1 Introduction

All arguments share certain key similarities: they have a goal and some support for the goal, although the form of the goal and support may vary dramatically. As an example, consider the following argument, which was posted by Clive Gibbons to the usenet group `sci.astro.amateur` on the 23rd of March, 2000. This argument was a reply to an earlier post asking why Meade's stock price had risen so much lately.

> *The recent spike has no doubt been caused by the March 20th airing of CBS Marketwatch Weekend, where TeraBeam CEO Dan Hesse was interviewed. He made a nice pitch about the future of high speed ISP technology. The show also mentioned that while TeraBeam wasn't a public company (yet), Meade Instruments *was* and they had just inked a new service alliance with TeraBeam. Presto! Guess whose stock got a big boost? ;)*

This argument, which was considered persuasive at the time it was posted (when Internet related stocks were booming), does not explicitly address the question. Instead, it puts forward certain propositions and invites the reader, via a rhetorical question, to put the pieces together.

Even such a simple example shows why argumentation is an interesting domain for dialogue systems, as it highlights aspects of dialogue which may be obscured in systems with simpler and more prescriptive interactions. For example, a system which looks up bus routes is unlikely to require a complex rebuttal-handling capability or a grasp of rhetoric. The requirements brought to light by argumentation systems are expected to help practitioners determine the features required by their particular dialogue systems, the level of support required for each feature, and the techniques that are appropriate for producing this level of support.

An ideal interactive argumentation system would have a long list of dialogue requirements. The fictitious argument in Figure 1 (U stands for User and S for System) illustrates a subset of them. The following list describes the capabilities that are necessary to fulfill these requirements.

Understand implicit goals (U1) – The user does not explicitly state the goal of the current argument, but the system determines that the user is interested in speculating about who will win the premiership, rather than discussing how well the Kangaroos have played during the pre-season matches.

Generate enthymematic arguments (S1) – These are arguments which omit easily inferable information [Zukerman et al., 1998]. As shown in our example, the system indirectly refutes the proposition that the Kangaroos will win by proposing Adelaide as a more likely candidate.

Understand enthymematic arguments (U2) – As for S1, the user's rebuttal is enthymematic in that it does not present a complete inference chain, and does not explicitly indicate whether he or she is attacking the system's last statement (S1) or reinforcing his or her own previous argument (U1).

Judiciously admit new information (S2) – The system concedes the user's point about the win/loss ratio (U2). If this information was previously unknown to the system, the system must decide whether it should admit this information, the belief it should associate with this new information, and how this information affects the argument.

U1: "The Northern Kangaroos will probably win the Australian Rules Football League premiership this year, since they have done well in the last few years and were in good form in the pre-season competition."

S1: "In each of the last two years Adelaide has fared poorly in the pre-season competition but has won the premiership."

U2: "But Adelaide has a worse win/loss ratio when playing matches at the MCG [the site where the premiership final is played] than the Kangaroos."

S2: "True, but if a club can win enough matches to make the finals, then their win/loss ratio for finals matches is more important than their win/loss ratio over the whole season."

U3: "Fair enough."

S3: "... and, over the last two years, Adelaide has had a better win/loss ratio than the Kangaroos when playing finals matches at the MCG."

U4: "What if Shaun Rehn [Adelaide's ruckman] needs more knee surgery?"

S4: "Adelaide would be somewhat less likely to win, but I would still favour them."

U5: "OK, what about Wayne Carey [Kangaroos' captain]?"

S5: "He is a good player, but not good enough to carry the Kangaroos."

U6: "That's not what I mean."

S6: "Wayne Carey has a shoulder injury at the moment."

Figure 1: Sample dialogue with an ideal argumentation system

Determine an argumentation strategy (S2) – The system uses a conditional argument to defeat the user's argument about Adelaide's win/loss ratio (U2). Context may make certain argumentation strategies more appealing than others. For instance, if the user has repeatedly mentioned a particular belief, the system may select an argumentation strategy that takes advantage of this (entrenched) belief.

Handle turn taking and interruptions (U3, S3) – This capability is asymmetrical, since the user should be allowed to interrupt the system, but usually the system would not interrupt the user. In this example, the system notes that the user has accepted its (partial) argument, but decides to present the remainder of its argument anyway.

Handle different dialogue acts (U4) – Conversational partners may change the topic of conversation or may probe each other's arguments to decide which parts of an argument to accept. This may be done by means of different types of sub-dialogues, e.g., information-sharing [Chu-Carroll and Carberry, 1995] or clarification [Litman and Allen, 1987]. In our example, the user poses an exploratory *what if* query to speculate on the effect of a particular event on the argument goal.

Select an interpretation (U5) – Utterance U5 may be interpreted as an exploratory query to consider the effect of Wayne Carey on the argument goal or as a shift in topic, where the focus is now on Carey's health. Both interpretations should be considered, and a preferred interpretation selected [Litman and Allen, 1987, Raskutti and Zukerman, 1991]. In this example, the former interpretation is (erroneously) selected.

Recover from misunderstandings (U6, S6) – U5 was previously interpreted as a query about the effect of Wayne Carey on the argument goal. However, when the user indicates that this is not his or her intent, the system adopts an alternative interpretation – that the user wants to shift the discussion to Wayne Carey's health.

This list of capabilities is not complete, nor is it unique to argumentation. Most of these capabilities are required by any dialogue system to some degree, but they are heavily called upon in argumentation systems.

During the initial design and construction of our argumentation system, NAG (*Nice Argument Generator*), our goals were broad: (1) to develop an architecture and system for producing arguments that are correct and, where possible, persuasive; and (2) to model certain aspects of human argumentation, such as selecting appropriate argumentation strategies and generating enthymematic arguments. Later we expanded our goals to include operators which enable a user to interact with the argumentation system. Additional work on understanding implicit goals and arguments was undertaken in a follow-on system called BIAS [Jitnah et al., 2000, Zukerman et al., 2000a]. In this paper, we focus on NAG's argument-generation process and on its dialogue capabilities. In particular, we demonstrate how NAG succeeds in providing some of the capabilities mentioned above and discuss why it fails in providing others.

This paper is organized as follows. In the next section we present a sample interaction with NAG. We then discuss NAG's knowledge representation scheme, which integrates belief representation with attentional focus to enable the consideration of contextual information during the argumentation process (Section 3). In Section 4 we consider the system's reasoning and argumentation capabilities, which support the generation and understanding of enthymematic arguments. We then discuss the probabilistic patterns and argument grammar that enable NAG to render its arguments in English (Section 5). Next we describe NAG's dialogue capabilities, which support the handling of exploratory queries as a first installment of a more comprehensive dialogue facility (Section 6). In Section 7 we examine NAG's performance with respect to the argumentation and dialogue capabilities listed above, and in Section 8 we present the results of our preliminary evaluations. Finally, we discuss related research (Section 9) and present concluding remarks (Section 10).

2 Sample Interaction

In this section, we describe an actual interaction with a user. Our domain is a fictitious murder scenario, which was chosen to minimize the influence of a user's pre-existing beliefs. The interaction started when the user was given the following preamble.

> *Scratchy, a notorious spy from Vulcan, was found murdered in his bedroom, which is in the second storey of his house.*
>
> *Indentations were found in the ground right outside Scratchy's window, and they were observed to be circular. In addition, one set of footprints was found outside Scratchy's window, but the bushes were undisturbed.*
>
> *Itchy and Scratchy were long-term enemies, and Itchy's fingerprints were found on the murder weapon (a gun discovered at the scene and registered to Itchy).*

Itchy, who is the head of the INS (Internal Naturalization Service), has a ladder with oblong supports, which he was planning to use to paint his house.

Poochie, the town's mayor, has a cylindrically-supported ladder which is available only to him.

The user was then asked for his degree of belief in the goal proposition [Itchy did not murder Scratchy] and in Itchy's means, opportunity and motive to kill Scratchy, and also for his confidence in his judgment regarding Itchy's guilt. The user indicated that it was *rather likely* that Itchy had a motive to kill Scratchy, *quite likely* that he had the means, *a little bit unlikely* that he had the opportunity, and *a little bit unlikely* that he killed Scratchy; his confidence in this last judgment was *a little bit unconfident.*[1] NAG then generated the initial argument displayed in Figure 2 to support the goal. This argument takes the user's beliefs into account, but the presented degrees of belief are NAG's.

After this argument, the user retained his belief in Itchy's motive to kill Scratchy, but reduced his belief in Itchy's means to kill Scratchy to *rather likely.* He also reduced his belief in Itchy's opportunity to murder Scratchy and Itchy's guilt to *quite unlikely,* which are closer to NAG's normative beliefs than the user's initial beliefs after the preamble. However, his confidence in his judgment regarding Itchy's guilt dropped to *quite unconfident.* After the user entered these values (also shown in Figure 2), NAG presented once more its argument for Itchy's innocence. This argument constituted the basis for the exploratory interaction.

The exploratory interaction started with a request to exclude the proposition [Itchy fired the gun], which resulted in a stronger argument for Itchy's innocence, where the paragraph regarding Itchy's means to murder Scratchy (second paragraph in the argument in Figure 2) was omitted, and the concluding paragraph was replaced as follows.

Even though Itchy very probably had a motive to kill Scratchy and Itchy could have had the means to murder Scratchy, the very probable lack of opportunity to murder Scratchy implies Itchy's almost certain innocence.

The user returned to the original argument (without retaining the exclusion), and asked NAG to support the proposition [Only one person was outside Scratchy's window] (Figure 2, paragraph 3). NAG generated the following sub-argument.

A single set of footprints outside Scratchy's window implies a single person being outside Scratchy's window.

The user then requested that this sub-argument be included in the main argument, and asked a *What if* question which explored the effect of a belief of *even chance* in the proposition [Itchy was outside Scratchy's window] (Figure 2, paragraphs 3 and 5). This resulted in an argument which yielded a weak belief in the goal.

Itchy and Scratchy very probably hated each other; therefore, Itchy very probably had a motive to kill Scratchy.

[1]The descriptive names for the probabilities are based on those described in [Elsaesser, 1987].

Itchy fired the gun and Itchy's gun was used to kill Scratchy; hence, Itchy had the means to murder Scratchy.

Circular indentations were found outside Scratchy's window; hence, Scratchy very probably was shot from outside the window.

If Itchy maybe was outside Scratchy's window and Scratchy very probably was shot from outside the window, Itchy possibly had no opportunity to murder Scratchy.

Despite a very probable motive to kill Scratchy and the means to murder Scratchy, the possible lack of opportunity to murder Scratchy implies Itchy's possible innocence.

After the exploratory interaction, the user was asked post-test questions regarding his belief in the goal proposition and in Itchy's means, opportunity and motive for killing Scratchy. The user did not change his beliefs or his confidence as a result of the interaction (but three of these beliefs were already close to NAG's).

3 Knowledge Representation

We define a *nice* argument to be both normatively correct and persuasive for the target audience. Generating a nice argument sometimes involves trade offs. For example, normatively correct information may be omitted if it is deemed not to be persuasive. NAG attempts to generate nice arguments, but if this is not possible, normative correctness is preferred to persuasiveness. The generation of correct arguments requires normative domain knowledge, while the generation of persuasive arguments requires a model of the audience's beliefs and inferential ability. The latter is also required for the production of enthymematic arguments, to support the omission of easily inferable information.

In this section we describe NAG's main knowledge representation formalisms: Bayesian networks (BNs) [Pearl, 1988], which are used for reasoning about arguments, and semantic networks, which incorporate contextual information.

3.1 Domain Knowledge and the User Model

When constructing or analyzing an argument, NAG relies on two collections of information: a normative model composed of different types of Knowledge Bases (KBs) which represent NAG's best understanding of the domain of the argument, and a user model also composed of different types of KBs which represent the user's presumed beliefs and inferences. A KB represents information in a single format, e.g., a semantic network (SN), a BN or a rule-based system. The KBs in the normative and user models are consulted by specialist Reasoning Agents, which are activated to fill gaps in a partial argument (Section 4). The KBs in the user model are consulted to make an argument persuasive for the target audience, while the normative KBs are consulted to generate a correct argument. The consultation of these models allows NAG to balance the normative correctness and persuasiveness of its arguments.

When reasoning about an argument, relevant material from several KBs may need to be combined into a common representation. NAG uses BNs for this purpose. One BN is used to combine information obtained from the user model KBs, and another BN is used to combine information obtained from the normative model KBs.

Bayesian networks – a brief overview. BNs have become a popular representation for reasoning under uncertainty, as they integrate a graphical representation of the relationships between propositions with a sound Bayesian foundation. BNs are directed acyclic graphs where nodes correspond to random variables. The nodes in a BN are connected by directed arcs, which may be thought of as causal or influence links; a node is influenced by its parents. The connections also specify independence assumptions between nodes, which allow the joint probability distribution of all the random variables to be specified by exponentially fewer probability values than the full joint distribution. A *conditional probability distribution* (CPD) is associated with each node. The CPD gives the probability of each node value for all combinations of the values of its parent nodes. The probability distribution for a node with no predecessors is its prior distribution. Given these priors and the CPDs, we can compute posterior probability distributions for all the nodes in a BN, which represent *beliefs* about the values of these nodes. The observation of specific values for nodes is called *evidence*. Beliefs are updated by re-computing the posterior probability distributions given the evidence. Belief propagation for singly-connected networks can be done efficiently using a message passing algorithm [Pearl, 1988]. When networks are multiply-connected (i.e., when there is a loop in the underlying undirected graph), simple belief propagation is not possible; informally, this is because we can no longer be sure whether evidence has already been counted at a node having arrived via another route. In such cases, inference algorithms based on clustering, conditioning or stochastic simulation may be used [Pearl, 1988]. Since belief propagation can take considerable time for large networks, NAG updates only the portion of each BN containing nodes relevant to the argument at hand. This saves time and also crudely models the way in which people operate, since people generally do not exhaustively explore every ramification of their plans.

Bayesian networks have been used in several user modeling applications, e.g., [Charniak and Goldman, 1993, Conati et al., 1997]. Charniak and Goldman (1993) used BNs for plan recognition in the framework of story understanding. They automatically generated a BN from a sequence of observations by applying rules which utilize plan knowledge to instantiate the network. The incorporation of prior probabilities into this network supports the selection of plausible explanations of observed actions. Similarly, Conati *et al.* (1997) applied the mechanism described in [Huber et al., 1994] to automatically construct a BN from the output of a rule-based physics problem solver that generates all the possible solutions to a given physics problem. This BN was then used to identify a student's problem-solving strategy and predict his or her next step. In NAG's case, the BN in the normative model and the BN in the user model are incrementally extended by the Reasoning Agents using information obtained from the KBs. The way in which a Reasoning Agent looks for relevant material in a KB varies according to the type of the KB. For instance, given a proposition to be supported, a Reasoning Agent that accesses a rule-based system focuses on rules that have this proposition as their consequent or antecedent. These rules are converted into sub-graphs whose nodes represent propositions and whose links represent inferences. These sub-graphs are then added to the BN where the argument is being built. That is, information sourced from the user model KBs is added to the user model BN, and information sourced from the normative model KBs is added to the normative model BN. Thus, the nodes in the BNs in the normative and user models and the beliefs in these nodes may change as NAG reasons about an argument (as a result of belief propagation or the incorporation into the BNs of additional information found in the KBs). This is in contrast to the KBs consulted by NAG, which are static.

The portions of the normative and user model BNs that are structurally common to both networks form an *Argument Graph*. For example, consider the situation depicted in Figure 3 where NAG's normative model believes that Poochie's ladder was outside Scratchy's window (P1) and that it is available only to Poochie (P2). In addition, the normative model has an inference pattern whereby these propositions have implications about Poochie being outside Scratchy's window (P3). Now, let us assume that the user model shares the normative model's belief in P1, is uncertain about P2, and also believes that Poochie was seen at the local toy store on the night of the murder (P4). Further, the user model has an inference pattern whereby P1, P2 and P4 affect the belief in P3. In this example, P1 and P2 and their links to P3 are included in the Argument Graph, but P4 is excluded since it does not appear in the normative model. When analyzing the Argument Graph, propagation is done twice, once for the BN in the normative model and once for that in the user model (in our example, the probability calculations for the user model are marginalized over P4 to match the inference pattern in the normative model). Thus, we determine the normative strength and persuasiveness of the same argument by consulting the normative model and the user model respectively, thereby determining the niceness of the argument. In this example, NAG's argument for P3 achieves a strong belief in this proposition in the normative model, but only a middling belief in the user model.

Constructing the Argument Graph from propositions in the user model reduces the likelihood that the argument will contain propositions with which the user disagrees. As a result, the user may be more inclined to believe the resulting argument. However, in order to maintain normative correctness, the propositions in NAG's arguments must also appear in the normative model with levels of belief that are compatible with those in the user model. If this is not the case, then it will not be possible to generate a persuasive argument that is also normatively correct. For instance, this might happen when NAG's user model contains insufficient domain knowledge to argue effectively, or when the beliefs in the user model largely contradict those in the normative model. We submit that in these cases, it is better for NAG to present some argument than to say nothing. Therefore, if an initial attempt to build a nice argument fails, NAG tries to complete the argument generation process using just the normative model. It is worth noting that by the time NAG gives up building a nice argument by drawing on both the normative model and the user model, it will have explored the relevant portions of the normative model quite thoroughly. Hence, the resulting argument will often begin with observable premises. Still, such an argument is probably less persuasive than one that is completely based on the user model, since the user may not accept the new or contradicting premises being presented. However, our current model does not account for this type of impact.

3.2 Incorporating Context

NAG uses contextual information during content planning to quickly find relevant information that should be included in the current argument, and during argument presentation to order the propositions in the argument and omit propositions the user is likely to infer from already mentioned material (Section 5).

NAG captures connections between the items mentioned in the discourse by using a *semantic-Bayesian network*, which is composed of a hierarchical semantic network built on top of a BN (Figure 4). Each of the normative model and the user model contains an instance of this structure. The semantic network portion (upper part of the pyramid in Figure 4) contains concepts that are connected to each other as well as to propositions in the BN. These connections represent associative relations. The BN portion (base of the pyramid) represents causal and evidential relations between propositions. The semantic-Bayesian network is used by NAG to simulate attentional focus in each model. During content planning, this simulation allows NAG to determine which propositions in its normative and user models are salient in the current context, and hence potentially useful for the argumentation process. For instance, talking about the proposition [Itchy murdered Scratchy] may remind both the user and the system of deaths and shootings, and in turn of guns, which prompts the investigation of propositions involving guns, such as [Itchy used the gun] or [Itchy's gun was used to kill Scratchy].

At the beginning of the argumentation process, NAG receives as input a goal proposition (with a target level of belief) and propositions containing background information. In the examples discussed in this paper, the goal proposition is [Itchy did not murder Scratchy], and the background propositions correspond to the preamble shown in Section 2. This information is used to establish an initial context for the argument, which consists of the goal proposition and salient concepts and propositions obtained from the goal and the preamble propositions. These salient concepts and propositions are obtained by performing *spreading activation* [Anderson, 1983] in the semantic-Bayesian network in each of the user model and the normative model. This process simulates the focusing of attention on concepts and propositions one is reminded of after seeing the goal proposition and the preamble.

The spread of activation starts from the goal proposition, the propositions in the preamble, and the concepts in these propositions. These concepts and propositions receive an initial level of activation, which is then passed through the semantic-Bayesian networks – each node being activated to the degree implied by the activation levels of its neighbours, the strength of association to those neighbours, and its immediately prior activation level (vitiated by a time-decay factor). For example, when the node representing murder is activated, the strongly connected nodes representing death, kill and shoot receive a high level of activation, while the more weakly connected nodes representing gun and blood receive a lower level of activation. Activation now spreads from the newly activated nodes, e.g., causing shoot to pass additional activation to gun, and so on. The spreading activation process continues until an activation cycle fails to activate any new node. At this time, the items in the semantic-Bayesian networks which achieve a threshold activation level are incorporated into the new context. In our examples, after the presentation of the preamble and the argument goal, the salient concepts are Itchy, Scratchy, ladder, murder and window, and the salient propositions are those which include more than two of these concepts.

As the content planning process progresses, the argument context gets expanded. During argument construction, it contains the concepts and propositions included in the current partial argument plus the concepts and propositions that just became salient. For example, as NAG attempts to create an argument about Itchy's innocence, it tries to demonstrate that Itchy lacked the means, motive or opportunity to commit the crime. Thus, the context is extended with these subgoals. In addition, the context may be expanded through the incorporation of associatively activated propositions that are not directly connected within the BNs. For instance, if the propositions [Scratchy's dead body found] and [Times newspaper reports Scratchy murdered] become activated, then the concepts death and time will also become active. These concepts will in turn activate other propositions, such as [Time of death was 11pm]. The user's interactions with the system also affect the context, e.g., when the user asks NAG to include a proposition in the current argument, that proposition is brought into the current context (Section 6).

The context extension and spreading activation processes provide NAG with a direct implementation of attentional focus, which is used to identify portions of the semantic-Bayesian networks that are relevant to the current argument [Zukerman et al., 1998].

4 The Argumentation Process

NAG's main components and the data flow between them are illustrated in Figure 5. The argumentation process consists of a series of focusing-generation-analysis cycles which are driven by the Argument Strategist. During argument generation, the Generator is invoked to build up NAG's argument, and the Analyzer is invoked to assess the correctness and persuasiveness of the argument.

The argument generation process starts with the receipt of an argument goal (and a target level of belief), which is then passed to the Strategist. At present, NAG is given the initial goal, and the user can select subsequent goals. Each focusing-generation-analysis cycle proceeds as follows.

Focusing – The Strategist first invokes the Attentional Mechanism (Section 3.2) to focus attention on propositions in the user model and the normative model that are likely to be useful in the argument. In the first cycle, this process generates an initial context, whose propositions are shown in Figure 6, and in later cycles this process extends the argumentation context.

Generation – The Strategist then calls the Generator to continue the argument building process. This is done by activating the Reasoning Agents in order to find additional information to be incorporated into the Argument Graph (Section 3.1). The Reasoning Agents perform this task by looking in the KBs for material pertaining to each previously unexamined proposition in the current context. In the first cycle, none of the propositions in the context have been previously examined, and so are passed to the Reasoning Agents; during later cycles, only newly added propositions are investigated. In our example, in the first cycle the generation process connects several of the propositions in the initial context and uncovers some new propositions, which are linked within the BN, e.g., [Itchy and Scratchy were enemies] and [Itchy used the gun]. This process yields the Argument Graph in Figure 7, where propositions in the initial context appear in dark grey and newly added propositions in light grey.

Analysis – The Argument Graph is returned to the Strategist, which invokes the Analyzer to determine the niceness of the argument. The Analyzer performs constrained Bayesian propagation on the portion of the Argument Graph that is connected to the goal. This is done once in each of the normative model and the user model to assess the normative strength and persuasiveness of the argument respectively.

If the resulting belief in the goal proposition fails to satisfy the given target level of belief in the normative model or the user model, then the Strategist again calls the Attentional Mechanism to expand the current context (which includes the Argument Graph), initiating another focusing-generation-analysis cycle. If the Generator reports to the Strategist that the argument cannot be enhanced further in the user model, then the Strategist instructs the Generator to continue collecting material from the normative model only.[2] If the Generator reports a failure to find new material in the normative model, then NAG attempts to argue for the opposite goal, e.g., if Itchy's innocence cannot be proven, perhaps his guilt can. NAG then presents to the user the argument which yields the stronger belief in its goal. This ability to reverse the belief in the goal enables the system to react appropriately when new information comes to light or when a hypothetical question is asked by the user. For example, the user may ask NAG a hypothetical question which contradicts a previously accepted belief in a crucial proposition, e.g., "What if Itchy *didn't* have a motive to murder Scratchy?" In this case, NAG may generate a new argument where the revised belief in the goal, in this case Itchy's innocence or guilt, is very different from the belief that was held before the hypothetical question was asked.

After producing an Argument Graph that satisfies the Analyzer in both the normative model and the user model (e.g., Figure 8, where propositions added in the last cycle appear in white), NAG determines an appropriate argument presentation strategy. The strategies being considered at present are: premise to goal, inference to the best explanation, *reductio ad absurdum* and argument by cases. However, the texts presented in this paper and those presented to the users in our evaluations were generated using only the premise to goal argumentation strategy. The generation of arguments using different strategies is discussed in [Zukerman et al., 2000b].

Once an Argument Graph corresponding to a nice argument has been generated and a presentation strategy selected, NAG tries to remove propositions that have only a small effect on the belief in their consequents. After suggesting the removal of a particular proposition, the Analyzer checks whether the argument still achieves its goal. If it does not, the proposition in question is reinstated. After no more propositions can be removed in this manner, the Argument Graph is passed to the Presenter/Interface. This module orders the propositions in the argument, removes easily inferred intermediate conclusions, renders the argument in hypertext form, and presents the argument to the user through a WWW interface [Zukerman et al., 1999].

An interesting feature of NAG's focusing-generation-analysis cycle is that it may also be used for understanding users' arguments, since most arguments generated by people are enthymematic. This would be done by

[2]This strategy may occasionally result in missed opportunities to argue. For example, this happens if earlier propositions from the normative model were omitted from the Argument Graph because they did not appear in the user model, but these propositions are essential for the completion of an argument.

calling the Attentional Mechanism to focus on propositions that are relevant to a user's argument and may be useful in filling small reasoning gaps in this argument; then using the Generator to construct small sub-arguments around these propositions in order to fill the reasoning gaps; and finally calling the Analyzer to check whether the augmented argument is acceptable. If the Generator could produce sub-arguments which readily repaired any gaps (e.g., with a few, "obvious" inferential steps), then the user's enthymematic argument would be understood by the system. BIAS (*Bayesian Interactive Argumentation System*) [Jitnah et al., 2000, Zukerman et al., 2000a] is a follow-on system which extends NAG's framework in this manner. BIAS calls NAG to create arguments and then interprets the user's rejoinders, which are constructed piece-wise with a WWW interface. In the future, we intend to enhance BIAS' capabilities so that it can interpret a wider variety of rejoinders and counter-arguments generated by a user – the ultimate goal being a fully-fledged argumentation system that can converse with the user indefinitely.

5 Argumentation Patterns and Grammar

The generation of natural language output from a BN requires the identification of probabilistic patterns which yield specific argumentation patterns. The patterns we have considered so far are: *explain-away* [Pearl, 1988], *neutralize, contradict, cause, evidence* and *imply* (which is a generic pattern that is used when the others are not applicable). The formulas presented below identify these probabilistic patterns with reference to the simple BN in Figure 9 (A and B may be groups of nodes).

Explain away. Reflects a situation where there are several potential explanations for a proposition, and finding out that one of these explanations is likely reduces the probability of the others. Its probabilistic requirements are: $P(C|A) > P(C)$ and $P(C|B) > P(C)$, which means that A and B are potential explanations for C; and $P(A|BC) < P(A|C)$ and $P(B|AC) < P(B|C)$, which means that given C, B explains away A and vice versa. Finally, we require $P(A|BC) < \mathit{threshold}$ and $P(B|AC) < \mathit{threshold}$, where *threshold* is a probability indicative of an unlikely event; this ensures that the explaining-away effect has a useful impact on the argument.

Neutralize. Reflects a situation where some of the antecedents of an implication undermine the effect of others on the consequent, and the posterior belief in the consequent remains largely unchanged from its prior. That is, B neutralizes A if $P(C|AB) \cong P(C)$ and $P(C|A) < P(C)$.

Contradict. Similar to *Neutralize*, but one of the antecedents wins. That is, $P(C|AB) > P(C)$ and $P(C|A) < P(C)$.

Imply. Reflects a situation where the current beliefs in the antecedents increase the belief in the consequent. That is, $P(C|AB) > P(C)$.

Cause. Like *Imply*, but the relations have a "cause" tag.

Evidence. Like *Imply*, but the relations run in the opposite direction to the links in the BN.

NAG looks for these patterns in the Argument Graph, and saves a list of the pattern instances it finds. In order to generate an argument from these pattern instances, NAG must determine an ordering of the propositions to be presented, subject to the constraints imposed by the selected argumentation strategy. For instance, the premises of an implication may be presented in more than one order, since this aspect is not specified by any of the argumentation strategies. In addition, NAG tries to remove easily inferred intermediate conclusions. Both tasks are performed by activating the Attentional Mechanism (Section 3.2) during a simulated presentation of the argument. This mechanism keeps track of changes in the activation of each proposition.

In order to determine the order of the propositions to be presented, NAG first simulates presenting the propositions in the found pattern instances in different orders. The 'mention' of a proposition causes its activation in the semantic-Bayesian network in NAG's user model. Activation is then spread from this proposition, which in turn increases the activation of the propositions that are semantically linked to it. This process crudely models the human propensity to anticipate what will be mentioned next while reading some text. After the possible orderings of the propositions in each pattern instance have been considered, NAG selects for presentation the ordering that minimizes the number of propositions which are mentioned "cold" (i.e., without first being partially activated by previous propositions). The rationale for this policy is that this ordering contains the smallest number of unexpected propositions.

After an order of presentation has been established, NAG simulates this presentation in order to determine whether there are easily inferred intermediate conclusions that may be omitted. These intermediate conclusions are propositions that have a high level of activation before their planned presentation and are strongly believed in both the normative and user models due to inferences from just-mentioned propositions. The feasibility of removing these propositions is checked by both the Attentional Mechanism and the Analyzer.

Finally, the ordered list of pattern instances and propositions is passed to our *Argument Grammar* to generate natural language output. This grammar is composed of productions that render probabilistic patterns in English. Two of these productions are illustrated in Figure 10 with reference to Figure 9; the words in sans serif are invocations of production rules. Explain-away takes as input two lists of propositions ($\{A\}$ and $\{B\}$), and one singleton (C), which is the pivot on which the explanation hinges. Contradict-short takes two lists of antecedents ($\{A\}$ against the consequents and $\{B\}$ in favour), and one list of consequents, $\{C\}$. This production is used when $\{A\}$ contains only one or two propositions.

The Argument Grammar is feature-based. The productions in Figure 10 illustrate the *type* feature, which indicates whether a proposition should be realized in sentential form or nominal form. For instance, the "despite" realization of Although requires a nominal continuation (e.g., "Despite Itchy's hatred for Scratchy"), while as shown in Figure 10, the "although" realization requires a sentential continuation. Our grammar assumes that each proposition has at least one realization (nominal or sentential); it can produce a nominal realization from a sentential one and vice versa by performing simple manipulations. To expedite the interaction, currently each node in the BN has a hand-generated sentential or nominal realization. In principle, these realizations may be replaced by appropriate calls to grammars such as SURGE [Elhadad and Robin, 1996]. The *direction?* and *cause?* parameters of the productions are obtained from tags in the BNs. For example, *cause?*=+ and *direction?*='forward' indicate a causal relation, while *direction?*='backward' indicates an evidential relation. The *strength?* parameter is obtained from the normative-model BN, since NAG must state its own degrees of belief.

6 Dialogue Facilities

An ideal argumentation system interface should be simple to use, support multimedia interaction (accept and produce speech or text accompanied with diagrams as necessary), and allow the user to interrupt the system at any time. As an evolutionary step towards such ideals, NAG features a hypertext interface where arguments are presented in English, and the user may respond by clicking on a portion of the argument or on one of NAG's dialogue options.

Figure 11 shows NAG's initial argument for Itchy's innocence, which was generated for a user who entered the following beliefs after seeing the preamble presented in Section 2: it is *very probable* that Itchy had a motive to kill Scratchy, *quite likely* that he had the means, *a little bit unlikely* that he had the opportunity, and *a little bit unlikely* that he killed Scratchy. After reading the argument, the user may respond as follows: (1) by clicking on one of the options presented at the bottom of the window; or (2) by clicking on a portion of the argument to focus upon a particular proposition, which in turn leads to the display of additional response options (Figure 12, where the user has clicked on the proposition [Itchy was not outside Scratchy's window]). After each user response, NAG tries to generate a possibly revised argument which incorporates the user's request. The user can either retain this argument or ignore it. The latter choice does not necessarily result in the reinstatement of the previous argument. Rather, a different argument may be generated, since the context has now changed as a result of the interaction.

We now describe the operations available to the user in NAG's interface. These operations allow a user to examine different aspects of an argument.

Select a new goal. This operation enables a user to change the proposition being argued for or against. The newly selected goal proposition is added to the reasoning context, and the Strategist activates the focusing-generation-analysis process (Section 4).

Since NAG can argue only for the propositions it knows, we need a means to restrict the user to these propositions when selecting a goal. At present, this is done by allowing the user to choose a proposition from a pull-down menu. This solution is not appropriate in the long run, since it requires the interface to know in advance all the propositions accessible to the argumentation process (this is not a reasonable requirement for a system that consults knowledge sources containing a large number of propositions). Alternative goal selection options are currently being investigated.

Ask NAG to argue for or against a proposition in the argument. The objective of this option is similar to that of the information-sharing sub-dialogues described in [Chu-Carroll and Carberry, 1995]. The proposition selected from the argument is added to the reasoning context, and the focusing-generation-analysis process is activated to produce a sub-argument for or against this proposition (Figure 13, where the proposition [Itchy's gun was used to kill Scratchy] is argued for). NAG presumes that a new sub-argument for a selected proposition must be stronger than the current sub-argument for that proposition. If NAG cannot generate a sub-argument for the selected proposition, it will attempt to generate a sub-argument for its negation, and inform the user of this fact. If both attempts to generate a sub-argument fail, then NAG reports this failure.

Include/exclude a proposition. As when selecting a new goal, a proposition to be included in the argument is selected from a pull-down menu. The proposition is added to the reasoning context, and the focusing-generation-analysis process is activated to produce a revised argument which includes this proposition. The resultant argument may differ substantially from the last argument, particularly if the included proposition has a substantial effect on the belief in the goal. For instance, a proposition which substantially reduces the belief in the goal requires the introduction of additional sub-arguments to maintain the desired level of belief in the goal. If, despite the inclusion of the selected proposition in the context, the resultant argument does not include this proposition, then NAG tries to link the proposition to the argument by performing additional focusing-generation-analysis cycles.[3] If the connection between the proposition and the argument is still not achieved, NAG reports its failure.

A proposition to be excluded is selected from the current argument. NAG tries to construct an argument without it by removing this proposition from the BNs (and those ancestors of this proposition that have no other connections to the BNs). This involves severing the links between this proposition and the nodes connected to it, and using as premises any children of this proposition that appear in the argument. Figure 14 illustrates the effect of excluding the proposition selected in Figure 12, [Itchy was not outside Scratchy's window], from NAG's argument. This results in a weaker argument for the goal, since NAG can no longer argue effectively for Itchy's lack of opportunity to murder Scratchy. Although any probabilistic impact of excluded propositions is avoided, the exclusion has the opposite effect on attentional processing: asking someone not to think of unicorns has the opposite effect. Thus, excluded propositions are *added* to the reasoning context. This may cause NAG to incorporate into the argument propositions that are related to the excluded propositions.

Consider the effect of a proposition (*what about*). This operation is similar to the *include proposition* operation. However, here NAG returns only the reasoning path which connects the selected proposition to the goal (rather than the entire argument), and reports on the effect of this proposition on the goal. This allows the user to investigate the effect of individual factors on the goal. To perform this operation NAG adds the selected proposition to the reasoning context and activates the focusing-generation-analysis process (two additional focusing-generation-analysis cycles may be performed if necessary, as done for the *include proposition* operation). The sub-graph that connects the selected proposition to the goal is then returned for presentation. The user can choose to revert to the previous (possibly modified) argument, or to incorporate the examined factor into the argument.

Consider a hypothetical belief (*what if*). In this operation, the selected proposition is assigned the belief stipulated by the user, it is then converted into a premise and added to the reasoning context in both the normative model and the user model. Figure 15 illustrates a situation where the user explores the effect of an absolute belief in the proposition [Oblong indentations were found outside Scratchy's window] on the goal proposition (this belief contradicts the observation that the found indentations were circular). This results in NAG no longer being able to argue for Itchy's innocence, and instead arguing for his guilt. In this operation, the changes in belief in the selected proposition (and its implications) are *temporary*,

[3] At present, the number of additional cycles is limited to two. This is because propositions that cannot be reached in two more cycles than those required to build an argument are unlikely to be relevant to the argument.

since hypothetical reasoning involves the introduction of beliefs that are not necessarily correct, and hence should not be perpetuated. After producing a revised argument in light of the hypothetical belief, NAG reinstates the original beliefs in both the normative model and the user model. It then returns to the previous (possibly modified) argument.

Undo changes. The user may undo the inclusion or exclusion of propositions and also the inclusion of supporting sub-arguments and counter-arguments. Each undo operation brings a proposition into focus. In the case of inclusions and exclusions, the proposition brought into focus is the included or excluded proposition respectively, while in the case of sub-arguments or counter-arguments, the proposition brought into focus is the subgoal of these portions of the argument. After each undo, the focusing-generation-analysis process is reactivated, and NAG presents the resulting argument.

Present a rebuttal to a short-form rejoinder. The follow-on system BIAS features a limited capability whereby a user can present short-form rejoinders, such as expressions of doubt (e.g., "but Poochie's ladder has cylindrical supports") or requests for considering a fact (e.g., "consider that Itchy was planning to paint his house"). This capability differs from the exploratory operations described above in that BIAS interprets the user's rejoinders [Zukerman et al., 2000a] and then generates rebuttals to these rejoinders [Jitnah et al., 2000] (as opposed to incorporating the effect of the user's exploratory operations into the system's arguments). The interpretation of a user's rejoinder consists of determining the intended effect of the rejoinder on the system's argument. This is done by applying a simplified version of the focusing-generation-analysis process described in Section 4 on the user model semantic-Bayesian network in order to fill gaps between the user's rejoinder and the Argument Graph. This process yields candidate reasoning paths from which a path is selected by the system if there is a clear winner. Otherwise, the candidate paths are presented to the user for selection. BIAS then considers the effect of the user's rejoinder on the argument (according to the normative and user models) in order to determine a rebuttal strategy.

7 Discussion

We now return to the capabilities of an ideal argumentation system mentioned in Section 1, and consider the extent to which NAG supports these capabilities. We also discuss some of the design trade-offs we made in order to build a practicable argumentation system.

Understand implicit goals and enthymematic arguments – As indicated in Section 6, the follow-on system BIAS has the ability to understand enthymematic short-form rejoinders. These rejoinders are enthymematic if more than one inference step is required to connect them with the system's argument. NAG side-steps the problem of understanding implicit goals by receiving direct instructions from the user regarding argumentation goals or subgoals. Further, the specification of a goal is made tractable at the interface level by offering to the user a list of propositions from which a selection can be made. However, as mentioned in Section 6, this solution is appropriate only for systems with relatively small KBs.

Generate enthymematic arguments – NAG generates enthymematic arguments by omitting information that can be easily inferred from the argument presented so far.

Admit new information – NAG is currently implemented as a closed system. It admits information presented by the user only if this information is already in its normative KBs. This policy is enforced by the interface, which presents to the user for selection only propositions from the normative KBs. The user may consider non-normative belief values when performing hypothetical reasoning (*whatif*), but this is done only on a temporary basis.

Determine an argumentation strategy – NAG can select from several argumentation strategies, e.g., premise to goal and argument by cases, when presenting its arguments. The strategy chosen for a particular argument depends on how well the information in the Argument Graph fits the requirements of the candidate strategies. For example, if NAG cannot show that a particular proposition is true or false, but can show that the argument goal follows regardless of the truth value of this proposition, then an argument by cases will be selected.

Handle turn taking and interruptions – NAG presents its arguments as a single turn in a conversation. The user cannot interrupt NAG during the presentation of an argument, thereby side-stepping the problem of handling interruptions. Handling conversational turn taking is simplified since each interaction step is composed of a request formulated by the user followed by a response from NAG.

Handle dialogue acts – Our exploratory queries constitute a direct implementation of different types of dialogue acts. For instance, the introduction of a new topic is implemented by NAG's operator for selecting a new goal, information-sharing sub-dialogues are implemented by the operator for arguing for or against a proposition in the argument, and a limited form of negotiation sub-dialogues is implemented by BIAS' short-form rejoinders. However, NAG does not update the conversational context. That is, after a sub-dialogue, the conversation returns to the main argument, which is the focus of the next interaction (the only exception to this mode of operation pertains to the introduction of a new goal, which remains until the user decides to change it). At the same time, our use of attentional focus implies that the attentional state can go only forward, i.e., even if the user chooses to ignore the outcome of a request, the change in context has already occurred, hence the interaction cannot return to a previous attentional state. In addition, the consideration of the user's request may have caused changes in beliefs in the user model or the normative model (e.g., due to the propagation of the beliefs in newly added propositions). Thus, if the user chooses to ignore the outcome of a request, the argument that is 'reinstated' may be different from the original one. Although at first glance this feature may appear disconcerting, it is worth noting that the people who participated in our trials did not comment on it. Still, further trials are required to determine the need for a capability that supports the return to earlier belief states.

Select an interpretation and recover from misunderstandings – For most of the operations described in Section 6 (except BIAS' short-form rejoinders), once an operation is selected by the user, it is simply executed.

Thus, the burden of recovering from misunderstandings is placed on the user. One type of misunderstanding that may take place pertains to the capabilities or operation of the system. For instance, a user may have to try different operations to extract the desired information from an argument. Another type of misunderstanding is due to forgetfulness (or from another point of view, due to the interface not being sufficiently helpful). Such a misunderstanding occurs when the user requests that NAG exclude a particular proposition from all arguments and later asks NAG to generate an argument for a different goal which requires the excluded proposition. In the future, NAG should incorporate a conversational strategy that assists the user in clarifying his or her intentions or recovering from misunderstandings.

A misunderstanding may occur in BIAS' interpretation of a user's short-form rejoinders if none of the candidate interpretations generated by BIAS is selected by the user, or if BIAS selects and generates a rebuttal for the wrong interpretation. At present, BIAS does not recover from such misunderstandings.

Although our interface is practical, it is clearly not as powerful or flexible as an ideal interface. It is mainly an exploratory interface in that it allows a user to probe the system's arguments, while giving the user only limited capabilities to present his or her own views. The limited scope of NAG's interface avoids certain problems that must be addressed by more ambitious conversational systems, such as dealing with a changing conversational context, understanding a user's intentions and recovering from misunderstandings. As a result, the user is burdened with these problems, in particular with error recovery. At the same time, exploratory operations give the user opportunities to thoroughly investigate an argument.

8 Evaluation

NAG's argumentation capability was evaluated in two separate experiments. In the first experiment, we evaluated NAG's arguments in isolation [Zukerman et al., 1998], and in the second experiment we evaluated both its arguments and its exploratory operations [Zukerman et al., 1999].

In the first experiment, we presented 32 participants with an argument generated by NAG for the proposition [Large asteroid struck Earth 65 million years BC], after conducting a pre-test which tested their beliefs in the goal proposition and in key propositions related to the goal, such as [Iridium is abundant in asteroids] and [There was a sudden cooling of Earth 65 million years BC]. The participants showed a clear tendency to shift their belief in the goal proposition and in the key propositions in response to NAG's argument.

In the second experiment, 16 participants used the interactive version of NAG with the murder scenario shown in this paper. The subjects reported their degrees of belief in Itchy's guilt and his means, motive and opportunity to kill Scratchy, and their confidence in their judgments immediately after reading the preamble, then after NAG's initial argument for Itchy's innocence, and after interacting with NAG using our exploratory operations. The most substantial changes in belief occurred after receiving NAG's initial argument, where the mean degree of belief in Itchy's guilt dropped from 49% to 41%, the mean belief in Itchy's opportunity to murder Scratchy dropped from 60% to 45%, and the mean belief in Itchy's motive increased from 70% to 76%. Surprisingly, there was a slight drop (from 83%

to 81%) in the mean belief in Itchy's means to murder Scratchy. After interacting with NAG, the largest change in belief concerned Itchy's guilt: eight subjects changed their beliefs to be closer to NAG's, seven subjects did not change their beliefs (three retained their belief in Itchy's probable innocence and four remained uncertain), and only one subject adopted a new belief that was farther from NAG's. The mean degree of belief dropped from 41% after the initial argument to 36% after interacting with NAG. Even though these results are not statistically significant (due to the small sample size and high variance), they suggest that NAG's arguments were persuasive and that our exploratory interactions were helpful in further persuading users to accept NAG's arguments. Interestingly, the users' confidence dropped after the initial argument (possibly because NAG's argument contradicted some of the users' intuitions), but increased overall after the exploratory interactions. In most cases, the increased confidence accompanied user beliefs which agreed with or moved towards the beliefs in NAG's normative model.

We speculate that the modest effect of the exploratory interactions was partly because the interactions had no clear goal (the users were just told to "play" with NAG) and partly due to NAG's limited domain knowledge. The lack of a clear goal led to a variation in the exploration strategies adopted by our subjects. Some subjects continued to explore NAG's argument until they were convinced or until they felt they would never be convinced about the truth of each major premise in the murder story; other subjects explored only those items that appeared in the preamble and seemed the most likely to convict or exonerate Itchy, e.g., why Poochie's ladder was only available to Poochie, and the details of the ownership of the murder weapon and the fingerprints on it. The limited scope of the domain was pointed out by most subjects. In particular, several users expressed disappointment when the system was unable to consider a scenario in which they were interested, e.g., Itchy and Poochie colluded to murder Scratchy, or Poochie framed Itchy. Several users also noted the awkwardness of NAG's English output and some found it repetitive. However, nobody attributed to the system intentions to deceive or hide information. Thus, the main hurdle for an open-ended dialogue system (in terms of the domain) seems to pertain to its knowledge representation. In contrast, dialogue systems that deal with restricted domains, e.g., systems that look up schedules or hotels, do not face this problem.

From the point of view of usability, most of the subjects learned how to navigate the interface quickly without guidance. However, a few subjects needed somebody to show them how to select a portion of an argument for exploration. The visual cue that in a web-browser an underlined hypertext link can be selected by clicking on it was obscured by the fact that the entire initial argument was underlined and selectable. Once these subjects were shown how to select a portion of the argument for investigation they continued without help. In addition, some subjects found it tedious to have to return to the main argument screen before starting each different line of inquiry, as required by NAG's interface. For example, after requesting a sub-argument about Itchy firing the gun, a user must return to the main argument before he or she can request a sub-argument for Scratchy being shot from outside the window. In contrast to these comments about improving the facilities for navigating the argument, no subject expressed a desire for a new dialogue option.

We also conducted informal comparisons between the arguments produced by NAG and those found in books and other media. These comparisons yielded the following results. Content-wise, unlike people, NAG does not omit from its arguments information that does not suit its goals. Rather, it includes "the whole truth" to the best of its knowledge.[4] NAG's arguments differ most from those generated by people in that NAG's arguments stand on their content only, while people use a variety of stylistic and rhetorical techniques to increase the persuasiveness of their arguments. For instance, people may use different turns of phrase to strengthen the impact of information that is useful to them and to weaken the impact of information they find counter-productive. They may also use intimidation or appeal to emotion either directly, e.g., by belittling their opponent, or indirectly, through the selection of lexical items and grammatical constructs (the relation between such stylistic parameters and emotional factors was studied in [Hovy, 1990, Marcu, 1996]). Finally, as observed by some of our survey participants, NAG's presentation style is quite repetitive, while people often vary their prose to make it more interesting.

9 Related Research

Two graphical techniques for analyzing arguments are the well-known Toulmin warrant structure [Toulmin, 1958] and Cohen's tree structures [Cohen, 1987]. The Toulmin warrant structure contains the following elements: (1) a *claim* – the argument goal, (2) *data* – the evidence for the claim, (3) *warrant* and *backing* – the reasoning used to link the data to the claim, (4) a *qualifier* – an adverb or phrase modifying the claim to indicate its strength, and (5) *reservations* – circumstances or conditions that undermine the argument. The premises NAG uses to ground its arguments are equivalent to Toulmin's data, while NAG's intermediate inferences and conclusions play an equivalent role to the warrant and backing. When presenting its arguments, NAG adds a qualifier to each step in the presentation. These qualifiers are based on the probability of the corresponding nodes in the normative-model BN. Since NAG must present normatively correct arguments, it must disclose any reservations it has about its own arguments. However, reservations are usually not mentioned together in a group as in the Toulmin structure. Instead, each reservation is normally mentioned close to the portion of the argument that it affects. The same effect can be generated from the Toulmin structure pattern if the argument is sufficiently broken down into its constituent sub-arguments.

Cohen's method of argument analysis uses linguistic clues and the ordering of the statements in an argument to build a tree structure that represents the argument [Cohen, 1987]. Each statement in the argument is represented by a node in the tree. The tree is formed so that each node or statement offers support for its parent in the tree. NAG's Argument Analyzer also performs argument analysis (on its own arguments as part of the focusing-generation-analysis cycle). However, NAG can also understand arguments that have a graph structure, not just a tree.

Our work extends traditional interaction and explanation capabilities for expert systems, e.g., [Buchanan and Shortliffe, 1985], in that it uses BNs as its main representation formalism and supports the exploration of arguments in ways that complement the justifications generated by earlier systems. The system described in [Moore and Swartout, 1990] features a

[4]The only grounds for the omission of information pertain to the discrepancy between the beliefs in the user model and those in the normative model. This may result in the omission of information which disadvantages NAG's arguments. However, this omission is not made in order to win an argument.

capability whereby users employ hypertext links to seek additional explanations for propositions presented in an initial explanation. This operation is similar to asking NAG to argue for or against a proposition.

NAG is similar in scope to the interactive argumentation system IACAS [Vreeswijk, 1994]. Like NAG, IACAS allows a user to add or remove information from its arguments. However, IACAS does not model attentional focus or tailor its arguments to the user. Instead, it keeps presenting supporting arguments for a goal proposition until the user is satisfied or chooses a new goal. Argument construction systems, such as Belvedere [Suthers et al., 1995] and Euclid [Smolensky et al., 1988] use a graphical argumentation language to assist users in the creation of an argument. However, these systems do not generate their own arguments.

Several researchers have considered different aspects of argument generation, e.g., [Flowers et al., 1982, Quilici, 1992, Restificar et al., 1999, Carenini and Moore, 2000, Grasso et al., 2000]. Flowers *et al.* presented a partial theory of argumentation which advocated the combination of distinct knowledge sources, in a similar way to NAG's consultation of several different KBs. However, Flower's implementation focused on recognizing and providing episodic justifications to historical events, while NAG focuses on generating probabilistic arguments. Quilici studied the generation and recognition of the justification for a proposal in a plan-based context. Both tasks were performed by applying a set of justification rules in backwards chaining mode from the proposal to known premises. Restificar *et al.* applied argument schemata to recognize a user's intentions from his or her rejoinders and to generate short rebuttals to these rejoinders. Both of these systems dealt with short exchanges, where each dialogue contribution consisted of a few propositions. This allowed them to use simple reasoning techniques. In contrast, NAG generates extended probabilistic arguments. The system described in [Carenini and Moore, 2000] generates evaluative arguments which mention information that is likely to be of interest to a user, and omit information that is not likely to be of interest. This system differs from NAG in two respects: (1) it does not model explicitly the persuasiveness of its arguments, and (2) NAG does not model the user's preferences or interests.[5] Grasso *et al.* described a system that provides nutrition advice using dialectic or informal arguments. Such arguments need not start from strictly true premises and may use inferences which are not logically valid. A large number of human arguments fall into this category, and indeed such arguments are often essential when the primary goal is to be persuasive. In principle, features which support the generation of informal arguments could be incorporated into NAG. However, thus far we have considered non-normative reasoning to a limited extent only, through our models of some human cognitive deficiencies [Korb et al., 1997].

Dialogue systems and specific dialogue phenomena have been considered by many researchers. Here we discuss the capabilities of some representative systems as an indication of the capabilities that should be incorporated into NAG in the future. The dialogue system described in [Jönsson, 1995], which operates in an information providing context, contains a dialogue manager which directs the interaction between a user and the system by taking advantage of observations of the user's behaviour during information-seeking interactions. In addition, Jönsson's system uses contextual infor-

[5]The incorporation of a model of preferences and interests in a system which generates arguments for the truth of a proposition (rather than evaluative arguments) requires the consideration of the extra-rational impact of these factors.

mation to further specify the user's requirements. In contrast, NAG considers a user's request and the immediately preceding argument to determine the focus of attention, which in turn affects the argument generation process. As indicated in Section 1, several researchers have considered specific dialogue phenomena in isolation. Clarification and information-sharing sub-dialogues have been considered in [Litman and Allen, 1987] and [Chu-Carroll and Carberry, 1995] respectively. Procedures for selecting an interpretation among candidate options are described for example in [Litman and Allen, 1987, Raskutti and Zukerman, 1991]. An abductive mechanism for identifying and recovering from misunderstandings is discussed in [McRoy and Hirst, 1995]. Specific aspects of understanding enthymematic discourse, such as understanding indirect speech acts and expressions of doubt, have been investigated in [Green and Carberry, 1992] and [Carberry and Lambert, 1999] respectively. The intention recognition capabilities of the BIAS follow-on system [Zukerman et al., 2000a] are closest to those of the system described in [Carberry and Lambert, 1999]. However, BIAS also generates rebuttals.

The generation-analysis part of NAG's focusing-generation-analysis cycle resembles Chu-Carroll and Carberry's propose-evaluate-modify cycle [Chu-Carroll and Carberry, 1995], which determines whether to engage in an information-sharing sub-dialogue. However, Chu-Carroll and Carberry focused on the determination of information-sharing strategies and the selection of propositions for discussion, while NAG uses its focusing-generation-analysis cycle to generate extended probabilistic arguments. NAG's generation-analysis process also resembles the abductive mechanism used in [Hobbs et al., 1993] to infer implicit information. However, there are two important differences between NAG and the work by Hobbs *et al.*: NAG is a system that reasons under uncertainty and NAG performs generation as well as analysis. A generative system based on the work of Hobbs *et al.* is described in [Thomason et al., 1996]. That system deals with what can be readily inferred, and so deleted, during communication, but the generated discourse does not present an argument in support of a proposition, as done by NAG.

Other mechanisms that omit easily inferred information from discourse are considered in [Zukerman and McConachy, 1993, Mehl, 1994, Horacek, 1997, Fehrer and Horacek, 1997]. Mehl's system turned explicit arguments into ones where easily inferred information was left implicit. However, this system required a complete argument as input, while NAG constructs its own arguments. The systems described in [Zukerman and McConachy, 1993, Horacek, 1997, Fehrer and Horacek, 1997] omit easily inferred information by applying hand-crafted inference rules which model a user's inferential patterns in different domains. In contrast, NAG uses both Bayesian propagation and semantic activation to choose which propositions it can omit from its arguments.

NAG and the systems described in [Huang and Fiedler, 1997, Reed and Long, 1997, Reed, 1999] consider focus of attention during argument presentation. Huang and Fiedler used a limited implementation of attentional focus to select the step of a mathematical proof to be mentioned next. Reed and Long considered attention in order to generate additional information that makes a concept salient, while the system described in [Reed, 1999] uses salience to decide which items to omit from its argument plans. In contrast, NAG combines both activation and strength of belief when deciding whether to omit a proposition. Further, NAG composes its arguments from BNs, while Reed's system uses logic derived planning operators to construct its arguments.

Finally, NAG's use of spreading activation to model attention and BNs for reasoning resembles the use of these mechanisms in the story understanding system described in [Charniak and Goldman, 1993]. Charniak and Goldman's system automatically built and incrementally extended a BN from propositions read in a story, so that the BN represented hypotheses that became plausible as the story unfolded. During this process, their system used marker passing (a form of spreading activation) to restrict the nodes included in the BN, so that only nodes that were relevant (as well as likely) in the current context were incorporated. Similarly, NAG uses spreading activation to focus on relevant propositions when expanding the BNs in its normative and user models. Bayesian propagation is then used to calculate beliefs in the propositions in the Argument Graph in each model.

10 Conclusion

An ideal argumentation system would support a full argumentative interaction with users, presenting its own arguments and allowing users to present different types of rejoinders, such as questions and counter-arguments, as well as their own original arguments, which in turn would lead to the system's rebuttals, and so on. In this paper, we have considered the dialogue and interface requirements of such a system, and discussed the performance of our system NAG with respect to these requirements.

We have discussed NAG's use of its normative and user models to yield arguments that are both correct and persuasive for a target audience. We have also shown how NAG uses the argument context to guide its choices during content planning, to focus on propositions that are useful for argument construction, and during argument presentation, to order propositions and omit easily inferable propositions. We have also described NAG's interactive capabilities, which are implemented in the form of exploratory operations a user can perform on NAG's arguments. Our preliminary evaluations suggest that both NAG's initial arguments and the exploratory operations can shift users' beliefs in the argument goal and in other relevant propositions.

Our current exploratory operations are restricted in scope and place the main burden of the interaction on the user (due to the limited interpretive ability of the system). However, these operations, which depart from a normal dialogue paradigm, enable the system to side-step certain problems that must be addressed by conversational systems and have not been satisfactorily resolved to date. At the same time, these operations afford the user increased opportunities to investigate the system's arguments. This indicates that in the absence of fully-interactive interfaces to complex reasoning systems, exploratory operations such as those presented here constitute a viable option; and even when such dialogue systems become available, exploratory operations may be retained as an alternative mode of interaction with these systems.

Acknowledgments

This work was supported in part by Australian Research Council grant A49531227. The authors are indebted to Kevin Korb for his work on NAG, and to Deborah Pickett for her implementation of NAG's interface. The authors also thank Nathalie Jitnah and Sarah George for their work on BIAS.

References

[Anderson, 1983] Anderson, J. R. (1983). *The Architecture of Cognition.* Harvard University Press, Cambridge, Massachusetts.

[Buchanan and Shortliffe, 1985] Buchanan, B. G. and Shortliffe, E. H. (1985). *Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project.* Addison-Wesley Publishing Company.

[Carberry and Lambert, 1999] Carberry, S. and Lambert, L. (1999). A process model for recognizing communicative acts and modeling negotiation subdialogues. *Computational Linguistics*, 25(1):1–53.

[Carenini and Moore, 2000] Carenini, G. and Moore, J. D. (2000). A strategy for generating evaluative arguments. In *INLG'2000 – Proceedings of the First International Conference on Natural Language Generation*, pages 47–54, Mitzpe Ramon, Israel.

[Charniak and Goldman, 1993] Charniak, E. and Goldman, R. P. (1993). A Bayesian model of plan recognition. *Artificial Intelligence*, 64(1):50–56.

[Chu-Carroll and Carberry, 1995] Chu-Carroll, J. and Carberry, S. (1995). Generating information-sharing subdialogues in expert-user consultation. In *IJCAI95 – Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 1243–1250.

[Cohen, 1987] Cohen, R. (1987). Analyzing the structure of argumentative discourse. *Computational Linguistics*, 13(1):11–24.

[Conati et al., 1997] Conati, C., Gertner, A. S., VanLehn, K., and Druzdzel, M. (1997). On-line student modeling for coached problem solving using Bayesian Networks. In *UM97 – Proceedings of the Sixth International Conference on User Modeling*, pages 231–242, Sardinia, Italy.

[Elhadad and Robin, 1996] Elhadad, M. and Robin, J. (1996). An overview of SURGE: A reusable comprehensive syntactic realization component. Technical Report 96-03, Department of Mathematics and Computer Science, Ben Gurion University, Beer Sheva, Israel.

[Elsaesser, 1987] Elsaesser, C. (1987). Explanation of probabilistic inference for decision support systems. In *Proceedings of the AAAI-87 Workshop on Uncertainty in Artificial Intelligence*, pages 394–403, Seattle, Washington.

[Fehrer and Horacek, 1997] Fehrer, D. and Horacek, H. (1997). Exploiting the addressee's inferential capabilities in presenting mathematical proofs. In *IJCAI97 – Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, pages 959–964, Nagoya, Japan.

[Flowers et al., 1982] Flowers, M., McGuire, R., and Birnbaum, L. (1982). Adversary arguments and the logic of personal attack. In *Strategies for Natural Language Processing*, pages 275–294. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

[Grasso et al., 2000] Grasso, F., Cawsey, A., and Jones, R. (2000). Dialetical argumentation to solve conflicts in advice giving: A case study in the promotion of healthy nutrition. *International Journal of Human-Computer Studies*, 53(6):1077–1115.

[Green and Carberry, 1992] Green, N. and Carberry, S. (1992). Conversational implicatures in indirect replies. In *Proceedings of the Thirtieth Annual Meeting of the Association for Computational Linguistics*, pages 64–71, Newark, Delaware.

[Hobbs et al., 1993] Hobbs, J. R., Stickel, M. E., Appelt, D. E., and Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, 63(1-2):69–142.

[Horacek, 1997] Horacek, H. (1997). A model for adapting explanations to the user's likely inferences. *User Modeling and User-Adapted Interaction*, 7(1):1–55.

[Hovy, 1990] Hovy, E. H. (1990). Pragmatics and natural language generation. *Artificial Intelligence*, 43(2):153–197.

[Huang and Fiedler, 1997] Huang, X. and Fiedler, A. (1997). Proof verbalization as an application of NLG. In *IJCAI97 – Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, pages 965–970, Nagoya, Japan.

[Huber et al., 1994] Huber, M. J., Durfee, E. H., and Wellman, M. P. (1994). The automated mapping of plans for plan recognition. In *UAI94 – Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, pages 344–350, Seattle, Washington.

[Jitnah et al., 2000] Jitnah, N., Zukerman, I., McConachy, R., and George, S. (2000). Towards the generation of rebuttals in a Bayesian argumentation system. In *Proceedings of the First International Natural Language Generation Conference*, pages 39–46, Mitzpe Ramon, Israel.

[Jönsson, 1995] Jönsson, A. (1995). Dialogue actions for natural language interfaces. In *IJCAI95 – Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 1405–1411, Montreal, Canada.

[Korb et al., 1997] Korb, K. B., McConachy, R., and Zukerman, I. (1997). A cognitive model of argumentation. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, pages 400–405, Stanford, California.

[Litman and Allen, 1987] Litman, D. and Allen, J. F. (1987). A plan recognition model for subdialogues in conversation. *Cognitive Science*, 11:163–200.

[Marcu, 1996] Marcu, D. (1996). The conceptual and linguistic facets of persuasive arguments. In *Proceedings of ECAI-96 Workshop – Gaps and Bridges: New Directions in Planning and NLG*, pages 43–46, Budapest, Hungary.

[McRoy and Hirst, 1995] McRoy, S. W. and Hirst, G. (1995). The repair of speech act misunderstandings by abductive inference. *Computational Linguistics*, 21(4):435–478.

[Mehl, 1994] Mehl, S. (1994). Forward inferences in text generation. In *ECAI94 – Proceedings of the Eleventh European Conference on Artificial Intelligence*, pages 525–529, Amsterdam, The Netherlands.

[Moore and Swartout, 1990] Moore, J. D. and Swartout, W. R. (1990). Pointing: A way toward explanation dialogue. In *AAAI90 – Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 457–464, Boston, Massachusetts.

[Pearl, 1988] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann Publishers, San Mateo, California.

[Quilici, 1992] Quilici, A. (1992). Arguing about planning alternatives. In *COLING-92 – Proceedings of the Fourteenth International Conference on Computational Linguistics*, pages 906–910, Nantes, France.

[Raskutti and Zukerman, 1991] Raskutti, B. and Zukerman, I. (1991). Generation and selection of likely interpretations during plan recognition. *User Modeling and User Adapted Interaction*, 1(4):323–353.

[Reed, 1999] Reed, C. (1999). The role of saliency in generating natural language arguments. In *IJCAI99 – Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 876–881, Stockholm, Sweden.

[Reed and Long, 1997] Reed, C. and Long, D. (1997). Content ordering in the generation of persuasive discourse. In *IJCAI97 – Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, pages 1022–1027, Nagoya, Japan.

[Restificar et al., 1999] Restificar, A., Syed, A., and McRoy, S. (1999). Arguer: Using argument schemas for argument detection and rebuttal in dialogs. In *UM99 – Proceedings of the Seventh International Conference on User Modeling*, pages 315–317, Banff, Canada.

[Smolensky et al., 1988] Smolensky, P., Fox, B., King, R., and Lewis, C. (1988). Computer-aided reasoned discourse, or, how to argue with a computer. In Guindon, R., editor, *Cognitive Science and its Applications for Human-computer Interaction*, pages 109–162. Lawrence Erlbaum Associates.

[Suthers et al., 1995] Suthers, D., Weiner, A., Connelly, J., and Paolucci, M. (1995). Belvedere: Engaging students in critical discussion of science and public policy issues. In *AIED-95 – Proceedings of the Seventh World Conference on Artificial Intelligence in Education*, Washington, D.C.

[Thomason et al., 1996] Thomason, R. H., Hobbs, J. R., and Moore, J. D. (1996). Communicative goals. In *Proceedings of ECAI-96 Workshop – Gaps and Bridges: New Directions in Planning and NLG*, pages 7–12, Budapest, Hungary.

[Toulmin, 1958] Toulmin, S. (1958). *Uses of Argument.* Cambridge University Press, Cambridge.

[Vreeswijk, 1994] Vreeswijk, G. (1994). IACAS: An interactive argumentation system. Technical Report CS 94-03, Department of Computer Science, University of Limburg.

[Zukerman et al., 2000a] Zukerman, I., Jitnah, N., McConachy, R., and George, S. (2000a). Recognizing intentions from rejoinders in a Bayesian interactive argumentation system. In *PRICAI2000 – Proceedings of the Sixth Pacific Rim International Conference on Artificial Intelligence*, pages 252–263, Melbourne, Australia.

[Zukerman and McConachy, 1993] Zukerman, I. and McConachy, R. (1993). Generating concise discourse that addresses a user's inferences. In *IJCAI93 – Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 1202–1207, Chambery, France.

[Zukerman et al., 1998] Zukerman, I., McConachy, R., and Korb, K. B. (1998). Bayesian reasoning in an abductive mechanism for argument generation and analysis. In *AAAI98 – Proceedings of the Fifteenth National Conference on Artificial Intelligence*, pages 833–838, Madison, Wisconsin.

[Zukerman et al., 2000b] Zukerman, I., McConachy, R., and Korb, K. B. (2000b). Using argumentation strategies in automated argument generation. In *Proceedings of the First International Natural*

Language Generation Conference, pages 55–68, Mitzpe Ramon, Israel.

[Zukerman et al., 1999] Zukerman, I., McConachy, R., Korb, K. B., and Pickett, D. A. (1999). Exploratory interaction with a Bayesian argumentation system. In *IJCAI99 – Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 1294–1299, Stockholm, Sweden.

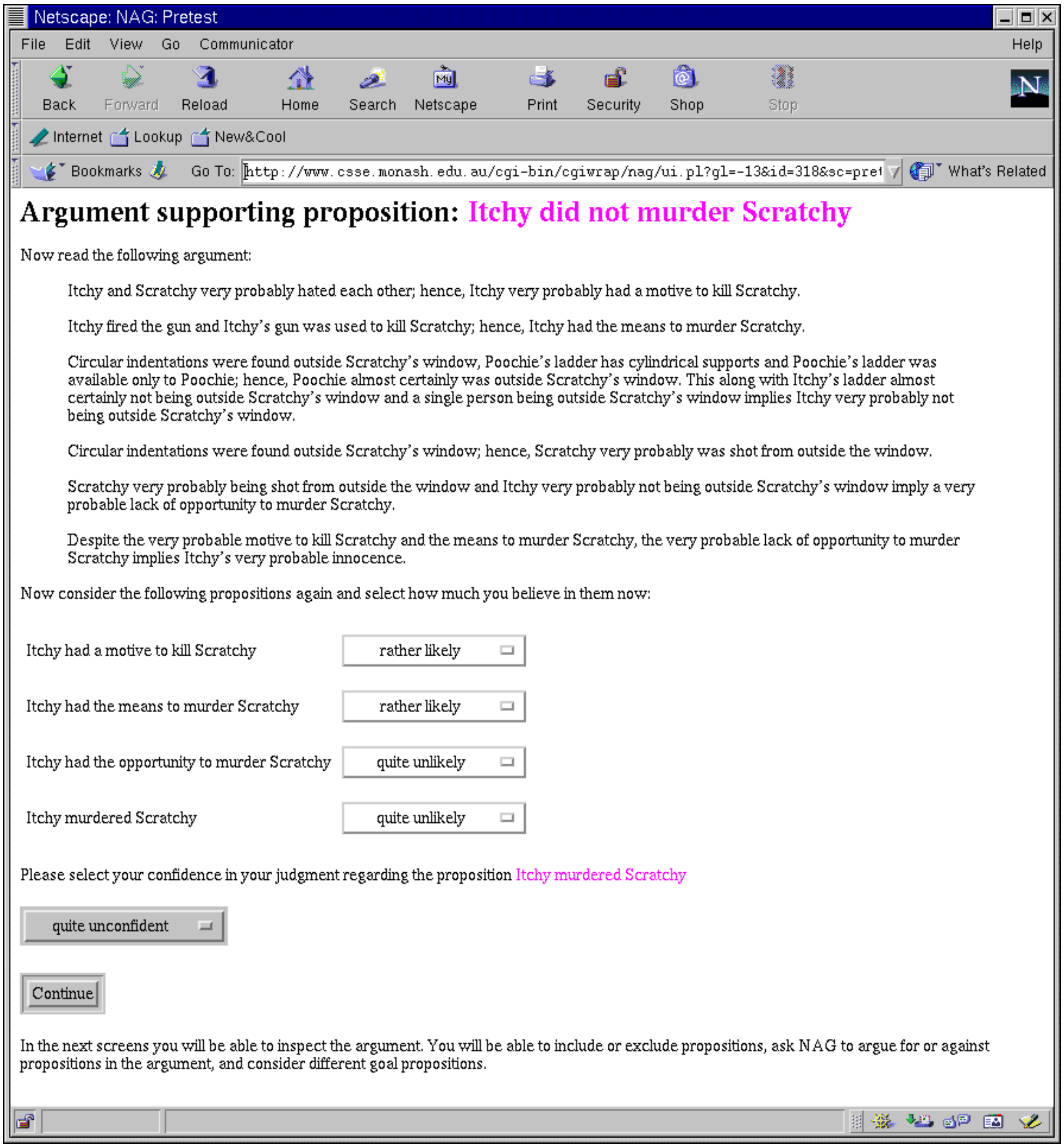

Figure 2: NAG's test interface and a sample argument

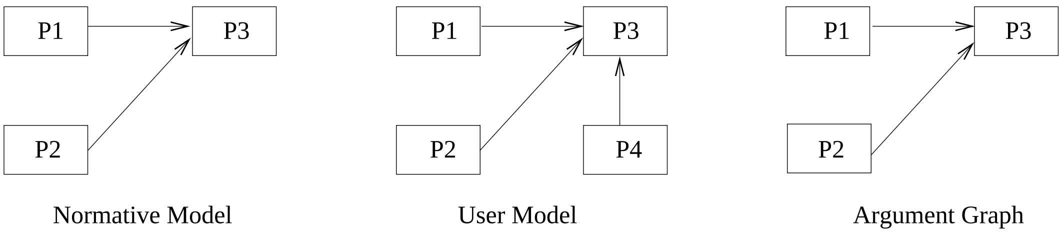

Figure 3: Combining propositions and links to form an Argument Graph

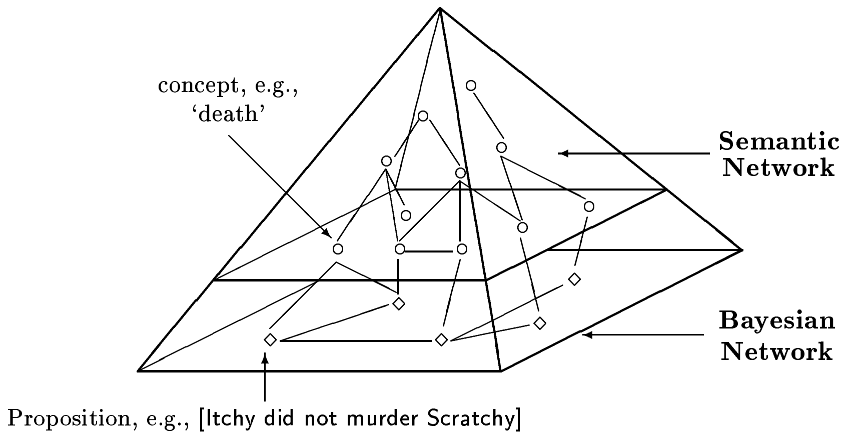

Figure 4: Semantic-Bayesian network

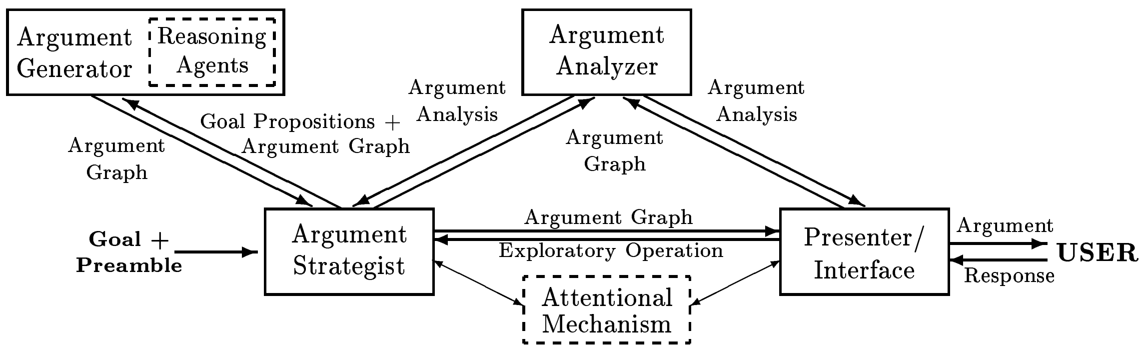

Figure 5: System architecture

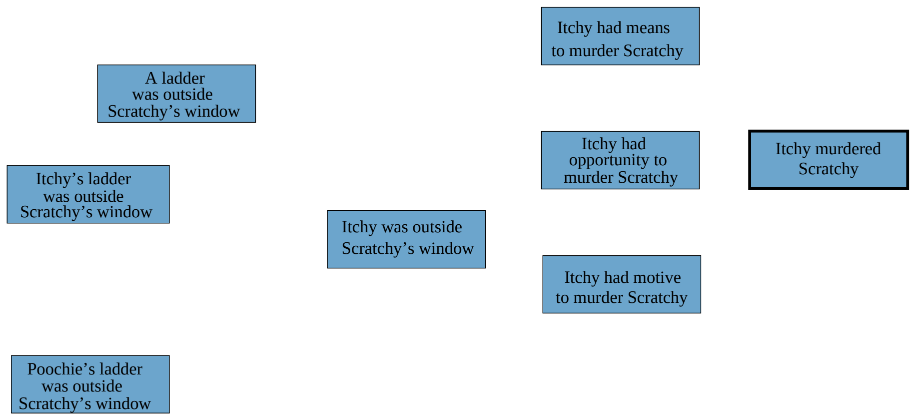

Figure 6: Propositions in the initial context for the murder example

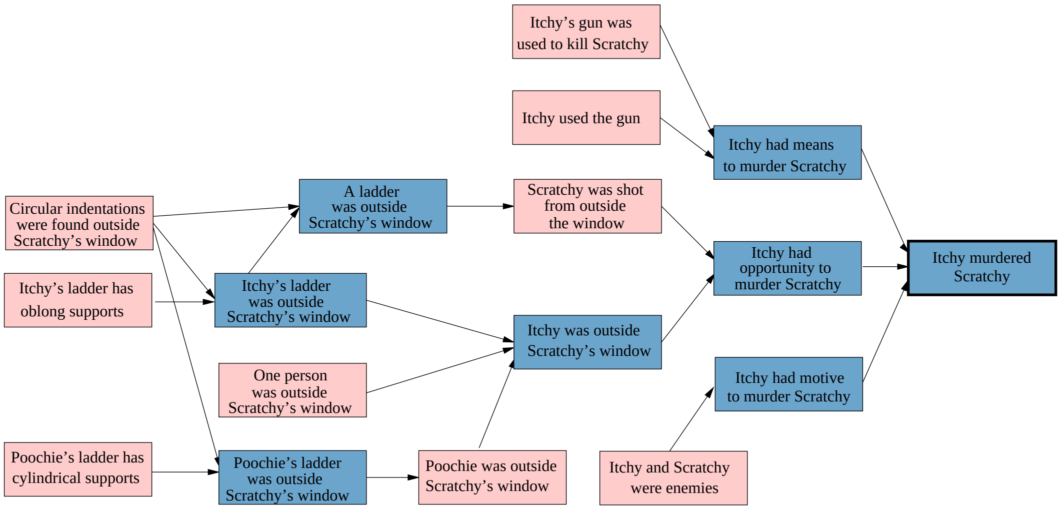

Figure 7: Argument Graph for the murder example after one round of focusing and generation

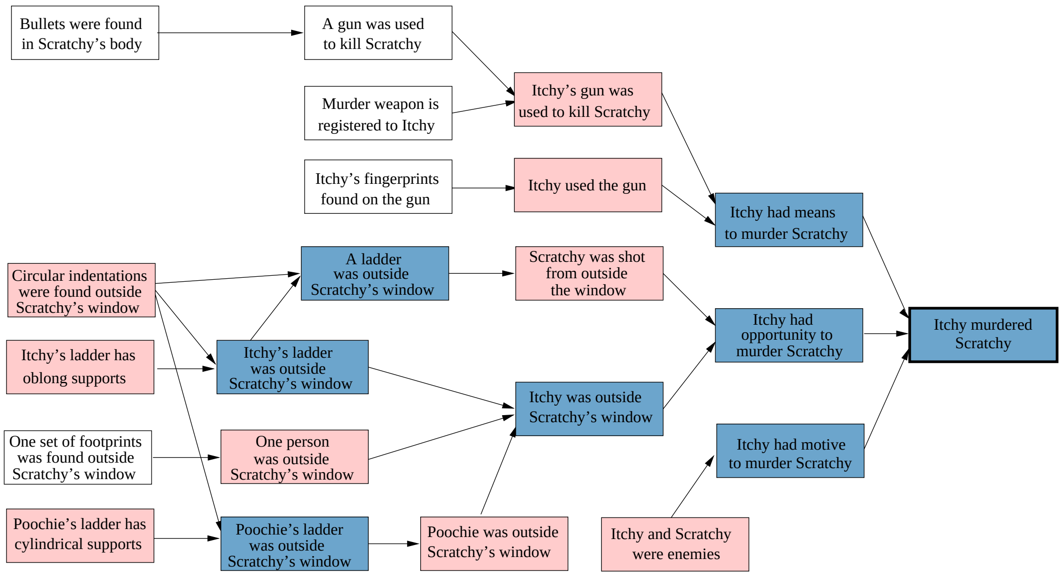

Figure 8: Final Argument Graph for the murder example

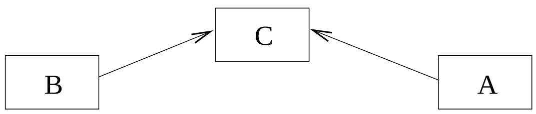

Figure 9: A simple Bayesian network

Explain-away($\{A\}, \{B\}, C$,*strength?*):
Pattern:
Given(*type$_a$?*) Antecedent(C,*type$_a$?*),
Antecedent($\{A\}$,nominal) Imply(*strength?,direction?,cause?,type$_c$?*)
Consequent($\{B\}$,*type$_c$?*).
Example:
Explain-away([Sprinklers were on last night], [It rained], [Grass is wet], probable)
`Given that the grass is wet, the fact that the sprinklers were on last night implies that it probably did not rain.`

Contradict-short($\{A\}, \{B\}, \{C\}$,*strength?,direction?,cause?*):
Pattern:
Although(*type$_a$?*) Antecedent($\{A\}$,*type$_a$?*),
Antecedent($\{B\}$,nominal) Imply(*strength?,direction?,cause?,type$_c$?*)
Consequent($\{C\}$,*type$_c$?*).
Example:
Contradict-short([Itchy's fingerprints found on gun], [Itchy has alibi], [Itchy is innocent], true, forward, $-$)
`Although Itchy's fingerprints were found on the gun, Itchy's alibi implies his innocence.`

Figure 10: Sample productions in the Argument Grammar: parameters, pattern and example

Figure 11: NAG's exploratory interface and a sample argument

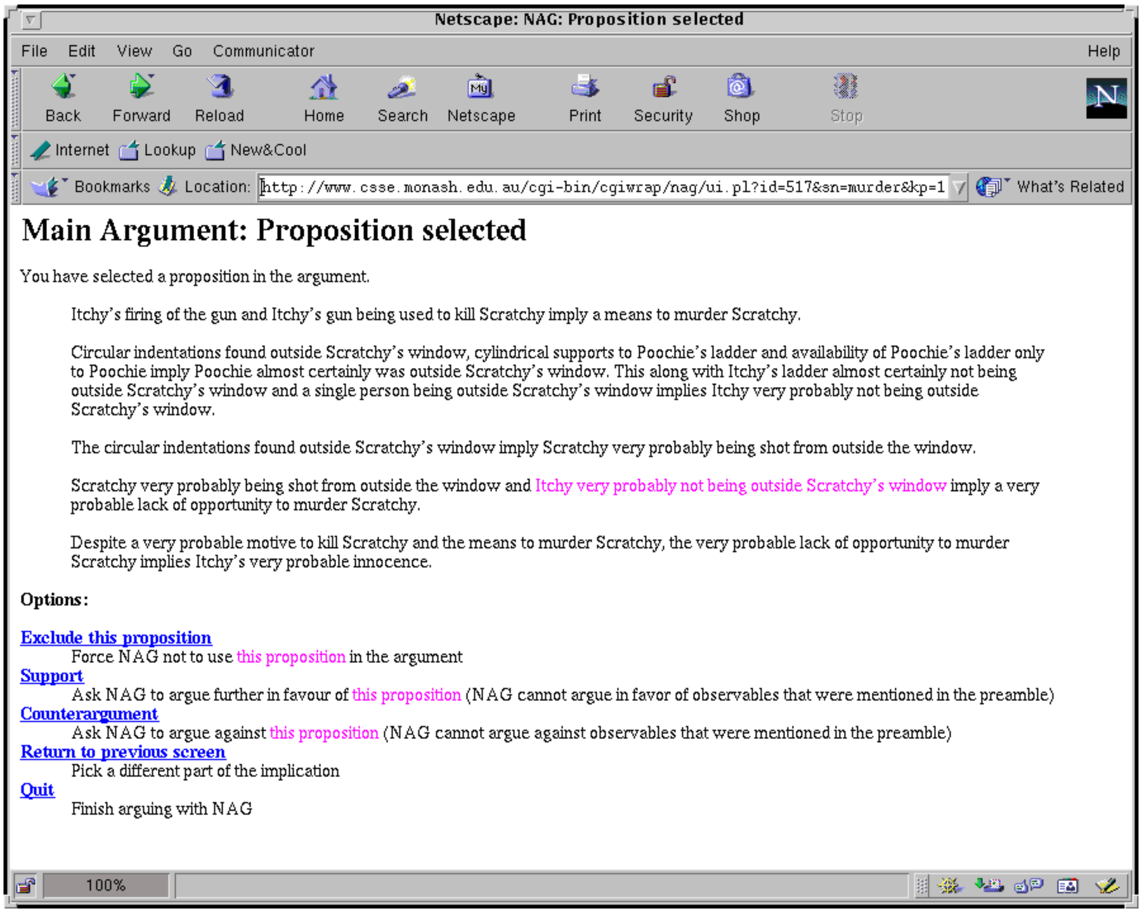

Figure 12: Selecting a proposition for inspection

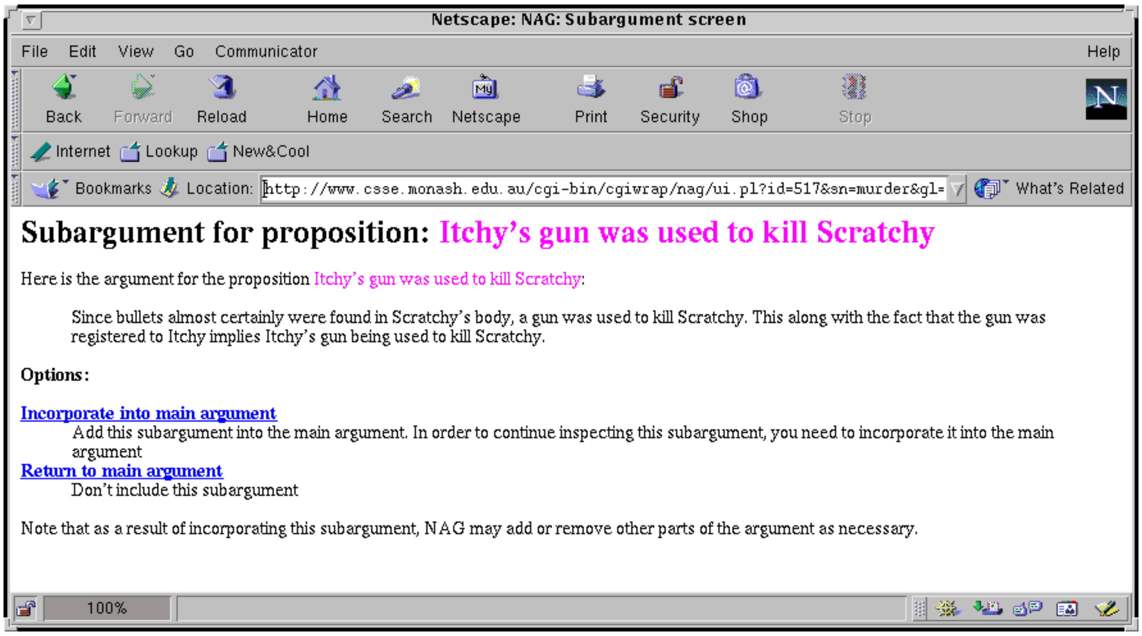

Figure 13: Generating a sub-argument for a proposition

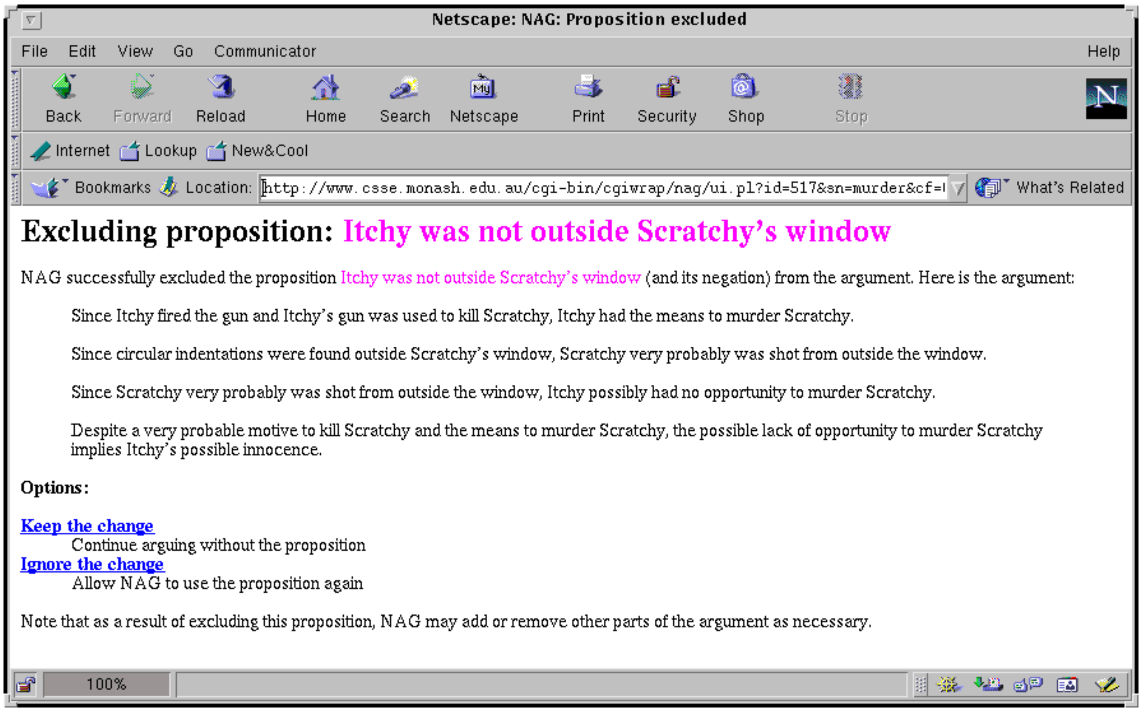

Figure 14: Excluding a proposition from the argument

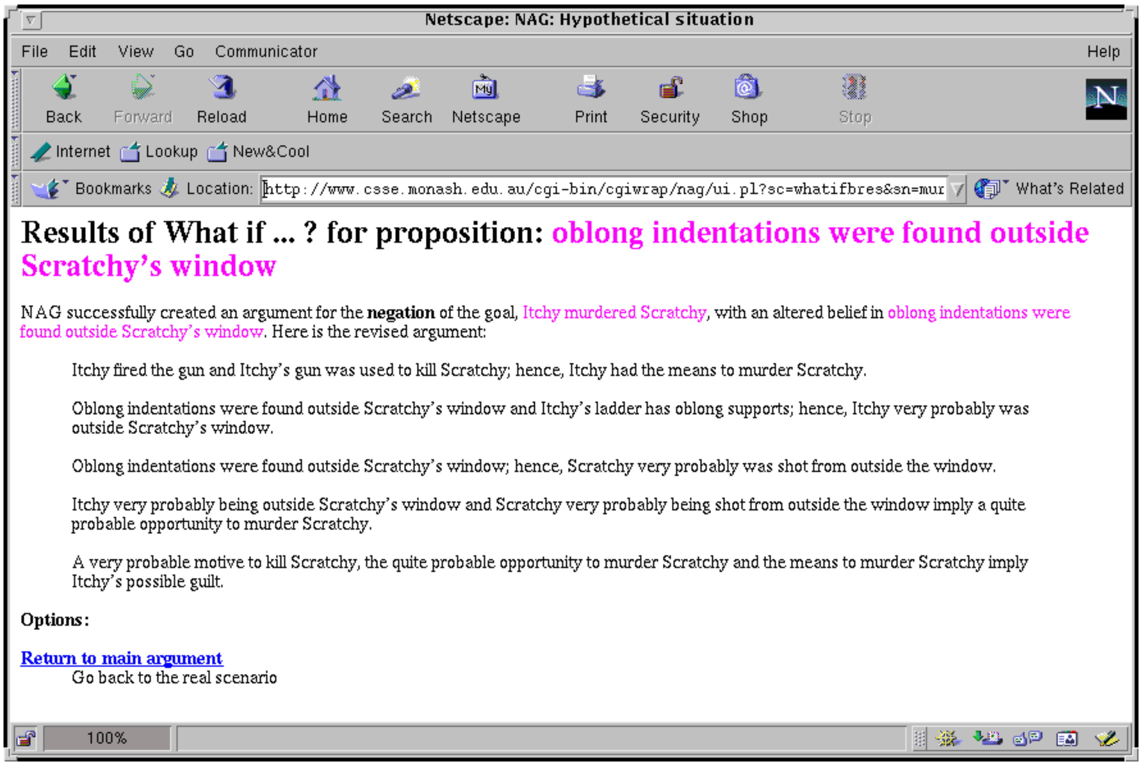

Figure 15: Reasoning hypothetically from a proposition