Contrasting and Corroborating Citations in Journal Articles

Adam Meyers
New York University
meyers@cs.nyu.edu

Abstract

Automatic recognition of CORROBORATE and CONTRAST relations between citations may enhance citation analysis. We describe a system that identifies these citation relations using predicate/argument and discourse structures.

1 Introduction

The citation of publications has been used to measure the impact of authors, publications, publishers, fields of study, etc., as represented in graphs in which nodes represent documents and edges connect documents if one refers to the other (citation graphs) or if a third document refers to both (co-citation graphs). NLP may be used to supplement these graphs with information about why documents are cited. While previous work (Teufel et al., 2009; Athar, 2011; Athar and Teufel, 2012) record positive and negative sentiment about cited work, we record information about how cited documents are compared to each other. CONTRASTing documents may describe different approaches or different opinions. Documents which CORROBORATE each other may follow a single approach. A document that is cited as corroborating with many other documents may be very salient. Some example instances of these CONTRAST/CORROBORATE relations are provided in Figure 1, the document containing the citations (represented as *we* or *this study*) contrasts with [1] and corroborates [2]; [3] and [4] contrast with each other; and [5] and [6] corroborate.

We use square brackets and numbers (IEEE style) to represent citations that are the object of this study (Figure 1). We use last name plus date (APA style) for works cited as part of this research effort. Examples in this paper (modified for brevity) are from the PubMed Central corpus[1]. As

1. In contrast to **[1]**, **we** found clear detrimental effects of prophylaxis.
2. **This study** corroborates a study **[2]** finding no evidence of cross-hemisphere invasions.
3. FluA and FluB viruses have a common origin **[3]**. Thus, it is expected that aa residues of PA are conserved between FluA and FluB **[4]**.
4. Some species shed lots of virus, yet suffer few damaging effects **[5]**. On the other hand, species of swan, show 100% mortality within days of inoculation with HPAIV (H5N1) **[6]**.

Figure 1: CORROBORATE and CONTRAST

in Figure 1, arguments of relations are in bold and signals are underlined. Other important elements (Figure 3) have boxes drawn around them.

In this paper, we explore the relations between sentence-internal predicate relations, intersentential discourse relations and relations between citations. Then we describe and evaluate a system which derives CORROBORATE and CONTRAST relations between citations.

2 How Citation Relations are Encoded

2.1 The phrase/citation connection

(Abu Jbara and Radev, 2012) describes a system for identifying the **referential scope** of each citation, a text fragment that the citation is semantically related to–multiple citations can share the same referential scope. We assume that arguments of grammatical and discourse relations are essentially the referential scopes of the citations that we are concerned with. Like (Abu Jbara and Radev, 2012), we cover 2 cases. In the first case, the citation is an argument of some predicate (Figure 1, citation 1). In the second case, the citation is parenthetically linked to a constituent (Figure 1, ci-

[1] http://www.ncbi.nlm.nih.gov/pmc/

Proceedings of Recent Advances in Natural Language Processing, pages 460–466,
Hissar, Bulgaria, 7-13 September 2013.

tations 2, 3 and 4) ((Abu Jbara and Radev, 2012) refers to this case as non-syntactic). We found the second case to be more common than the first for CORROBORATE and CONTRAST.

Authors use *we*, *our*, *this research*, and other phrases which we call **self-citations** to refer to their own work. These self-citations participate in the same citation relations as conventional citations. Thus, CORROBORATE and CONTRAST relations can have a self-citation arguments, e.g., Figure 1, **we** and **This study** in ex. 1 and 2. Self-citations occur in regular noun phrase positions (subject, object, etc.), but not parenthetically.

2.2 Citations, Discourse and Grammar

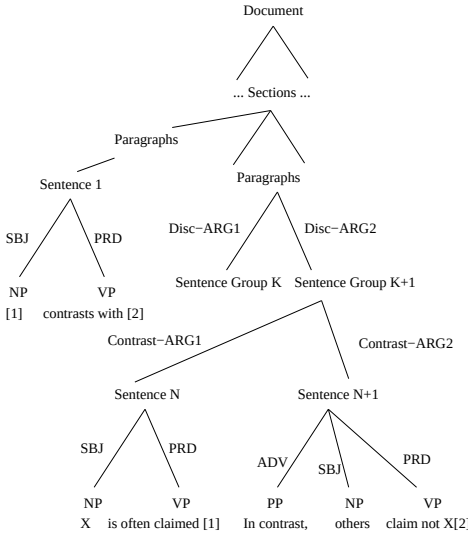

Figure 2: Constituent Structure of a Document

In Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) and related approaches (Marcu, 2000), the discourse structure of a document forms a tree, with the root representing the document; internal nodes representing sections, paragraphs and multi-clause sentences, and leaves representing single clauses. The edge labels on the set of outgoing branches from a node collectively represent relations among the children of that node. As in Figure 2, substituting the leaves of a discourse tree with predicate argument representations (or parses) results in a rooted graph for a document with words as leaves. If the referential scopes of a pair of citations correspond to a pair of siblings at any level in this graph, the relation represented at the parent node can correspond to a citation relation. The sentence "*[1] contrasts with [2] regarding whether X is or is not true*" is analogous to the sentence "*X is often claimed [1]. In contrast, others claim not X [2]*" because the subject and object of the verb *contrast* correspond to the discourse arguments of *in contrast*. Taking this approach, we assume that discourse units and grammatical arguments are the referential scopes of citations. Furthermore, we limit our attention to citations scopes that are no more than a few sentences long (as in the Penn Discourse Treebank (PDTB) (Miltsakaki et al., 2004)).

We assume that: (1) there is a grammatical or discourse relation corresponding to each CORROBORATE and CONTRAST citation relation; (2) each such grammatical/discourse relation takes 2 arguments, each argument being a sequence of sentences, a sentence, a phrase or a word; and (3) more than one citation can be associated with each argument. Given these assumptions we seek to identify: (a) the candidate grammatical/discourse relation and its arguments; and (b) the sets of citations that correspond to these arguments. We then hypothesize the corresponding citation relation for each ARG1/ARG2 pair in the Cartesian product of the set of citations in the ARG1 domain and the set of arguments in the ARG2 domain.

This means that to identify CONTRAST and CORROBORATE citation relations, we need to identify syntactic and discourse signals that would imply that citations hold these relations. In the case of syntactic predicates, these turn out to be a list of words (*contrast, corroborate, endorse, ...*) that are idiosyncratic to this task. For discourse connectives, we can use some previous classifications: causal discourse connectives (*thus, therefore*) tend to be linked to CORROBORATE citation relations; and CONTRAST discourse connectives (*in contrast, on the other hand, however*) tend to be linked to CONTRAST citation relations. While all the cited work on discourse relations posit these same relations for consecutive sentences with no explicit connectives, we have not explored this avenue yet.[2]

2.3 Multi-Sentence Units

We recognize a third DISCOURSE relation, **EXPAND**, which does not directly link to either of our citation relations. Rather, 2 sentences in an EXPAND relation are treated as a single unit, which can, itself be a discourse argument. Fur-

[2](Prasad et al., 2007) reports annotating 16053 implicit and 18459 explicit discourse relations in their corpus.

A) Prior to the HPAI H5N1 virus epidemics, wild bird mortality from AI virus infection had been rare **[7]**, **[8]**.

B) In contrast, HPAIV H5N1 is unusual as high mortality rates have occurred in wild birds [9], [10].

C) Passerine birds have been naturally affected by HPAI H5N1 viruses [11]–[14].

D) Experimental infections of passerine [15], [16] characterized these birds as vulnerable.

E) Additionally, Gronesova et al., 2008 found that 18% of samples from 12 passeriform species tested positive for influenza A viral genome in a surveillance study [17].

Figure 3: Multi-S ARG2: (Kalthoff et al., 2009)

thermore, EXPAND relations can (transitively) link such units to additional sentences, producing larger multi unit chunks. Citations in 2 multi-sentence units can be in CORROBOATE or CONTRAST relations in the same manner as the single sentence cases described above. Figure 3 contains one such example. Sentences B, C, D and E are linked together into one unit by means of 3 EXPAND relations, holding between sentence pairs {B,C}, {C,D} and {D,E}. The discourse connective *In contrast* takes sentence A as one argument and the unit B through E as a second argument. Based on this CONTRAST discourse relation, we deduce that the citations in A (7 and 8) are in contrast with the citations in B through E (9–17), resulting in 18 CONTRAST citation relations.

Our EXPAND discourse relation approximately corresponds to several discourse relations in other frameworks (EXAMPLE, ELABORATION, LIST and others in (Marcu, 2000); CONJUNCTION, INSTANTIATION and others in PDTB). We collapse these relations in order to simplify the task. 2 mechanisms for identifying EXPAND relations both of which are evident in Figure 3: (i) using discourse connectives, e.g, the EXPAND relation between D and E is signaled by the connective *Additionally*; and (ii) based on the *cohesion* between 2 sentences – this is the case for the links connecting sentences {B,C} and {C,D}. Following (Marcu, 2000) (and others), cohesion can be determined by elements that indicate continuity between sentences such as anaphoric words (the demonstrative *these*) or repeated words from the previous sentence. In Figure 3, *birds, have, H5N1* and *viruses* are repeated in C, after occurring in B. D contains the demonstrative *these*, and repeats the word *birds* that is found in C. Our system takes these cohesive signals as evidence that an EXPAND relation holds between 2 consecutive sentences. Our Expand relations are used to approximate the larger citation context. (Athar and Teufel, 2012) uses similar methodology in their citation sentiment system.

3 DocRelate

3.1 Our Approach

Each file in our corpus of PubMed scientific articles has the citations premarked. Preprocessing includes the marking of all sentence and paragraph boundaries. Our citation relation system, DocRelate processes each document from beginning to end, one sentence at a time. All processing is based on regular expressions and simple string matches and is therefore both faster and less accurate than a syntactically-sophisticated approach would be. Nevertheless, we expect that aspects of DocRelate that deal with relations across sentence boundaries to be essentially the same as they would be in systems using deeper processing.

For each sentence, we find: (a) lexical signals; (b) sentence dividers (semi-colons, coordinate/subordinate conjunctions); and (c) citations (conventional and self-citations). For each lexical signal, we establish a clause$_1$ and a clause$_2$. If the lexical signal follows a sentence divider, clause$_1$ is the portion of the current sentence preceding the sentence divider, whereas clause$_2$ follows the sentence divider, e.g., the sentence divider is *and* in

The public considers frequently reported infectious diseases, to be the most severe **[18]** *and therefore people's anxiety correlated with a negative perception of the disease* **[19]**.

Otherwise, the previous and current sentences are clause$_1$ and clause$_2$ (Figure 1 ex. 3 and 4).

We maintain a dictionary of lexical signals, which includes: surface forms, local disambiguating information, part of speech (POS) and CITATION relation (EXPAND, CORROBORATE or CONTRAST). For lexical signals with POS of 'adverb', clause$_1$ is assumed to contain the ARG1 citations and clause$_2$ is assumed to contain the

ARG2 citations: most examples discussed in this paper follow this pattern, e.g., Figure 1 ex. 3 and 4. For other POS: (preposition, verb, adjective and subordinate conjunction (SCONJ)) both arguments are inside clause$_1$. For most sentence-internal cases, the signal divides the clause into 2 parts: citations preceding the signal are candidate ARG1s and those following the signal are ARG2s, e.g., Figure 1 ex. 1 and 2. When an SCONJ occurs at the beginning of a clause, the clause must be divided at a centrally-located comma. Citations preceding that comma are ARG2 candidates, whereas citations following the comma are ARG1s. The 2 cases of SCONJ are:

Case 1 Limited studies have suggested dsRNA is an activator of NLRP3 inflammasome **[20]** although this has been disputed **[21]**.

Case 2 Although influenza strains resistant to NA inhibitors are less prevalent **[22]**, resistance to oseltamivir has been reported **[23]**,**[24]**.

For Case 1, ARG1 is 20 and ARG2 is 21; for case 2, both 23 and 24 are ARG1s and 22 is ARG2. Our approach to SCONJ is essentially the same as that of (Marcu, 2000) among others.

In the absence of discourse connectives, we can hypothesize an EXPAND relation between 2 sentences if the second sentence refers back to the first, as determined by: (a) the proportion of words in the second sentence also occurring in the first, ignoring a list of stop words; (b) the presence of abbreviations in the second sentence corresponding to word sequences in the first; (c) the presence of referring expressions found in the second sentence (*this, these, those, another, it, they, them, itself, themselves, their, here, latter*); and (d) the occurrence of self-citations in both the current sentence and the previous one.

Algorithm 1 is our approach for finding citation relations in an article. After each sentence is processed, citations that are not embedded in a sentence-internal ARG2 are recorded as potential ARG1s for the next sentence (the **cites** function) and each EXPAND relation causes the previous set of ARG1 citations to be stored as well (in St_ARG1). This makes analyses like that of Figure 4 possible: the (ARG1) citations preceding *However* are contrasted with the (ARG2) citations in the sentence containing *However*. The citations in sentence A are stored due to the Expand relation

foreach *sentence in document* **do**
  $S \longleftarrow Sentence$
  $P \longleftarrow PreviousSentence$
  $SLink \longleftarrow DiscRel(S, P)$
  Output Sentence-internal relations
  if $SLink \in \{CONTRA, CORROB\}$ **then**
    $ARG2 \longleftarrow cites(S)$
    $ARG1 \longleftarrow St\_ARG1 \cup cites(P)$
    if *Satisfy_Constraints(ARG1,ARG2)* **then**
      Output SLink Relation for
      $\forall\{a_1, a_2\} \in ARG1 \times ARG2$
    end
    Empty St_ARG1
    if $ARG1 \neq \emptyset$ **then**
      $St\_SLink \longleftarrow SLink$
      $St\_ARG1 \longleftarrow ARG1$
    end
  end
  else if *SLink = EXP* **then**
    $ARG1 \longleftarrow cites(P)$
    $ARG2 \longleftarrow cites(S)$
    if $St\_SLink \neq \emptyset \wedge ARG2 \neq \emptyset$ **then**
      Output: St_Link relation for
      $\forall\{a_1, a_2\} \in St\_ARG1 \times ARG2$
    end
    else
      add cites(S) to st_ARG1
    end
  end
  else
    Empty St_ARG1 and St_SLink
  end
end

Algorithm 1: Identify Citation Relations

between sentences A and B, motivated by the referring expression *These* and abbreviation *HPAIV* (*highly pathogenic avian influenza virus*). When the procedure evaluates sentence C, the citations in A are potential ARG1s. However, as there is not an EXPAND relation between sentences B and C, these potential ARG1s are not stored for connectives in subsequent sentences.

Cross-sentence CONTRAST and CORROBORATE signals (**St_SLink** in Figure 1) are stored in addition to previous ARG1s up to that point. As long as there is a continuous sequence of EXPAND relations linking the subsequent sentences, citations in those sentences can fill the ARG2 slot for **St_SLink**. Figure 3 is one such example: the citations in the the sentence preceding *In contrast* are ARG1s and the citations following the signal are ARG2s: both citations in the sentence and in subsequent sentences. Storage of these elements is emptied in the absence of EXPAND.

We have implemented the following constraints on these procedures: (1) 2 clauses cannot be linked by multiple discourse relations. Conflicts favor the relations CONTRAST and CAUSE over EXPAND (where discourse CAUSE relations imply

A Recently, an H9N2 AIV was isolated from pigs in several provinces in China **[25]**,**[26]**,**[27]**, and a H5N1 HPAIV was identified in pigs in Asian countries **[28]**.

B These observations have led to the conclusion that swine can serve as direct and intermediate hosts for many subtypes of AIVs including the HPAIV of the H5 and H7 subtypes.

C However, there is recent evidence that domestic pigs show only low susceptibility to H5N1 HPAIV **[29]**, **[30]**.

Figure 4: Multi-S ARG1: (Ma et al., 2008)

CORROBORATE citation relations). This creates separate multi-sentence units for CONTRAST and CORROBORATE relations, since all storage is emptied in the absence of cross-sentence EXPAND relations; (2) the sets of citations for proposed ARG1 and ARG2 cannot have a member in common – this rules out relations that do not make sense (a document contrasting with itself) or that are uninformative (a document corroborating with itself). may be due to failures

3.2 Lexical Entries for Signals

We manually constructed a dictionary of signals licensing EXPAND, CORROBORATE and CONTRAST relations. The CORROBORATE and CONTRAST entries are signals which license citation relations, the entries being based on their roles in syntax and discourse structure. The EXPAND entries are signals that license EXPAND discourse relations. We have 246 entries for EXPAND, 48 for CONTRAST and 31 for CORROBORATE. This was feasible because there are a small number of these signals that cover most cases. We based our dictionary on previous work. We examined entries in COMLEX Syntax (Macleod et al., 1996) including the 7 coordinate conjunctions, 108 SCONJ, 96 adverbs marked as (META-ADV :CONJ T), and a few other adverb classes as well. We examined the set of discourse connectives marked in PDTB and classified in its manual (Prasad et al., 2007). We also did some manual annotation (unpublished work) and examined files from our training corpus while creating the system. Sample lexical entries (Figure 5) include: base forms, parts of speech (POS), relation licensed, and constraints. Multi-word expressions and POS disambiguation is implemented by requiring or excluding certain words before/after the key words. For example, *contrast* can be a verb; one of the multi-word prepositions {*in contrast with*, *in contrast to*, *by contrast to*}; or one of the multi-word adverbs {*in contrast*, *by contrast*}. Choice of POS for *contrast* determines the relative positions of ARG1 and ARG2, e.g., for the adverb, it is in the previous clause. Another constraint is the sentence-initial requirement, since some adverbs connect clauses when they occur initially, but not when they occur elsewhere in the sentence. For example, sentence-initial *roughly* can introduce a sentence that elaborates some aspect of the previous sentence (EXPAND), e.g., *Roughly, the chance that this would happen was 8 to 1*. However, the non-initial use of *roughly* still means something like *approximately*, but the connection with the previous sentence is no longer there, e.g., *The odds were roughly 8 to 1*.

POS	BASE	Variants	Function
VERB	*support*	+ed/s/ing	CORROB
Constraint: not after *the*|*a*; not before *vector*; not in FUNDING/ACKNOW Section			
VERB	*contrast*	+ed/s/ing	CONTRA
Constraint: before *with*; not after *by*|*in*			
PREP	*contrast*		CONTRA
Constraint: before *with*|*to*; after *in*|*by*			
ADV	*additionally*		EXPAND
ADV	*contrast*		CONTRA
Constraint: after *in*|*by*; not before *with*|*to*			
ADV	*roughly*		EXPAND
Constraint: sentence-initial only			
ADV	*thus*		CORROB

Figure 5: Sample Lexical Entries

3.3 System Evaluation

We ran DocRelate on a 20 document held-out test corpus. Figure 6 represent 216 correct answers out of 291 relations in the answer key (manual annotation by the author after the system was completed). We evaluated (CORROBORATE, CONTRAST) relations between citations, but not discourse relations between sentences (CAUSE, CONTRAST, EXPAND) or predicate argument relations.

4 Concluding Remarks

We achieve the highest accuracy for relations linking citations across adjacent sentences. Long-

Relation	Instances in Answer Key	All		Same Sentence		Next Sentence		2 or More Apart	
		Prec	Rec	Prec	Rec	Prec	Rec	Prec	Rec
Contra	156	90%	67%	75%	71%	99%	69%	53%	47%
Corrob	135	94%	83%	83%	57%	93%	92%	100%	79%
All	291	92%	74%	88%	76%	95%	66%	92%	74%
Instances in Answer Key		291		102		157		75	

Figure 6: Precision/Recall for Citation Relations

1. Dynamic models of epidemics are widely accepted **[31]**. So stochastic methods have emerged as the best way to model infectious diseases data **[32]**. [Missing CORROBORATE: *so* **NOT in lexicon**]

2. Age is not considered in **our model**, though it may affect behavior and, thus, risk of becoming infected [33]. [Marked CORROBORATE (*thus*) instead of CONTRAST (*though*)]

3. Fibroblasts transfected with ANGPT1 reduced expression of endothelial-selective adhesion molecules **[34]**. However, in **these studies** gene transfer was performed prior to lung injury. [Incorrect CONTRAST: *these studies* is not self-citation]

Figure 7: Example Sources of Error

distance citation relations were more difficult because: (a) they depend on additional (EXPAND) discourse relations; and (b) their relative rarity posed a challenge for evaluation (there were 17 contrast and 58 corroborate long-distance relations). While the single-sentence case is similar to the 1-sentence-apart case, our results for the latter case were lower because: (a) the inventory of same-sentence signals is larger and many are missing from our lexicon; (b) these signals are less reliable; and (c) our pattern-based approximations of syntactic rules did not work for the sentence-internal case – using parsing based rules would have helped. Our false negatives exceed our false positives. We observed errors due to the following: (1) missing entries in our dictionary; (2) defects in our sentence-internal syntactic analysis; and (3) false positives for self-citations. Some sample errors are provided as figure 7.

We presented an analysis of how authors of technical documents depict corroborations and contrasts between documents We have presented a syntactically naive system, that accounts for most aspects of this analysis. We showed that it was possible in many cases to derive relations between citations from predicate and discourse relations among the constituents that those citations link to. Our current system achieves accuracy of 92% precision and 74% recall for CORROBORATE/CONTRAST relations, with some variation based on relation type and the distance between the citations in terms of sentences. for citations not in adjacent sentences. The main contribution of this paper is the working out of the details of how to identify citation relations. Towards this goal, we described a robust system using simple, manually written string-based rules. In future work, we plan to identify properties of additional discourse structures that impact the problem of identifying citation relations. It is likely that a more elaborate system would achieve better results. Such systems could include features based on parsing, semantic role labeling and other text processing, thus making more precise rules available. Systems based on Machine Learning approaches could also be created based on the features described here, as well as text-processing-based features.

Acknowledgments

Supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D11PC20154. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government. Approved for Public Release; Distribution Unlimited.

References

A. Abu Jbara and D. Radev. 2012. Reference scope identification in citing sentences. In *Proc. of NAACL:HLT 2012*.

A. Athar and S. Teufel. 2012. Detection of implicit citations for sentiment detection. In *ACL 2012 Workshop on Detecting Structure in Scholarly Discourse*, pages 18–26.

A. Athar. 2011. Sentiment analysis of citations using sentence structure-based features. In *Proc. of the ACL 2011 Student Session*.

D. Kalthoff, A. Breithaupt, B. Helm, J.P. Teifke, and M. Beer. 2009. Migratory status is not related to the susceptibility to hpaiv h5n1 in an insectivorous passerine species. *PLoS One*, 4(7):e6170.

W. Ma, R.E. Kahn, and J.A. Richt. 2008. The pig as a mixing vessel for influenza viruses: Human and veterinary implications. *J Mol Genet Med*, 3(1):158–66.

C. Macleod, A. Meyers, and R. Grishman. 1996. COMLEX Syntax: An On-Line Dictionary for Natural Language Processing. In *Proceedings of Euralex96*.

W. C. Mann and S. A. Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

D. Marcu. 2000. *The Theory and Practice of Discourse Parsing and Summarization*. The MIT Press, Cambridge.

E. Miltsakaki, A. Joshi, R. Prasad, and B. Webber. 2004. Annotating discourse connectives and their arguments. In A. Meyers, editor, *NAACL/HLT 2004 Workshop: Frontiers in Corpus Annotation*, pages 9–16, Boston, Massachusetts, USA, May 2 - May 7. Association for Computational Linguistics.

R. Prasad, E. Miltsakaki, N. Dinesh, A. Lee, and A. Joshi. 2007. The Penn Discourse Treebank 2.0 Annotation Manual. `www.seas.upenn.edu/~pdtb/PDTBAPI/pdtb-annotation-manual.pdf`.

S. Teufel, A. Siddharthan, and C. Batchelor. 2009. Towards domain-independent argumentative zoning: Evidence from chemistry and computational linguistics. In *EMNLP-2009*.