



February 2008

## Sense Annotation in the Penn Discourse Treebank

Eleni Miltsakaki

*University of Pennsylvania*, [elenimi@seas.upenn.edu](mailto:elenimi@seas.upenn.edu)

Livio Robaldo

*University of Turin*

Alan Lee

*University of Pennsylvania*, [aleewk@babel.ling.upenn.edu](mailto:aleewk@babel.ling.upenn.edu)

Aravind K. Joshi

*University of Pennsylvania*, [joshi@cis.upenn.edu](mailto:joshi@cis.upenn.edu)

Follow this and additional works at: [https://repository.upenn.edu/cis\\_papers](https://repository.upenn.edu/cis_papers)

---

### Recommended Citation

Eleni Miltsakaki, Livio Robaldo, Alan Lee, and Aravind K. Joshi, "Sense Annotation in the Penn Discourse Treebank", . February 2008.

Postprint version. Published in *Lecture Notes in Computer Science*, Volume 4919, Computational Linguistics and Intelligent Text, 2008, pages 275-286.

Publisher URL: [http://dx.doi.org/10.1007/978-3-540-78135-6\\_23](http://dx.doi.org/10.1007/978-3-540-78135-6_23)

This paper is posted at ScholarlyCommons. [https://repository.upenn.edu/cis\\_papers/388](https://repository.upenn.edu/cis_papers/388)

For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

## Sense Annotation in the Penn Discourse Treebank

### Abstract

An important aspect of discourse understanding and generation involves the recognition and processing of discourse relations. These are conveyed by discourse connectives, i.e., lexical items like *because* and *as a result* or implicit connectives expressing an inferred discourse relation. The Penn Discourse TreeBank (PDTB) provides annotations of the argument structure, attribution and semantics of discourse connectives. In this paper, we provide the rationale of the tagset, detailed descriptions of the senses with corpus examples, simple semantic definitions of each type of sense tags as well as informal descriptions of the inferences allowed at each level.

### Comments

Postprint version. Published in *Lecture Notes in Computer Science*, Volume 4919, Computational Linguistics and Intelligent Text, 2008, pages 275-286.

Publisher URL: [http://dx.doi.org/10.1007/978-3-540-78135-6\\_23](http://dx.doi.org/10.1007/978-3-540-78135-6_23)

# Sense Annotation in the Penn Discourse Treebank

Eleni Miltsakaki<sup>1</sup>, Livio Robaldo<sup>2</sup>, Alan Lee<sup>1</sup>, and Aravind Joshi<sup>1</sup>

<sup>1</sup> Institute for Research in Cognitive Science, University of Pennsylvania  
{elenimi, aleewk, joshi}@linc.cis.upenn.edu

<sup>2</sup> Department of Computer Science, University of Turin  
robaldo@di.unito.it

**Abstract.** An important aspect of discourse understanding and generation involves the recognition and processing of discourse relations. These are conveyed by discourse connectives, i.e., lexical items like *because* and *as a result* or implicit connectives expressing an inferred discourse relation. The Penn Discourse TreeBank (PDTB) provides annotations of the argument structure, attribution and semantics of discourse connectives. In this paper, we provide the rationale of the tagset, detailed descriptions of the senses with corpus examples, simple semantic definitions of each type of sense tags as well as informal descriptions of the inferences allowed at each level.

## 1 Introduction

Large scale annotated corpora have played and continue to play a critical role in natural language processing. The continuously growing demand for more powerful and sophisticated NLP applications is evident in recent efforts to produce corpora with richer annotations [6], including annotations at the discourse level [2], [8], [4]. The Penn Discourse Treebank is, to date, the largest annotation effort at the discourse level, providing annotations of explicit and implicit connectives. The design of this annotation effort is based on the view that discourse connectives are predicates taking clausal arguments. In Spring 2006, the first version of the Penn Discourse Treebank was released, making available thousands annotations of discourse connectives and the textual spans that they relate.

Discourse connectives, however, like verbs, can have more than one meaning. Being able to correctly identify the intended sense of connectives is crucial for every natural language task which relies on understanding relationships between events or situations in the discourse. The accuracy of information retrieval from text can be significantly impaired if, for example, a temporal relation anchored on the connective *since* is interpreted as causal.

A well-known issue in sense annotations is identifying the appropriate level of granularity and meaning refinement as well as identifying consistent criteria for making sense distinctions. Even if an ‘appropriate’ level of granularity can be identified responding to the demands of a specific application, creating a flat set of sense tag is limiting in many ways. Our approach to the annotation of sense

tag in PDTB is to define a small hierarchy of sense tags containing coarse sense distinctions at the top and finer at the bottom. This schema is flexible enough to allow the annotators to choose a tag from a level that is comfortable to them. In addition, it allows the user of the corpus to pick the level that is useful for his or her purposes or even add levels of annotation if finer distinctions are desirable.

In this paper, we present our work on adding sense annotations to all the explicit and implicit connectives in the Penn Discourse Treebank (approx. 35,000 tokens). In Section (2), we give a broad overview of the Penn Discourse Treebank, detailing the types of connectives that have been annotated. In Section (3), we present the tagset used for the annotation of senses of connectives in the Penn Discourse Treebank, its hierarchical organization, and simple formal semantic descriptions for each tag. In Section (4), we present a small set of pragmatic tags that we used to capture rhetorical uses of connectives.

## 2 The Penn Discourse Treebank

Following the views toward discourse structure in [12] and [3], the Penn Discourse Treebank treats discourse connectives as discourse-level predicates that take two abstract objects such as events, states, and propositions [1] as their arguments. It provides annotations of the argument structure, attribution and semantics of discourse connectives. The PDTB annotations are done on the Wall Street Journal (WSJ) articles in the Penn TreeBank (PTB) II corpus [7]. Each annotation relates a discourse connective with its two<sup>1</sup> arguments, labelled as **Arg2**, for the argument that appears in the clause that is syntactically bound to the connective, and **Arg1**, for the other argument.

Discourse connectives in the PDTB are distinguished primarily into **Explicit** discourse connectives, that include a set of lexical items drawn from well-defined syntactic classes, and **Implicit** discourse connectives, which are inserted between paragraph-internal adjacent sentence-pairs not related explicitly by any of the syntactically-defined set of **Explicit** connectives. In the latter case, the reader must attempt to infer a discourse relation between the adjacent sentences, and ‘annotation’ consists of *inserting* a connective expression that *best* conveys the inferred relation. Multiple discourse relations can also be inferred, and are annotated by inserting multiple **Implicit** connectives. In (1), we show three examples that respectively involve an **Explicit** connective (1.a), an **Implicit** connective (1.b), and multiple **Implicit** connectives (1.c). In all examples reported below, **Arg1** is shown in italics, **Arg2** in boldface, and the discourse connective(s) underlined.

- (1) a. *She hasn't played any music* since **the earthquake hit**.  
b. *They stopped delivering junk mail*. [**Implicit=so**] **Now thousands of mailers go straight into the trash**.

---

<sup>1</sup> The assumption of the arity constraint on a connective's arguments has been upheld in all the annotation done thus far. Discourse-level predicate-argument structures are therefore unlike the predicate-argument structures of verbs at the sentence-level, where verbs can take any number of arguments.

- c. *The small, wiry Mr. Morishita comes across as an outspoken man of the world.* [Implicit=when, Implicit=for example] **He lectures a visitor about the way to sell American real estate and boasts about his friendship with Margaret Thatcher’s son.**

Adjacent sentence-pairs between which **Implicit** connectives cannot be inserted are further distinguished and annotated as three types: **AltLex**, for when a discourse relation is inferred, but insertion of an **Implicit** connective leads to a redundancy in the expression of the relation due to the relation being alternatively lexicalized by some ‘non-connective’ expression (as in (2.a)); **EntRel**, for when no discourse relation can be inferred and where the second sentence only serves to provide some further description of an entity in the first sentence (as in (2.b)), and **NoRel**, for when no discourse relation or entity-based coherence relation can be inferred between the adjacent sentences (as in (2.c)).

- (2) a. *So Seita has introduced blonde cigarettes under the Gauloises label, and intends to relaunch the unsuccessful Gitanes Blondes in new packaging.* [AltLex=the aim is] **The aim is to win market share from imported cigarettes, and to persuade smokers who are switching to blonde cigarettes to keep buying French..**
- b. *Proceeds from the offering are expected to be used for remodeling the company’s Desert Inn resort in Las Vegas, refurbishing certain aircraft of the MGM Grand Air unit, and to acquire the property for the new resort.* EntRel **The company said it estimates the Desert Inn remodeling will cost about \$32 million, and the refurbishment of the three DC-8-62 aircraft, made by McDonnell Douglas Corp., will cost around \$24.5 million.**
- c. *Jacobs is an international engineering and construction concern.* NoRel **Total capital investment at the site could be as much as \$400 million, according to Intel.**

The PDTB has been used as a resource for Natural Language Generation [11], and for Sense Disambiguation [10]. This paper focuses on PDTB sense annotation, and describes the tagset used to annotate the discourse connectives. The reader interested in the overall annotation is addressed to [9] and [13].

### 3 Annotation of Senses in the PDTB

The Penn Discourse Treebank provides sense tags for the **Explicit**, **Implicit** and **AltLex** connectives. Depending on the context, the content of the arguments and possibly other factors, discourse connectives, just like verbs, can have more than one meaning. For example, *since* seems to have three different senses, one purely ‘Temporal’ (as in (3.a)), another purely ‘Causal’ (as in (3.b)) and a third both ‘Causal’ and ‘Temporal’ (as in (3.c)).

- (3) a. *The Mountain View, Calif., company has been receiving 1,000 calls a day about the product since it was demonstrated at a computer publishing conference several weeks ago.*
- b. *It was a far safer deal for lenders since NWA had a healthier cash flow and more collateral on hand.*
- c. *Domestic car sales have plunged 19% since the Big Three ended many of their programs Sept. 30.*

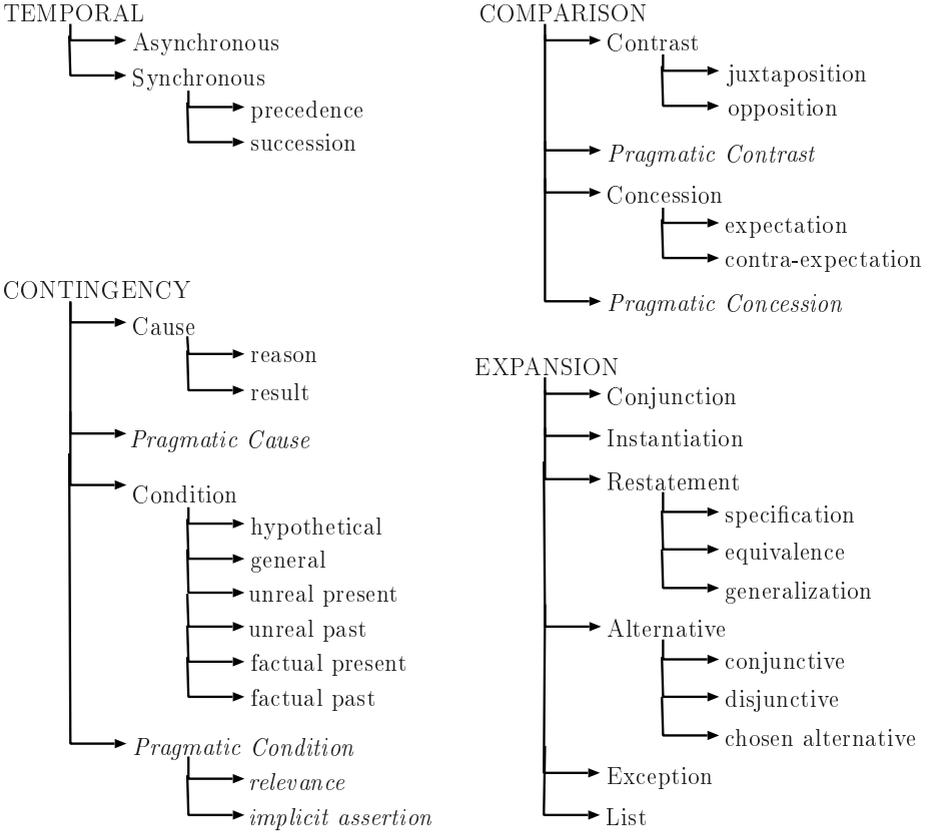
Sense annotations in PDTB provide tags specifying the sense of the connective in cases of ambiguity, and in every case they provide a semantic description of the relation between the arguments of connectives. When annotators identify more than one simultaneous interpretations, multiple sense tags are provided. Sense annotations specify one or more, but not necessarily all the semantic relations that may hold between the arguments of the connectives.

The tagset of senses is organized hierarchically (shown in Figure 1) and comprises three levels: *class*, *type* and *subtype*. The top level, or *class level* of the hierarchy represents four major semantic classes: ‘TEMPORAL’, ‘CONTINGENCY’, ‘COMPARISON’ and ‘EXPANSION’. For each class, a second level of *types* is defined to further refine the semantics of the class levels. For example, ‘CONTINGENCY’ has two types, ‘Cause’ (relating two situations via a direct cause-effect relation) and ‘Condition’ (relating a hypothetical scenario with its possible consequences). A third level of *subtype* specifies the semantic contribution of each argument. For ‘CONTINGENCY’, its ‘Cause’ type has two subtypes, ‘reason’ and ‘result’, which specify which argument is interpreted as the cause of the other. A typical connective labelled as ‘reason’ is *because* and a typical connective labelled as ‘result’ is *as a result*.

For most types and subtypes, we also provide some hints about their possible semantics. In doing so, we do not attempt to represent the internal meaning of *Arg1* and *Arg2*, but simply refer to them as  $\| Arg1 \|$  and  $\| Arg2 \|$  respectively. We believe that roughing out the semantics of the sense tags provides a starting point for the definition of an integrated logical framework able to deal with the semantics of discourse connectives but it also helps the annotators in choosing the proper sense tag.

The hierarchical organization of the sense tags serves two purposes. First, it efficiently addresses well-known issues regarding inter-annotator reliability, by allowing the annotators to select a tag from a level that is comfortable to them. Sense annotators in PDTB are not forced to make fine semantic distinctions when they are not confident that their world knowledge or discourse context can support more specific interpretations. Secondly, the hierarchical organization of tags also allows useful inferences at all levels. For example, (1) illustrates a case where neither the text nor the annotators’ world knowledge has been sufficient to enable them to provide a sense tag at the level of subtype. Instead, they have provided one at the level of type.

- (1) *Besides, to a large extent, Mr. Jones may already be getting what he wants out of the team, even though it keeps losing.*



**Fig. 1.** Hierarchy of sense tags

Connectives can also be used to relate the *use* of the arguments of a connective to one another or the use of one argument with the sense of the other. For these *rhetorical* or *pragmatic* uses of connectives, we have defined *pragmatic* sense tags - specifically ‘Pragmatic Cause’, ‘Pragmatic Condition’, ‘Pragmatic Contrast’ and ‘Pragmatic Concession’.

In what follows, we provide descriptions for all the semantic labels of the sense hierarchy.

### 3.1 Class ‘TEMPORAL’

‘TEMPORAL’ is used when the situations described in the arguments are related temporally. The class level tag ‘TEMPORAL’ does not specify if the situations are temporally ordered or overlapping. Two types are defined for ‘TEMPORAL’: ‘Asynchronous’ (i.e., temporally ordered) and ‘Synchronous’ (i.e., temporally overlapping). ‘Asynchronous’ has two subtypes, ‘precedence’ and ‘succession’, which specify which situation takes place before the other one. The tag ‘precedence’ is

used when the connective indicates that the situation in **Arg1** precedes the situation described in **Arg2**, as *before* does in (2). The tag ‘succession’ is used when the connective indicates that the situation described in **Arg1** follows the situation described in **Arg2**, as *after* does in (3).

- (2) But a Soviet bank here would be crippled unless Moscow found a way to settle the \$188 million debt, *which was lent to the country’s short-lived democratic Kerensky government* before **the Communists seized power in 1917**.
- (3) No matter who owns PS of New Hampshire, after it emerges from bankruptcy proceedings *its rates will be among the highest in the nation*, he said.

The tag ‘Synchronous’ applies when the connective indicates that the situations described in **Arg1** and **Arg2** overlap. The type ‘Synchronous’ does not specify the form of overlap, i.e., whether the two situations started and ended at the same time, whether one was temporally embedded in the other, or whether the two crossed. Typical connectives tagged as ‘Synchronous’ are *while* and *when*, the latter shown in (4).

- (4) *Knowing a tasty – and free – meal* when they eat one, the executives gave the chefs a standing ovation. (TEMPORAL:Synchrony) (0010)

### 3.2 Class ‘CONTINGENCY’

‘CONTINGENCY’ is used when the situations described in the arguments are causally influenced. It has two types, ‘Cause’ and ‘Condition’. The main difference between the two is that in ‘Cause’ the connective expressing the relation does not have any impact on whether the arguments are taken to hold or not. For instance, in (4.a), the situations specified in **Arg1** and **Arg2** are taken to hold true independently of the connective. It is true that the use of dispersants was approved, that a test on the third day showed positive results, and that the latter caused the former. In this case, the directionality of causality, i.e., that  $\|\text{Arg2}\|$  is the cause and  $\|\text{Arg1}\|$  the effect, is specified with the subtype ‘reason’. Formally, we represent the semantics of ‘reason’ as  $\|\text{Arg1}\| < \|\text{Arg2}\| \wedge \|\text{Arg1}\| \wedge \|\text{Arg2}\|$ , where  $<$  is a logical operator taken from [5].  $\text{Arg1} < \text{Arg2}$  is intended to model the causal law<sup>2</sup> ‘**Arg1** causes **Arg2**’. The reverse case, i.e., when  $\|\text{Arg1}\|$  is the cause and  $\|\text{Arg2}\|$  the effect, is labelled with the subtype ‘result’.

The type ‘Condition’ is used to describe all subtypes of conditional relations. In addition to causal influence, ‘Condition’ allows some basic inferences about the semantic contribution of the arguments. Specifically, the situation in **Arg2** is taken to be the condition and the situation described in **Arg1** is taken to be the consequence, i.e., the situation that holds when the condition is true. Unlike ‘Cause’, however, the truth value of the arguments of a ‘Condition’ relation cannot be determined independently of the connective. For this reason, we introduce

---

<sup>2</sup> As largely discussed in the literature, causality cannot be modeled via the logical implication ‘ $\rightarrow$ ’. ‘ $\rightarrow$ ’ will be used to handle the semantic of ‘Restatement’ (see below).

some branching-time logic operators into our rough description of the semantics of ‘Condition’ subtypes:  $A$ ,  $F$ , and  $G$ .  $A$  universally quantifies over all possible futures; therefore,  $A\beta$  is true iff  $\beta$  is true in all possible futures.  $F$  and  $G$  are respectively existential and universal quantifiers over instants in a single future:  $F\alpha$  is true iff  $\alpha$  is true in some instant in a possible future, while  $G\alpha$  is true iff  $\alpha$  is true in every instant in a possible future.

The sense hierarchy includes six basic subtypes of ‘Condition’. Example (4.b) is marked with the subtype ‘general’ because the sentence-pair describes a generic truth about the world (or a statement that describes a regular outcome everytime the condition holds true). We formalize its semantics as  $AG(\|\text{Arg2}\| < \|\text{Arg1}\|)$ , i.e., when a sentence-pair is tagged as ‘general’, in all possible futures, it is always the case that  $\|\text{Arg2}\|$  causes  $\|\text{Arg1}\|$ . Compare with ‘hypothetical’ which marks a causal relation that holds true only at the moment when the sentence is uttered. An example is shown in (4.c); this is a case of ‘hypothetical’ in that, in the future, even if the negotiators start to focus on those areas, the talks may be unsuccessful (i.e., in the future, there may be other factors that affect the performance of the talks). The formal semantics of ‘hypothetical’ is  $\|\text{Arg2}\| < AF \|\text{Arg1}\|$ : if  $\|\text{Arg2}\|$  holds true,  $\|\text{Arg1}\|$  is caused to hold too at some instant in all possible subsequent futures. Examples (4.d) and (4.e) are respectively marked as ‘factual present’ and ‘unreal present’. The tag ‘factual present’ applies when  $\text{Arg2}$  denotes a situation that has either been presented as a fact in the prior discourse or is believed by somebody other than the speaker or writer. From a formal point of view, we add a conjunct stating that  $|\text{Arg2}|$  is true or it is believed to hold true. The subtype ‘unreal present’ applies when  $\text{Arg2}$  describes a condition that either does not hold at present or considered unlikely to hold. In such a case, we assert the formula  $\|\text{Arg2}\| < AF \|\text{Arg1}\| \wedge \sim\|\text{Arg2}\|$ <sup>3</sup>. The other two subtypes are ‘factual past’ and ‘unreal past’, which are respectively similar to ‘factual present’ and ‘unreal present’ except that in this case the first argument refers to a situation that is assumed to have taken place at a time in the past.

- (4) a. *Use of dispersants was approved when a test on the third day showed some positive results.*  
 b. *They won't buy if the quality is not here.*  
 c. Both sides have agreed *that the talks will be most successful if negotiators start by focusing on the areas that can be most easily changed.*  
 d. *If that's true, Orange County has to be at least 10% of that.*  
 e. *If the film contained dialogue, Mr.Lane's Artist would be called a homeless person.*

### 3.3 Class ‘COMPARISON’

The class tag ‘COMPARISON’ applies when a discourse relation is established between  $\text{Arg1}$  and  $\text{Arg2}$  in order to highlight prominent differences between the

<sup>3</sup>  $\sim\|\text{Arg2}\|$  means that  $\|\text{Arg}\|$  does not hold or not expected to hold.

two situations. Semantically, the truth of both arguments is independent of the connective or the established relation. ‘COMPARISON’ has two types to further specify its semantics. In some cases, *Arg1* and *Arg2* share a predicate or a property and the difference is highlighted with respect to the values assigned to this property. This interpretation is tagged with the type ‘Contrast’. There are also cases in which the highlighted differences are related to expectations raised by one argument which are then denied by the other. This interpretation is tagged with the Type ‘Concession’. In ‘Contrast’ both arguments describe a situation that is not asserted on the basis of the other one. In this sense, there is no directionality in the interpretation of the arguments. This is an important difference between the interpretation of ‘Contrast’ and ‘Concession’. Two subtypes of ‘Contrast’ are defined to further specify the type of values that are compared: ‘juxtaposition’ (weak contrast) and ‘opposition’ (strong contrast). The latter is applied when the values assigned to some shared property are taken from the extremes of a gradable scale (e.g., tall-short, accept-reject, etc.), the former otherwise. For example, (5.a) is tagged as ‘opposition’ because the two arguments describe opposite performances of the banks. Example (5.b) is tagged as ‘juxtaposition’ because the shared predicate *rose* or *jumped* takes two different values (69% and 85%) and the shared predicate *rose to X amount* takes two individual entities (the net operating venue and the net internet bill).

- (5) a. *Its bank in Texas also reported a loss of \$23.5 million for the quarter but that its consumer banks in Oregon, California, Nevada and Washington performed well during the quarter.*  
 b. *Operating revenue rose 69% to A\$8.48 billion from A\$5.01 billion but the net interest bill jumped 85% to A \$686.7 million from A \$371.1 million.*

‘Concession’ also has two subtypes: ‘expectation’ and ‘contra-expectation’. The subtype ‘expectation’ applies when *Arg2* describes a situation *A* which causes another situation *C*, and *Arg1* asserts (or implies) the situation  $\neg C$  (i.e.  $\|Arg2\| < C \wedge \|Arg1\| \rightarrow \neg C$ ), as in (6.a). The subtype ‘contra-expectation’ applies when *Arg1* causes *C* and *Arg1* denies it, as in (6.b).

- (6) a. Although the purchasing managers’ index continues to indicate a slowing economy, *it isn’t signaling an imminent recession.*  
 b. *The Texas oilman has acquired a 26.2% stake valued at more than \$1.2 billion in an automotive-lighting company, Koito Manufacturing Co. But he has failed to gain any influence at the company.*

Some times in the discourse the intended ‘juxtaposition’ or ‘opposition’ is clear and sometimes it is not. When it is not, the sense of the connective is considered ambiguous and the higher level tag ‘Contrast’ applies. In fact, the gradable scale with respect to which we discriminate between ‘juxtaposition’ and ‘opposition’ strongly depends on the context where the sentence is uttered. For example, consider the pair *black-white*. These two concepts are usually taken to be antonyms.

Therefore, it seems that whenever **Arg1** assigns *black* and **Arg2** assigns *white* to a shared property (e.g. *Mary is black whereas John is white*), the discourse connective has to be labelled as ‘opposition’. Nevertheless, in many contexts *black* and *white* are just two of the colors that may be assigned to the shared property (e.g., imagine *Mary bought a black hat whereas John bought a white one* uttered in a shop that sell red, yellow and blue hats as well). In such a case, they are not antonyms, and the discourse connective has to be labelled as ‘juxtaposition’.

### 3.4 Class ‘EXPANSION’

Under the class ‘EXPANSION’ we group all the relations which expand the discourse and move forward its narrative or exposition. ‘EXPANSION’ includes several types which refine its semantics. The type ‘Conjunction’ is used when the situation described in **Arg2** provides additional, discourse new, information that is related to the situation described in **Arg1**. It is inferred that the information described in **Arg2** is not related to **Arg1** in any of the ways described in the other types of ‘EXPANSION’ ( $\|\mathbf{Arg1}\| \wedge \|\mathbf{Arg2}\|$ ). The tag ‘Instantiation’ is used when  $\|\mathbf{Arg1}\|$  evokes a set of events and  $\|\mathbf{Arg2}\|$  picks up one of these events and describes it in further detail; in this case, besides the logical conjunction of the arguments, we assert *exemplify'*( $\|\mathbf{Arg2}\|, \lambda x. x \in g(\|\mathbf{Arg1}\|)$ ), where *exemplify'* is a predicate taken from [3], *g* a function that ‘extracts’ the set of events from the semantics of **Arg1**, and *x* is a variable ranging over them. *exemplify'* asserts that **Arg2** further describes one element in the extracted set. A connective is marked as ‘Restatement’ when the semantics of **Arg2** restates the semantics of **Arg1**. It is inferred that the situations described in **Arg1** and **Arg2** hold true at the same time. The subtypes ‘specification’, ‘generalization’, and ‘equivalence’ further specify the ways in which **Arg2** restates **Arg1**. In particular, besides the conjunction of the two arguments, we assert a logical implication ( $\rightarrow$ ) between the arguments:  $\|\mathbf{Arg1}\| \rightarrow \|\mathbf{Arg2}\|$  (generalization),  $\|\mathbf{Arg1}\| \leftarrow \|\mathbf{Arg2}\|$  (specification), and  $\|\mathbf{Arg1}\| \leftrightarrow \|\mathbf{Arg2}\|$  (equivalence). The type ‘Alternative’ applies when the two arguments of the connective evoke two alternative situations. The type ‘Alternative’ is further specified with the subtypes ‘conjunctive’, ‘disjunctive’ and ‘chosen alternative’. The ‘conjunctive’ subtype is used when both alternatives are possible ( $\|\mathbf{Arg1}\| \vee \|\mathbf{Arg2}\|$ ), ‘disjunctive’ when two situations are evoked in the discourse but only one of the two holds ( $\|\mathbf{Arg1}\| \text{ xor } \|\mathbf{Arg2}\|$ ), and ‘chosen alternative’ when two alternatives are evoked in the discourse and the one denoted by **Arg2** is taken ( $\|\mathbf{Arg1}\| \text{ xor } \|\mathbf{Arg2}\| \wedge \|\mathbf{Arg2}\|$ ). The type ‘Exception’ applies when **Arg2** evokes a situation which makes **Arg1** not fully be true. In other words, in case of ‘Exception’, **Arg1** is false, **Arg2** is true and if **Arg2** were false, **Arg1** would be true. So, the formal semantics of ‘Exception’ is  $\neg \|\mathbf{Arg1}\| \wedge \|\mathbf{Arg2}\| \wedge \neg \|\mathbf{Arg2}\| \rightarrow \|\mathbf{Arg1}\|$ . Finally, the type ‘List’ applies when the events or states expressed in **Arg1** and **Arg2** are members of a list of events or states enumerated in the discourse. It is possible that semantically **Arg1** and **Arg2** are not related. For the appropriate interpretation of the discourse all the elements of the list must be retrieved. The

predicate of the list (i.e., what **Arg1** and **Arg2** are elements of) must be retrieved from the prior discourse.

In (7) we provide examples of ‘Instantiation’ (7.a), ‘specification’ (7.b), ‘chosen alternative’ (7.c), and ‘Exception’ (7.c). Unfortunately, for space constraints we cannot provide an example of each type and subtype of ‘EXPANSION’.

- (7) a. *Hypertext books are clearly superior to normal books. For example, they have database cross-referencing facilities ordinary volumes lack.*  
b. *I never gamble too far. [Implicit=In particular] I quit after one try, whether I win or lose.*  
c. *It isn't allowed to share in the continuing proceeds when the reruns are sold to local stations. Instead ABC will have to sell off the rights for a one-time fee.*  
d. *Boston Co. officials declined to comment on the unit's financial performance this year except to deny a published report that outside accountants had discovered evidence of significant accounting errors in the first three quarters' results.*

## 4 Pragmatic Uses of Connectives

The PDTB contains instances of rhetorical or pragmatic uses of connectives. For these instances, we define a small set of sense tags called *pragmatic*. We have found instances of ‘Pragmatic cause’, ‘Pragmatic Condition’, ‘Pragmatic Concession’, and ‘Pragmatic Contrast’. For instance, the tag ‘Pragmatic Condition’ is used for instances of conditional constructions whose interpretation deviates from that of the semantics of ‘Condition’. Specifically, these are cases of explicit *if* tokens with **Arg1** and **Arg2** not being causally related. In all cases, **Arg1** holds true independently of **Arg2**. Two subtypes of ‘Pragmatic Condition’ have been defined: ‘relevance’ and ‘implicit assertion’. The former applies where **Arg2** provides the context in which the description of the situation in **Arg1** is relevant. An example of ‘relevance’ is shown in (8.a); note that there is no causal relation between the two arguments. The pragmatic tag ‘implicit assertion’ applies in special rhetorical uses of *if*-constructions when the interpretation of the conditional construction is an implicit assertion. In (8.b), for example, **Arg1**, *O'Connor is your man* is not a consequent state that will result if the condition expressed in **Arg2** holds true. Instead, the conditional construction in this case implicitly asserts that O'Connor will keep the crime rates high.

- (8) a. If anyone has difficulty imagining a world in which history went merrily on without us, *Mr. Gould sketches several.*  
b. If you want to keep the crime rates high, *O'Connor is your man.*  
c. *Mrs Yeargin is lying. [Implicit=because] They found students in an advanced class a year earlier who said she gave them similar help.*

‘Pragmatic Cause’ is used when **Arg1** expresses a claim and **Arg2** provides justification for this claim, as shown in (8.c). The situations described in **Arg1** and **Arg2** are not causally influenced. Epistemic uses fall under this category.

## 5 Inter-annotator Agreement and Adjudication

The PDTB corpus was sense annotated by two annotators. Class level inter-annotator agreement was 92% and subtype level agreement (the most refined level) was 77%. Class level disagreement was adjudicated by a team of three experts. Disagreement at lower levels was resolved by providing a sense tag from the immediately higher level. For example, if one annotator tagged a token with the type ‘Concession’ and the other, with the type ‘Contrast’, we resolved the disagreement by providing the the class level tag ‘Comparison’ based on the assumption that both a concessive and contrastive interpretation could be construed.

## 6 Summary and Future Work

The Penn Discourse Treebank provides annotations of discourse connectives and their arguments. Discourse connectives, like verbs, can have more than one sense. Here, we presented the tagset that we used to annotate the senses of connectives. The tagset is organized hierarchically in three levels, with coarse sense distinctions made at the top level and finer distinctions provided in lower levels. The PDTB corpus contains instances of rhetorical uses of connectives. These instances are annotated with a small set of pragmatic tags. For each sense tag we provided a simple formal description of its semantics. We are currently studying the distribution of senses per connective and looking more closely at semantic features of the arguments in order to develop empirically motivated descriptions of the semantic roles of the arguments and also in order to identify useful features for models of automatic sense disambiguation.

## Acknowledgments

We are grateful to Nikhil Dinesh, Rashmi Prasad and Bonnie Webber for their insightful comments on this work. The PDTB project is partially supported by NSF Grant: Research Resources, EIA 02-24417 to the University of Pennsylvania (PI: Aravind Joshi). The first release (PDTB-1.0) is distributed free of charge under terms of the GNU General Public License. The corpus and associated software is available for download from <http://www.seas.upenn.edu/pdtb>. The second release, including sense annotations, is scheduled for January 2008.

## References

1. Asher, N.: Reference to Abstract Objects. Kluwer, Dordrecht (1993)
2. Carlson, L., Marcu, D., Okurowski, M.: Current Directions in Discourse and Dialogue, Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory. Kluwer Academic Publishers, Dordrecht (2003)
3. Forbes-Riley, K., Webber, B., Joshi, A.: Computing discourse semantics: The predicate-argument semantics of discourse connectives in D-LTAG. *Journal of Semantics* 23, 55–106 (2006)

4. Gaizauskas, R., et al.: The timebank corpus. In: *Corpus Linguistics 2003*. Lancaster, U.K (2003)
5. Giordano, L., Schwind, C.: Conditional logic of actions and causation. *Artificial Intelligence* 157, 239–279 (2004)
6. Kingsbury, P., Palmer, M.: From Treebank to Propbank. In: *Third International Conference on Language Resources and Evaluation, LREC-2002*, Las Palmas, Canary Islands, Spain (2002)
7. Marcus, M.P., Santorini, B., Marcinkiewicz, M.A.: Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* 19(2), 313–330 (1993)
8. Mitkov, R., et al.: Coreference and anaphora: Developing annotating tools, annotated resources and annotation strategies. In: *Proc. of the Discourse Anaphora and Anaphora Resolution Colloquium (DAARC 2000)*, Lancaster, U.K. (2000)
9. Miltsakaki, E., et al.: The Penn Discourse Treebank. In: *Proc. of the 4th International Conference on Language Resources and Evaluation (LREC 2004)* (2004)
10. Miltsakaki, E., et al.: Experiments on sense annotation and sense disambiguation of discourse connectives. In: *Proc. of the 4th Workshop on Treebanks and Linguistic Theories (TLT2005)* (2005)
11. Prasad, R., et al.: Discourse TreeBank as a resource for natural language generation. In: *Proc. of the Corpus Linguistics Workshop on Using Corpora for NLG* (2005)
12. Webber, B., et al.: Anaphora and discourse structure. *Computational Linguistics* 29(4), 545–587 (2003)
13. Webber, B., et al.: A short introduction to the PDTB. In: *Copenhagen Working Papers in Language and Speech Processing* (2005)