

# Towards the Generation of Rebuttals in a Bayesian Argumentation System

Nathalie Jitnah, Ingrid Zukerman, Richard McConachy and Sarah George

School of Computer Science and Software Engineering

Monash University

Clayton, Victoria 3800, AUSTRALIA

email: {njitnah,ingrid,ricky,sarahg}@csse.monash.edu.au

## Abstract

We describe a mechanism which generates rebuttals to a user’s rejoinders in the context of arguments generated from Bayesian networks. This mechanism is implemented in an interactive argumentation system. Given an argument generated by the system and an interpretation of a user’s rejoinder, the generation of the rebuttal takes into account the intended effect of the user’s rejoinder, determined on a model of the user’s beliefs, and its actual effect, determined on a model of the system’s beliefs. We consider three main rebuttal strategies: refute the user’s rejoinder, strengthen the argument goal, and dismiss the user’s line of reasoning.

## 1 Introduction

During argumentation, conversational partners often use expressions of doubt, such as “But the victim was stabbed”, and requests for the consideration of additional facts they consider relevant, such as “What about the fingerprints found on the gun?”. In this paper, we describe a mechanism which generates rebuttals to such rejoinders in the context of arguments generated from Bayesian networks (BNs) (Pearl, 1988). This mechanism is implemented in a system called BIAS (*Bayesian Interactive Argumentation System*). Given an argument produced by BIAS and a follow-up rejoinder posed by a user, our mechanism generates a rebuttal on the basis of a line of reasoning identified by BIAS from the user’s rejoinder. These capabilities constitute a significant step towards allowing a user to interact freely with an argumentation system and to improve the explanation capability of Bayesian systems.

Normal arguments are unconstrained in the sense that they can use whatever means are available to justify a goal proposition. In contrast, rebuttals are constrained, since they must address the point through which the conversational partner attempted to undermine or question a previous argument. To illustrate the operation of BIAS and its rebuttal capability, consider the exchange in Figure 1, which consists of a preamble that contains background information, followed by an argument generated by

BIAS, a user’s rejoinder and BIAS’ rebuttal.<sup>1</sup> The domain of implementation is a murder investigation where the question under consideration (the goal proposition) is “Did Mr Green murder Mr Body?”, and both the user and the system have access to evidence. After the presentation of the argument where BIAS contends Mr Green’s possible innocence,<sup>2</sup> the user presents a rejoinder which requests that BIAS consider a fact that was omitted from the argument: *The found gun is available only to Mr Green*. BIAS infers from this rejoinder that the user is adding support to Mr Green having the means to kill Mr Body, and hence to Mr Green’s guilt, through the following line of reasoning, which is determined as described in (Zukerman et al., 2000): *The gun being available only to Mr Green*  $\rightarrow$  *The gun was fired by Mr Green*  $\rightarrow$  *Mr Green had the means to kill Mr Body*  $\rightarrow$  *Mr Green killed Mr Body*. BIAS finds that it does not share the user’s belief in the rejoinder proposition, and that in addition, the effect of this proposition on the goal is rather weak. This prompts the generation of a rebuttal of the form *Deny–Dismiss–Follow*, whereby the rejoinder proposition is denied, its effect on the goal proposition is dismissed, and its implications are followed hypothetically until they break down due to the marginal effect of the rejoinder on Mr Green’s guilt.

In the next section, we present our knowledge representation formalism, followed by an outline of our procedure for determining a user’s line of reasoning. In Section 4, we describe our algorithm for rebuttal generation and discuss our results. We then review related work and present concluding remarks.

## 2 Knowledge Representation

During the argumentation process, BIAS maintains two models of belief: a normative model and a user model, each of which is represented as a BN. The normative model contains information gathered directly by BIAS from the murder scenario, while the user model stores propositions that are presumed to

<sup>1</sup>The argument and rebuttals shown in this paper are realized in English as described in (Zukerman et al., 1999).

<sup>2</sup>The mechanism which generates this argument is described in (Zukerman et al., 1998).

<p><b>Preamble:</b> Mr. Body was found dead in his bedroom, which is in the second story of his house. Bullet wounds were found in Mr. Body’s body. The bedroom window was broken and broken glass was found inside the window. A gun was found on the premises, and some fingerprints were found on the gun. In addition, inspection of the grounds revealed footprints in the garden and circular indentations in the ground outside the bedroom window.</p>
<p><b>BIAS’ argument:</b> Bullets being found in Mr Body’s body implies Mr Body was almost certainly shot. This implies he was almost certainly murdered. Forensics matching the bullets with the found gun implies the gun is almost certainly the murder weapon. Forensics matching the fingerprints with Mr Green implies Mr Green probably fired the gun. This together with the gun almost certainly being the murder weapon implies Mr Green probably fired the murder weapon, which implies he very probably had the means to murder Mr Body. The Bayesian Times reporting Mr Body took Mr Green’s girlfriend implies Mr Green and Mr Body possibly were enemies, which implies Mr Green possibly had a motive to murder Mr Body. The neighbour reporting Mr Green not being in the garden at 11 implies Mr Green very probably wasn’t in the garden at 11. Forensics reporting the time of death being 11 and the forensic analysis of the time of death being reliable implies the time of death was probably 11, which together with Mr Green very probably not being in the garden at 11 implies he probably wasn’t in the garden at the time of death. This implies he probably didn’t have the opportunity to murder Mr Body. Even though Mr Green very probably had the means to murder Mr Body and he possibly had a motive to murder Mr Body, Mr Green probably not having the opportunity to murder Mr Body implies he probably didn’t murder Mr Body.</p>
<p><b>User’s rejoinder:</b> Consider that the found gun is available only to Mr Green.</p>
<p><b>BIAS’ rebuttal:</b> Actually, it is very improbable that the found gun is available only to Mr Green. However, even if it was available only to Mr Green, this would have only a small effect on the likelihood that Mr Green murdered Mr Body. This is for the following reason. The found gun being available only to Mr Green implies it is more likely that Mr Green fired the gun, making it almost certain. This implies it is more likely that he fired the murder weapon, making it almost certain, which implies it is even more likely that he had the means to murder Mr Body. This implies it is only slightly more likely that he murdered Mr Body.</p>

Figure 1: Sample Argument, Rejoinder and Rebuttal

be believed by the user. These propositions may be obtained from a variety of sources, e.g., they may have been inspected by the user in the murder scenario (by means of a WWW interface), or appear in BIAS’ previous arguments or the user’s rejoinders. Arguments generated by BIAS are represented by means of an *Argument Graph* – a sub-network of the normative model BN which ideally also contains nodes from the user model BN.

The interpretation process, where BIAS identifies the reasoning path intended by the user, takes place in the user model, since BIAS tries to “make sense” of what the user is saying according to the system’s view of the user’s beliefs (Zukerman et al., 2000). In contrast, the processes for generating the initial argument and the rebuttals consult both the user model and the normative model to produce arguments that rely on beliefs held by both BIAS and the user if possible. Further, during rebuttal gener-

ation, the choice of a rebuttal strategy depends on the intended effect of the user’s argument (according to the user model) and its actual effect (according to the normative model).

### 3 Determining a User’s Line of Reasoning

Our procedure for recognizing a user’s intended line of reasoning from his/her rejoinder receives two inputs: a linguistic clue (“but” or “consider”) and a rejoinder proposition ( $\mathbf{R}$ ), e.g., “but Mr Green was in the garden”. It then finds paths in the user model BN that connect  $\mathbf{R}$  to the goal proposition (Zukerman et al., 2000). During this process, BIAS copes with inference patterns that are different from its own by allowing inferred paths to contain a small “gap” composed of propositions that did not exist previously in the user model. Figure 2(a) illustrates an Argument Graph, a rejoinder  $\mathbf{R}$ , and

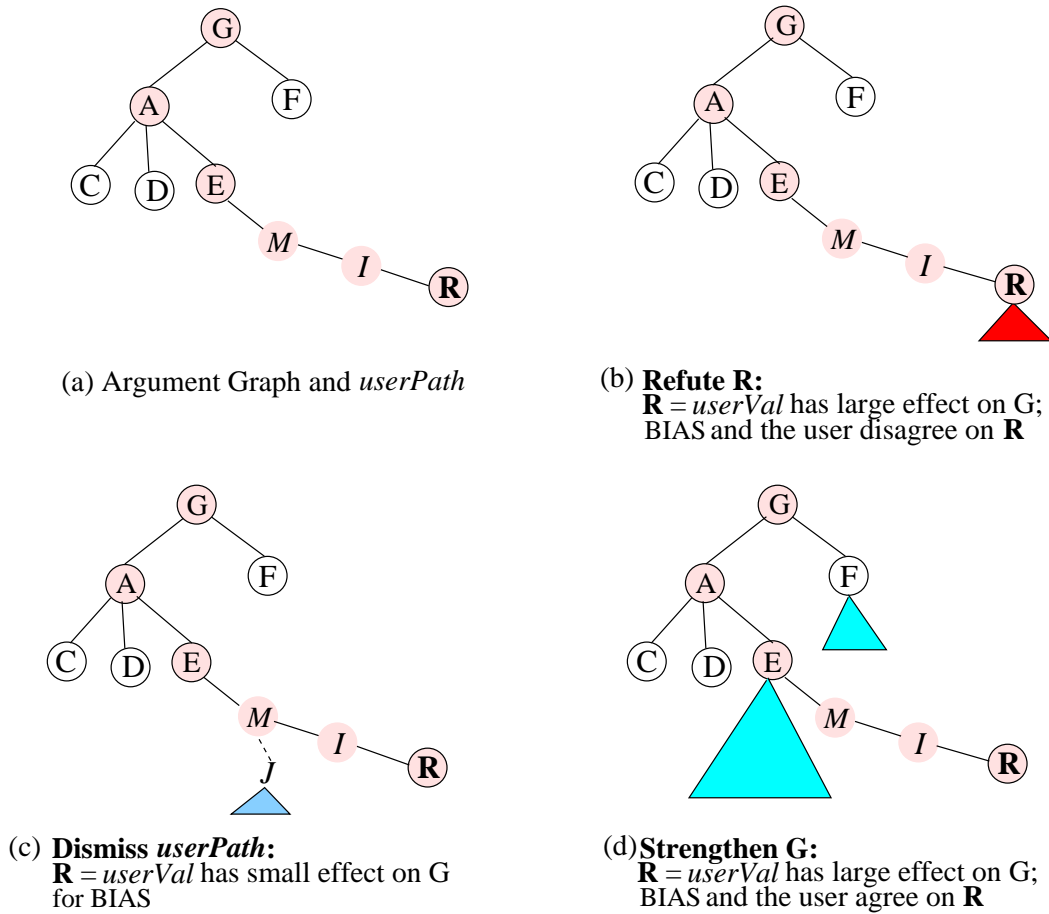


Figure 2: Sample Argument Graph and Rejoinder Strategies

path  $\mathbf{R}$ - $I$ - $M$ - $E$ - $A$ - $\mathbf{G}$  between them (composed of grey nodes). This path, called *userPath*, represents the line of reasoning intended by the user. The gap in this path contains nodes  $I$  and  $M$  (in italics), which means that the user inferred  $E$  directly from  $\mathbf{R}$ .

Each path is assigned a score based on the following factors: the impact of  $\mathbf{R}$  on BIAS' argument along this path, whether path nodes are in the user's attentional focus, and BIAS' confidence in this path (determined from the information source of the nodes in this path, e.g., whether the user has seen the propositions in the path, asserted a belief about them or read them in BIAS' arguments). BIAS then selects the highest-scoring path. If several paths have a high score, the user is asked to choose one of them. Typically, BIAS returns a single path, and sometimes it returns two or three paths. Hence, presenting them to the user for selection is a reasonable course of action.

## 4 Rebuttal Generation

Given a user's rejoinder proposition  $\mathbf{R}$ , we consider three main types of rebuttals: (1) refute  $\mathbf{R}$ , (2) dis-

miss the line of reasoning intended by the user (*userPath*), and (3) strengthen the argument goal  $\mathbf{G}$ . Diagrammatic representations of these rebuttal strategies and abridged versions of their applicability conditions appear in Figure 2(b-d). These conditions, which are specified in the following sections, depend on (1) whether the rejoinder affects the system's argument directly or indirectly, (2) the beliefs in  $\mathbf{R}$  in the normative and user models, and (3) the impact of  $\mathbf{R}$  on the goal proposition along *userPath* in the normative and user models.

### 4.1 Refute the rejoinder

This strategy consists of generating an argument against the user's belief in the rejoinder proposition  $\mathbf{R}$ . This strategy is applicable under the following conditions:

- (R1) The beliefs in  $\mathbf{R}$  in the user model and the normative model differ significantly (the user's belief in  $\mathbf{R}$  contradicts BIAS' belief); and
- (R2) Either
  - (a)  $\mathbf{R}$  was stated or implied in BIAS' argument ( $\mathbf{R}$  appears in the Argument Graph), or

- (b) The belief in  $\mathbf{R}$  stated by the user has a significant effect on the goal  $G$  in the normative model in the same direction as its effect on  $G$  in the user model.

For example, if the user’s rejoinder to the argument in Figure 1 was “But Mr Green and Mr Body were not enemies”, then conditions R1 and R2a would be satisfied, since the rejoinder directly contradicts what was stated by BIAS in the argument. If the user’s rejoinder was “But the neighbour saw Mr Green shoot Mr Body”, then conditions R1 and R2b would be satisfied, since an inference from this rejoinder contradicts BIAS’ belief in Mr Green’s lack of opportunity to kill Mr Body (and consequently in Mr Green’s guilt). The argument schema for the *refute the rejoinder* strategy and a sample rebuttal produced with this schema are shown in Figure 3.<sup>3</sup> The sub-argument that argues against the rejoinder proposition is generated by activating our Bayesian argument generator (Zukerman et al., 1998) with the proposition *Mr Green and Mr Body were enemies* as the goal. In this case, the belief in the rejoinder node resulting from the sub-argument differs from that stated in the initial argument, owing to the additional information included in the sub-argument. Hence, the implications from the rejoinder node are followed. The procedure for following these implications is described in Section 4.2.

#### 4.2 Dismiss the user’s line of reasoning

This strategy consists of showing the user how his/her argument fails to achieve its intended effect. We distinguish between concessive and contradictory dismissals. The former is used when the system agrees with the rejoinder proposition  $\mathbf{R}$ , and the latter when the system disagrees with  $\mathbf{R}$ . This strategy is applicable under the following condition:

- (D)  $\mathbf{R}$  does not significantly affect the belief in  $G$  in the normative model.

This condition is illustrated by the rejoinder to the argument in Figure 1, “Consider that the found gun was available only to Mr Green”, which purports to increase the belief in Mr Green’s means to kill Mr Body, and hence Mr Green’s guilt. However, since this increment is quite small, BIAS adopts the dismissal strategy, which follows the effect of the user’s rejoinder through the user’s line of reasoning, pointing out how the effect of the rejoinder differs from its intended effect. It is worth noting that the main difference between a dismissal and a strengthening of the goal is that BIAS decides to generate a dismissal when its current beliefs are sufficient to invalidate the user’s line of reasoning, whereas it decides to

<sup>3</sup>The rejoinders shown in this paper are posed by the user immediately after the argument in Figure 1.

<p><b>Refute <math>\mathbf{R}</math>:</b></p> <ol style="list-style-type: none"> <li>1. Deny the belief in <math>\mathbf{R}</math> stated by the user.</li> <li>2. Present a sub-argument for the normative belief in <math>\mathbf{R}</math>.</li> <li>3. If <math>\mathbf{R}</math> is not in the Argument Graph or the belief in <math>\mathbf{R}</math> as a result of the sub-argument differs from that originally stated by BIAS, then follow the effect of <math>\mathbf{R}</math> along <i>userPath</i> up to the first node in the Argument Graph whose belief is the same as that stated in the initial argument.</li> </ol>
<p><b>Rejoinder:</b> But Mr Green and Mr Body were not enemies.</p> <p><b>Rebuttal:</b> Actually, it is quite likely that Mr Green and Mr Body were enemies. This is for the following reason. The forensic analysis of the blue paint being reliable and forensics having found some blue paint which they estimate is one week old implies a blue car was here last week. This together with Mr Green having a blue car implies Mr Green’s car was almost certainly here last week, which implies Mr Green almost certainly visited Mr Body last week. The neighbour being sober implies she is very probably reliable. This together with the neighbour reporting Mr Green arguing with Mr Body last week implies the neighbour very probably heard Mr Green arguing with Mr Body last week, which together with Mr Green almost certainly visiting Mr Body last week implies he almost certainly argued with Mr Body. The Bayesian Times reporting Mr Body took Mr Green’s girlfriend implies Mr Body probably seduced Mr Green’s girlfriend. This together with Mr Green almost certainly arguing with Mr Body implies Mr Green and Mr Body probably were enemies. Let’s now go back to the main argument. Mr Green and Mr Body probably being enemies implies it is more likely that Mr Green had a motive to murder Mr Body, making it rather likely. This implies it is only slightly more likely that he murdered Mr Body.</p>

Figure 3: *Refute the rejoinder* Schema and Example

strengthen the goal when additional information is required to defeat the impact of the user’s rejoinder.

Our algorithm for dismissing the user’s line of reasoning follows *userPath* until it reaches a point where the user’s line of reasoning fails, i.e., it has no effect on a proposition on *userPath* in the Argument Graph. It is necessary for the rebuttal to reach the

Argument Graph even if the failure of the rejoinder occurs earlier in *userPath*, because the user’s rejoinder refers to the argument, hence at least one proposition in the argument must be mentioned when addressing the impact of this rejoinder.

The user’s line of reasoning may fail due to the following factors: (1) s/he did not consider propositions that have a significant effect on the propositions in *userPath*; or (2) his/her belief in one or more of the propositions s/he did consider differs significantly from that in the normative model, and this proposition has a substantial effect on a proposition in *userPath*. Propositions of the first type are included in a set called *SIGneighbours*, and propositions of the second type are included in *DIFFneighbours*. Our dismissal algorithm calls our Bayesian argument generator to generate sub-arguments for the propositions in *DIFFneighbours*, but simply presents the propositions in *SIGneighbours* without arguing for them.

**Algorithm** *DismissUserReasoning*(*userPath*)

Let *userPath* be composed of propositions  $\mathbf{R} = P_0 \rightarrow P_1 \rightarrow P_2 \rightarrow \dots \rightarrow P_n = \mathbf{G}$ .

1. For  $i = 1$  to  $n$  do:
  - (a) Set *SIGneighbours*( $P_i$ ) to the nodes that are linked to  $P_i$  in the normative model but not in the user model and have a significant effect on the belief in  $P_i$ .
  - (b) If the belief in  $P_i$  in the user model differs significantly from the belief in  $P_i$  in the normative model, then set *DIFFneighbours*( $P_i$ ) to the nodes that are linked to  $P_i$  in both the user model and the normative model and which have a different belief in the user model from that in the normative model.
  - (c) For each node  $P_j \in \text{DIFFneighbours}(P_i)$  generate a sub-argument for the normative belief in  $P_j$ .
2. Present the resulting rebuttal using the appropriate schema, *DismissContradict* or *DismissConcede* (Figures 4 and 5 and respectively).

Our concessive schema differs from our contradictory schema in two respects. Firstly, the former acknowledges the user’s rejoinder, while the latter denies it. In addition, the concessive schema follows the user’s line of reasoning starting from the normative belief in the rejoinder proposition (which is close to the belief indicated by the user), while the contradictory schema follows a hypothetical line of reasoning starting from the user’s belief in the rejoinder proposition (which differs substantially from the normative belief). In both cases the user’s line of reasoning fizzles out, due to its small effect on the

<p><b>DismissContradict</b> <i>userPath</i>:</p> <ol style="list-style-type: none"> <li>1. Deny the belief in <math>\mathbf{R}</math> stated by the user, and dismiss its hypothetical effect on the goal proposition.</li> <li>2. Present the sub-arguments for the nodes in <i>DIFFneighbours</i>.</li> <li>3. FollowPath <i>userPath</i> from the rejoinder proposition to the goal.</li> </ol>
<p><b>FollowPath</b> <i>userPath</i>  For <math>i = 0</math> to <math>n - 1</math> (where <math>n</math> is the number of nodes in <i>userPath</i>) do:</p> <ol style="list-style-type: none"> <li>1. If <math>P_{i+1}</math> is not in the Argument Graph or <math>\text{DIFFneighbours}(P_{i+1}) \neq \emptyset</math>, then present an implication from <math>P_i</math> to <math>P_{i+1}</math> which includes the nodes linked to <math>P_{i+1}</math> in the user model plus the nodes in <i>SIGneighbours</i>(<math>P_{i+1}</math>).  Else present an implication which reflects only the relative impact of <math>P_i</math> on <math>P_{i+1}</math>.</li> <li>2. If the resulting belief in <math>P_{i+1}</math> is the same as that stated in the initial argument, then stop.</li> </ol>

Figure 4: *DismissContradict* Schema and *FollowPath* Procedure

goal according to the normative model irrespective of its truth value.

Both schemas follow *userPath* from the rejoinder proposition to the goal using procedure *FollowPath* (Figure 4). This procedure distinguishes between propositions in *userPath* for which the main influencing factors (*DIFFneighbours* and *SIGneighbours*) should be presented, and those which require only information regarding the relative impact of the preceding proposition in *userPath*. The latter propositions are characterized as follows: (1) they appear in the Argument Graph; and (2) the user’s beliefs in the nodes outside *userPath* that have a significant effect on these propositions are consistent with the normative beliefs in these nodes. For instance, the rebuttal in Figure 1, which is generated by means of the *DismissContradict* schema, presents the relative influence of *Mr Green fired the gun* on *Mr Green fired the murder weapon*, since the user and BIAS hold consistent beliefs regarding the gun being the murder weapon.

To illustrate the operation of the dismissal algorithm, let us consider the rejoinder “But the time of death was 11”, which yields the following line of reasoning: *The time of death was 11* ( $\&$  *Mr Green was in the garden at 11*)  $\rightarrow$  *Mr Green was in the garden at the time of death*  $\rightarrow$  *Mr Green had the opportunity to kill Mr Body*  $\rightarrow$  *Mr Green killed Mr Body*. *DIFFneighbours* includes only one proposition, *Mr*

**DismissConcede *userPath*:**

1. Acknowledge the belief in **R** stated by the user, and dismiss its effect on the goal proposition.
2. Present the sub-arguments for the nodes in *DIFFneighbours*.
3. FollowPath *userPath* from the rejoinder proposition to the goal.

**Rejoinder:** But the time of death was 11.**Rebuttal:**

Indeed, it is quite likely but not entirely certain that the time of death was 11. However, this has only a small effect on the likelihood that Mr Green murdered Mr Body.

I will show that Mr Green almost certainly wasn't in the garden at 11.

Mr Green's witness not being related to Mr Green implies she is very probably reliable. This together with Mr Green's witness reporting Mr Green being at the football at 10:30 implies Mr Green was almost certainly at the football at 10:30.

The neighbour being sober implies she is almost certainly reliable. This together with the neighbour reporting Mr Green not being in the garden at 11 implies the neighbour never saw Mr Green in the garden at 11, which together with Mr Green almost certainly being at the football at 10:30 implies he almost certainly wasn't in the garden at 11.

Let's now go back to the main argument.

Even though the time of death was probably 11, Mr Green almost certainly not being in the garden at 11 implies it is only slightly less likely that he was in the garden at the time of death. This implies it is only slightly less likely that he had the opportunity to murder Mr Body, which implies it is only slightly less likely that he murdered Mr Body.

Figure 5: *DismissConcede* Schema and Example

*Green was in the garden at 11*, since the belief in it in the normative model differs from that in the user model, thereby prompting the generation of a sub-argument for this proposition. This sub-argument is stronger than that incorporated in the initial argument, yielding a belief in Mr Green being in the garden at 11 that is lower than the belief indicated in the original argument, which in turn reduces the belief in Mr Green being in the garden at the time of death, Mr Green having the opportunity to kill Mr Body, and Mr Green actually murdering Mr Body. The resulting rebuttal, which is presented by means of the *DismissConcede* schema, appears in Figure 5.

**4.3 Strengthen the goal**

This strategy consists of generating a stronger argument for the original goal proposition **G**, bringing to bear information that did not appear in the initial argument (either because **BIAS** was unaware of it or because **BIAS** chose to exclude it from the argument). This strategy is applicable under the following conditions:

- (G1) The beliefs in **R** in the normative and user models are consistent; and
- (G2) **R** has a substantial detrimental effect on the belief in **G** in the normative model. This change in belief should be in the same direction as the change occurring in the user model.

These conditions represent a situation where the system did not take into account a particular fact, but when this fact comes to its attention the system realizes the effect of this fact on the goal. For instance, if the user discovers new evidence that places Mr Green in the garden at 11, a rejoinder which presents this proposition will increase the belief in Mr Green's opportunity to kill Mr Body along the following line of reasoning: *Mr Green was in the garden at 11* → *Mr Green was in the garden at the time of death* → *Mr Green had the opportunity to kill Mr Body* → *Mr Green killed Mr Body*. In this case, **BIAS** tries to strengthen the argument for Mr Green's innocence by arguing separately against propositions along this line of reasoning (other than the rejoinder node, which is true in this example). If no sub-argument can be generated for these nodes or the generated sub-arguments do not significantly affect the goal, then **BIAS** agrees with the user.

Our algorithm for strengthening the goal searches along *userPath* for propositions that have been affected by the rejoinder, but that will reinforce **BIAS**' goal proposition if their belief is changed. It then tries to generate sub-arguments that change the beliefs in these propositions. In order to localize the effect of the user's rejoinder, the search and sub-argument generation processes start at **R** and proceed towards the goal. The presentation of the rebuttal is also done in this order, using a procedure which is similar to the *FollowPath* procedure described in Section 4.2.

**Algorithm *StrengthenGoal*(*userPath*)**

Let *userPath* be composed of propositions  $\mathbf{R}=P_0 \rightarrow P_1 \rightarrow P_2 \rightarrow \dots \rightarrow P_n=\mathbf{G}$ .

1. For  $i = 1$  to  $n$ , while the belief in **G** is not as intended by **BIAS**, do:
  - (a) Determine which belief in  $P_i$  will move the belief in **G** in the normative model in the direction intended by **BIAS**.
  - (b) If this belief in  $P_i$  differs from the current belief in  $P_i$ , then

- i. Generate a sub-argument for the desired belief in  $P_i$ .
  - ii. If the sub-argument yields a significant change in the belief in  $P_i$  or in the belief in  $G$  then store the sub-argument in  $SubAG(P_i)$ .
2. Present the resulting rebuttal (composed of the user’s line of reasoning and intervening sub-arguments) using the StrengthenGoal schema in Figure 6.

To illustrate the operation of this algorithm, let us reconsider the rejoinder “Consider Mr Green was in the garden at 11”, and let us assume that the rejoinder proposition is true. Inspection of the propositions affected by this rejoinder reveals that if Mr Green was not in the garden at the time of death, then the belief in the goal would be closer to that intended by BIAS. However, an argument for this proposition cannot be generated. Hence, BIAS proceeds to the next proposition, *Mr Green had the opportunity to murder Mr Body*, and calls our Bayesian argument generator to generate an argument that contradicts this proposition. The Bayesian generator produces an argument which reduces the belief in this proposition. However, this belief cannot be reduced to the extent that it exculpates Mr Green. Thus, BIAS attempts to generate an argument for the goal node (by trying to reduce the belief in Mr Green’s means and motive to kill Mr Body). However, this attempt also fails, leaving BIAS with a moderate belief in Mr Green’s guilt.<sup>4</sup>

It is important to note that although BIAS’ immediate objective is to strengthen its belief in the goal proposition, its primary purpose is to “tell the truth” to the best of its knowledge (which may contradict its initial beliefs), rather than to win the argument at all costs. Our algorithm supports this attitude by retaining any sub-argument which has a significant impact on the goal or on a proposition on *userPath*. We use this disjunctive condition on impacts in order to address a situation where a proposition  $P_j$  on *userPath* has been affected by a sub-argument, but does not affect the goal because of an inaccurate belief in a proposition  $P_k$  which appears later on *userPath* (recall that the propositions are inspected from  $\mathbf{R}$  towards the goal). However, once a sub-argument for  $P_k$  is presented, then  $P_j$  affects the goal. If BIAS accepted only sub-arguments for propositions which have a significant impact on the goal, then in this case it would miss the opportunity to strengthen the goal.

<sup>4</sup>The resulting argument has not been included owing to space limitations.

**StrengthenGoal *userPath*:**

1. Acknowledge the belief in  $\mathbf{R}$  stated by the user, and set *lastProposition* to  $\mathbf{R}$ .
2. Until the goal proposition is reached do:
  - (a) If after *lastProposition* there is a proposition  $P_i \in userPath$  for which a sub-argument was generated ( $SubAG(P_i) \neq \emptyset$ ), then
    - i. Follow *userPath* from *lastProposition* to  $P_i$ .
    - ii. Present the sub-argument for  $P_i$ .
    - iii. Set *lastProposition* to  $P_i$ .
  - (b) Else follow the remainder of *userPath*.

Figure 6: *Strengthen the goal* Schema

## 5 Related Research

Our research builds on work described in (Zukerman et al., 1998), which generated arguments from BNs, and (Zukerman et al., 1999), which enabled a user to explore the impact of different propositions on the generated arguments. The former system only generated arguments, while the latter received instructions from a user (through a menu) about modifications to be performed to a previously generated argument, e.g., including or excluding a proposition, and then generated a new argument in response to these instructions. Neither of these systems generates rebuttals which take into account a user’s intentions, as done by BIAS.

Several researchers have dealt with different aspects of argumentation, e.g., (Flowers et al., 1982; Quilici, 1992; Chu-Carroll and Carberry, 1995; Carberry and Lambert, 1999). Like BIAS, the system described in Carberry and Lambert (1999) combined linguistic and contextual knowledge to recognize a user’s intentions from rejoinders. However, their system did not generate rebuttals. Chu-Carroll and Carberry (1995) provided a comprehensive approach for proposal evaluation which focused on dialogue strategies rather than argumentation strategies. In addition, they considered exchanges where each participant utters one or two propositions in each conversational turn. In contrast, we focus on strategies for the generation of extended probabilistic rebuttals to individual rejoinders. In the future, our strategies will be combined with dialogue strategies in a complete argumentation system.

Flowers et al. (1982) presented a partial theory of argumentation which advocated the combination of distinct knowledge sources; their implementation focused on recognizing and providing episodic justifications to historical events. Our focus on the generation of rebuttals in the context of BNs allows us

to provide an operational definition for the broad argumentation strategies discussed in the literature, e.g., attack the main point directly or attack the supporting evidence (Flowers et al., 1982).<sup>5</sup>

The argumentation system described in (Quilici, 1992) used a plan-based model of the user's beliefs to recognize the justification for a user's proposal and provide its own justifications. However, the rebuttals generated by this system were based on a single strategy: applying backwards chaining using a set of justification rules. This strategy is a special case of the more general rebuttal schemas presented here.

## 6 Conclusion and Future Work

We have offered a mechanism for generating rebuttals to a user's rejoinders in the context of arguments generated by a Bayesian argumentation system. We have implemented three main argumentation strategies: refuting the rejoinder, strengthening the argument goal, and dismissing the user's line of reasoning. For each strategy we have identified applicability conditions, proposed a procedure which determines the information to be included in a rebuttal, and defined a presentation schema.

An interesting area of future research pertains to the omission of information from an argument. There are different types of information which may be omitted from an argument, such as (1) easily inferred information and information which has a small effect on the argument; (2) information which is required for representational reasons, but makes the resulting argument more confusing; (3) probabilistic information which, although correct, makes the resulting argument more tedious; and (4) previously stated information. Our previous research deals with the first type of information (Zukerman et al., 1998), and in this paper we have identified some conditions for the omission of previously stated information when expressing the relative impact of a proposition. Another factor that affects the omission of information is the trade off between accuracy and conciseness. The omission of information affects the belief in the conclusion(s) presented in an argument. Stating beliefs that differ from a system's own beliefs may cause the system to appear inconsistent or even deceitful, while presenting all the relevant factors may yield a verbose argument. A mechanism which addresses these issues will support the generation of better arguments and rebuttals.

The evaluation of this research encompasses several components: (1) the WWW interface, (2) the path-finding mechanism, and (3) the rebuttal-generation mechanism. A preliminary evaluation of

---

<sup>5</sup>We do not handle the "attack the claim that the evidence gives support for the main point" strategy, as this involves inferring Conditional Probability Matrices for the user model, which is outside the scope of this research.

the path-finding mechanism has yielded encouraging results (Zukerman et al., 2000). Two types of evaluation are envisaged for the rebuttal-generation mechanism. A whole-system evaluation, where users interact freely with BIAS, may be used to determine whether users are satisfied with the system as a whole. In contrast, a specific evaluation of rebuttals would be restricted to showing users rejoinder-rebuttal pairs (after showing an initial argument), and eliciting the users' reactions regarding the appropriateness of the rebuttals.

## 7 Acknowledgments

This work was supported in part by Australian Research Council grant A49927212.

## References

- Carberry, S. and Lambert, L. (1999). A process model for recognizing communicative acts and modeling negotiation subdialogues. *Computational Linguistics*, 25(1):1–53.
- 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.
- 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.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann Publishers, San Mateo, California.
- Quilici, A. (1992). Arguing about planning alternatives. In *COLING-92 – Proceedings of the Fourteenth International Conference on Computational Linguistics*, pages 906–910, Nantes, France.
- Zukerman, I., Jitnah, N., McConachy, R., and George, S. (2000). Recognizing intentions from rejoinders in a Bayesian interactive argumentation system. To appear in *PRICAI2000 – Proceedings of the Sixth Pacific Rim International Conference on Artificial Intelligence*, Melbourne, Australia.
- 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, 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.