Mapping Different Rhetorical Relation Annotations: A Proposal

Farah Benamara
Institut de Recherche en Informatique de Toulouse
Université Paul Sabatier
Toulouse, France
benamara@irit.fr

Maite Taboada
Department of Linguistics
Simon Fraser University
Vancouver, Canada
mtaboada@sfu.ca

Abstract

Annotation efforts have resulted in the availability of a number of corpora with rhetorical relation information. The corpora, unfortunately, are annotated under different theoretical approaches and have different hierarchies of relations. In addition, new sets of rhetorical relations have been proposed to account for language variation. The types of relations, however, tend to overlap or be related in specific ways. We believe that differences across approaches are minimal, and a unified set of relations that works across languages is possible. This paper details a new taxonomy of relations organized in four top level-classes with a total of 26 relations. We propose a mapping between existing annotations and show that our taxonomy is robust across theories, and can be applied to multiple languages.

1 Motivation

The annotation of discourse relations in language can be broadly characterized as falling under two main approaches: the *lexically grounded approach* and an approach that aims at *complete discourse coverage*. Perhaps the best example of the first approach is the Penn Discourse Treebank (Prasad et al., 2008). The annotation starts with specific lexical items, most of them conjunctions, and includes two arguments for each conjunction. This leads to partial discourse coverage, as there is no guarantee that the entire text is annotated, since parts of the text not related through a conjunction are excluded. On the positive side, such annotations tend to be reliable. PDTB-style annotations have been carried out in a variety of languages (Arabic, Chinese, Czech, Danish, Dutch, French, Hindi and Turkish), and in some cases the taxonomy of relations had to be modified, by adding or merging relations (Prasad et al., 2014).

Complete discourse coverage requires annotation of the entire text, with most of the propositions in the text integrated in a structure. It includes work from two theoretical perspectives, either intentionally driven, such as RST (Mann and Thompson, 1988) or semantically driven, such as SDRT (Asher and Lascarides, 2003). RST proposes a tree-based representation, with relations between adjacent segments, and emphasizes a differential status for discourse components (the nucleus vs. satellite distinction). Annotated resources exist in Basque, Dutch, German, English, Portuguese and Spanish. Captured in a graph-based representation, with long-distance attachments, SDRT proposes relations between abstract objects using a relatively small set of relations. Corpora following SDRT exist in Arabic, French and English.

Manually annotated resources have contributed to a number of applications, most notably discourse segmentation into elementary discourse units, identification of explicit and implicit relations for the purpose of discourse parsing, and development of end-to-end discourse parsers (Hernault et al., 2010; Feng and Hirst, 2014; Joty et al., 2015). These parsers have been successfully deployed in NLP applications including machine translation, sentiment analysis and automatic summarization (Thione et al., 2004; Heerschop et al., 2011; Hardmeier, 2013).

Each approach has its own hierarchy of discourse relations, but relations tend to overlap or be related in a few specific ways. We suggest there are four general ways of mapping relations across approaches: (1) *Specialization*, where a relation R in one approach can correspond to several relations in another approach, (2) *Generalization*, in which several relations in one approach correspond to one relation in another approach, (3) *Omission* involves a relation defined in one approach, but not taken into account in another. Finally, in (4), *Definition*, relations have similar names, but different definitions.

We propose here a unified hierarchy of discourse relations. Our proposal has several motivations. First of all, with the wide availability of annotated corpora, it would be beneficial to have a system for mapping relations across approaches. In particular for classification tasks such as discourse parsing, access to larger amounts of data is likely to yield better results. Secondly, and from a more theoretical point of view, we would like to propose that differences across approaches are minimal, and a unified set of relations is possible. This would facilitate the work of discourse analysts and would also result in better annotation efforts. Third, a unified set of discourse relations would allow us to compile a list of discourse markers and other signals for those relations, which would also benefit discourse annotation. Finally, this is a first step towards multilingual discourse analysis. Many studies have compared the use of discourse markers across languages, and how they differ in translation (Degand, 2009; Zufferey and Degand, 2014). We would like to contribute to that area of study by unifying and integrating the types of relations that markers can signal.

Merging different discourse relation taxonomies involves, in our view, different steps, having to do with: (1) segmentation, (2) unifying the set of relations, (3) proposing possible signals, (4) unifying discourse structures, and (5) providing a language for merging annotations. We focus here on step (2). For proposals for steps (4) and (5), see Venant et al. (2013), and Chiarcos (2014), respectively.

2 Methodology

Our first focus are the two theories that we are most familiar with, RST and SDRT. We next plan to find correspondences between our unified RST-SDRT hierarchy and the PDTB taxonomy.

The first step consists of grouping relations in top-level classes. Our goal is to minimize the number of top-level classes and, at the same time, reduce the number at the fine-grained level, avoiding the proliferation of relations seen in the RST Discourse Tree-Bank (Carlson et al., 2003).

Two main criteria were used in creating the hierarchy. First of all, the proposed hierarchy should be stable enough for language variation. By this we mean that the main classes at the top level should remain constant. We believe that there is little cross-linguistic variation when it comes to the higher-level classification of discourse relations. The second, related criterion in our organization, is that the hierarchy has to be open to modification at the low level. This is where previous research has observed variation due to language and genre.

Definitions of relations are based on three further criteria. First of all, we do not define relations on the basis of the status of their arguments. The nucleus-satellite distinction in RST is not relevant for our basic definition of relations. Secondly, we focus on the effect that a relation has on meaning, and not on how it is lexically triggered by a discourse marker or lexical device. Finally, we provide intentional effects when needed. Our taxonomy is both intentionally and semantically driven, motivated by our desire to find a balance between RST and SDRT.

Our starting points are the set of RST definitions from the RST website[1] and the definitions provided within the RST-DT (Carlson et al., 2003). For SDRT, we considered the relations defined in the SDRT literature (Asher and Lascarides, 2003), plus the adaptations created when annotating data in different projects: Discor (Reese et al., 2007), Annodis (Afantenos et al., 2012), plus the classification proposed for Arabic (Keskes et al., 2014).

3 Towards a unified hierarchy

We built a hierarchy with four top-level classes: TEMPORAL, STRUCTURAL, THEMATIC and CAUSAL-ARGUMENTATIVE, organized in three levels with a total of 26 relations. We have taken into account all SDRT, PDTB, RST-DT and RST relations, with the exception of the following relations from the RST-DT: Topic Change (topic-shift,

[1] http://www.sfu.ca/rst/

topic-drift), Textual Organization and Topic-Comment (problem-solution, question-answer, statement-response, topic-comment, comment-topic, rhetorical-question). We believe that some of those relations structure topics, but are not necessarily discourse relations. Table 1 summarizes the inventory of the proposed relations at each level.

3.1 Temporal class

A class of relations indicating relations which set events in terms of time or a similar frame is a component of most hierarchies. In some classifications, this group includes relations of background or framing, but we prefer an exclusively temporal class.

Arguments in the temporal class need to share the same topic, and the relations always express co-temporal constraints, i.e., temporal ordering between the main eventualities e_1 and e_2 introduced respectively in their arguments. The class includes three relations: SEQUENCE, INVERTED SEQUENCE and SYNCHRONOUS.

3.2 Thematic class

Thematic is a broad class which includes relations among the content of the propositions. They structure and organize information in the discourse, and can be divided into three different subclasses:

(1) ELABORATION. A group of discourse relations that connect utterances describing the same state of affairs. Further classified into: PARTITIVE, GENERALITY, OBJECT, SUMMARY, RESTATEMENT, and MEANS. For most purposes, the further specificity is not necessary, and in some cases it may be difficult to distinguish among the subclasses. We believe, however, that these more specific relations may be useful when analyzing certain genres, or for particular applications.

(2) FRAMING. This class includes relations that provide a framework for understanding the content of the situation described in the discourse segment. It includes two relations: FRAME and BACKGROUND. FRAME holds when a is a frame and b is in the scope of that frame, generally when a is at the beginning of a sentence. Several cases are possible: temporal, spatial or domain frames. This relation has no direct equivalence in RST. BACKGROUND is equivalent to the RST relations Background and Circumstance. It is used to capture a specific spatio-temporal structure, to accommodate presuppositions in discourse, or to set the stage of a story.

(3) ATTRIBUTION. Attribution relates a communicative agent in the first argument and the content of a communicative act introduced in the second. Both the RST-DT and SDRT take Attribution as a discourse relation. PDTB, on the other hand, treats it as orthogonal to discourse annotation. In our case, we follow Asher et al.'s (2006) position on reportative constructions in discourse, who consider that the treatment of these verbs is necessary for a correct analysis of the semantics and discourse structure of stories in news corpora. We agree, however, with the PDTB, that it is not a fully-fledged relation, with the same intentional effects, but we do believe that it should be annotated.

3.3 Structuring class

This class contains relations of textual organization at a high level, which organize the structure of the information in terms of themes or topics (but are rhetorical, not relations of topic management). ALTERNATION holds when there is a disjunction between a and b. PARALLEL occurs when a and b have similar semantic and syntactic structures, and it requires a and b to share a common theme. It has the same semantics as List in RST-DT. CONTINUATION holds between two segments when they both elaborate or provide background to the same segment. It also occurs in cases where there is no clear rhetorical relation between the segments. Equivalent to the RST-DT relation Elaboration-additional and to Continuation in SDRT.

3.4 Causal-argumentative class

This class contains two broad classes, one causal and one argumentative. We see them as related to each other, as conjunctions and other discourse markers can be present to indicate a causal relation or be more abstract in an argumentative use (e.g., *I'm only saying this because I care.*).

3.4.1 Causal

We distinguish between CAUSE/RESULT and PURPOSE. Within the first sub-class, REASON holds when the main eventuality of the second argument is understood as the cause of the eventuality in the first argument. RESULT relates a cause to its effect: the

TEMPORAL	SEQUENCE, INVERTED SEQUENCE, SYNCHRONOUS
THEMATIC	ELABORATION → Partitive, Generality, Object, Summary, Restatement, Means
	FRAMING → Frame, Background
	ATTRIBUTION
STRUCTURING	ALTERNATION, PARALLEL, CONTINUATION
CAUSAL-ARGUMENTATIVE	CAUSAL Cause/Result → Reason, Result, General Condition Purpose
	ARGUMENTATIVE Support → Motivation, Evidence/Justification, Evaluation/Interpretation Opposition → Contrast, Concession, Antithesis

Table 1: Inventory of proposed relations in the unified hierarchy

main eventuality of *a* caused the eventuality given by *b*. GENERAL CONDITION holds when the first segment is a hypothesis and the second the consequence. PURPOSE holds when the second segment (*b*) describes the aim, the goal or the purpose of the event described in the first segment (*a*). Most often, it can be paraphrased as "*a* in order to *b*."

3.4.2 Argumentative

This class includes the SUPPORT and OPPOSITION sub-classes, which are used to advance an argument. SUPPORT mainly captures justification, explanation (not causal), evaluation and evidence. OPPOSITION groups relations where the segments have similar semantic structures, but contrasting themes, i.e., sentence topics, or when one constituent negates a default consequence of the other.

4 Mapping RST and SDRT annotations

To test the stability of our proposed hierarchy across both theoretical and language variations, we mapped it to annotations in three corpora: the RST-DT English corpus (Carlson et al., 2003), the SDRT Annodis French corpus (Afantenos et al., 2012), and the RST Spanish Treebank (RST-ST) (da Cunha et al., 2011). The taxonomies in these corpora respectively contain 78, 17 and 28 relations. The total number of annotated information in terms of frequency of relations is 18,255 for RST-DT, 3,345 for Annodis, and 3,115 for RST-ST. Tables 2, 3, 4, and 5 provide frequency of our relations in each of the three corpora above in the four main classes. The distribution of our four classes across the corpora is respectively 3.61%, 25.55%, 47.80%, and 23.08%. The proportions are quite similar to the original distribution in each corpus, taking into account the slightly different structures of each taxonomy.

	RST-DT	Annodis	RST-ST
Seq.	224	350	74
Sync.	160		
Inv. Seq.	59	27	
Total TEMPORAL: **894**			

Table 2: Temporal class mapping

	RST-DT	Annodis	RST-ST	Total
Alter.	21	18	9	48
Paral.	1,211	59		1,270
Conti.	4,144	682	171	4,997
Total STRUCTURING: **6,315**				

Table 3: Structuring class mapping

Most mappings were relatively straightforward, except for some relations that can be either *missing* or *too specific*. The first case concerns the relations FRAME, PARALLEL and ATTRIBUTION that were not annotated in the RST-ST, as well as the relation FRAME for RST-DT. The second case is the most frequent and occurs when at least two relations having specific semantics or intentional effects need to be merged to find their corresponding instances in a given corpus. For example, the Temporal relations SEQUENCE and SYNCHRONOUS in Annodis were annotated using the same Narration relation. The SDRT Elaboration captures most of our ELABORATION relations (except for OBJECT). On the other hand, the RST-ST corpus considers only one Temporal relation, namely Sequence, and RST-ST Elaboration includes OBJECT, PARTITIVE, and GENERALITY. Another interesting case concerns the Support relations MOTIVATION and EVIDENCE/JUSTIFICATION. SDRT does not distinguish between the causal and the epistemic use of causal relations and considers only two causal relations, Explanation and Result. This is why we counted SDRT causal relations only in the Causal/Result sub-class (this is marked by (*) and (+) in Table 5). Finally, the RST-DT multinuclear relations Cause-result and Consequence are similar and can correspond to either REASON or RESULT. We only counted them in the total column (see † in Table 5).

	RST-DT	Annodis	RST-ST	Total
Obj.	2,698	525	1,444	6,995
Part.	176	614		
Gen.	884			
Sum.	83		8	
Rest.	140		22	
Means	226		175	
Frame		225		225
Backg.	937	157	344	1,438
Attr.	3,070	74		3,144
Total THEMATIC: **11,802**				

Table 4: Thematic class mapping

	RST-DT	Annodis	RST-ST	Total
Reason	52	128 (+)	77	1245†
Result	159	162 (*)	193	
Cond.	285	20	53	406
Excep.	43		5	
Purp.	568	94	127	789
Motiv.	206	cf. (*)	28	234
Evid.	780	cf. (+)	98	878
Eval.	600	75	99	774
Cont.	352	143	58	1,378
Conc.	293		50	
Antith.	402		80	
Total CAUS.ARG.: **5,704**				

Table 5: Causal-Argumentative class mapping

5 Discussion and conclusions

We have presented a unified taxonomy for discourse relations which can be used to map existing annotations, and to annotate new corpora. We believe our taxonomy is robust across theoretical approaches, and can be applied to multiple languages.

We would like to mention some further aspects of our work. First of all, this is not (or not yet) an annotation task. Therefore, we do not provide any kappa values, or any other measure of agreement across annotators. We believe the original corpora were sufficiently tested in that regard, and we trust that their agreement is good enough for us to map from the corpora. Secondly, we have skirted the issue of signalling of relations, in particular by discourse markers. How to recognize relations, by the average hearer/reader, or by annotators, is an active topic of research (Koornneef and Sanders, 2013).

A number of issues are outstanding, the first with regard to segmentation. Different corpora have followed different segmentation methodologies, sometimes impacting the types of relations present in the taxonomy, as is the case with the multiple subtypes of Elaboration relations in the RST-DT corpus. Our intention is to provide coarse and fine-grained segmentation options, so that either can be adopted, depending on the goals of the research.

More crucial to the task of unifying annotations is the issue of the structure of the discourse. RST, RST-DT and RST-ST all take trees as the fundamental structure. SDRT, however, postulates graphs as the basic structure. On the other hand, mapping to a lexically-grounded approach, like that in the PDTB, is certainly possible. We believe that mapping and predicting relations can be a theory-independent task, and that the rich annotations in PDTB are useful for tasks such as discourse parsing.

Two further practical aspects remain unresolved. First, relations with no correspondence across taxonomies need to be considered. One solution is to ignore them and then predict a partial structure for some texts. The second issue is the task of manual annotation. As we have pointed out, in cases where one original taxonomy is more detailed, mapping relations from that taxonomy onto an existing corpus may require further annotation. If annotation is undertaken, then that could solve our first practical problem, because then there is an opportunity to annotate some of the relations with no mapping.

Finally, an excellent test of the usefulness of the taxonomy would be to carry out experiments in discourse parsing. We would like to merge annotated corpora, and test whether the larger size of the training data improves the results of a discourse parser.

References

[Afantenos et al.2012] Stergos Afantenos, Nicholas Asher, Farah Benamara, Myriam Bras, Cecile Fabre, Mai Ho-Dac, Anne Le Draoulec, Philippe Muller, Marie-Paule Pery-Woodley, Laurent Prevot, Josette Rebeyrolles, Ludovic Tanguy, Marianne Vergez-Couret, and Laure Vieu. 2012. An empirical resource for discovering cognitive principles of discourse organisation: The ANNODIS corpus. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*.

[Asher and Lascarides2003] Nicholas Asher and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.

[Asher et al.2006] Nicholas Asher, Julie Hunter, Pascal Denis, and Brian Reese. 2006. Evidentiality and intensionality: Two uses of reportative constructions in discourse. In *Workshop on Constraints in Discourse*.

[Carlson et al.2003] Lynn Carlson, Daniel Marcu, and Mary Ellen Okurowski. 2003. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In Jan van Kuppevelt and Ronnie Smith, editors, *Current Directions in Discourse and Dialogue*, pages 85–112. Kluwer.

[Chiarcos2014] Christian Chiarcos. 2014. Towards interoperable discourse annotation. discourse features in the ontologies of linguistic annotation. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*.

[da Cunha et al.2011] Iria da Cunha, Juan-Manuel Torres-Moreno, and Gerardo Sierra. 2011. On the Development of the RST Spanish Treebank. In *Proceedings of the 5th Linguistic Annotation Workshop*, pages 1–10.

[Degand2009] Liesbeth Degand. 2009. Describing polysemous discourse markers: What does translation add to the picture? In Stef Slembrouch, Miriam Taverniers, and Mieke Van Herreweghe, editors, *From will to well. Studies in Linguistics offered to Anne-Marie Simon-Vandenbergen*. Academia Press, Gent.

[Feng and Hirst2014] Vanessa Wei Feng and Graeme Hirst. 2014. A linear-time bottom-up discourse parser with constraints and post-editing. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 511–521.

[Hardmeier2013] Christian Hardmeier. 2013. Discourse in statistical machine translation: A survey and a case study. *Discours*, 11.

[Heerschop et al.2011] Bas Heerschop, Frank Goossen, Alexander Hogenboom, Flavius Frasincar, Uzay Kaymak, and Franciska de Jong. 2011. Polarity analysis of texts using discourse structure. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pages 1061–1070.

[Hernault et al.2010] Hugo Hernault, Helmut Prendinger, David A. duVerle, and Mitsuru Ishizuka. 2010. HILDA: A discourse parser using Support Vector Machine classification. *Dialogue and Discourse*, 1(3).

[Joty et al.2015] Shafiq Joty, Giuseppe Carenini, and Raymond Ng. 2015. CODRA: A novel discriminative framework for rhetorical analysis. *Computational Linguistics*, page in press.

[Keskes et al.2014] Iskandar Keskes, Farah Benamara, and Lamia Belguith Hadrich. 2014. Learning Explicit and Implicit Arabic Discourse Relations. *Journal of King Saud University Computer and Information Sciences: Special Issue on Arabic NLP: Current State and Future Challenges*.

[Koornneef and Sanders2013] Arnout W. Koornneef and Ted Sanders. 2013. Establishing coherence relations in discourse: The influence of implicit causality and connectives on pronoun resolution. *Language and Cognitive Processes*, 28:1169–1206.

[Mann and Thompson1988] William C. Mann and Sandra A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

[Prasad et al.2008] Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The Penn Discourse Treebank 2.0. In *6th International Conference on Language Resources and Evaluation (LREC)*.

[Prasad et al.2014] Rashmi Prasad, Bonnie Webber, and Aravind Joshi. 2014. Reflections on the penn discourse treebank, comparable corpora, and complementary annotation. *Computational Linguistics*, 40(4):921–950.

[Reese et al.2007] B. Reese, J. Hunter, P. Denis, N. Asher, and J. Baldridge. 2007. Reference manual for the analysis and annotation of rhetorical structure. Technical report, Department of Linguistics, The University of Texas, Austin.

[Thione et al.2004] Gian Lorenzo Thione, Martin Van den Berg, Livia Polanyi, and Chris Culy. 2004. Hybrid text summarization: Combining external relevance measures with structural analysis. In Stan Szpakowicz Marie-Francine Moens, editor, *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*.

[Venant et al.2013] Antoine Venant, Nicholas Asher, Philippe Muller, Pascal Denis, and Stergos Afantenos. 2013. Expressivity and comparison of models of discourse structure. In *Proceedings of the SIGDIAL 2013 Conference*, Metz, France, August.

[Zufferey and Degand2014] Sandrine Zufferey and Liesbeth Degand. 2014. Annotating the meaning of discourse connectives in multilingual corpora. *Corpus Linguistics and Linguistic Theory*, page in press.