Generating the Structure of Argument

Chris Reed
Department of IS and Computing
Brunel University
Middlesex, Uxbridge UB8 3PH, England
Chris.Reed@brunel.ac.uk

Derek Long
Department of Computer Science
University of Durham
South Road, Durham DH1 3LE, England
D.P.Long@dur.ac.uk

Abstract

This paper demonstrates that generating arguments in natural language requires planning at an abstract level, and that the appropriate abstraction cannot be captured by approaches based solely upon coherence relations. An abstraction based planning system is presented which employs operators motivated by empirical study and rhetorical maxims. These operators include a subset of traditional deductive rules of inference, argumentation theoretic rules of refutation, and inductive reasoning patterns. The paper presents a unified system in which the various argument forms are employed in generating rich, complex structures for persuasive text.

Introduction

The ability to generate arguments in natural language is attracting wide–ranging research interest, and it is becoming clear that the problem is also stimulating investigation of a number of problems of importance to natural language generation (NLG) as a whole (Reed and Long, 1997a). Argumentation is particularly appropriate as an NLG problem both because it is more highly structured than other forms of natural language, and because there are a variety of established metrics developed in rhetoric and social psychology for judging the resultant quality of a text. These advantages are being exploited in the design and implementation of the *Rhetorica* system, of which the current work forms a part.

1 Problems with RST

A number of limitations of the generative capacity of Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) have recently been identified – most notably, its inability to adequately handle intention (Moore and Paris, 1994). Through investigation of a particular genre – persuasive text – it has become clear that RST suffers from a much wider catalogue of crippling restrictions, severely limiting its applicability to generation in this genre and questioning its suitability elsewhere.

Mann and Thompson discuss the key role played by the notion of *nuclearity* – that relations hold between one nucleus and one satellite. They do, however, concede (p269) that there are a few cases in which nuclearity breaks down – and these they regard as rather unusual. The two types of multi−nuclear constructs they identify are *enveloping structures* – "texts with conventional openings and closings" and *parallel structures* – "texts in which parallelism is the dominant organising pattern". Both of these are not just common in argument, but form key components. Enveloping structures are precisely what are described by, for example, Blair (1838) (citing Cicero), when presenting the dissection of argument into introduction, proposition, division, narration, argumentative, pathetic and conclusion (these are by no means obligatory in every argument, nor is there any great consensus over this particular characterisation; most authors, however, would agree that some such gross structure, usually involving introduction and conclusion, is appropriate). These structures are found with great frequency in natural argument, and cannot, therefore, be ignored. Parallel structures form the very basis of argument, since only the most trivial will involve lines of reasoning in which a single premise supports a single conclusion. Multiple subarguments conjoined to support a conclusion are the norm (see for example, Cohen (1987), Reed and Long, 1997b), and these necessarily form parallel structures.

Another shortcoming is highlighted by a dissonance between RST and argument analysis (see Eemeren *et al*. (1996) for a review). A given text may be amenable to multiple RST analyses – not just as a result of ambiguity, but because there are, at a fundamental level, "multiple compatible analyses". This contrasts with the view in argumentation theory, where one argument has a single, unequivocal structure. There may, of course, be practical problems in identifying this structure, and two analysts may disagree on the most appropriate analysis (and indeed this latter has a close parallel in RST, since different analysts are at liberty to make different 'plausibility judgements' as to the aims of the speaker). The presence of these problems, however, is not equivalent to claiming that arguments may simply have more than one structure, a claim which would pose insurmountable problems to the evaluation process (– argumentation theory aims to determine a means of classifying an argument as either good or bad, and the presence of inherent structural multiplicity would present the possibility of an argument being simultaneously good and bad).

Finally, there is a more intuitive problem with RST, highlighted by analysing argument structure. Although there is much debate over the number and range of rhetorical relations (e.g. Hovy, 1993) there seems to be no way of dealing with the idea of argumentative support. In the first place, as Snoeck–Henkemans (1997) points out, Motivation, Evidence, Justification, Cause, Solutionhood and other relations could all be used argumentatively (as well, of course, as being applicable in non–argumentative situations). Elhadad (1992) draws a similar conclusion (though his list of potentially argumentative relations is somewhat shorter). Thus it is impossible to identify an argumentative relation on the basis of RST alone. Secondly, RST offers no way of capturing higher level organisational units, such as Modus Ponens, Modus Tollens, and so on. For although their structure (or at least the structure of any one instance) can be represented in RST – and, given Marcu's (1996) elegant extensions, even their hierarchical use in larger units – adopting this approach necessitates a lower level view. It becomes impossible to represent and employ a Modus Tollens subargument supporting the antecedent of a Modus Ponens – rather, the situation can only be characterised as P supporting through one of the potentially argumentative RST relations Q, and showing that ~Q, so ~P, and ~P then supporting through one of the potentially argumentative RST relations R, therefore R. Apart from being obviously cumbersome, the representation has lost the abstract structure of the argument altogether, and is not generalisable and comparable to other similar argument structures. (It could perhaps be maintained that such structures could be represented as RST schemas, but there are several problems with such an approach: in the first place, schemas cannot abstract from individual relations, so there would need to be a separate 'Modus Ponens' schema for each possible argumentative support relation; furthermore, the optionality and repetition rules of schema application (Mann and Thomson, 1988, p248) are not suited to argument, as they license the creation of incoherent argument structure).

It is for these reasons, and particularly, the last, that although RST plays an important role in the current work, it is subsumed by a layer which explicitly represents argumentative constructs. These constructs can be mapped on to the most appropriate set of RST relations (thus, for example, the implicature in an MP may be realised into any one of the potentially argumentative relations mentioned above). The approach thus maintains the generative capabilities of RST (particularly when extended along the lines of Marcu (1996) to ensure coherency through adducement of canonical ordering constraints), whilst embracing the intuitive argumentative relationships at a more abstract level. It is these latter relationships which characterise the structure of the argument (i.e. the structure which argumentation theory strives to determine). The relationships are also unambiguous: a single argument has exactly one structure at this level abstraction (though multiplicity is not thereby prevented at the RST level). Further, parallelism occurs only at the higher level of abstraction (multiple subarguments contribute to a conclusion, but each subargument is mononucleaic), and similarly, enveloping structures are also characterised only at the higher level (thus the RST is restricted to a predominantly mononucleaic structure). Finally, complete argument texts are not obliged to have complete RST trees. For although most parts of a text are likely to have unifying RST analyses, and although there must be a single overarching structure at the highest level of abstraction, the refinement to RST need not enforce the introduction of rhetorical relations between parts. This expands the flexibility and generative capacity of the system encompassing a greater proportion of coherent arguments (including, for example, those found in laws and contracts).

2 Abstraction–Based Planning

The structure of argument is thus planned at a level more abstract than RST. To exploit the intrinsic hierarchical structuring of argument, the current work makes use of AbNLP (Fox and Long, 1995), a hierarchical planner based upon the concept of *encapsulation*, whereby the body of an abstract operator contains goals rather than operators, and further, that the body of an operator is not opened up until an entire abstract plan has been completed (i.e. there are no goals left unfulfilled at that level of abstraction). On completion of an abstract plan (which can be seen, in discourse planning, as a skeletal outline of what is to be communicated), the *refinement* operation opens up all the abstract operator bodies, such that the structure and constraints determined at one level of abstraction are propagated to the next level down. As a consequence, many choices which might have been considered during planning of an argument at the detailed level can be pruned as they become inconsistent with the abstract plan. Such an approach has the potential to considerably improve upon the performance of a classical planner, (Bacchus and Yang, 1992). The use of AbNLP in a framework for argumentative discourse planning is discussed in more detail in Reed *et al*. (1996).

The operators employed by AbNLP utilise a highly parsimonious set of intentional goals. Belief goals are used to build the content of an argument (as in much other NLG work); saliency goals to express the intention to convey information to the hearer (following a notion of saliency similar to that proposed in Walker, 1996); and topic manipulation goals to control the focus of attention through the discourse. The roles of these goals and their interrelationships are explored in relation to the information–intention–attention model of Grosz and Sidner (1986) in more detail in Reed and Long (1997a).

3 Deductive Operators

The choice of operators implemented in the *Rhetorica* system has been influenced by a number of factors. The rules of inference are clear candidates for operationalisation: moves such as Modus Ponens are clearly vital components of any argument – though, as noted in Grosz and Sidner (1986), p201, it is inappropriate to view the implication step as one of conventional material implication. The relationship is rather one of *support* – the hearer must be brought to believe that (given the current context and domain of discourse) the first proposition warrants, in part, concluding the second. Even given this weaker, predicate–based reading of a Modus Ponens argument, it is still unclear that any of the other rules of inference (which are, after all, formally redundant) should be necessary. The answer lies in the second consideration, which is entirely empirical – the reason that the argument planning needs to be able to employ other rules of inference is that such argument forms occur naturally. Modus Tollens, for example, is perfectly common, with numerous (real world) examples in argumentation texts such as Fisher (1988). Further, there is a variety of evidence which suggests that Modus Tollens in fact occupies a crucial position in human reasoning (Ohlsson and Robin (1994) cite examples not only from psychology, artificial intelligence and empirical observation, but also by reference to classic examples of Euclid, Galileo, etc.)

Disjunctive Syllogisms are also found reasonably often, but the remaining rules of inference are found very rarely. For this reason, only the three logical argument forms, MP, MT and DS, are currently implemented.

The three deductive operators are shown in Figure 1 (the orderings constraining the body goals – enforcing initial and terminal positions for `PUSH_TOPIC` and `POP_TOPIC` respectively – are omitted for clarity). The preconditions on each operator act as filters on their applicability (so that, for example, the version of MP shown is only applicable for the situation in which the hearer believes the negation of the conclusion). The body is bounded by the topic manipulators which take as a parameter the topic for the current argument, of the form `arg(X,P)`, a generic expression representing an argument concluding P using a premise X (used to abstract from Modus Ponens, Modus Tollens, etc.). The saliency goals also employ the same context parameter: this is used at a later stage to ensure a basic level of coherency (by placing utterances within the appropriate focus space) – discussion of this mechanism is beyond the scope of this paper. Finally, the operator bodies also include goals of belief which are satisfied (after refinement) by further applications of the operators (e.g. the goal `t1` in a given MP could be fulfilled by an MT)

4 Refutation operators

In addition to these deductive operators, *Rhetorica* also employs pseudo–deductive operators, by means of which counter–counterargumentation structures can be developed. The importance of including such refutation in an argument has been conclusively demonstrated in social psychology (Hovland, 1957). The operators required to effect the generation of such structure are closely related to the notions of conflict explored by Haggith (1996), and draw upon the distinction between rebutting and undercutting counterarguments, identified in (Toulmin, 1958). Given the situation portrayed in Figure 2, in which the speaker believes *p* because of *a*, and also disbelieves *b* because of *d* and *e*, and the hearer believes *~p* supported by *b* and *c*, a number of options are available to the speaker. The conventional MP operator discussed above can be employed to support *p* by *a* – this is rebuttal. In addition, the hearer's belief in *~p* can be undercut by arguing against one of its supports, namely, *b*.

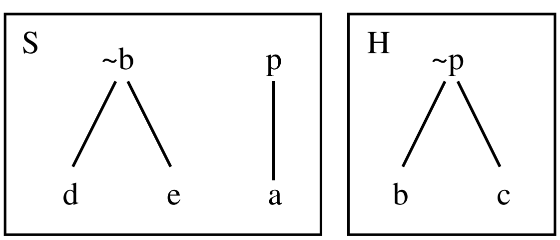

Figure 2. Sample scenario

There are thus no new operators for rebutting, since those in Figure 1 already fulfil that role. Undercutting, however, requires two new operators, one which characterises a refutation of a premise (UCP), and one which characterises a refutation of the validity of an inference (UCI). The operator definitions are shown below in Figure 3.

There are several points to note about these definitions. First, that they are fairly loose, since the speaker is not obliged to believe the falsity of the hearer's premise, merely be able to persuade the hearer of that falsity (though the speaker is somewhat constrained by rules of terminal goal fulfilment – in particular, the `BEL(H, P)` goal is prohibited from fulfilment by any means other than substantial support). Secondly, in the UCP operator, it is necessary to make sure that the hearer is aware that the premise supports his conclusion – but clearly, the speaker doesn't want to offer any further support for the inference, hence the absence of a belief goal corresponding to `IS_SALIENT(Ag, X → P, arg(~X,P))`. Lastly, a similar issue faces the UCI operator – the `t1` goal expresses the need to make the premise salient before attacking it. Indeed, stating counterarguments is the key to counter–counterargumentation: just mentioning a counterargument can bolster a claim. The goal is particularly interesting both from a realisation point of view (where information can be exploited that *x* is being made salient in the context of an argument from *~x*), and an ordering point of view (whether or not statement should precede refutation, and then whether or not UCP/I argumentation should precede pro support is a major issue of debate in psychology, Hass and Linder, 1972).

The deductive operators, however, do not offer the full range of argument forms found in natural text. One major omission is the class of inductive operators, including analogy, inductive generalisation, and causal relation. The framework is designed to admit all these operators, but the current work concentrates upon inductive generalisation.

Inductive generalisation (IG) is of particular interest for a number of reasons. The first is the frequency with which various naively statistical and probabilistic arguments are employed in natural language. More importantly, though, are the problems faced in argumentation theoretic analyses of inductive generalisation. Freeman (1991) examines the problems in depth, and, building on Toulmin's work (1958), and its criticisms, comes to a well justified conclusion that IG should be treated as a convergent arrangement. His argument rests largely on the distinction between the ‘ground adequacy’ and ‘relevance’ questions: in analysing any argument as dialogical, the analyst can look at any two premises and infer that some imaginary opponent had asked a question after the first premise to elicit the second. If that question was ‘Can you give me another reason?’ (ground adequacy), the resulting structure is convergent, whereas if that question was ‘Why does the premise support the conclusion?’ (relevance), the resulting structure is linked. An inductive generalisation is thus based on a number of premises between which an imaginary opponent continually asks the ground adequacy question. The reason, Freeman claims, that inductive generalisation may be intuitively mistaken for a linked structure is that each premise in itself lends only very weak support to the conclusion, and that this generally results in assumption of linkage. Freeman’s work and its relation to other accounts of linked and convergent argumentation is explored more fully in (Reed and Long, 1997b).

In following Freeman’s attractive account of IG, it may appear that the required convergent structure can be fully accounted for in the existing framework, by allowing the standard iterative fulfilment of goals of belief discussed (Reed and Long, 1997a). However, Freeman’s account, because it is analytic, omits the rather obvious fact that premises in an IG have something in common with each other and with the conclusion. That a premise in an IG is related to the conclusion *in some respect* cannot be handled simply by iterating through all available supports for an argument, since there is no way to select all and only those premises which support the conclusion in the given respect. Furthermore, it is important that the IG itself is seen as a unit, since it is quite inappropriate for subsequent ordering heuristics to be at liberty to intersperse various deductive premises for a conclusion in the midst of the inductive premises (or further, that if there exist two or more IGs supporting the same conclusion – each employing a different common attribute – it is inappropriate to mix premises from the various arguments). Seeing the whole IG as a unit enables appropriate scoping for reordering: the premises within the unit can be reordered wholly within the unit, and the unit itself can be moved around wholesale with respect to the other premises. An IG is thus viewed in the current work as a premise. This is illustrated in the diagrammatic argument notation as a *phantom node*, as shown in Figure 4.

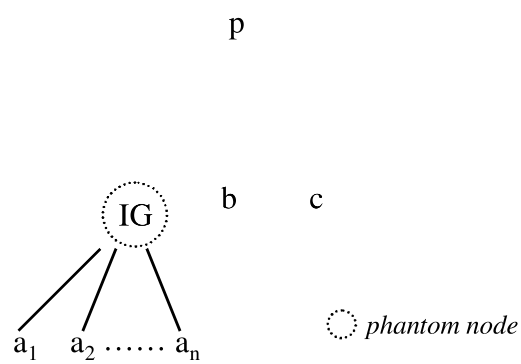

Figure 4. Inductive Generalisation

Thus the IG premise phantom node is generated along with all the other premises for a given conclusion. Then, after refinement, the individual premises within the inductive argument are determined, occurring concurrently with identification of supports for the other premises which are at the level of the inductive generalisation. In the scenario illustrated in Figure 4, for example, the first round of planning identifies that there are three supports for the conclusion *p*, namely, a Modus Ponens argument from each of *b* and *c*, and an inductive generalisation. After an appropriate order is determined for these three, refinement opens up the bodies of the operators, and the supports for *b* and *c* are identified, and the inductive generalisation is fleshed out to include a_1 through a_n. The process of building an inductive generalisation thus involves two different operators: the IG operator, which identifies that an inductive generalisation is appropriate, and the ISUP operator, which is used to select each inductive premise. To prevent an inductive generalisation from being considered at every turn, the precondition list on IG states that there must exist at least one premise which can be used inductively – this is a bare minimum since an inductive generalisation employing a single premise is clearly very weak. Strengthening the notion of inductive generalisation is a trivial task of increasing the minimum number of premises which must exist for the application of

IG to be licensed.

The complete definitions for IG and ISUP are given below in Figure 5.

A single new function is required to express the common feature of the premises and conclusion which license the inductive generalisation – this is implemented as a simple function call to `HAS_PROPERTY` which determines whether or not a given property holds for a given proposition (this functionality is encapsulated in an 'oracle' following [Cohen87]). In both IG and ISUP, the notion of 'support' is thus eschewed altogether and simply remains implicit in the fact that propositions are the same in respect R. It is not necessary to introduce a new notion of support.

Conclusion

This paper has presented a number of features of the *Rhetorica* system, and has introduced the deductive, refutation and inductive generalisation operators which are employed to generate the abstract structure of an argument. In related work, this abstract structure is often lost – certainly in coherence relation based NLG (such as operational RST), but also in (Elhadad, 1992) (which captures some, but not all of the commonly found argument structures) and in (Maybury, 1993) (which fails to capture the hierarchical nature of argument).

Evaluation of non−task−oriented NLG is difficult, particularly when the output is not text, but a plan of primitive operators. However, several evaluative observations support the approach. First, though only touched upon here, the planning process produces a partially specified plan in which the underspecification is precisely that licensed by Cohen−like constraints on argument coherency (Reed and Long, 1997a) appropriated from empirical studies in argumentation theory. Furthermore, the approach enables these coherency constraints to be expressed in a tractable way. Finally, a comparison of system output with natural arguments (of equivalent propositional content) in a small corpus suggests that the constraints of coherency discussed here do indeed ensure the generation of coherent argument structures, and that the interplay between them and constraints of persuasive effect facilitate the construction of natural language arguments which are both coherent and effective.

References

Bacchus, F. & Yang, Q. (1992) *The Expected Value of Hierarchical Problem−Solving*, AAAI−92 pp. 369—374

Blair, H. (1838) *Lectures on Rhetoric and Belles Lettres*, Charles Daly, London

Cohen, R. (1987) *Analyzing the Structure of Argumentative Discourse*, Computational Linguistics 13/1, pp. 11—24

Eemeren, F.H. van, Grootendorst, R. & Snoeck−Henkemans, F. (1996) *Fundamentals of Argumentation Theory*, Lawrence Erlbaum

Elhadad, M. (1992) *Generating Coherent Argument Paragraphs*, COLING'92, Nantes, pp.638—644

Fisher, A. (1988) *The Logic of Real Arguments*, Cambridge University Press, Cambridge, UK

Fox, M. & Long., D. (1995) *Hierarchical Planning using Abstraction*, IEE Proceedings on Control Theory and Applications 142/3, pp.197—210

Freeman, J.B. (1991) *Dialectics and the Macrostructure of Arguments*, Foris, Dordrecht

Grosz, B.J. & Sidner, C.L. (1986) *Attention, Intentions and the Structure of Discourse*, Computational Linguistics 12/3, pp.175—204

Haggith, M. (1996) "A meta−level argumentation framework for representing and reasoning about disagreement", PhD Thesis, Univ. of Edinburgh

Hass, R.G. & Linder, D.E. (1972) *Counterargument availability and the effects of message structure on persuasion*, Journal of Personality and Social Psychology 23/2 pp.219−233

Hovland, C.I. (1957) *The Order of Presentation in Persuasion*, Yale University Press, New Haven

Hovy, E. H. (1993) *Automated Discourse Generation Using Discourse Structure Relations*, Artificial Intelligence 63, pp. 341—385

Mann, W.C., Thompson, S.A. (1988) *Rhetorical structure theory*, Text 8/3, pp. 243—281

Marcu, D. (1996) *Building Up Rhetorical Structure Trees* AAAI'96, Portland, Oregon

Maybury, M.T. (1993) *Communicative Acts for Generative Natural Language Arguments* AAAI−93 pp.357—364

Moore, J.D. & Paris, C.L. (1994) *Planning Text For Advisory Dialogues*, Computational Linguistics 19/4, pp. 651—694

Ohlsson, S. & Robin, N. (1994) *The Power of Negative Thinking: The Central Role of Modus Tollens in Human Cognition* in "Proc. of the16th Conf. of the Cognitive Science Soc.", pp.681—686

Reed, C.A., Long, D.P. & Fox, M. (1996) *An*

Architecture for Argumentative Dialogue Planning, in Gabbay, D. & Ohlbach, H.J. (eds) "Practical Reasoning", Springer Verlag, pp. 555–566

Reed, C.A. & Long, D.P. (1997a) *Content Ordering in the Generation of Persuasive Discourse* IJCAI'97, Nagoya, Japan, pp. 1022–1027

Reed, C.A. & Long, D.P. (1997b) *Multiple Subarguments in Logic, Argumentation, Rhetoric and Text Generation* in Gabbay, D.M., et al. (eds) "Qualitative and Quantitative Reasoning", Springer Verlag, pp.496–510

Snoeck–Henkemans, F. (1997) *Verbal indicators of argumentation and explanation* in "Proc. of the OSSA Conf on Argument and Rhetoric", St. Catharines, Ontario

Toulmin, S. E. (1958) *The Uses of Argument*, Cambridge University Press, Cambridge, UK

Walker, M.A. (1996) *The effect of resource limits and task complexity on collaborative planning in dialogue*, Artificial Intelligence 85, pp.181–243