Towards Computational Rhetoric

FLORIANA GRASSO   *University of Liverpool*

Abstract: The notions of argument and argumentation have become increasingly ubiquitous in Artificial Intelligence research, with various application and interpretations. Less attention has been, however, specifically devoted to rhetorical argument. The work presented in this paper aims at bridging this gap, by proposing a framework for characterising rhetorical argumentation, based on Perelman and Olbrechts-Tyteca's New Rhetoric. The paper provides an overview of the state of the art of computational work based on, or dealing with, rhetorical aspects of argumentation, before presenting the characterisation proposed, corroborated by walked-through examples.

Résumé: On rencontre de plus en plus souvent les notions d'argument et d'argumentation ainsi que leurs diverses applications dans les recherches en intelligence artificielle. Toutefois moins d'attention a été accordée aux arguments rhétoriques. Dans cet artice on tente de combler cette lacune en traçant une structure, fondée sur la Nouvelle Rhétorique de Perelman et d'Obrecht-Tytéca, qui décrit l'argumentation rhétorique. Cette description est précédée d'un survol des travaux courants en programmation qui reposent sur des aspects théoriques de l'argumentation, ou qui ont affaire avec ceux-ci.

Keywords: Computational Linguistics, Artificial Agents, *New Rhetoric*

1. Introduction

In the past decade, the notion of argument has been more and more extensively used in computation, or more specifically in Artificial Intelligence (AI). Very diverse are the applications, among which the most natural is perhaps to the field of Artificial Intelligence and Law (Bench-Capon 1997; Prakken 1997; Jones & Sergot 1993). The metaphors of argumentation and negotiation have become widespread in the model of artificial *agents* (Wooldridge 2000), to represent how they communicate, and possibly come to an agreement for the execution of either a common or a competitive task (Parsons & Jennings 1996; Karacapilidis & Papadias 2001; Sillince 1994; Kraus, Sycara, & Evenchik 1998). Argumentation has been used as a way to represent uncertain and defeasible reasoning, where the acceptance or otherwise of a proposition entails more a weighing of arguments in favour and against, than a logical or probabilistic process of going from some premises to the conclusion (Amgoud & Cayrol 1998; Dung 1995; Krause *et al.* 1995; Vreeswijk 1993). Some other applications concern the use of arguments for computer supported collaborative learning (Lajoie, Greer, Munsie, Wilkie, Guerrara, & Aleong, 1995; Pilkington, Hartley, Hintze, & Moore, 1992; Ravenscroft, 2000), or to improve the interaction with expert systems (André and Rist 2000), and even in software engineering

(Carbogim, Robertson, and Lee 2000).

The work we present here is concerned in particular with "rhetorical argumentation," which on the other hand has enjoyed consideration to a somewhat lesser extent. Only a handful of researchers have been undertaking computational works that focus on the rhetorical aspect of arguments, or on the structure of rhetorical discourse. By rhetorical arguments we denote arguments which are both heavily based on the audience's perception of the world, and concerned more with evaluative judgements than with establishing the truth or the falsity of a proposition.

Rather than a logic focus on argumentative reasoning, or a pure computational linguistic focus on modelling discourse which happens to be argumentative, we place ourselves halfway, and specifically focus on the characterisation of rhetorical argumentative discourse. Methodologically, we do this by explicitly drawing upon the philosophy of argument.

This paper proposes an all encompassing framework for the formalisation of rhetorical argumentation, inspired by a well-established philosophical theory, the *New Rhetoric*. The work is admittedly biased towards the generation, rather than understanding, of argumentative discourse, and we see, and indeed we built, our formalisation to be comfortably used in a system for generating argumentative discourse based on AI planning. At the same time, we see our work as a contribution to typical pragmatic issues of providing a characterisation of what the speaker has to take into account when producing a "good" argument, and therefore what the hearer can ascribe to the speaker. The framework is well grounded on theoretical foundations, and indeed our proposal coherently integrates work in the philosophy of argument, in linguistics, in social psychology and in artificial intelligence.

Before proceeding with the description of our framework, we feel it is important to present a brief *excursus* on the ways in which research in rhetoric has found application in artificial intelligence, and especially in computational linguistics.

1. Computational Approaches to Rhetoric

In our overview, we find it is useful to distinguish three main applications of the notion of rhetorical argumentation: as a paradigm for discourse modelling, for the production of persuasive messages, and as an aid to knowledge representation.

1.1 Discourse Processing

In talking about computational perspectives, and in consideration of the communicative dimension of rhetorical argumentation, we have first and foremost to refer to the field of computational linguistics, and in particular to those researchers who are concentrated on the study of discourse structure.

In the field, three main theories emerge of organisation of discourse: the schemata organisation (McKeown 1985), the approach by Grosz and Sidner 1986 and Rhetorical Structure Theory (RST) (Mann and Thompson 1988).

McKeown (1985) defines schemata as "a representation of standard pattern of discourse structure which efficiently encodes the set of communicative techniques that a speaker can use for a particular discourse purpose" (p. 20). Apart from the computational slant, the main difference in the schema philosophy, when compared with Perelman and Olbrechts-Tyteca's approach, is that McKeown's schema comprises a whole set rhetorical predicates, "compiled" in a routine that can be used on the basis of a very general goal (e.g., "compare and contrast objects X and Y"). This, underlined by Moore and Paris (1992), makes all rhetorical processes transparent to the discourse generator, and therefore impossible to modify when the context is changed. This lack of explicitness makes, in fact, these schemata unsuitable for rhetorical argumentation, where one should be sensible to any change in the audience.

The approach of Grosz and Sidner (1986) is generally an analytical one, aimed at evaluating whether a given discourse, typically a dialogue, is coherent. The evaluation is based on three aspects of discourse: intentional structure, attentional structure, and linguistic structure. The approach postulates that, behind each discourse segment, identified by means of the linguistic structure, one should be able to see an intentional purpose, which relates to those of other discourse segments in one of two ways: satisfaction precedence (one must be satisfied before the other) or dominance (one is a sub-purpose of the other). The whole intentional structure cannot however be successful if one does not account for the attentional state of the hearer, which represents what is relevant at each stage. Perhaps because of the generic nature of the relations described, this approach is mainly used for analysis rather than production of discourse, a task for which the most widely used approach is the RST.

The thesis of RST is that a text is coherent if and only if it can be recursively decomposed into pairs of non overlapping, adjacent clauses, or *spans*, which are related one to the other by one of a set of rhetorical relations. A text can therefore be mapped into a tree, whose root, with the rhetorical relation associated to it, represents its main purpose. The rhetorical dimension of the theory lies exactly in the definition of the relations: Mann and Thompson (1988) list a set of relations, found on the basis of empirical observations, for each of which they postulate a set of constraints that have to hold for the relation to be successful, summarising what the speaker and the hearer have to believe about the world. The two components of a rhetorical relation typically have an asymmetric role: one (the nucleus) is more fundamental than the other (the satellite) to the success of the relation itself. The set of relations is open, and has in fact being augmented by many authors (e.g., Hovy 1993 incremented the original collection into a taxonomy of about 120 relations), but it is not meant to cover all types of discourse: it is not possible, for instance, to represent "enveloping" structures like opening and closing (although Marcu 1997 defines *ad hoc* relations to capture them), "parallel" structures (e.g. comparisons), laws or poetry. Nevertheless, the theory is widely used for the analysis of hierarchically organised text, and has also had some variations able to account, for instance,

for dialogues (Stent and Allen, 2000). Moreover, the hierarchical hypothesis makes RST particularly suitable for automated generation of text by means of an AI planner (Hovy 1993; Moore and Paris 1993; Marcu 1996; de Rosis, Grasso, and Berry 1997).

As pointed out by Reed and Long (1998), RST is, however, unsuitable for dealing with arguments, because of a series of problems. First, as also identified by Moore & Pollack (1992), RST does not adequately handle intentions, an important shortcoming for rhetorical argumentation, although this problem has been partially solved by Marcu (2000a) by coupling RST with intentions *à la* Grosz and Sidner. Second, RST considers as exceptional the cases in which nuclearity breaks down to give place to parallel structures, structures that are the rule rather than the exception in argumentation. Finally, RST does not account for argumentative relations as such, nor for high level organisations such as *Modus Ponens*, or, most importantly, for structured combinations of higher level units.

1.1.1 Implemented Systems

One of the earliest work in the computational study of argumentative discourse is perhaps the one by Cohen (1987), with an emphasis more on argument understanding than generation. The approach stressed the importance of a theory of coherence structure, as well as a theory of the relationships among the speaker's and hearer's beliefs, which should be coupled with the analysis of linguistic clues in order to be able to understand arguments.

The research by Maybury (1993) focuses on the application of argumentation to a natural language generation system, representing argument as a series of communicative acts that have the purpose of performing some communicative goal. The aim of a communicative act is to affect the knowledge, the beliefs or the goals of the addressee. Different "argumentative" communicative act are identified, not surprisingly, in an Aristotelian fashion: "deductive arguments" are used to affect the beliefs of the addressee by providing a "proof" of a claim (such as the various kind of syllogism); "inductive arguments" are used to convince of some claim by providing some examples or evidences (such as illustrations, analogy, etc).; whereas "persuading arguments" are used to affect the goals of the addressee (such as motivation, indicating the desirable consequences of an action, how the action is part of a goal, etc). The communicative acts are used in a hierarchical planner *à la* Sacerdoti (1977), as in many other works in natural language generation, such as Cawsey (1992). Here, "plan operators" define the constraints and the preconditions that must hold before a communicative act applies, its intended effects on the addressee's cognitive state, represented by means of intentional operators expressing knowledge and desires, and how the act can be decomposed into subacts. However, in Maybury's work argumentation theory is used simply to define new communicative acts for generating a monologue, so no measure is possible to calculate how effective the act was. Nor is there an accurate description of the

addressee's cognitive state, necessary to trigger the appropriate act. The focus of the system by Reed, Long, Fox, and Garagnani (1996) is on the representation of *logical arguments*. A classification of beliefs, which we find fundamentally rhetorical, distinguishes among *factual beliefs*, testable or definitional, *opinions* based on moral and aesthetic judgement, and therefore personal and not provable, and *cultural beliefs*, based upon "sociocultural maxims" as defined by the authors, but which can easily relate to the classical concept of *Loci*, or *Topoi*. The knowledge of the system consists of a set of *events*, which can be related by means of a relationship of *support*: all the events, or perceived beliefs, which are supported by held events are themselves held. Events with no supporting events are called *grounded*: these can be due to external sources, internal motivations and value-systems (another rhetorical concept) or *a priori* beliefs about the ontology of the construction of the beliefs, which are necessary because both agents must agree on the description of the knowledge. An argument consists of one or more premises, which can in turn be sub-arguments, and exactly one conclusion. Premises can be linked to the conclusion by standard logical relations, e.g. rules of classical logic, by inductively reasoned implications, such as inductive generalisation, causal implication or analogy, or by rhetorical fallacies, which are, however, not fully investigated by the authors. An attempt is also made to structure the argumentative discourse by specifying the position that the statements comprising the argument hold in the sentence: "Conclusion-First," when the premises are examples, or when the conclusion is deliberately provocative; "Conclusion-Last," used for longer, more complex or less convincing arguments; or "Conclusion-Sandwich," typically used when the speaker has completed a Conclusion-Last argument which has not been accepted and therefore further support is required. The arguments are generated with an abstraction-based planner (Fox and Long 1995), where the operators implement rules of logical inference, such as *Modus Ponens* or *Modus Tollens*, plus a few refutation rules (Reed and Long, 1998). The approach is therefore dialectical, rather than rhetorical, in its nature, more focussed on the way claims support each other than on the communicative, audience oriented dimension of the argument proposed. The work is, however, complete in its treatment, for instance, of ordering (Reed and Long 1997) and saliency (Reed 1999).

Another important work to mention is the one by Elhadad (1995), important because it, like ours, relies on an established theory, namely Anscombre and Ducrot's "Argumentation Within Language" (AWL) (Anscombre and Ducrot 1983). AWL is, in fact, a linguistic theory, describing the ways in which sentences are expressed to support a set of conclusions rather than the objects used as premises. A fundamental notion is the "sentence argumentation orientation" which gives an indication of what conclusions can be supported by a sentence. Two sentences are said to have the same orientation if they lead to similar conclusions. This is independent from the content of the sentence itself. So for instance the sentence *Tom is as tall as Mary* can have the same orientation both of *Tom is tall* and of *Tom is short*, depending on which sentence follows describing the height of Mary. Elhadad showed how

AWL could be used in several phases of the text generation process, from content determination and organisation, to lexical choice. With the aim of advising students about the courses to follow, the device of *topos* is used to express rules to perform *evaluations*. An evaluation is simply an answer to the question *How P is X?* , where P is a scale and X is an entity, and originates rules of the form *the more/less X is P, the more/less Y is Q*. A topos is then formed by four primitive relations:

1. an evaluation of the entity X on the scale P (for example a class on the scale of difficulty);
2. an evaluation of the entity Y on the scale Q (for example a student on the scale of desiring to take a course);
3. the expression of a gradual relation between P and Q (*the harder the class, the less desire*);
4. the expression of a topical relation between X and Y.

The topoi are used to select information from the knowledge base, which could lead to the conclusion desired (to take/not to take a course). They are used in process of content organisation, interpreted as rhetorical relations, in that they explain the connection between two text spans. Topoi are also used in the lexical choice of determiners (*many, few* in the phrase *the AI course has many/few assignments*), adjectives (*difficult, interesting*), verbs (*require, enjoy*) and connectives (*but, therefore*).

De Rosis, Grasso, and Berry (1997) compared computer produced and human produced texts in a drug-prescription domain. Although the computer generated explanations reproduce the overall characteristics of the human texts, several elements, among which most importantly argumentative techniques, contribute to increase the *argumentative strength* of the human messages, and are not reproduced in the artificial ones. While claiming that an ideal solution to this problem would only be possible with a more fine grained representation of goals and beliefs of Speaker and Hearer, together with more sophisticated text planning techniques, a more practical strategy is proposed, employing a simple top-down planner to collect and order the information content of the explanation, and subsequently refining the plan by applying some rhetorical and argumentative strategies. No definition of effectiveness of an argumentation strategy in terms of changes in the beliefs of the user is given, however, and the argumentative rules are the mere application of empirical observations.

A major emphasis on argumentative text can be found in the work by Carenini and Moore (2001), who concentrates on "evaluative arguments," that is arguments advising on the appropriateness of a particular object or course of action. Arguments are tailored to users' values and preferences to the objects of discussion, so that a numerical measure can be obtained of how valuable an entity is to the user, a measure based on utility theory. Such a measure helps shape the produced piece of advice, by deciding whether to mention specific features of the entity, changing the order of presentation of the features, or changing the emphasis of the evaluation (e.g., "good" as opposed to "excellent"). The actual generation of the piece

of advice uses a pipelined architecture, a standard mechanism in natural language generation (Reiter and Dale, 1997), including a "text planner," aimed at organising the general structure of the text, a "microplanner," shaping the text at a sentence level, and a "sentence realiser," aimed at actually producing the sentence in the chosen language. While the major emphasis is put on audience's modelling, the work does not, however, address the question of using rhetorical devices/schemata in the argument generation.

1.2 Generation of Arguments

We list here works that, despite a smaller emphasis on discourse production, are addressing the issue of producing persuasive (to an audience) arguments, and therefore can be included in the rhetorical category.

Research specifically aimed at designing a system capable of analysing and composing argument is due to Zukerman, Korb, and McConachy (1996). A system is discussed that, given a proposition and a user model, is able to produce an argument supporting the proposition on the basis of the user model characteristics, in order to bring the user to an established degree of belief in the proposition. Especially interesting is the authors' definition of *nice* arguments, that is arguments that achieve a compromise between what is justifiable and what persuades, sometimes preferring what is believable to what is correct. The system can also analyse arguments proposed by the user, and agree with them or counter-argue. The architecture of the system comprises four modules: (1) an argument strategist, (2) an argument generator, (3) an argument analyser, and (4) a user interface. The argument generator takes as input the proposition to be argued for, the degree of belief to be achieved and a system *attitude*, which determine the degree of fairness of the system, (it can go from "orthodox," which allows only normative inferences, to "licentious," which allows any action to increase the effectiveness of the argument). The output of this stage is an *argument graph* in a Bayesian network (Pearl 1991), which assigns to each inference a strength value, based on the conditional probability to achieve the given degree of belief. The analyser evaluates the argument graphs, guided by several parameters, such as the argument's normative strength, (whether it is sound, or it is an inductive inference, and so forth) and its effectiveness, (determined by the expected effect on the belief model of the user). The argument strategist selects the approach for presentation of the argument, with the principal aim to make it as concise as possible; it selects the presentation method, (for example, whether to present it completely or selecting branches of the graph; which style is the most appropriate, e.g., concessive or hypothetical, and so on); and it also selects the strategy to respond to the user's argument. Despite the attention to persuasiveness rather than soundness, however, the only concession granted to these *nice* arguments is their having some steps missing to their logic chains. Values and persuasiveness techniques are not considered here, so these arguments cannot be defined "rhetorical."

PERSUADER (Sycara 1990) is a model of persuasive argumentation, in a scenario representing the negotiation process in a settlement. Here assumptions are made which are similar to ours: the argument generation process is always guided by argumentation goals, aiming at changing the importance the *persuadee* attaches to concepts, and argumentation strategies are used to achieve them. The theoretical basis on which the strategies are defined, however, is not made explicit. An order is defined among a set of arguments in the labour mediation domain, according to their presumed strength, for instance a threat or a promise is meant to be the strongest argument, whilst an appeal to universal principles is described as the weakest one in this domain. The choice of the argument depends on the current goal of the persuader, that is changing the importance of a goal/issue, changing the perception of an issue's value, or aiming at goal abandonment on the part of the persuadee via threats or promises, even if the latter goal in our view can be included in one of the former two, as a threat or a promise is in fact meant to change the persuadee's priorities with respect to the issue in consideration (see also Castelfranchi 2000 for a discussion on an agent's internal and external conflicts).

Finally, it is worth mentioning the work by Saeedi and Sillince (1999), which implemented "rhetorical" rules, such as fairness, reciprocity and deterrent, to simulate argumentation and debate. The model is based on Toulmin's schema, and is meant to provide a framework to facilitate collaborative decision-making. The rules have, however, a high degree of domain dependence, and are not defined in a way that encourage discrimination among them from case to case.

1.3 Representing knowledge as argument

A final aspect in the model of rhetorical arguments that is worth mentioning in our survey concerns the representation of knowledge. We have discussed more extensively about knowledge representation elsewhere (Grasso 2000), but here we would like to mention those systems which adopt a rhetorical view in modelling their knowledge structure, by explicitly representing preferences and values.

One of the earliest works on this respect is Lowe's *SYNVIEW* system (Lowe 1985), which proposes a method for collecting and synthesising large quantity of information and presenting them to a user in a manageable structure. Of particular interest from our point of view is the way *SYNVIEW* combines multiple viewpoints in a unique structure. The user is presented with an overview of topics related to a lead topic, ranked in order of their "importance" for a general understanding of the lead topic. A user can then suggest modification or addition at many levels. At the simplest level, a user can modify the attribution of the importance score. Or, the user can suggest a different verbalisation of the topic, or even add new subtopics. If a topic is a statement of a fact, this can open a structured debate. In a debate, the subtopics are further statements in support or against the topic of the debate. The explicit representation of arguments of a debate follows Toulmin's general structure, but is simpler: at the first level an argument is represented by a ranked list of topics

for and against it. There is no explicit consideration of the warrant to the links of the subtopic to the conclusion.

Optimist (Clark and Wilkes-Gibbs 1986) was designed with the aim of helping geologists in oil explorations. The hypothesis of using a process of argumentation for this purpose is that "argumentation" among experts is useful to pool knowledge together and enforce consistency, and that, as experts frequently disagree, it is useful to collect evidence from similar cases which form a focus for a discussion. An argument is defined as an inference (or chain of inferences) whose validity is open to question. To reason about arguments, some meta-rules can be used, which are represented as meta-level knowledge. In order to represent different opinions, several knowledge-bases, labelled with the owner's name, are necessary. To keep the knowledge base manageable, a *skeleton* + *certainty sets* representation is used: a single rule-base is developed and then the body of the rules is separated from the certainty measures attached to them. In this way, several certainty values can be stored under different names.

More recently, Karacapilidis (1996) suggested an argumentation-based frame-work in which the basic elements are considered the *Positions, Issues* and *Arguments* of different agents during an interaction. Positions are any data the agents assert (declarations, justifications, etc.), the issues are the decisions to be made or the goals to be achieved and the arguments are assertions about the positions regard-ing their attributes that can be used for or against them. Comparison of alternative positions is made according to the value of *importance* for the agents. Preference relations are introduced between intervals of importance of arguments. As explained by Karacapilidis, Trousse, and Papadias (1997), broader *discussion domains* are associated with every viewpoint of a case. The association proceeds "top down" from the more general topics to their subtopics.

A different use of viewpoints is the one by McCoy (1989), where perspectives are used to dynamically highlight selected attributes of a knowledge base. Simi-larly, Jameson *et al.* (1994) use a notion of perspective in order to provide "biased" information, where the bias consists in focusing the user's attention on aspects of the knowledge base which are thought to be of particular interest, diverting atten-tion from more "dangerous" ones. In both cases, therefore, a perspective is more a lens hiding portions of the knowledge base, rather than a rhetorical *locus* device, to attribute evaluations to concepts.

We will see in the next sections how an explicit representation of perspectives is fundamental in building a framework for rhetorical argumentation.

2. A Formal Model of Rhetorical Argumentation

In this section we present a characterisation of rhetorical argumentation based on Perelman and Olbrechts-Tyteca's *New Rhetoric*[1]. We should perhaps stress imme-diately that the formalisation, albeit given in logical (or quasi-logical?) terms, is by no means advocating a logical approach to rhetoric, as this would miss the point of a

rhetorical argument altogether. The effort is only aimed at providing a well-defined semantics for our concepts. In fact, the logicalisation of rhetorical argumentation is only part of the model: this section provides a definition for rhetorical argument, while Section 3 will add to this definition its rhetorical dimension by coupling it with the concept of a rhetorical schema. The notion of schema will of course draw from Perelman and Olbrechts-Tyteca's definition, but will also be inspired by theories on discourse structure, as rhetorical argumentation cannot ignore language and discourse as its prime means for being conveyed.

Before giving our model for a rhetorical argument we need to introduce the primitive objects of our language, and in particular, after a general assumption on our domain ontology, we introduce a notion peculiar to rhetorical argumentation, and to Perelman and Olbrechts-Tyteca's *New Rhetoric*: the notion of *value*. This notion, together with the one of *perspective*, will serve to build up the definition of a rhetorical argument. We will then conclude with an example of what it means to argue rhetorically, showing the advantages and the idiosyncrasies of our model.

2.1 Objects of Discourse

We will assume that the argumentation process can take place on a certain set of objects of discourse, forming an ontology. We have proposed in (Grasso 2000) a more formal definition of the ontology we need, but now it will suffice to say that we need a set of concepts, C_o, of an ontology O, and a set of relationships among these concepts, R_o, again in the ontology O. We think of a relationship, in this context, as any of the possible ways in which two concepts may be linked in the ontology: one might be the generalisation of the other, they may both contribute to define a third object, they may be an action and one of its roles, they may have identical values for one of the attributes, and so on and so forth. We express the fact that there exists a relationship r_k among the objects c_i and c_j in an ontology O by writing that:

$$r_k(c_i, c_j)$$

with

$$r_k \in R_o, \quad c_i, c_j \in C_o$$

For the sake of simplicity in the notation, we will consider, without loss of generality, only binary relations, but the definitions below can be naturally extended to consider n-ary relations.

2.2 Values and Perspectives

We have hinted before that a rhetorical argument is, by and large, a means to *pass value* from one object of discourse to another: a good value if arguing in favour or a bad one if arguing against. The notion of value is therefore crucial to our formalisation. In order to argue rhetorically, we need to express in some way that "X has good value" or "X has bad value," where X is a concept, as one of our primitive notions.