relations. Our search for new tokens uses the latest version of the PropBank layer of the OntoNotes v5.0. corpus, since this PropBank version contains additional tokens for copular verbs and their argument structure, as well as modifications to tokens from PropBank-I. However, since only about 75% of the PTB is included in OntoNotes, the remaining 25% was taken from PropBank-I. Clausal adjuncts identified from these two versions of PropBank were then divided (using the syntactic trees in the PTB) into those that had an explicit subordinating form and those that did not. The former set was filtered to retain tokens not already annotated in PDTB 2.0. Most of these contain subordinators as the connecting elements (as in Ex. 8-9) that we will consider as signals of explicit intra-sentential discourse relations. The latter set yielded free adjuncts (Ex. 7), both present participles and past participles.

Altogether, over 5000 tokens signaling potential intra-sentential relations have been seeded in this way for further annotation, semi-automatically or manually. In the set comprising free adjuncts, approximately 75% were found to be assigned the ARGM-ADV role in Propbank while approximately 19% are assigned the similarly underspecified ARGM-PRD role. This leaves only 6% assigned to the Purpose, Causal, Manner and Temporal roles which, as discussed above, have strong correspondences with PDTB senses. Because of this, the annotation of free adjuncts is being done manually. The annotation guidelines extend directly from those used in annotating PDTB 2.0 while some new senses and refinements have also been introduced, including the addition of Purpose and Manner senses.

Unlike the free adjuncts, we see less underspecification with adjuncts that are subordinated by some explicit form. Adjuncts with the Purpose role are the most frequent, at 50%, followed by ARGM-MNR (26%). ARGM-ADV continues to appear in this set, although less frequently (18%). All other roles account for the remaining 6% of the tokens. We expect that these tokens can be annotated semi-automatically. As noted earlier, because the PDTB senses are in some cases more refined than Propbank roles (for example the Tem-

ARGM-ADV	58.11
ARGM-CAU	60.67
ARGM-TMP	76.42
ARGM-PRP ²	52.40
ARGM-MNR	62.10
Overall	77.44

Table 2: Performance of ASSERT on ARGMs in OntoNotes v5.0.

poral sense, is further distinguished between Temporal.Synchrony and Temporal.Asynchrony) and because the argument spans needs to be consistent with PDTB guidelines, each token will need to be looked at manually. But the overall time and effort spent on annotation will be reduced while consistency with an existing complementary annotation layer is enhanced.

4 On Enrichment of ARGM Roles in PropBank

The focus in PropBank is on the predicate argument structure of verbs, and while some of the clausal adjunct arguments such as ARGM-ADV, ARGM-CAU, ARGM-MNR, ARGM-PNC and ARGM-TMP can signal intra-sentential discourse relations, distinguishing subtle discourse-specific nuances for such adjuncts was not its primary goal. It is therefore not surprising to find as a semantic role, ARGM-ADV, which was devised to capture adverbials that could not clearly be labeled with one of the more specific adjunct roles. However, many semantic role labeling methods that utilize PropBank-style annotations do try to differentiate between the various ARGM roles. Of interest then are the performance results of semantic role labelers on the task of predicting ARGM roles. As shown in Table 2, although the overall F₁-score of a semantic role labeler ASSERT (Pradhan et al., 2004) on PropBank is in the high 70s (across a diverse set of genres annotated in the OntoNotes v5.0 test set (Pradhan et al., 2013)), the F₁-scores for all ARGM roles, except for ARGM-TMP, is in the low 60s or below 60.

As discussed in Section 2, both ARGM-ADV and ARGM-TMP exhibit the most variability in PDTB senses, the former in terms of the number of PDTB senses associated with it in significant proportions, and both in terms of how often they

²ARGM-PRP is the label to which ARGM-PNC was changed, over the course of the OntoNotes project.

are associated with multiple PDTB senses. Therefore, it might be worth enriching these ARGM roles, and possibly others such as ARGM-CAU, in PropBank with the sense distinctions found for them in the PDTB. Not only would this allow semantic role labelers to learn the finer distinctions that are currently lumped into a coarse-grained category, it would also make for a better integrated resource. We can compare this overloading, primarily of ARGM-ADV, with similar overloading of numbered arguments Arg2..Arg5 in PropBank. A study by Yi et al. (2007) demonstrated that refining these numbered arguments with a more fine-grained set of thematic roles, using a mapping from VerbNet (Schuler, 2005), improves classifier performance. In the case of Arg2 (the largest of these numbered argument classes), the F_1 -score improved by an absolute 10% points. We feel optimistic that through the proposed PDTB-informed refinement, we can get a significant performance boost in the prediction of ARGMs currently labeled as ARGM-ADV and possibly for other ARGM types as well. We plan to explore this in our future work. The finer distinctions could potentially also allow for prediction of multiple equally plausible labels, thus allowing more accurate evaluation of semantic role labelers that might, for example, learn to annotate some instances as ARGM-CAU rather than ARGM-TMP but are currently being needlessly penalized.

5 Conclusion and Future Work

A complete and well-defined annotation of intra-sentential discourse relations can have far-reaching benefits for Language Technology applications. We have explored the possibility of leveraging the sentential semantics represented in PropBank for closing the gap in the PDTB annotations of such relations. We have also suggested the converse benefit of the PDTB discourse-level semantics for enriching Propbank semantic roles, with beneficial consequences for the semantic role labeling methods that utilize PropBank.

In future work, we plan to examine another PropBank annotated element (ARGM-DIS) as a further source for seeding the PDTB. ARGM-DIS is meant to be used to annotate words or phrases that “connect a sentence to a preceding sentence” (Bonial et al., 2010). While many of these are either irrelevant to the PDTB, such as *vocatives* (“Guys”), *interjections* (“Well”, “of course”), *par-*

entheticals (“not to be crass”), and *attitudinal phrases* (“Clearly”, “Maybe”), or already systematically annotated in the PDTB (*coordinating conjunctions*, *discourse adverbials* and *attributive phrases* (“he said”)), among the over 6500 tokens of ARGM-DIS annotated in PropBank, some may be *alternative lexicalizations* of the discourse relations annotated in the PDTB (Prasad et al., 2010), including “after all”, “at the very least”, and “in effect”. Those we will examine as possible seeds for inter-sentential and intra-sentential relations.

The current study was limited to arguments of verb predicates. However, we plan to also consider arguments of eventive noun predicates as annotated in NomBank (Meyers et al., 2004). We also plan to explore the use of PDTB–Propbank overlaps to identify annotation inconsistencies in one or the other corpus, following recent work on annotation consistency control (Frank et al., 2012).

Acknowledgments

This work was partially supported by NSF grants IIS-1422186 and IIS-1421067.

References

- Asher, N. 1993. *Reference to Abstract Objects in Discourse*. Kluwer, Boston MA.
- Bonial, C., Babko-Malaya, O., Choi, J., Hwang, J., and Palmer, M. 2010. PropBank annotation guidelines (version 3.0). Technical Report, Center for Computational Language and Education Research, University of Colorado at Boulder.
- Dinesh, N., Lee, A., Miltsakaki, E., Prasad, R., Joshi, A., and Webber, B. 2005. Attribution and the (non)-alignment of syntactic and discourse arguments of connectives. In *Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*, pages 29–36.
- Feng, V. W. 2014. *RST-style discourse parsing and its applications in discourse analysis*. PhD thesis, Department of Computer Science, University of Toronto.
- Frank, A., Bögel, T., Hellwig, O., and Reiter, N. 2012. Semantic annotation for the digital humanities: Using Markov Logic Networks for annotation consistency control. In *Linguistic Issues in Language Technology*, 7(1), pages 1–21.
- Joty, S., Carenini, G., and Ng, R. 2012. A novel discriminative framework for sentence-level discourse analysis. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP/CoNLL)*, pages 904–915.

- Lin, Z., Ng, H. T., and Kan, M.-Y. 2012. A PDTB-styled end-to-end discourse parser. *Natural Language Engineering*, 20(02), pages 151–184.
- Marcus, M., Santorini, B., and Marcinkiewicz, M. A. 1993. Building a large scale annotated corpus of English: The Penn TreeBank. *Computational Linguistics*, 19(2), pages 313–330.
- Meyers, A., Reeves, R., Macleod, C., Szekely, R., Zielinska, V., Young, B., and Grishman, R. 2004. The NomBank project: An interim report. *Proceedings of the HLT/NAACL Workshop: Frontiers in corpus annotation*, pages 24–31.
- Palmer, M., Gildea, D., and Kingbury, P. 2005. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1), pages 71–105.
- Pradhan, S., Ward, W., Hacioglu, K., Martin, J., and Jurafsky, D. 2004. Shallow semantic parsing using Support Vector Machines. In *Proceedings of the Human Language Technology Conference/North American chapter of the Association of Computational Linguistics (HLT/NAACL)*, pages 233–240.
- Pradhan, S., Moschitti, A., Xue, N., Ng, H. T., Björkelund, A., Uryupina, O., Zhang, Y., and Zhong Z. 2013. Towards robust linguistic analysis using OntoNotes. In *Proceedings of the 17th Conference on Computational Natural Language Learning (CoNLL)*, pages 143–152.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., and Webber, B. 2008. The Penn Discourse TreeBank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*, pages 2961–2968.
- Prasad, R., Joshi, A., & Webber, B. 2010. Realization of discourse relations by other means: Alternative Lexicalizations. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING): Posters*, pages 1023–1031.
- Prasad, R., Webber, B., and Joshi, A. 2014. Reflections on the Penn Discourse Treebank, comparable corpora and complementary annotation. *Computational Linguistics*, 40(4), pages 921–950.
- Schuler, K. K. 2005. VerbNet: A broad-coverage, comprehensive verb lexicon. *Ph.D. Thesis*, University of Pennsylvania.
- Yi, S. T., Loper, E., and Palmer, M. 2007. Can semantic roles generalize across genres? *Proceedings of HLT/NAACL 2007*, pages 548–555.
- Weischedel, R., Palmer, M., Marcus, M., Hovy, E., Pradhan, S., Ramshaw, L., Xue, N., Taylor, A., Kaufman, J., Franchini, M., El-Bachouti, M., Belvin, R., and Houston, A. 2013. OntoNotes release 5.0. Technical report. Linguistic Data Consortium, <https://catalog.ldc.upenn.edu/LDC2013T19>.