

# Deriving rhetorical relationships from semantic content

**Richard Power**

Department of Computing  
Open University, UK  
r.power@open.ac.uk

## Abstract

This paper investigates to what extent rhetorical relations can be assigned purely on the basis of propositional content, without any reference to speaker goals or other pragmatic information. This task confronts any NLG system designed to generate coherent text from a set of formally represented statements; we consider it here in the context of an ontology verbaliser, for which the input is a set of axioms encoded in the web ontology language OWL. A simple set-theoretical model of the possible semantic relationships between two statements is proposed; this model allows 46 logically consistent relationships, of which we hypothesise that 11 are rhetorically coherent. This hypothesis is tested through an empirical survey which also provides evidence on how the coherent patterns are expressed linguistically.

## 1 Introduction

Perhaps the murkiest area in the language sciences is the issue of how statements are combined in a discourse. Much research has been based (more or less strictly) on Rhetorical Structure Theory (RST), (Mann and Thompson, 1987), a theory grounded in intuitions about naturally occurring texts and more concerned with comprehensive coverage than formal adequacy. Categories like ‘concession’ and ‘elaboration’ have to be assigned through human judgement, and remain somewhat subjective despite efforts to refine them and clarify their definitions (Carlson and Marcu, 2001). Other researchers have

looked more deeply into the meaning of the relations, analysing them through rhetorical features (Hobbs, 1985; Sanders et al., 1992; Knott, 1996) with more emphasis on theory than on the requirements of practical annotation.

In this paper we attack the problem from a new direction. Instead of starting from naturally occurring texts, and human judgements thereupon, we consider the far more restricted issue of how a rhetorical relationship could be assigned to two axioms drawn from an ontology, and hence to the sentences generated from these axioms by an ontology verbaliser,<sup>1</sup> using only information that is internal to the ontology. This means that we accept the strict limitations of the ontology formalism, assumed to be OWL-DL (Horrocks et al., 2003); statements that cannot be represented in this formalism are excluded. It also means that the ontology is the *only* source of knowledge about the domain, and that no pragmatic information is available at all, beyond the implicit fact that each axiom has been asserted. This is precisely the situation that confronts an NLG (Natural Language Generation) system that aims to generate a coherent text from an OWL ontology, using only generic methods (i.e., methods that require no additional domain knowledge). How can such a system decide whether two statements from the ontology are related, and if so, classify the relationship in a way that guides their linguistic realisation?

An example will clarify both the exact task, and how it might be approached. Suppose that an on-

---

<sup>1</sup>Examples of ontology verbalisers are SWAT Tools (SWAT Project, 2011), described by Williams et al. (2011), ACE (Attempto Project, 2011), and OntoVerbal (Liang et al., 2011).

	OWL statement	Example of verbalisation
1	$ClassAssertion(C,I)$	Butch is a dog
2	$ObjectPropertyAssertion(P,I,J)$	Mary owns Butch
3	$ClassAssertion(ObjectSomeValuesFrom(P,C),I)$	Butch lives in a kennel
4	$SubClassOf(C,D)$	Every dog is a canine
5	$SubClassOf(C, ObjectHasValue(P,I))$	Every dog likes Mary
6	$SubClassOf(C, ObjectSomeValuesFrom(P,D))$	Every dog lives in a kennel
7	$DisjointClasses(C,D)$	No dog is a cat
8	$EquivalentClasses(C,D)$	A dog is defined as a domestic canine

Table 1: Common axiom patterns in OWL. A study of over 200 ontologies indicated that these patterns comprise over 95% of all axioms (Power and Third, 2010). Variables  $C$ ,  $D$  denote classes,  $I$ ,  $J$  denote individuals, and  $P$  denotes a property.

tology contains the axioms in table 1, and that we are interested in the relationship between the axioms numbered 6 and 3:

Every dog lives in a kennel.  
Butch lives in a kennel.

Even within our restricted formulation of the task, there are two sources of evidence that can be exploited here. First, confining our attention to these two statements alone, we may note that they share the same predicate term ‘lives in a kennel’ (corresponding to a constructed class in OWL). On this basis we might presume that the statements are related, and propose a neutral method of linking them such as ‘Every dog lives in a kennel; so does Butch’. However, we can go further if we exploit the second source of evidence, *the other statements in the ontology* — and in particular statement 1 which connects the terms ‘Butch’ and ‘dog’. Taking this into account, we could interpret the second statement in our pair (3) as an implication of the first statement (6), or perhaps as an example of it:

Every dog lives in a kennel; therefore so does Butch.  
Every dog, including Butch for example, lives in a kennel.

The purpose of this paper is to model systematically the patterns of relationship among the terms in two statements, and to show through an empirical study which pairs are judged to be rhetorically related, and which are not. For pairs judged to be related, we also present evidence on how the statements could be combined linguistically (e.g., using aggregation and/or discourse connectives).

## 2 Coherence model

We begin by constructing a simple model which covers OWL statements based on three axiom functors: *ClassAssertion*, *ObjectPropertyAssertion*, and *SubClassOf*.<sup>2</sup> The commonest patterns are shown in axioms 1-6 of table 1, along with sample English realisations conforming to most verbalisers (Kaljurand and Fuchs., 2007; Hart et al., 2008; Schwitter and Meyer, 2007).

For the axioms considered, we can give a simple uniform semantics in which each statement links two sets, one denoted by the subject, the other by the predicate; the meaning of the statement is that the predicate set contains the subject set. To accommodate individuals within this scheme we can replace them by enumerated classes with only one member (in OWL these can be constructed using the functor *OneOf*). Thus ‘Butch is a dog’ means that the set containing only Butch is a subset of the set of dogs; ‘Butch lives in a kennel’ means that the set containing only Butch is a subset of the set of things that live in kennels, and so forth. Both statements in a pair can then be reduced to a pair of sets  $SP$ , where  $S$  is the subject set and  $P$  is the predicate set, the structure of the pair being  $S_1P_1 + S_2P_2$ .<sup>3</sup> With four sets we now have six potential relationships to consider:  $S_1P_1$ ,  $S_1S_2$ ,  $S_1P_2$ ,  $P_1S_2$ ,  $P_1P_2$ , and  $S_2P_2$ . Two of these ( $S_1P_1$ ,  $S_2P_2$ ) correspond to the original statements; the other four may be addressed elsewhere in

<sup>2</sup>Elsewhere (Power and Third, 2010) we have shown that these functors cover around 80% of all axioms.

<sup>3</sup>Note that this semantics is derived from the underlying OWL formulas, and would not be applicable to some sentences in English (e.g., ones expressing existential statements such as ‘At least one dog likes Mary’).

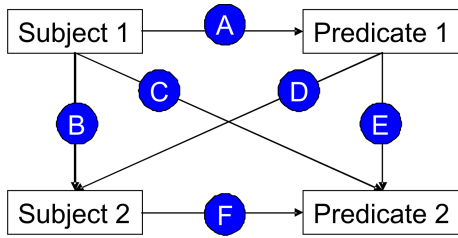


Figure 1: Subject-Predicate relations for two statements

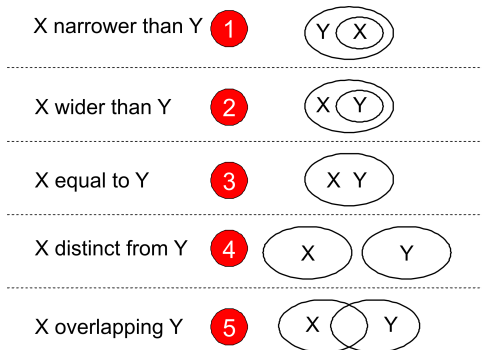


Figure 2: Relations among two sets

the ontology, thus providing additional information on whether and how the statements are rhetorically related. The six relationships are shown diagrammatically in figure 1 by the arrows labelled A–F.

The next question is how these relationships among sets should be classified. Among various possibilities, a plausible method is shown in figure 2: given two sets  $X$  and  $Y$ , either  $X$  will be narrower than  $Y$ , or wider, or equal, or distinct, or overlapping. These relations are represented in OWL as follows: (1) narrower by *SubClassOf*( $X,Y$ ); (2) wider by *SubClassOf*( $Y,X$ ); (3) equal by *EquivalentClasses*( $X,Y$ ); (4) distinct by *disjointClasses*( $X,Y$ ); and (5) overlapping, implicitly, by absence of the above. A similar set of relations has been proposed by MacCartney and Manning (2009) for the textual entailment task.<sup>4</sup>

With this model, the rhetorical relationship between two statements can be profiled by assigning an integer from 1–5 (figure 2) to each of the relation-

<sup>4</sup>MacCartney and Manning actually use seven relations, because they distinguish as a separate case disjoint and overlap relations in which the classes  $X$  and  $Y$  cover all entities in the domain (i.e., every entity must belong either to  $X$  or to  $Y$  or both). This refinement is not relevant for our purposes.

ships A–F (figure 1); to represent such assignments succinctly we will use a six-number code such as 131231 meaning  $A=1, B=3, C=1, D=2, E=3, F=1$ . If we assume that subjects are always narrower than predicates, two of these relations (A and F in figure 1) will always be 1. This leaves a potential  $5^4$  or 625 combinations for the other four relations (B to E in figure 1). However, most of these combinations are contradictory; by writing a Prolog program<sup>5</sup> which applies consistency constraints, we have shown that the consistent combinations number only 46 (Power, 2011a). These are presented with handcrafted examples suggesting that some of the patterns are rhetorically coherent, while others, although logically consistent, are not. On the basis of these examples, the author judged that 11 pairs out of 46 were rhetorically related and 35/46 were not. The list was then given to two colleagues who picked out exactly the same eleven patterns, illustrated in table 2. In this table the patterns are also grouped, and given names which we hope are intuitively easier to grasp than their codes. On inspection, it turns out that a simple rule explains our selections: we judged a pattern coherent either if the two statements had a set in common (i.e., if the cross-statement relations B–E contained relation 3), or if all cross-statement relations were disjoint (i.e., 144441).

### 3 Empirical validation

The empirical study described in this section has two aims. First, we seek firmer evidence regarding the division of the 46 logically consistent patterns into coherent and incoherent (i.e., rhetorically related and unrelated). However intuitive this division, it is interesting not only to confirm it, but to see whether there are degrees of coherence both within the sheep (so to speak) and the goats. Secondly, where people judge that two statements have sufficient affinity to be presented together, we are interested in how they combine them linguistically, and whether each pattern is associated with characteristic discourse connectives or syntactic configurations.

To generate examples for testing, it is convenient to construct an ontology that contains just enough material to produce at least one example for each

<sup>5</sup>The program can be downloaded from the website at Power (2011b).

N	Code	Name	Example
1	131211	Widening Elaboration	Dogs are canines; dogs are vertebrates
2	131221	Narrowing Elaboration	Dogs are vertebrates; dogs are canines
3	131251	Additive Elaboration	Dogs are canines; dogs are domestic mammals
4	111231	Widening Comparison	Dogs are vertebrates; canines are vertebrates
5	121231	Narrowing Comparison	Canines are vertebrates; dogs are vertebrates
6	141231	Disjoint Comparison	Dogs are vertebrates; cats are vertebrates
7	151231	Additive Comparison	Canines are vertebrates; domestic mammals are vertebrates
8	111311	Forward Reasoning	Dogs are canines; canines are vertebrates
9	123221	Backward Reasoning	Canines are vertebrates; dogs are canines
10	144441	Contrast	Dogs are canines; cats are felines
11	131231	Restatement	Dogs are canines; dogs are canines

Table 2: Classification of the coherent patterns

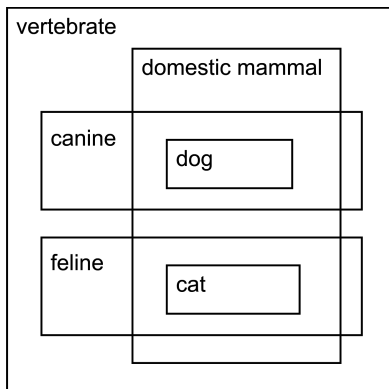


Figure 3: Minimal ontology for coherent patterns

pattern. For the eleven patterns hypothesised to be coherent, the *minimal* such ontology is shown diagrammatically in figure 3. The important feature of this diagram is not the names of the classes, but their relationships; by varying the names it would be possible to generate test examples in different domains. Note that to generate examples for all 46 consistent patterns, we would have to add more classes, the main reason being that incoherent patterns like 155551 require several classes that partially overlap one another (corresponding to weakly related concepts). However, using only the minimal ontology it is possible to generate 10 examples that were *not* selected as coherent. It is therefore convenient to test the proposed coherence partition using only material generated from the minimal ontology: this ensures that the concepts used in all patterns are as similar as possible, and also yields two groups of roughly equal size. In fact, by eliminating the arguably trivial restatement pattern, in which the two statements

in the pair are exactly the same, we obtain exactly ten patterns in the group presumed coherent, and ten in the group presumed incoherent; all of these patterns are shown in table 3. To save space this table uses an abbreviated wording in the ‘Example’ column; the wording actually used is illustrated in figure 4, with ‘Dogs’ replaced by ‘A dog’, a formulation preferred by subjects in an evaluation of the SWAT verbaliser (Stevens et al., 2011).<sup>6</sup>

To present each participant with a conveniently brief task (in our experience, anything over five minutes yields a high drop-out rate), two surveys were compiled from the patterns in table 3, each composed of five patterns from the coherent group and five from the incoherent group, arbitrarily selected and then arranged in a random order (the same for all subjects doing a given survey). Survey I was sent to the SIGDIAL mailing list, Survey II to the SIGGEN list. When uptake proved much greater for Survey II, we also sent Survey I to a local departmental list, and invited people on the SIGGEN list to do Survey I as well as (or instead of) II; since the questions were all different, no duplication resulted if a participant did both surveys. Overall 45 participants completed Survey I and 52 completed Survey II.<sup>7</sup>

A snapshot from Survey I is shown in figure 4.

<sup>6</sup>The issue of how best to word a universal statement requires further research. ‘Every X is a Y’ is perhaps most precise, but sometimes sounds unnatural; ‘Xs are Ys’ and ‘an X is a Y’ are more natural but more open to other interpretations. For the statements in the survey we assume it was obvious that a generic interpretation was intended, and no subjects commented that the sentences were ambiguous or in any way unclear.

<sup>7</sup>It can therefore be inferred from table 3 which questions belonged to which survey.

N	Code	Freq	%	Example
1	WiEl 131211	25/45	56%	Dogs are canines; dogs have backbones
2	NaEl 131221	37/52	71%	Dogs have backbones; dogs are canines
3	AdEl 131251	30/45	67%	Dogs are canines; dogs are domestic mammals
4	WiCp 111231	37/45	82%	Dogs have backbones; canines have backbones
5	NaCp 121231	34/45	76%	Canines have backbones; dogs have backbones
6	DiCp 141231	50/52	96%	Canines have backbones; felines have backbones
7	AdCp 151231	24/52	46%	Canines have backbones; domestic mammals have backbones
8	FwRe 111311	51/52	98%	Dogs are canines; canines have backbones
9	BwRe 123221	42/52	81%	Canines have backbones; dogs are canines
10	CoRe 144441	43/45	96%	Dogs are canines; cats are felines
Total		373/485	77%	
11	Incoh 111511	0/45	0%	Dogs are canines; domestic mammals have backbones
12	Incoh 125221	9/52	17%	Canines have backbones; dogs are domestic mammals
13	Incoh 141211	2/45	4%	Dogs are domestic mammals; cats have backbones
14	Incoh 141221	4/52	8%	Dogs have backbones; cats are domestic mammals
15	Incoh 141411	2/45	4%	Dogs are canines; felines have backbones
16	Incoh 141451	12/52	23%	Dogs are canines; cats are domestic mammals
17	Incoh 141511	0/45	0%	Dogs are domestic mammals; felines have backbones
18	Incoh 144221	2/52	4%	Dogs have backbones; cats are felines
19	Incoh 144251	2/45	4%	Dogs are domestic mammals; cats are felines
20	Incoh 145221	3/52	6%	Canines have backbones; cats are domestic mammals
Total		36/485	7%	

Table 3: Coherence judgements for each pattern. Subjects were asked to judge whether the statements in each pair could be appropriately presented together. The data are the number of ‘Yes’ responses to this question. Patterns 1–10 were hypothesised coherent, patterns 11–20 incoherent.

**Question 1**

\* In this example, "canine" is a family that includes wolves, foxes etc. as well as dogs.

Do you think it would be appropriate to link the following statements by presenting them in the same sentence, or consecutive sentences, in the order given?

**(a) A dog belongs to the canine family (b) A dog is a domesticated mammal**

Yes

No

If your answer is yes, how would you combine the two statements (possibly adding words to clarify their relationship)?

Figure 4: First question in Survey I

Participants were asked to judge whether it would be appropriate to link the two statements in a text (in the given order), by presenting them either in the same sentence or in consecutive sentences; if they answered this question in the affirmative, there was an optional follow-up question asking them to indicate, by typing freely into a text box, how they might combine them. To score these responses, we counted four features:

**And:** The statements were combined neutrally using ‘and’, or a full stop or a semicolon, without any discourse connective.

**Con:** A discourse connective was employed (possibly in addition to ‘and’).

**Agg:** Either the subject or predicate terms of the statements were aggregated.

**Rel:** One statement was expressed as a relative clause inside the other.

The resulting counts for the coherent patterns are shown in table 4. Frequencies for specific discourse connectives (excluding ‘and’) are shown in table 5.

Pattern	Connectives
Widening Elaboration	therefore (4), hence (1), so (1), which means that (1)
Narrowing Elaboration	also (2), because (2), in addition (2), more specifically (2), furthermore (1), moreover (1)
Additive Elaboration	however (1), therefore (1)
Widening Comparison	because (6), in fact (5), like (5), as (4), as do (2), since (2), so (2), as does (1), for example (1), in general (1), as well as (1), more generally (1), therefore (1)
Narrowing Comparison	therefore (8), so (5), hence (4), for example (3), as (2), like (2), as does (1), as well (1), in particular (1), including (1), ipso facto (1), such as (1)
Disjoint Comparison	also (6), too (6), as well as (2), so does (2), just like (1), similarly (1)
Additive Comparison	as do (2), actually (1), also (1), for (1), in general (1), like (1), so do (1), too (1)
Forward Reasoning	also (1), therefore (1)
Backward Reasoning	for example (5), e.g. (2), example is (2), as (1), by the way (1), therefore (1)
Contrast	whereas (11), while (10), but (3), however (1), as (1)

Table 5: Connectives suggested for each pattern, with their frequencies

Pattern	And	Con	Agg	Rel
Widening Elaboration	14	7	8	3
Narrowing Elaboration	27	10	9	0
Additive Elaboration	19	2	6	11
Widening Comparison	2	32	12	2
Narrowing Comparison	2	30	9	1
Disjoint Comparison	30	18	22	0
Additive Comparison	12	9	10	0
Forward Reasoning	36	2	2	1
Backward Reasoning	23	12	0	2
Contrast	16	26	0	0

Table 4: Frequencies of various devices for combining the statements in the ten coherent patterns presented. ‘And’ = Linked only by ‘and’ or punctuation; ‘Con’ = Connective; ‘Agg’ = Aggregation; ‘Rel’ = Relative clause.

## 4 Analysis of results

### 4.1 Coherent and incoherent

The first question is whether the results confirm our intuitive classification of the patterns into coherent and incoherent. Table 3 demonstrates clearly that they do. Summing across all subjects, we obtained 373/485 (77%) positive responses for patterns that satisfied our coherence criterion (upper half of table 3), compared with 36/485 (7%) positive responses for patterns that did not satisfy this criterion (lower half) – obviously a highly significant association<sup>8</sup>. Overall, judgements were fairly evenly divided be-

<sup>8</sup>On a 2x2  $\chi^2$  test for association between pattern (coherent vs incoherent) and judgement (positive vs negative) we obtain  $\chi^2=480$  with  $df=1$ , two-tailed  $p < 0.00001$ .

tween positive and negative, with 409 ‘Yes’ answers against 561 ‘No’ answers.

Looking in more detail at the coherent group, we found clear differences in degree, with several patterns obtaining positive responses of 95% and over, with others not far above the 50% level (and one just below). On a two-tailed binomial test assuming equal *a priori* probabilities for ‘Yes’ and ‘No’, frequencies over 70% are significant at the  $p < 0.01$  level and frequencies over 75% at the  $p < 0.001$  level; thus we have three patterns (widening elaboration, additive elaboration, additive comparison) for which there is not a clear consensus that the statements are related closely enough to be combined in a discourse.

### 4.2 Distinctive realisation

The second question is whether we find evidence that the coherent patterns are *distinctive*, as shown by the linguistic devices by which they are combined. Here table 4 shows that the realisation profiles for the ten patterns differ sharply. With relatively few responses these results should be seen only as suggestive, but several trends are already apparent:

- For widening and narrowing comparison, a discourse connective is almost always used; for the other patterns, combinations using only ‘and’ or a full stop are common.
- Conversely, for additive elaboration and forward reasoning a discourse connective is almost never

used; for the other patterns connectives other than ‘and’ are common.

- Aggregation is commonly used for comparisons, and especially for disjoint comparison (e.g., ‘Canines and felines have backbones’).
- Relative clause combinations are commonly used only for one pattern, additive elaboration (e.g., ‘Dogs, which are domesticated mammals, belong to the canine family’).

### 4.3 Discourse connectives

The final question is whether the discourse connectives proposed for the coherent patterns are distinctive, and linked to familiar rhetorical relations such as EVIDENCE and EXAMPLE. Here again the results are only suggestive, but consistent themes do emerge from the subjects’ choices. For widening and narrowing elaboration these choices signal the EVIDENCE relation (‘therefore’, ‘because’) as well as ELABORATION (‘also’, ‘moreover’). For widening and narrowing comparison EVIDENCE is also common, with more signs of sensitivity to generalising or specifying (‘more generally’, ‘in particular’), a rhetorical move somewhat neglected in RST and other theories. For all comparisons, but especially disjoint comparison, the connectives often signal SIMILARITY. Backward reasoning is the only pattern for which choices often signalled EXAMPLE. Finally, choices for our contrast pattern were dominated by ‘whereas’ and ‘while’, marking as one would expect the CONTRAST relation.

## 5 Discussion

### 5.1 Comparison with other approaches

The most similar work, both in spirit and substance, is the taxonomy of coherence relations proposed by Sanders et al. (1992), who also aim to cover a restricted set of relations using relatively precise theoretical concepts. Their fundamental distinction is between *causal* and *additive* relations, where ‘cause’ is defined (oddly) as an *implication* between two discourse segments: thus if one statement implies the other we have a causal relation; if not we have an additive one. Causal relationships are further distinguished by order of presentation: if antecedent precedes consequent the order is *basic*, otherwise *non-*

*basic*. The theory also distinguishes whether the relation is semantic or pragmatic, and whether statements are presented in positive or negative polarity; these features are not distinguished in our model which is restricted to semantic relations and positive polarity. Combining the values of their four features, Sanders et al. list 12 patterns of which three are comparable with ours: (1) Causal-Semantic-Basic-Positive, (2) Causal-Semantic-Nonbasic-Positive, and (3) Additive-Semantic-Positive; the first two are labelled ‘Cause-consequence’ and the third ‘List’.

In our model, the causal-additive distinction is easily made for the elaboration patterns (i.e., those with equivalent subject terms): if the predicate terms are widening or narrowing the relation is ‘causal’, if they overlap it is ‘additive’ (hence our choice of that word). The basic order for elaboration is widening elaboration (e.g., ‘dogs are canines’ implies ‘dogs are vertebrates’); narrowing elaboration is non-basic. For comparison patterns (those with equivalent predicate terms) the same distinctions hold, except that this time the basic order is narrowing comparison, and widening comparison is non-basic. Note however that we find no evidence that the basic order is preferred: on the contrary, positive coherence judgements were more common for the non-basic orders both for the elaboration and comparison patterns (although the differences are not large). We also find quite different realisation profiles for widening elaboration and narrowing comparison (both Causal-Basic in Sanders et al.’s taxonomy), and for narrowing elaboration and widening comparison (both Causal-Nonbasic). In line with Sanders et al. we obtain discourse connectives signalling implication (‘therefore’, ‘since’ etc.) for all these ‘causal’ patterns, but we also obtain connectives signalling generalisation or specification (‘more generally’, ‘in particular’) and exemplification (‘for example’) that depend on our more detailed classification.

Comparing our classification with RST is harder since the approaches are so different. Unlike Sanders et al., RST is not concerned with order of presentation, and has instead an asymmetry in the *importance* of the two statements, most relations having a ‘nucleus’ and a ‘satellite’. At present we have no way of assigning importance levels from the information encoded in an OWL ontology. Regard-

ing coverage, we can informally link our patterns to the following RST relations (Carlson and Marcu, 2001): comparison, contrast, elaboration-additional, elaboration-general-specific, example, and restatement.<sup>9</sup> On the other hand we cover some relations apparently missing from RST, which lacks any notion of co-premise (found in our forward and backward reasoning patterns), or of moving from specific to general or vice-versa (our distinction between widening and narrowing).

## 5.2 Limitations

As already mentioned, the methods proposed here are bounded by characteristics of the ontology verbalisation task: since the OWL standard (Horrocks et al., 2003) lacks any representation of pragmatics, or time, or causal relations between events, or modality, or probability, many relations dependent on these concepts lie outside our compass. However, even within this restriction of coverage, the theory and evaluation described here are far from complete.

Recall first of all that we have covered only those patterns in which the subject of each statement denotes a *subclass* of the predicate (relation number 1 in our code). Thus we cover ‘every dog is a canine’ (dogs are a subclass of canines), but not the following sentence patterns:

- (2) Only canines are dogs (subject is superclass of predicate)
- (3) A dog is defined as a domestic canine (subject and predicate are equivalent)
- (4) No dog is a cat (subject and predicate are disjoint)
- (5) Some pets are canines<sup>10</sup> (subject and predicate overlap)

In verbalising ontologies it would be unnecessary to cover pattern (2), which is merely an awkward inversion of *SubClassOf*, or (5), which is represented in OWL only indirectly. However, patterns (3) and (4) should be covered, since they correspond to the OWL functors *DisjointClasses* and *EquivalentClasses*, and their inclusion would raise the total

<sup>9</sup>The restatement pattern 131231 was deemed too trivial for inclusion in the survey, but might plausibly occur either for emphasis or to explain technical terms – for instance ‘Corgis are domestic canines, that is, they are dogs’.

<sup>10</sup>Actually this sentence is an oversimplified rendition of overlap, which would also require that some pets are not canines and some canines are not pets.

number of patterns from 46 to 297, and the subset conforming to our coherence rule from 11 to 62.<sup>11</sup>

A second limitation concerns the empirical validation, which addresses only a single very small content domain. Looking at a wider set of examples, it might emerge that the fivefold classification of semantic relations used here is oversimple, and that the taxonomic information in ontologies can be put to better use. To take just one example, the coherence of the disjoint comparison pattern might plausibly depend on the subject terms being not only disjoint, but also *siblings* in the taxonomy (Milosavljevic, 1997) – i.e., concepts at the same level of generality: subjects might be less inclined to judge the example coherent if canines were compared with kittens rather than felines, even though canines and kittens are also disjoint.

Next, we could probably produce a more flexible and generally applicable model if the semantic relations among sets were relaxed so that they allowed exceptions. In particular, by enforcing strict consistency we lose the pattern 131241, disjoint elaboration, in which a subject term is assigned to two incompatible predicates (e.g., ‘Butch is a wolf; Butch is a pet’). If we defined relations 1-4 in a way that allowed a little leeway (e.g., X is nearly a subclass of Y; X and Y are nearly disjoint; etc.), the repertoire of ‘consistent’ patterns could be expanded, and we would obtain a plausible context for the relations typically signalled by ‘but’ and ‘however’ (e.g., CONCESSION). Such a model would be useful for a system generating from data, which might find a few instances of wolf pets in a dataset where nearly all wolves are non-pets and nearly all pets are non-wolves, and thus generate ‘Butch is a pet even though he is a wolf’.

Finally, we have considered only how a rhetorical relationship could be assigned to a *pair* of statements, ignoring the issue of how a globally coherent text could be planned from pairwise assignments. However, this topic is already addressed in the literature, for instance by Marcu’s (1997) bottom-up planning algorithms.

<sup>11</sup>For details on how these numbers are computed see Power (2011b).



## 6 Conclusion

We have sketched a model through which an NLG system could decide whether two formally encoded statements are rhetorically related, and if so how, by examining cross-statement semantic relations evidenced by other statements in the knowledge base. Although in its early stages, the work suggests that a formal basis for assigning rhetorical relations is possible, at least for some relations. As well as guiding NLG systems that generate from ontologies and/or data, our method might prove useful in automatically detecting rhetorical relations in naturally-occurring text; in fact it has already been applied successfully to the task of textual entailment (MacCartney and Manning, 2009), which could be regarded as a special case in which the only rhetorical relation of interest is CONSEQUENCE.

## Acknowledgments

The SWAT project (Semantic Web Authoring Tool) is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grants G033579/1 (Open University) and G032459/1 (University of Manchester). Thanks are also due to colleagues and reviewers for their comments and suggestions.

## References

- Attempto Project. 2011. ACE OWL Verbalizer. Website. <http://attempto.ifi.uzh.ch/site/tools/>.
- Lynn Carlson and Daniel Marcu. 2001. Discourse tagging manual. Technical report, ISI Tech Report ISI-TR-545.
- Glen Hart, Martina Johnson, and Catherine Dolbear. 2008. Rabbit: Developing a control natural language for authoring ontologies. In Manfred Hauswirth, Manolis Koubarakis, and Sean Bechhofer, editors, *Proceedings of the 5th European Semantic Web Conference*, pages 348–360.
- Jerry Hobbs. 1985. On the coherence and structure of discourse. Technical report, Stanford University.
- Ian Horrocks, Peter F. Patel-Schneider, and Frank Van Harmelen. 2003. From shiq and rdf to owl: The making of a web ontology language. *Journal of Web Semantics*, 1:2003.
- Kaarel Kaljurand and Norbert Fuchs. 2007. Verbalizing OWL in Attempto Controlled English. In *OWLED: OWL Experiences and Directions*.
- A. Knott. 1996. A data-driven methodology for motivating a set of coherence relations. Technical report, University of Edinburgh. Ph.D. thesis.
- Shao Fen Liang, Donia Scott, Robert Stevens, and Alan Rector. 2011. Unlocking Medical Ontologies for Non-Ontology Experts. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*.
- Bill MacCartney and Christopher D. Manning. 2009. An extended model of natural logic. In *Proceedings of the Eighth International Conference on Computational Semantics, IWCS-8 '09*, pages 140–156, Stroudsburg, PA, USA. Association for Computational Linguistics.
- W. Mann and S. Thompson. 1987. Rhetorical structure theory: a theory of text organization. In L. Polyani, editor, *The structure of discourse*, Norwood, NJ. Ablex.
- Daniel Marcu. 1997. From Local to Global Coherence: A Bottom-Up Approach to Text Planning. In *AAAI/IAAI*, pages 629–635.
- Maria Milosavljevic. 1997. Augmenting the user's knowledge via comparison. In *In Proceedings of the 6th International Conference on User Modelling*, pages 119–130.
- Richard Power and Allan Third. 2010. Expressing OWL axioms by English sentences: dubious in theory, feasible in practice. In *Proceedings of the 23rd International Conference on Computational Linguistics*.
- Richard Power. 2011a. Coherence relations in ontologies. Technical Report Technical Report TR2011-01, ISSN 1744-1986, Department of Computing, Open University.
- Richard Power. 2011b. Rhetorical Coherence Patterns. Website. <http://mcs.open.ac.uk/rp3242/rhetoric.html>.
- T. Sanders, W. Spooren, and L. Noordman. 1992. Toward a taxonomy of coherence relations. *Discourse Processes*, (15):1–35.
- Rolf Schwitler and Thomas Meyer. 2007. Sydney OWL Syntax - towards a Controlled Natural Language Syntax for OWL 1.1. In *OWLED: OWL Experiences and Directions*.
- Robert Stevens, James Malone, Sandra Williams, Richard Power, and Allan Third. 2011. Automating generation of textual class definitions from owl to english. *Journal of Biomedical Semantics*, 2(S 2:S5), May.
- SWAT Project. 2011. SWAT Natural Language Tools. Website. <http://swat.open.ac.uk/tools/>.
- Sandra Williams, Allan Third, and Richard Power. 2011. Levels of organisation in ontology verbalisation. In *Proceedings of the 13th European Workshop on Natural Language Generation*, Nancy, France.