Persuasion as a Form of Inter-Agent Negotiation*

Chris Reed[1], Derek Long[2], Maria Fox[2] and Max Garagnani[2]

[1] Department of Computer Science,
University College London,
Gower St.,
London
WC1E 6BT
C.Reed@cs.ucl.ac.uk
http://www.cs.ucl.ac.uk/staff/C.Reed

[2] Department of Computer Science,
Durham University,
South Road,
Durham
{D.P.Long, M.Fox,
M.Garagnani}@dur.ac.uk

Abstract. Agents in a multi-agent environment must often cooperate to achieve their objectives. In this paper an agent, B, cooperates with another agent, A, if B adopts a goal that furthers A's objectives in the environment. If agents are independent and motivated by their own interests, cooperation cannot be relied upon and it may be necessary for A to *persuade* B to adopt a cooperative goal. This paper is concerned with the organisation and construction of persuasive argument, and examines how a rational agent comes to hold a belief, and thus, how new beliefs might be engendered and existing beliefs altered, through the process of argumentation. Argument represents an opportunity for an agent to convince a possibly sceptical or resistant audience of the veracity of its own beliefs. This ability is a vital component of rich communication, facilitating explanation, instruction, cooperation and conflict resolution. An architecture is described in which a hierarchical planner is used to develop discourse plans which can be realised in natural language using the LOLITA system. Planning is concerned with the intentional, contextual and pragmatic aspects of discourse structure as well as with the logical form of the argument and its stylistic organisation. In this paper attention is restricted to the planning of persuasive discourse, or monologue.

Keywords: agent communication, argumentation theory, rhetoric, belief modelling, planning.

*This work has been partly funded by EPSRC grant no. 94313824.

1. Introduction

In this paper an architecture is presented for the autonomous construction of argument, outlining the components required for persuasive communication. Argument plays a crucial role in multi-agent worlds in which agents are motivated to pursue their own interests. The focus of this work is restricted to logical argument, and the assumption is made that agents will adopt, or reject, certain beliefs if they are presented with convincing reasons to do so. Rhetorical argument, which has been pursued in the context of multi-agent communication [20], is difficult to model because its success depends on the inability of the hearer to recognise the fallacy it depends upon. It is assumed in this work that agents can both construct and follow logical argument.

The architecture described here is intended as a model of communication between two artificial agents, although it also models a restricted form of inter-human communication and human-computer interaction. The realisation of the planned discourses must result in their expression in some language. In this work, natural language has been chosen as an example because generation of natural language texts from planned monologues has been explored successfully in the LOLITA system [18]. The proposed architecture is not inextricably bound to the use of natural language - the planning levels are independent of the language in which plans will be expressed and the intentional structures produced at these levels are sufficiently abstract to remain unchanged by a shift to some more simple and restrictive artificial language.

2. Overview

The architecture described in this paper is hierarchical in structure, reflecting the distinct, though inter-related, levels of structure within arguments. The part of argument synthesis which is concerned with the resolution of syntax, expression and morphology comprises the lowest level of the architecture and represents the interface to the LOLITA system. Above this lowest level sits functionality based upon Mann & Thompsons's Rhetorical Structure Theory [11], and then above this is a hierarchy of levels concerned with the planning and presentation of the discourse. The highest level of abstraction in the architecture is the Argument Objectives (AO) level which determines the overall form of argument to be constructed. Three forms are considered: an agent can *explain*, *inform* or *persuade*. Informing the hearer of a proposition simply involves stating the proposition without the support of argument. It is assumed that the hearer had no knowledge of the proposition prior to being informed. Explaining a proposition involves providing support for the proposition through argument, under the assumption that the hearer knew of the proposition but was undecided about its truth value prior to the explanation. Persuading a hearer to accept a proposition involves undermining the hearer's current beliefs in the falsity of the proposition and then the provision of support for belief in the truth of the proposition. It is assumed that the hearer believes in the converse before the persuasion takes place. As this account suggests, a four-valued model of belief is used in the architecture described here. An agent either believes, disbelieves, is unaware of or is undecided about a proposition. This model is described and justified in §3.

The persuade and explain argument forms are developed at the Argument Structure (AS) level which produces the logical form of the argument, employing purely intentional data structures, by the application of operators modelling logical inference rules. This form is then augmented and modified by the subordinate Eloquence Generation (EG) level which is concerned with properties of a speech such as its length, detail, meter, ordering of sub-arguments, grouping of sub-arguments, enthymeme contraction, use of repetition and so on. The architecture is summarised in Fig. 1. A collection of operators describing the actions which may be employed at the different levels in argument construction are supplied to a hierarchical planner which plans firstly an intentional structure (an abstracted model of the intended discourse), which is then supplemented with further details to form a deeper content structure, linking the intentions to rhetorical structures for realisation. These structures are supplemented with the addition of constraints on textual realisation, including vocabulary restrictions, rhythm and mood of text and so on. This is finally realised as text.

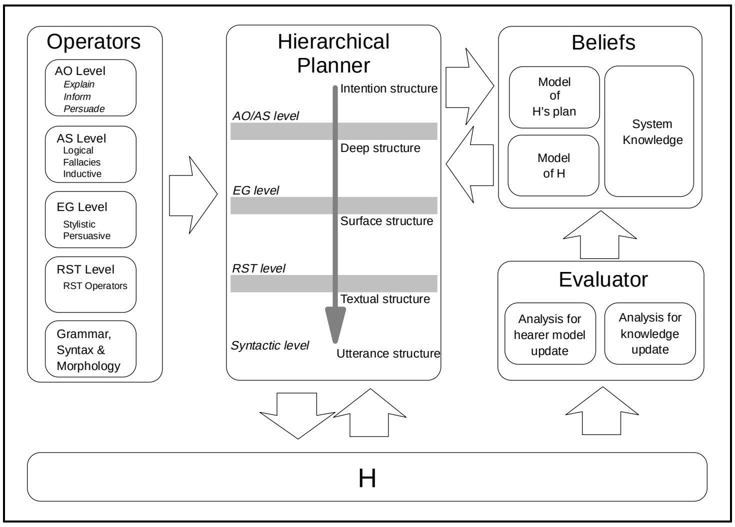

Fig. 1. System architecture overview

3. Belief Modelling

3.1. Issues

It has become clear that both AS and EG levels require access to the belief model of the audience, in addition, of course, to the beliefs of the system. The AS level needs to be able to assess in advance how different structures are likely to be received, and to take account of the beliefs of the audience in producing structure. The EG level employs the belief model to pitch an argument at the right level of detail, and to track the saliency of beliefs used during argumentation. A more detailed examination of the relationship between the higher levels and belief is described below in §4 and §5.

There are a number of issues involved in analysing belief. There are different kinds of belief. These will be referred to in this paper as factual beliefs, which are either testable (at least in theory) or are definitional (and include beliefs based on sensory experiences); opinions, which are based on moral and aesthetic judgement and are thus ultimately personal and unprovable (since there is no universally accepted and provably correct aesthetic-moral framework - and it is hard to conceive how there could be); and cultural beliefs, which are based upon sociocultural maxims (such as that which states that living into old age is desirable). This tripartite distinction has been based upon the views expressed implicitly by [3], when talking about general types of argument, but there are also numerous other divisions which could be detailed (see [1] for a number of examples). The way in which beliefs become manifest in argumentative dialogue - both their individual expression, and their interrelation - is chiefly dependent upon their class, so competent representation and recognition of these distinct belief types represents an important problem. A further discussion of these divisions and their implications is pursued in §3.2.

Another major problem is how to resolve two seemingly contradictory views generated by introspection, that of whether beliefs are best represented as dichotomous or scalar. Various problems associated with the concept of 'strength of belief' are discussed in [7]. A pragmatic resolution to this issue is in itself crucial to competent belief modelling.

A simple, but powerful, way to model beliefs is to say that an agent believes a proposition in a given context if that agent would be prepared to act on that proposition in that context. Action might take different forms reflecting different levels of commitment on the part of the agent to the implications of believing in the proposition. For example, action on the belief that lottery winnings should be distributed amongst the poor can range from simply claiming this belief to taking actions in the world which bring about the redistribution of individuals' winnings. Using this model, given a context, an agent can be said to believe, disbelieve or be undecided about a given proposition with respect to a given action.

In the current work, agents co-existing within an environment communicate in order to engage in co-operative activity. An agent, A, will attempt to persuade an agent, B, to adopt a certain goal. This objective can be achieved if A can persuade B to adopt beliefs which will result in B generating the goal A intends. In simple cases A will only need to make B believe that A wishes the goal to be adopted. This will occur in situations in which adoption of the goal will imply little planning and execution effort on the part of B, with no conflicts with B's existing goals. When achievement of the goal demands significant effort from B, or conflicts with B's existing goals, more complex beliefs might need to be induced. For example, B might need to be persuaded that achievement of the goal will result in benefits for B as well as for A. In the current work, monologue, and not dialogue, is being considered, and hence the contexts in which beliefs are held and the goal which A hopes that B will adopt are fixed so that beliefs can be treated as three-valued (believed, disbelieved or undecided), as proposed above.

Since A plans its monologue relative to its own model of B's beliefs, the beliefs that B actually holds are not accessible to A and cannot be considered relevant in the construction of A's monologue. Furthermore, it is possible for A to believe that B is ignorant of certain propositions, which precludes the possibility of B having any belief attitude (that is, any of the three values identified above) towards these propositions in A's model of B. In this case, A will believe B to be *unaware* of the proposition - a different relationship to belief, disbelief or indecision.

Another property of belief which is manifest in argument is that of saliency. Enthymeme contraction (where an implicit premise or conclusion is omitted) relies heavily upon saliency and so too does the process of focusing: keeping the argument to the point (Cf. the lower level focusing of [8]). The belief model used must therefore be able to competently handle the concept of saliency.

Argument modelling also relies on a treatment of the complex phenomenon of mutual belief: an argument is based upon common ground - a set of mutual beliefs (ie. which both parties hold, and which both parties also know the other to hold). Mutual belief is defined in terms of an infinite regress of nested beliefs. That is, A believes that A and B mutually believe a proposition, P, if A believes (i) P and (ii) that B believes P and (iii) that A and B mutually believe that A and B mutually believe P. The problem is to pragmatically choose a level of nesting beyond which 'mutual' belief is to be assumed. In making this choice, it is understood that no matter how many levels a system can cope with, it is always possible to construct a (highly convoluted) example which exceeds the capabilities of that system. From a psychological (and intuitive) point of view, choosing some arbitrary level of nesting by which to define mutuality seems rather implausible. In humans, it would appear that belief nesting is a resource bounded operation with no known limit, and it is possible to construct deeply nested examples which present remarkably little difficulty (such as the example in Fig. 2, below), though handling further complexity rapidly becomes extremely difficult and time consuming. It may be possible to utilise

In the classic Hitchcock film "North by Northwest", Cary Grant is the central character in a case of mistaken identity - he is mistaken for a goodguy agent by the evil James Mason. Towards the end, a scene occurs in which CG masquerades as the goodguy agent he is thought to be. As part of a goodguy ploy, he informs JM of his wish to defect, at which he is shot by EveMarie Saint, a goodguy agent working undercover as accomplice to JM. In her role as a villain, she needs to stop the defection, believing him to be a traitor to the goodguys. Thus with the proposition P that 'ES is a goodguy',

`BEL(ES, P)`	She knows she's a goodguy
`BEL(CG, P)`	He knows she's a goodguy ...
`BEL(ES, BEL(CG, ~P))`	... but she doesn't realise that ...
`BEL(CG, BEL(ES, BEL(CG, ~P)))`	... and he knows she doesn't realise ...
`BEL(Audience, BEL(CG, BEL(ES, BEL(CG, ~P))))`	... or at least that's what the audience thinks!

After JM has left however, CG gets up - the shooting had been faked. CG and ES must have been in league with each other after all. At this point the belief held must be changed to

`BEL(Audience, BMB(CG, ES, P))))`	They both know she's a goodguy, and both know that they both know it.

Hitchcock envisaged the whole situation and realised that it was unusual and interesting. There were in total, five levels of nested beliefs.

Fig. 2. Deep nesting of beliefs

this evidence and allow for a similar process in an implementation, so that some default operator (say, BMB, following [4]) is employed using a naively shallow level of nesting, but, in the light of new evidence, this may be replaced (or supplemented) with a more sophisticated nesting and appropriate operator.

3.2. Grounding of Beliefs

The knowledge representation adopted as the basis for this system has as a fundamental unit the concept of an *event*. An event is a piece of information representing a proposition which is perceived to be a part of reality: for example, "the sky is blue". The constituent elements of an event are the subject, the predicate and the object. It is possible for the object to be absent as in "the bomb exploded". If necessary, further features can be added, such as the time or the location of the event. This form of knowledge representation constitutes the basis for the semantic network on which the LOLITA system depends [18]. An example of such a network is given below in Fig. 3. The representation is reified so that events can themselves be the subject of events, with transformation being possible of the links between events (which represent various possible relations, such as implication or causation) so that these are also accessible as events. The events encoded within the knowledge base of an agent represent the beliefs held by that agent. The events can be representations of (the agents beliefs about) another agent's beliefs, allowing one agent to model the beliefs of another. An account of the semantics of the basic framework of the representation can be found in [17].

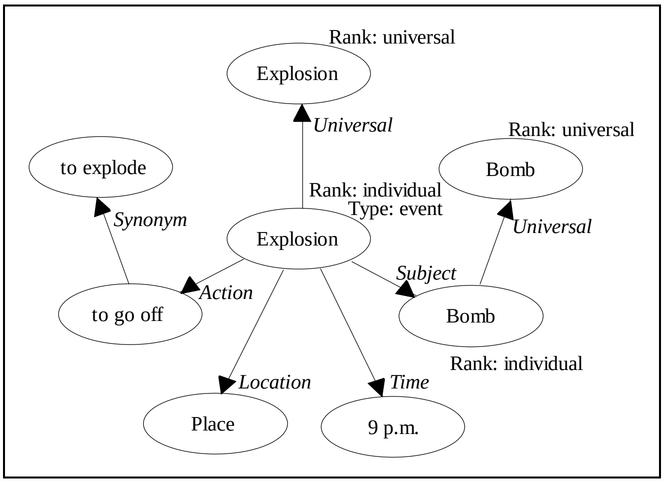

Fig. 3. “The bomb exploded”

In order to persuade another agent to adopt a belief, it is important to understand what might lead to the adoption. Beliefs are not held without reason: the reasons for an agent holding a belief are captured as the *supports* for that belief. In the LOLITA representation, this can be seen as a special event relationship holding between an event, E, and its supporter, E', namely: Supports(E', E). The presence of the pair of events in a network corresponding to the beliefs, or perceived beliefs, of an agent, E' and Supports(E', E), can be seen as sanctioning the inference that the event E holds.

The supports for a belief are of great importance, since it is the supports which allow an agent to decompose an argument into sub-arguments, through the identification of the supporting components which must be introduced into the main argument. Supporting relationships are established in the initial knowledge of the agent - the current work is not concerned with attempting to infer them from observations of the world or even from implicit relationships within the initial knowledge of the agent. The supports can be related to the event they support in different ways, but the two that have been considered so far are *implicational* relationships and *causal* relationships. In order for the recursive structure of events and their supports to terminate it is clear that there must be events which do not require support. These events are of great interest, since it is apparent that they represent premises of an argument and could prove difficult to persuade other agents to accept. It is therefore worth attempting to identify the nature of these events and determine why they might have no supports.

Events which have no supporting events are referred to in this work as *grounded* events. Grounded events can be grounded for several reasons. One key reason is that an agent will trust its own sensory input (which is not the same as the interpretation of that input). So, if the agent believes it has experienced some particular sensory experience, then the event representing that experience will be unsupported - it is grounded by virtue of the sensory experience itself. These events are a special case of a more general type of grounded event: the events which are grounded through external sources. In the case of sensory experience the external source is the sensory capacity of the agent itself, but, more generally, the source might be another agent. Of course, events which are sourced by another agent are (certainly for human agents, at least) only admitted into an agent's beliefs via its senses and it is possible to argue that the grounding of all events which have external sources must therefore be through sensory experiences. In practice it is not particularly helpful to distinguish the sensory event from the literal interpretation of that event when considering information provided by another agent: the question of the status of the content of the information (whether or not the content is accepted as a belief or not) is determined on the basis of an estimation of the reliability of the source agent, not on whether the senses that conveyed the message are considered reliable.

There is a further category of grounded events: some beliefs are held not because of a direct external source, but because of internal motivations and value-systems of an agent. An agent might believe that it is deemed punishable to carry out some action - this would be an externally sourced belief, grounded through the source from which it was obtained - but a belief that the action is wrong can only be based on the agent's own moral code. Events which are believed because of the agent's own value-system are considered to be grounded through that value-system. The distinction between events which are grounded through the agent's experience of the world and those which are grounded through the agent's value-system reflects the division of beliefs discussed earlier into factual beliefs and opinions.

Many beliefs are supported by a combination of grounded beliefs and causal relationships which reflect the way in which an agent believes the world works. For example, a belief that a ball is under a table might be supported by the grounded belief that the agent saw the ball roll slowly under the table and a (complex) belief that objects which are rolling slowly will quickly stop and stay where they come to rest. The first belief is grounded in sensory perception. The second belief (which might, in fact, be a structure of several interlocking beliefs) is a belief about the way that the world works. It is more difficult to assess the status of this belief: it is not grounded in experience, since a causal relationship cannot be experienced. It is unlikely to be sourced by a single information source: if it were, it might prove dangerously volatile, being susceptible to possible future revisions of the reliability of the source. Beliefs about the physical laws of the world in which an agent works are neither sourced by individual sensory experiences (although they can, of course, be undermined by single sensory experiences), nor is it likely that they are sourced by another agent. Further, these beliefs are not opinions - they do not reflect a value-judgement. Therefore, it must be possible for an agent to ground beliefs in one further way: beliefs about physical laws are grounded as hypotheses. That is, some beliefs are adopted because they represent working hypotheses from which deductions about the way the world will behave can be drawn, predictions made and control of the environment derived. In the tradition of falsificationism [13], these hypotheses survive until they can be bettered, or until demonstrated to be false.

Finally, there are beliefs which are a priori beliefs about the ontological status of the terms used in the construction of the beliefs. For example, agents will have beliefs about the meaning of elements such as part-of relationships. These beliefs are partly expressed in a declarative form, and, more importantly, partly procedurally, within the algorithms which manipulate the knowledge. Beliefs of this form will be assumed to be shared knowledge agreed upon by all parties to a monologue, since misunderstandings about the meaning of terms cannot be resolved in the scope of a monologue.

Having considered the various ways in which events can be grounded, it is now possible to consider the consequences of these alternatives for argument construction. Any unsupported event can be proposed to another agent by simply asserting it, in the hope that the agent will adopt it as an event sourced by the asserting agent. The adoption will depend on the assessment of the source agent's reliability on the information content according to the listening agent. In some cases this will prove adequate, but when adoption of the belief by the listening agent implies the generation and achievement of goals the reliability of the source agent might not be sufficient to lead to adoption of the belief. In this case, various strategies are possible according to the grounding of the event in the source agent's beliefs. If the event is grounded through being sourced by a third party, then an "appeal to authority" is possible, in which the third party is invoked as a source in the hope that this source will have the necessary reliability in the assessment of the listener to cause adoption of the belief. If the event is grounded through sensory experience then providing more detail of the context for the sensory experience can serve to improve the assessment of the reliability of the agent as a source, leading to acceptance of the veracity of the experience and the event it yields. If the event is a hypothesis about a physical law then an "appeal to the people" might be appropriate, if the law is a widely accepted one. Otherwise, demonstration of the predictive power of the law and a description of the testing to which it has been subjected will provide support for its adoption. Finally, opinions which rest purely on the value-system of the agent will prove the most difficult to persuade another agent to adopt. At present, these beliefs are handled by simply proposing them as premises to be accepted or rejected entirely at the listener's discretion. It is considered that argument founded on moral or ethical judgements are amongst the most difficult to uphold and the most bitterly contested of arguments.

4. Argument Objectives and Structure

The objectives of an argument will depend on the perceived beliefs of the hearer with respect to the belief that it is intended to convey. If the hearer is ignorant of the event, then simply informing the hearer of the event might be sufficient to have it adopted. If the hearer is aware of the event, but is undecided about whether to believe it, then it will be necessary to provide the hearer with supports from which to build a belief in the target event. Finally, if the hearer disbelieves the event then the objective of the argument will be to persuade the hearer to accept the event. If the hearer disbelieves the target event then the hearer must believe at least one event which is in contradiction with belief in the target event - this is an event-complement. This will be achieved by identifying supports that the hearer has for the beliefs in the event-complements (in the speaker's model of the hearer's beliefs) and systematically undermining these, before building supports for the belief that is intended to supplant the complementary belief originally held by the hearer. The new supports can be provided through explanation, unless the hearer believes in the complement of these events, in which case persuasive argument must again be employed. This coarse structure forms the foundation on which the argument structure can be built, and reflects the division into subarguments components developed in the AS level.

Through an analysis of a corpus of arguments drawn from a number of the sources mentioned in §2, the following structure has been identified:

(1) an argument consists of one or more premises and exactly one conclusion
(2) a premise can be a subargument (which itself consists of one or more premises and exactly one conclusion: the conclusion then stands as the premise in the superargument)
(3) a subargument is an integral unit whose components cannot be referred to from elsewhere, nor can the conclusion of a subargument rest upon premises extraneous to that subargument
(4) the only exception to (3) is where a conclusion in a distant subargument is restated locally as a premise

Analyses based on similar theories are performed in many texts: see for example, [25]. It is this structure which the AS level constructs, linking premises to conclusions through the use of two groups of operators: standard logical relations and inductively reasoned implications.

The first group comprises the standard rules of inference of classical logic. In the second group there are the inductive operators which are of three types: inductive generalisation, causal and analogical (eg. [10]). All three will have preconditions that are again tightly specified for context and belief, but with the additional constraint of the 'criteria of inductive strength', ie. the requirements for their application to produce an inductively strong argument.

A third group, not considered in this work, is the rhetorical fallacies, such as red herrings, straw men or false dichotomies. Although these fallacies have a role in natural language argument generation, they complicate the model of argument significantly, not least because of the necessary accompanying assumption that the audience does not have the powers of analysis to unmask the fallacies, leading to questions about the more general abilities of the audience to follow an argument where it is correctly structured. Further discussion of these fallacies can be found in [14].

5. Eloquence Generation

Although the EG level also employs a number of operators, the bulk of its functionality is based on the application of heuristics. The first task of the EG level is to control stylistic presentation, comprising a number of relatively low level modifications such as the vocabulary range, syntactic construction and the frequency of specific devices such as repetition. There are a large number of such factors detailed in the stylistic literature - see [16] for an overview, although these are intended to apply to human audiences with sophisticated appreciation of the stylistic phenomena. To date, such factors have been controlled by quite artificial and simplistic means, typically, the user setting a number of variables to particular values (eg. [22]). For the instantiations to be made and altered 'on-the-fly', the EG level must refer to a body of parameters, in addition to the belief model of the audience and context, all of which are modified dynamically. Within LOLITA, a powerful low-level plan realisation system exists which realises components of the semantic network as text, governed by stylistic parameters. This system is flexible and robust, allowing a wide-ranging control over text-generation while separating stylistic and content issues from the problems of grammatical correctness and syntactic construction [19].

One vitally important parameter affecting the argument is the relationship which the speaker wishes to create or maintain with the hearer. This relationship is established through stylistic rather than structural means, and is not necessarily divorced from other aims: if a hearer accepts the speaker's authoritative stance, for example, the speaker may be able to use the relationship to reinforce his statements (by increasing the hearer's assessment of the agent's reliability). Attempts at instigating different relationships between speaker and hearer account (at least in part) for a number of complex phenomena - humour, for example, is frequently used to establish the speaker as a friend (and therefore reliable and truthful). Such phenomena do not fall inside the scope of this work.

Some of the EG heuristics rely on the general principles of argument formation (though others make use of factors such as tone and rhythm). There are three possible statement orderings in an argument: pre-order (C P*), post-order (P* C), and hybrid order (P* C P*). The first is usually used where the P* are examples (in the case of factual conclusions: for opinions, the P* would often be analogies), especially when the hearer is being led to a hasty generalisation from the P* to the C. Pre-order is also used when the initial conclusion is deliberately provocative - the construction being used to draw attention to the argument (and consequently it is unusual to find a weak argument structured with the pre-order ordering). It can also be forced, should a premise not be accepted by the hearer, and require a sub-argument in its support. Post-order is the usual choice for longer, more complex or less convincing arguments (indeed some arguments are less convincing precisely because they are longer or more complex). It is also used for 'thin end of the wedge' argument and for grouping together premises which individually lend only very weak support to the conclusion. The hybrid construction is rarer, usually occurring when the speaker has completed an post-order argument which has not been accepted and therefore requires further support.

Finally, there are a number of features controlled by the EG level which lie somewhere between the stylistic and the rhetoric: artifices such as repetition and alliteration can prove extremely effective in constructing eloquent and compelling argument. It is interesting to note that almost all of EG heuristics are devices listed by classical texts as figures and tropes. Indeed this is a uniquely interesting property of the EG level: the emphasis it puts in utilisation of ideas posited in precomputational treatises, especially the classical texts of Cicero and Aristotle ([6], for example) and the ideas developed in the late middle ages and renaissance ([23] and [3] have been used as source material for much of the analysis). This fact poses unique challenges as well as affording unique advantages: on the one hand, the ideas are clear and unbiased, whilst on the other, may be difficult to transcribe to implementation.

6. Intention

Recent research has clearly shown that intention, or motivation, plays a crucial role in the generation of discourse, due to its key role in justification, explanation, failure recovery, self referential discourse and response to follow-up questioning [12], [26]. The framework proposed here employs intentions at every level of processing. At the highest level, intention is represented through the use of communicative goals (following Moore and Paris, [12]). An entire argument may ultimately rest upon a single intention such as believes(H, P). At the AS level, then, an abstract plan is produced which fulfils this top level goal. In contrast to Moore and Paris, communicative goals are not necessarily resolved into linguistic (ie. non-intention) goals. Rather, intentions can give rise both to linguistic structures and more refined intention structures. For example, the stylised description of the Modus Ponens operator, as used at the AS level, fulfils the postcondition believes(H, P) by introducing two new communicative goals, believes(H, X) and believes(H, implies(X, P)).

Modus Ponens		
Shell:	Preconditions:	believes(H, X)
		believes(H, implies (X, P))
	Add:	believes(H, P)
	Delete:	disbelieves(H, P)

Fig 4. Modus Ponens Operator Shell

It would be a dangerous oversimplification to assume that the MP step of an argument is simply a matter of stating the two premises followed by the conclusion, as might be inferred from Fig. 4. At a high level, there do exist the intentions for the hearer to believe X and implies(X, P) - this is part of the 'intention structure' described in Fig. 1. However, that intention structure undergoes heavy modification before being communicated.

At the AS level, then, intentions are substrate to the planning process; they form goals which are used in pre- and post-condition lists. At the EG level, most intentions play a rather different role, due to the difference in processing styles between the AS and EG. In particular, planning at the EG level primarily involves the use of maintenance goals, which affect the plan structure in a heuristic manner (as opposed to the more conventional achievement goals planned for and fulfilled at the AS level). For example, a typical EG level heuristic (implemented as a maintenance goal intention) blocks plan fragments which tell the hearer something they already know. In the example above, where the AS level suggests (in the body of the MP operator) uttering X, implies(X, P) and P, the EG level might remove the first premise if the hearer belief model indicated that the hearer already believed X.

The RST level also needs to maintain a representation of intention, and the proposed framework facilitates the adoption of the approach put forward by [12], wherein intentions map on a one-to-one basis with presentational relations and on a many-to-many basis with subject matter relations.

The hierarchical nature of the planning process means that the final surface structure communicated to the hearer is fully justified in terms of high level intentions, so that should explanation, justification or recovery be required, the system has recourse back to appropriate intentions with which to replan.

7. Planning Discourse

If the structure of knowledge alone is considered when planning discourse, it would appear that the planning task is not particularly difficult. This is because, given the initial assumption that the agent is not to deliberately mislead its audience and that its knowledge is consistent (or, rather, that once an inconsistency is discovered no argument will proceed founded on the inconsistent knowledge), operators that describe the argument process have a monotonic behaviour. That is, operators can be applied to shift the beliefs of an audience in only one direction: from disbelief through uncertainty to belief, or, conversely, from belief towards disbelief.

The role that planning plays in discourse construction is in the development and maintenance of *coherence* as the discourse develops. A speaker must present the components of an argument in a way that allows related elements to be connected in the hearer's knowledge. An argument cannot proceed by giving pieces of information in a disconnected and arbitrary way. It is necessary to organise the argument in a way that emphasises its structure and allows the hearer to identify the inferential structure and draw the appropriate inferences. This observation is reflected in the design of planning operators for the discourse planning domain. Operators affect not only the models of the state of the knowledge of the hearer but also the model of the hearer's view of the monologue, including its current focus, direction and intended purpose.

The planner used by the system described here is a hierarchical planner based on the use of *operator abstraction* and *encapsulation* [5]. It works by constructing a hierarchy of abstract plans, connected together by a special operation called *refinement*. The hierarchical structure of the plan development process corresponds to the organisation of the AO, AS and EG levels within the argument-planning framework. Most of the planning effort is concentrated within the AO and AS levels, with plan-post-processing being carried out during eloquence generation, which involves resolving previously unconstrained orderings between subarguments and the addition of stylistic elements to the plan structure based on the contextual information inherited from the AO and AS levels encoded within the plan structure. Abstract plans developed initially at the AO level and refined during the AS planning phase, shape the development of the final discourse by imposing increasing structure on the inter-relationships between components of the developing discourse. An abstract plan can be seen as a skeletal discourse, in which the overall structure is in place but none of the details of the argument have been planned. Refinement to more detailed planning levels, which is only performed when an abstract plan has been completed (that is, has no outstanding goals), must preserve the structure provided by the abstract plan. As a consequence of this, 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. Abstraction therefore assists in the planning of a complex argument by converging on a detailed argument from above. There is a great deal of evidence in the literature that this form of abstraction can considerably improve the performance of a classical planner [2].

Planners based on operator abstraction are characterised by the use of abstract operators which achieve an abstract effect and can be refined into the more detailed components that contribute towards these effects. Abstract operators contain some internal structure indicating what goals have to be achieved, or actions added, to achieve these abstract effects. The first planner to use operator abstraction was NOAH, [15], in which a plan was constructed as a procedural network composed of nodes representing operators and goals, and arcs representing temporal relations. When an operator is applied, its internal structure, consisting of goals to be achieved and actions to be applied, is added immediately to the network. NOAH is widely used as a basis for discourse and dialogue planning [9], and similar approaches were also adopted in NONLIN, [21], and SIPE, [24]. In the system described here, operator abstraction is used together with encapsulation which prevents the internal structures of abstract operators from being revealed before the contexts in which they will be interpreted have been constructed. Here, abstract operators are added to a plan as if they were primitive and only refined to reveal their hidden structures when the plan itself is completed and ready to be refined.

Operators have a *shell* and a *body*. The shell is visible at the planning level in which the operator is applied, and contains the preconditions and effects of the operator. The body contains a partial order on a set of goals which, when achieved after refinement of the plan, will combine to achieve the effects of the abstract operator. When a plan is refined it is necessary to tie together the local contents of the individual bodies to ensure that they maintain any relationships between the abstract operators in the plan prior to refinement. Thus, if the operators P and Q were ordered so that P precedes Q in the abstract plan, and P achieves a precondition for Q and, after refinement, P has been replaced by the goals p1 and p2 and Q has been replaced by the goals q1 and q2, it is necessary to have some way of insisting that the effect of p2 must last beyond the point in time at which Q was applied in the abstract plan. This coordination is achieved by the refinement operation of the planner and involves only linear work in the size of the abstract plan. Bodies of abstract operators contain only goals: not including plan steps in the bodies of operators ensures a fully flexible exploitation of context after refinement, so that unnecessary plan steps are not introduced.

An important feature of the abstraction mechanism offered by the planner used here is that different levels of abstraction can be characterised by different granularities of description. This allows the world, and the effects upon it which an agent wishes to bring about, to be expressed in terms of coarse-grained categories and relationships. For example: in the Towers of Hanoi domain towers of discs can be moved from peg to peg, and these moves will refine into many disc-moving operations at a later point in the development of the plan. In the blocks-world, towers can be transformed into different towers in two steps - destruction of the existing tower and construction of the new tower. These two steps will eventually refine into many carefully ordered block unstacking and stacking operations. In the construction of argument a coarse-grained argument to persuade a hearer to accept a proposition P might take the following form: Identify supports for Not(P), undermine one or more supports, explain reasons for belief in supports for P, inform about the causal link between supports and P. Following many levels of refinement this will result in a detailed argument in support of P, ready for realisation in an expressive language. It is possible for propositions of different granularities to co-exist within a plan, since the planner does not follow a rigidly hierarchical pattern of development within the AO, AS and EG levels, and the planner is equipped with mechanisms to detect and control their interaction [5].

The following operator is intended to give an example of how the beliefs of a speaker and hearer can be used to direct the planning process in the construction of a persuasive argument. The operator is not complete, since it does not include information concerning the current focus of the argument. In addition, the operator description language used by the hierarchical planner rests on a semantics based on information loss following the application of abstract operators. This makes operators syntactically more complex than this example and explanation of the complete encoding of a domain in this language is beyond the scope of this paper.

Persuade(H::agent, E::event)

Shell:	Preconditions:	disbelieves(H, E)
		supports(E'::event, E)
		disbelieves(H, supports(E', E))
	Add:	believes(H, E)
	Delete:	disbelieves(H, E)
		disbelieves(H, supports(E', E))
Body:	Goals:	t1: believes(H, E')
		t2: believes(H, supports(E', E))
		t3: believes(H, E)
	Ordering:	t1 < t3, t2 < t3

Fig 5. Persuade Operator

In this example, all propositions are evaluated with respect to the speaker's beliefs. The operator is abstract and contains sub-goals in its body. The preconditions are not sub-goals but conditions: the operator cannot be applied unless the speaker believes the preconditions in the state of application. The goals in the body are ordered with respect to one another. These goals and the ordering between them are revealed when the abstract plan containing this operator is refined. The operator shows how sub-goals involved in the achievement of an effect can be encapsulated so that the details of achieving them do not arise until the context in which they can be most usefully considered has been created. The principle upon which the hierarchical planning approach described in this paper is built is that, by postponing decisions, many of the choices that would normally arise in resolving them can be excluded. This is partly achieved by the ability to order goals in the bodies of abstract operators which are unordered when they appear as preconditions in STRIPS-style operators.

A similar persuade operator can be applied when the hearer disbelieves E' itself, but possibly accepts that E' supports E. If the hearer disbelieves E' then the hearer can be inferred to believe NOT(E'), since failure to believe in either E or E' implies that the hearer is undecided about both. An operator is required to undermine the hearer's beliefs in NOT(E') so that the hearer will be ready to accept reasoned arguments for belief in E'. This operator works by identifying the hearer's supports for E' and making the hearer undecided about them. This can be done at each stage either by undermining the hearer's belief in the supporting event or by undermining the supporting connection between two events. The process terminates more quickly if the supporting link is broken. If the hearer's beliefs in the events themselves are tackled the process can involve recursive invocation of the undermining process.

The persuade, explain and inform operators, of which this one is a simplified example, are applied at the AO level of the planning process. Operators embodying the principles of logical argument, such as Modus Ponens, are introduced at the AS level once the foundational and intentional structure of the argument is in place. The operations performed at the EG level are not encoded as planning operators since this level is a post-processing phase.

8. Conclusion

This paper has described an architecture for constructing extended arguments from the highest level of pragmatic, intention-rich goals to a string of utterances, using a core hierarchical planner employing operators embodying the principles of argument construction. The planner is responsible for all levels of the construction and organisation of the argument. The output of the planner is a detailed discourse plan complete with annotations to indicate emphasis, mood, rhythm, and other features characterising effective argument. This structure is given as input to the text generation level of the system. This component of the architecture is required to realise the abstract intentional structures into natural language utterances and is provided by LOLITA, a large scale, domain independent natural language system, in which a natural language generation algorithm is already implemented, [19], which subsumes responsibility for solving certain low-level text generation planning problems. The work described in this paper is still in progress and is currently concentrated on the representation of discourse operators for use by the hierarchical planner, the development of techniques applied at the EG level and the effects on argument construction of the speaker's partial knowledge of the hearer's beliefs. Future work will include the extension of the current framework to consider dialogue planning between motivated agents.

References

[1] Ackermann, R.J. *Belief and Knowledge*, Anchor, New York (1972)

[2] Baccus F. & Yang Q. The expected value of hierarchical problem-solving, in *Proceedings of the National Conference on Artifical Intelligence* (1992)

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

[4] Cohen, P.R., Levesque, H.J. "Rational Interaction as the Basis for Communication" in Cohen, P.R., Morgan, J., Pollack, M.E., (eds), *Intentions in Communication*, MIT Press, Boston (1990) 221-255

[5] Fox, M. & Long, D.P. “Hierarchical Planning using Abstraction”, *IEE Proc on Control Theory and Applications* **142** (3) (1995) 197-210

[6] Freese, J.H. (trans), Aristotle *The Art of Rhetoric*, Heinmann, London (1926)

[7] Galliers, J.R. "Autonomous belief revision and communication" in Gardenfors, P., (ed), *Belief Revision*, Cambridge University Press, Cambridge (1992) 220-246

[8] Grosz, B., Sidner, C.L "Plans for Discourse" in Cohen, P.R., Morgan, J. & Pollack, M.E., (eds), *Intentions in Communication*, MIT Press, Boston (1990) 418-444

[9] Hovy, E. H. "Automated Discourse Generation Using Discourse Structure Relations", *Artificial Intelligence* **63** (1993) 341-385

[10] Long, D., Garigliano, R. *Reasoning by Analogy and Causality: A Model and Application*, Ellis Horwood (1994)

[11] Mann, W.C., Thompson, S.A. "Rhetorical structure theory: description and construction of text structures" in Kempen, G., *Natural Language Generation: New Results in AI, Psychology and Linguistics*, Kluwer (1986) 279-300

[12] Moore, J.D., Paris, C.L. “Planning Text for Advisory Dialogues: Capturing Intentional and

Rhetorical Information", *Computational Linguistics* **19** (4) (1994) 651-694
[13] Popper, K.R. *The Logic of Scientific Discovery*, Hutchinson Education (1959)
[14] Reed, C.A., Long, D.P. & Fox, M. "An Architecture for Argumentative Dialogue Planning", in Gabbay, D. & Ohlbach, H.J. *Practical Reasoning*, Lecture Notes in AI Vol. 1085, Springer Verlag (1996) 555-566
[15] Sacerdoti, E.D. *A structure for plans and behaviour*, Elsevier, North Holland (1977)
[16] Sandell, R. *Linguistic Style and Persuasion*, Academic Press, London (1977)
[17] Shiu, S., Luo, Z., Garigliano, R. "A type theoretic semantics for SemNet", in Gabbay, D. & Ohlbach, H.J. *Practical Reasoning*, Lecture Notes in AI Vol. 1085, Springer Verlag (1996) 582-595
[18] Smith, M.H., Garigliano, R., Morgan, R.C. "Generation in the LOLITA system: An engineering approach" in *Proceedings of the 7th International Workshop on Natural Language Generation*, Kennebunkport, Maine (1994)
[19] Smith, M.H. *Natural Language Generation in LOLITA*, PhD Thesis, Durham University (1996)
[20] Sycara, K. "Argumentation: Planning Other Agent's Plans" in *Proceedings of the International Joint Conference on Artificial Intelligence* (1989) 517-523
[21] Tate, A. *Project planning using a hierarchic non-linear planner*, Technical Report, Department of Artifical Intelligence, University of Edinburgh (1976)
[22] Vreeswijk, G. "IACAS: an implementation of Chisolm's Principles of Knowledge", *Proceedings of the 2nd Dutch/German Workshop on Nonmonotonic Reasoning*, Utrecht, Witteveen, C. et al. (eds) (1995) 225-234
[23] Whately, R. *Logic*, Richard Griffin, London (1855)
[24] Wilkins D. *Practical planning: Extending the classical AI paradigm*, Addison-Wesley (1988)
[25] Wilson, B.A. *The Anatomy of Argument*, University Press of America, Washington (1980)
[26] Young, R.M., Moore, J.D. "DPOCL: A principled approach to discourse planning" in *Proceedings of the 7th International Workshop on Natural Language Generation*, Kennebunkport, Maine (1994) 13-20