# Cooperative Replies to Unbelievable Assertions: A Dialogue Protocol Based on Logical Interpolation


**M. Nykänen**
School of Computing
University of Eastern Finland
Finland

**S. Eloranta** and **O. Niinivaara**
Department of Computer Science
University of Helsinki
Finland

**R. Hakli**
Helsinki Institute for
Information Technology
Finland



Keywords: belief revision, dialogue protocols, convictions, integrity constraints, argumentation



Abstract: We propose a dialogue protocol for situations in which an agent makes to another agent an assertion that the other agent finds impossible to believe. In this interaction, unbelievable assertions are rejected using explanations formed by logical interpolation and new assertions are being made such that all previous rebuttals are taken into account.


## 1 Introduction

When two agents carry out a conversation with each other, one of them may well assert something which the other cannot believe for some reason. Witness the following example (Hansson, 1991):

*Conversation* 1.

**Amy:** Last summer I saw a three-toed woodpecker just outside my window. I could clearly see its red forehead and its red rump.

**Bob:** You must be mistaken. A three-toed woodpecker does not have a red forehead or a red rump.

**Amy:** You make me uncertain. Thinking about it, the only thing I am certain of is that the bird had a red forehead.

We study here such conversations: Both agents have beliefs that they are certain of and that they are not willing to give up during the conversation. Here, Bob's ornithological knowledge is one example, and Amy's certainty of seeing a bird with a red forehead is another. When hearing an unbelievable assertion, Bob faces the task of helping Amy by offering informative rebuttals. Amy then faces the task of generating another assertion while taking Bob's rebuttal into account. This interaction continues until either Amy comes up with an assertion which Bob can consider possible or she concludes that they have irreconcilable differences, at least as far as this conversation is concerned.

We consider these conversations in the context of *belief revision* (Alchourrón et al., 1985) in the presence of what we call *convictions*. By these convictions we mean those beliefs the agent refuses to give up, at least during the current conversation. In several fields there are important concepts that can be interpreted as convictions: In computer science, *integrity constraints* (Reiter, 1988) are needed to ensure consistency of databases. In philosophy, the properties of *knowledge* differ from those of belief (Hintikka, 1962), and people take a different stand on what they take to know and not merely believe. In nonprioritized belief revision, *core beliefs* are immune to revision (Hansson, 1999, gives a survey). In theories of argumentation, agents have *dark-side commitments*, which are their fundamental commitments that they find extremely hard to retract once stated in a conversation (Walton and Krabbe, 1995, pp. 11–12).

We encounter the problem of what an agent should do when another agent asserts something that conflicts with his convictions. In this paper we propose a solution, in which the agents carry out a conversation as an *interactive preparatory phase* before belief revision. In this phase they seek together a final assertion which does not conflict with either agent's convictions. In Conversation 1, Amy's second assertion might serve as something which they both might be able to believe. We do not consider what happens after this preparatory phase, that is, we do not concern ourselves whether the agents actually revise their epistemic states or not.

Focusing on these assertion-rebuttal conversations raises immediately three questions: First, how can an agent form his rebuttal to the unbelievable assertion? Second, how should the other agent form her[1] next assertion on the basis of her epistemic state while taking into account the new information in the rebuttal? And third, how can the conversation stay focused on its original subject?

Our answer to the first question is to use logical *interpolation*, since it gives a formula which is entailed by the convictions of the agent and entails the negation of the unbelievable assertion. Moreover, it can be read as a description how or an explanation why (Hintikka and Halonen, 1999) the assertion conflicted with the convictions of the agent, thereby bringing some currently relevant part of his convictions into light. Our answer to the second question is to use *hypothetical* thinking, as if the agent thought: "if I were to believe this rebuttal, then I would have this belief about my topic instead of the one I expressed before". She will form her next assertion as the least disbelieved alternative to the topic given the new information. Our answer to the third question is to require that their utterances remain relevant to it in the *letter-sharing* sense (Makinson, 2009, Definition 1.1), which is guaranteed by interpolation.

In related work, there are some approaches in which agents have convictions but do not use interaction for conflict resolution. These include such approaches to nonprioritized belief revision that secure some beliefs from revision. For instance, in Accommodative Belief Revision (Eloranta et al., 2008), the agent tries to guess what the other would have said, had she had his knowledge. In our solution, the agent does not have to guess what the other agent would believe, instead he gives her a chance to tell it.

Then there are approaches in which interaction is used as a preprocessing step before belief revision, but the possibility of agents having their private convictions is not considered. These include such merging approaches as mutual belief revision (Jin et al., 2007) and belief negotiation (Booth, 2006) in which all the agents' beliefs are weakened until they no longer contradict each other. As opposed to that, our solution is asymmetric. We have one agent, who is eager to inform another agent about some of her beliefs, whereas the other agent is willing to reply and share some of his convictions in case he finds the original assertion unbelievable. Application areas with such a setting include knowledge base systems in which some agents (either human beings or software agents) collect information and send it to one agent acting as a knowledge base with integrity constraints.

[1]We adhere to the convention that the asserting agent (such as Amy in Conversation 1) is female, whereas the rebutting agent (such as Bob) is male, and refer to them as "she" and "he" as well as by name.

Certain types of argumentation-based dialogues (Walton and Krabbe, 1995; Parsons et al., 2003) can also be viewed as preparatory phases for belief revision: They aim at finding out whether a particular assertion should be believed by exchanging information about arguments that either support or undermine it. In our approach, however, the goal is to find out *what* could be believed about the topic when the agents' convictions are taken into account, not *whether* a particular proposition should be believed or not.

For example, van Veenen and Prakken (2006) included asking "Why did you rebut my assertion?" among the moves in their negotiation protocol as an embedded persuasion game. However, their idea is to bring the grounds for the rebuttal to light so that they too can be subject to further scrutiny by the other agent within this conversation. In contrast, the purpose of our dialogues is not to persuade the other agent to accept the original assertion, but to find an alternative assertion that is acceptable.

Our aim is that the agents' assertions in the dialogues satisfy *the Cooperative Principle* presented by Grice (1989, Chapters 2 and 3) to govern conversations between cooperative agents. These maxims rule out the naive extremities to deal with an unacceptable input, that is, either to terminate the dialogue or to reply with all one knows about the subject. Something more is needed, thus we will propose the use of logical interpolation as a cooperative reply to an assertion that an agent is convinced to be false.

The paper is organized as follows. In section 2, we will introduce our notations and present the interpolation principle as a tool for generating cooperative replies to unbelievable assertions. In section 3, we will propose guidelines for generating a new modified assertion based on this reply. Section 4 presents the conversation protocol driven by these interpolants and shows that the conversations will always end with a rational outcome. In section 5, we will give conclusions and propose some directions for future research.

## 2 Cooperative replies

We will consider dialogues, that is, conversations between two agents. We assume that these two agents, named $A$ and $B$, such as Amy or Bob in Conversation 1, have *epistemic states*, which we denote with $\mathcal{A}$ and $\mathcal{B}$ correspondingly. These states contain the *belief sets* consisting of all the beliefs they currently hold; these sets we denote with $\mathbb{B}(\mathcal{A})$ and $\mathbb{B}(\mathcal{B})$. In these belief sets, beliefs are expressed with formulas of classical propositional logic; that is, they are beliefs about the actual state of affairs, and not about for instance beliefs about each other's beliefs. On the one hand, each agent may be willing to give up some of these beliefs given new evidence to the contrary. On the other hand, (s)he may regard some of them as convictions which (s)he will hold on to, regardless of any such new evidence. We denote the sets of convictions for agents $A$ and $B$ with $\mathbb{C}(\mathcal{A})$ and $\mathbb{C}(\mathcal{B})$ correspondingly. We assume that the sets of beliefs and the sets of convictions are non-contradictory and deductively closed, that is, $\mathbb{B}(\mathcal{A}) = \mathrm{Cn}(\mathbb{B}(\mathcal{A}))$, etc. We also assume that what an agent is convinced of, (s)he also believes, that is, $\mathbb{C}(\mathcal{A}) \subseteq \mathbb{B}(\mathcal{A})$ and $\mathbb{C}(\mathcal{B}) \subseteq \mathbb{B}(\mathcal{B})$.

Agent $A$ is the initiator of the conversation. We are focusing entirely on the situation, in which agent $B$ is convinced that the assertion made by $A$ is false. We are not concerned whether agent $B$ will give priority to the assertion if it does not contradict his convictions.

We use lower-case Greek letters to denote propositional formulas. Agent $A$ initiates the conversation with her initial assertion $\varphi$. Whenever the assertion is acceptable by agent $B$, the dialogue ends. Therefore we shall presume this initial assertion to be unbelievable, that is, $\mathbb{C}(\mathcal{B}) \models \neg\varphi$. As already mentioned in Section 1, we propose for agent $B$ to use the interpolant to this entailment as a cooperative reply in this situation, since (i) it is entailed by $\mathbb{C}(\mathcal{B})$ and (ii) it entails $\neg\varphi$ and (iii) it only employs vocabulary that appears in the original assertion (in terms of propositional variables).

Let $\mathbb{V}$ denote all the propositional variables and $\mathrm{Voc}(\alpha) \subset \mathbb{V}$ those appearing in the formula $\alpha$. Let us recall what interpolation is:

**Theorem 1** (Craig interpolation, for propositional logic)**.** *Given two propositional formulas* $\alpha$ *and* $\beta$, *if* $\alpha \models \beta$, *then there is some* interpolant $\theta$ *such that (i)* $\alpha \models \theta$, *(ii)* $\theta \models \beta$, *and (iii)* $\mathrm{Voc}(\theta) \subseteq \mathrm{Voc}(\alpha) \cap \mathrm{Voc}(\beta)$.

In the full paper (Nykänen et al., 2011), we prove this theorem for the system G3cp of sequent calculus (Negri and von Plato, 2001, Chapter 3.1) in a way which provides an explicit algorithmic construction for the interpolant $\theta$ given the formulas $\alpha$ and $\beta$.

In Conversation 1, $\alpha$ is a formula representing (some relevant part of) Bob's convictions $\mathbb{C}(\mathcal{B})$, $\beta$ is the negation $\neg\varphi$ of Amy's initial assertion $\varphi$, and $\theta$ is a suggestion for an explanation why $\alpha$ rules out $\beta$.

*Example* 1. Consider Conversation 1, and let the propositional variable $p$ stand for "Amy saw a three-toed woodpecker", $q$ stand for "Amy saw a bird with a red forehead", $r$ stand for "Amy saw a bird with a red rump", $s$ stand for "Amy saw a lark". The conversation starts when Amy asserts $p \wedge q \wedge r$. Assume Bob's convictions include $(p \vee s) \rightarrow (\neg q \wedge \neg r)$. Then $(p \vee s) \rightarrow (\neg q \wedge \neg r)$ entails an interpolant $p \rightarrow \neg q \wedge \neg r$, which again entails the negation of Amy's assertion, $\neg(p \wedge q \wedge r)$.

Let us consider how well this suggestion fares in light of *Grice's Maxims* (Grice, 1989, Chapters 2 and 3). These maxims elaborated his general Cooperative Principle into more specific conversational rules which the participants can be expected to observe:

**Maxim of Quantity:** (i) Make your contribution as informative as required (for the current purposes of the exchange). (ii) Do not make your contribution more informative than is required.

**Maxim of Quality:** Try to make a contribution which is true. More specifically: (i) Do not say what you believe to be false. (ii) Do not say that for which you lack adequate evidence.

**Maxim of Relevance:** Be relevant.

**Maxim of Manner:** Be clear.

Using an interpolant as a reply conforms to part (i) of the maxim of Quantity, because it conveys information that agent $A$ supposedly was not aware of, since it entails the negation of what she said. In part (ii), the amount of informativeness can be controlled by the selection of a suitable interpolation formula.

Using an interpolant as the reply conforms also to the maxim of Quality: not only does agent $B$ reply with a belief of his, but with a conviction, and we assume here that a rational agent does not obtain convictions without proper evidence.

According to the maxim of Relevance, a reply should somehow be related to the preceding conversation. A natural syntactic concept is letter-sharing: two formulas are relevant to each other, if they share some propositional variable (Makinson, 2009, Definition 1.1). In this regard, an interpolant is an extremely relevant reply, since it consists only of variables in both $\mathbb{C}(\mathcal{B})$ and the current suggestion by agent $A$, by property (iii) of Theorem 1. Makinson (2009) notes that although letter-sharing is not wholly unproblematic as a notion of relevance, it does have its uses in computational contexts. Hence we define here the *topic* of the conversation to be the variables $\mathrm{Voc}(\varphi)$ about which agent $A$ wants to have a conversation with agent $B$.

Another more refined concept of relevance could be to split the beliefs of an agent into disjoint parts, where each part consists of the agent's beliefs about a particular subject matter (Parikh, 1999). Since Theorem 1 can be extended to such split theories

(Kourosias and Makinson, 2007, Theorem 1.1), we anticipate our approach to apply in this setting as well.

There are also more semantic accounts of relevance. In the theory of Sperber and Wilson (2004), the main determinant of relevance of a reply is how much positive cognitive effect (such as learning, settling doubt, or correcting mistaken assumptions) it creates in the recipient. In this respect too, a reply by interpolant fares well since it contradicts with what agent *A* supposedly believes and should thus invoke a process of belief revision.

Regarding the maxim of Manner, one could argue that it does not concern conversations such as ours, where the messages exchanged are formulated in logic instead of in a natural language. We note, however, that even in our conversations agent *B* can tailor the form of his interpolant to enhance its clarity to agent *A*.

## 3 On generating new assertions

Let us contemplate on what agent *A* should do when she learns a formula $\theta$ that tells her why her previous assertion was unbelievable.

If $\theta$ conflicts with the convictions of agent *A*, we take that this dialogue should fail. If the rebuttal is not unbelievable to agent *A*, then she can either (i) accept the input as a conviction (since it was *B*'s conviction), (ii) accept the input as a belief, or (iii) treat the input conditionally. The treatment may depend, e.g., on how reliable the agent considers the other agent. By the maxim of Quality, agent *A* then continues the dialog with a new assertion $\psi$, which she accordingly (i) believes (now that she is convinced that $\theta$), (ii) believes (now that she has come to believe that $\theta$), or (iii) would believe, if she were to believe that $\theta$.

Denoting the doxastic conditional "if I were to believe $\theta$, then I would also believe $\psi$" as $\theta \circ\!\!\rightarrow \psi$, our requirement of the maxim of Quality becomes $\mathcal{A} \models \theta \circ\!\!\rightarrow \psi$, where this entailment is defined through the *Ramsey test:* if agent *A* were to revise her epistemic state $\mathcal{A}$ with $\theta$, then would $\psi$ be believed in the resulting state $\mathcal{A} \circ \gamma$? Thus in all three alternatives, *A* may anwer $\psi$ if and only if $\psi \in \mathbb{B}(\mathcal{A} \circ \theta)$. Note that this revision might be only tentative: the actual epistemic state of agent *A* might still be $\mathcal{A}$.

Now let us assume that the revision operator '$\circ$' that agent *A* uses (either when revising her epistemic state or when evaluating conditionals) satisfies the basic rationality criteria (R1)–(R4) for belief revision (Alchourrón et al., 1985) and the rationality criterion (IR1) for iterated belief revision (Darwiche and Pearl, 1997). That is:

$$\alpha \in \mathbb{B}(\mathcal{A} \circ \alpha). \tag{R1}$$

$$\text{If } \neg\alpha \notin \mathbb{B}(\mathcal{A}) \text{ then } \mathbb{B}(\mathcal{A} \circ \alpha) = \mathrm{Cn}(\mathbb{B}(\mathcal{A}) \cup \{\alpha\}). \tag{R2}$$

$$\text{If } \alpha \text{ is satisfiable then } \mathbb{B}(\mathcal{A} \circ \alpha) \text{ is consistent.} \tag{R3}$$

$$\text{If } \alpha \equiv \beta \text{ then } \mathbb{B}(\mathcal{A} \circ \alpha) = \mathbb{B}(\mathcal{A} \circ \beta). \tag{R4}$$

$$\text{If } \alpha \models \beta \text{ then } \mathbb{B}(\mathcal{A} \circ \alpha) = \mathbb{B}((\mathcal{A} \circ \beta) \circ \alpha). \tag{IR1}$$

Postulate (R1) says that the new piece of information is accepted, that is, the insertion succeeds. Postulate (R2) says that if the new piece of information is compatible with the old beliefs, neither is any of them discarded nor is anything not entailed by the old beliefs and the new information added to the belief set. Postulate (R3) says that adding a satisfiable formula to the belief set must not make it inconsistent. Postulate (R4) calls for syntax independence. Postulate (IR1) says that if $\alpha \models \beta$, then the beliefs in the epistemic state obtained when learning first $\beta$ and then $\alpha$ are the same as when learning just $\alpha$ in the first place.

By the maxim of Relevance, agent *A* must bear in mind (and take into account) all the rebutting formulas during the dialog. Let $\gamma$ denote the conjunction of those formulas. Thus in the scenarios, we must use $\gamma$ instead of $\theta$.

By the basic postulates for belief revision, we have $\mathcal{A} \models \theta \circ\!\!\rightarrow \psi$ if and only if $\mathcal{A} \circ \theta \models \theta \circ\!\!\rightarrow \psi$. Now assume that after *n* rebuttals, $\gamma_n = \theta_1 \wedge \theta_2 \wedge \ldots \wedge \theta_n$. Then by postulate (IR1), we have $\mathcal{A} \models \gamma_n \circ\!\!\rightarrow \psi$ if and only if $\mathcal{A} \circ \gamma_1 \circ \gamma_2 \circ \ldots \circ \gamma_n \models \gamma_n \circ\!\!\rightarrow \psi$. Thus the truth value of the conditional does not depend on whether the agent has actually revised her epistemic state on the way or not: all three alternatives for *A*'s actions remain equivalent in this respect.

Our framework allows for several methods for constructing a formula $\psi$ satisfying this requirement. Some methods are discussed in the full paper (Nykänen et al., 2011).

## 4 A conversation protocol

We will now give a conversation protocol in which agent *B* uses interpolation to create rebuttals. In our protocol, agent *A* starts with an assertion $\varphi$, which also fixes the topic of the conversation. When agent *B* receives an assertion which conflicts with $\mathbb{C}(\mathcal{B})$, he answers with an interpolant $\theta$. In the protocol, $\psi$ contains the most recent assertion made by agent *A*, and $\gamma$ is the conjunction of all the rebuttals made by agent *B* to her previous assertions, as above. The protocol is depicted as follows:

CONVERSATION PROTOCOL

1 $\psi \leftarrow \varphi$; $\gamma \leftarrow \top$
2 $A$ asserts $\psi$
3 **while** $\mathbb{C}(\mathcal{B}) \models \neg\psi$ with some interpolant $\theta$
4 **do** $B$ replies that he is convinced of $\theta$
5 $\gamma \leftarrow \gamma \wedge \theta$
6 **if** $\mathbb{C}(\mathcal{A}) \models \neg\gamma$
7 **then** $A$ says that their convictions conflict with each other
8 **return** FAIL
9 $\psi \leftarrow$ some formula chosen by $A$ such that $\mathcal{A} \models \gamma \circ\!\!\rightarrow \psi$
10 $A$ asserts $\psi$
11 $B$ replies that he too considers this $\psi$ believable
12 **return** SUCCESS with $\psi$

Selecting the next assertion $\psi$ on line 9 is possible if and only if $\mathbb{C}(\mathcal{A}) \cup \{\gamma\}$ is consistent, and this is guaranteed by line 6. If our conversation protocol terminates successfully (on line 12) then we do have an agreement: both agents could believe the final assertion $\psi$. For agent $B$, this follows by line 3. For agent $A$, this is an invariant of the **while** loop: it holds before the loop by line 1 and the maxim of Quality (i), and it continues to hold after each execution of the loop body by line 9.

Notice the algorithm uses the current epistemic states of the agents. As to the convictions the agents have, this causes no problems, because the agents do not give up their convictions. The beliefs of agent $B$ are not used in the protocol. As to the beliefs of agent $A$, whether she has actually revised her epistemic state with $\gamma$ or not does not affect the truth value of the conditional, as discussed in section 3.

The protocol could generate a conversation like the following:

*Conversation* 2.

**Amy:** I saw a bird with a red forehead and a red rump. It was a three-toed woodpecker!

**Bob:** A three-toed woodpecker does not have a red forehead.

**Amy** *(thinking to herself):* In that case I would have to give up believing either that I saw a three-toed woodpecker or a red forehead. I prefer to keep believing the former. Hence I can take Bob's rebuttal into account by giving up the latter. I can also keep believing in the red rump, since he did not challenge that part.

*(aloud to Bob):* In that case I think I saw a three-toed woodpecker, which had a red rump but no red forehead.

**Bob:** A three-toed woodpecker does not have a red rump either.

**Amy** *(after similar thinking):* Well, in that case I think I saw a three-toed woodpecker, but it had neither a red rump nor a red forehead.

**Bob:** Now *that* is something I could believe!

If our protocol terminates with failure instead (on line 8) then this outcome is warranted as well: On the one hand, $\mathbb{C}(\mathcal{B}) \models \gamma$ by property (i) of Theorem 1 and lines 3 and 5. On the other hand, $\mathbb{C}(\mathcal{A}) \models \neg\gamma$ by line 6. Agent $A$ can even explain this conflict between their convictions with an interpolant corresponding to the entailment on line 6, thereby expressing her own convictions about the topic. This could be useful if the agents attempt to reconcile their convictions somehow, but we do not consider such attempts here.

If we assume that agent $A$ keeps her assertions $\psi$ relevant to the topic in our chosen letter-sharing sense, then our protocol becomes finite, albeit $O(2^{|\mathrm{Voc}(\varphi)|})$ in the worst case.

**Theorem 2** (Finiteness)**.** *Assume that the assertions $\psi$ by agent A in our conversation protocol satisfy* $\mathrm{Voc}(\psi) \subseteq \mathrm{Voc}(\varphi)$. *Then the maximum number of times its **while** loop can be executed is*[2]

$$|\{w \restriction \mathrm{Voc}(\varphi) \colon w \in \mathrm{Mod}(\mathbb{C}(\mathcal{A})) \setminus \mathrm{Mod}(\mathbb{C}(\mathcal{B}))\}|\,.$$

*Proof.* See the full paper (Nykänen et al., 2011). □

However, we expect actual conversations to terminate in much fewer rounds than the pessimal upper bound in Theorem 2, since the more precise explanations $\theta$ agent $B$ gives for his rejections, the fewer assertions agent $A$ needs. And indeed, maxim of Quantity (i) directs agent $B$ towards such precise explanations $\theta$. However, we leave constructing precise interpolants $\theta$ for later study.

## 5 Conclusions and future work

We considered dialogues as a preparatory phase for belief revision and we presented a dialog protocol for resolving conflicts resulting from unbelievable assertions. Depending on the result of the dialogue, the agents either have found out that their convictions are in conflict with each other, or they have found a formula that neither of them finds unbelievable. During the dialogue, the unbelievable assertions do not cause the rebutting agent to change his epistemic state, but the asserting agent might (or might not) change her epistemic state due to the rebuttals. In general, we allow the agents to change their epistemic states during the dialog.

[2]Here $f \restriction D = \{\langle x, f(x)\rangle : x \in D\}$ denotes the restriction of the function $f$ into the domain $D$.

Our protocol can terminate successfully even when there is a conflict, since it may never surface during the protocol. Suppose for instance that $A$ asserts some $a \vee b$, where $A$ can believe the first disjunct but not the second disjunct, and vice versa for $B$. We leave such pseudoagreements to further study, since avoiding them would require continuing the conversation further even after finding this first mutually believable formula.

The work can be extended to several directions, such as adding the possibility to extend the topic with new literals, adding the possibility to agree to restrict the topic, adding new utterance types to the agents (for instance, for making the protocol symmetric), and considering more expressive languages as is done e.g. in cooperative query answering (Gaasterland et al., 1992).

## REFERENCES


Alchourrón, C. E., Gärdenfors, P., and Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530.

Booth, R. (2006). Social contraction and belief negotiation. *Information Fusion*, 7:19–34.

Darwiche, A. and Pearl, J. (1997). On the logic of iterated belief revision. *Artificial Intelligence*, 89(1-2):1–29.

Eloranta, S., Hakli, R., Niinivaara, O., and Nykänen, M. (2008). Accommodative belief revision. In Hölldobler, S., Cutz, C., and Wansing, H., editors, *11th European Conference on Logics in Artificial Intelligence (JELIA 2008)*, number 5293 in Lecture Notes in Artificial Intelligence (LNAI), pages 180–191. Springer.

Gaasterland, T., Godfrey, P., and Minker, J. (1992). An overview of cooperative answering. *Journal of Intelligent Information Systems*, 1:123–157.

Grice, P. (1989). *Studies in the Way of Words*. Harvard University Press.

Hansson, S. O. (1991). Belief contraction without recovery. *Studia Logica*, 50:251–260.

Hansson, S. O. (1999). A survey of non-prioritized belief revision. *Erkenntnis*, 50(2-3):413–427.

Hintikka, J. (1962). *Knowledge and Belief*. Cornell University Press.

Hintikka, J. and Halonen, I. (1999). Interpolation as explanation. *Philosophy of Science*, 66 (Proceedings):S414–S423.

Jin, Y., Thielscher, M., and Zhang, D. (2007). Mutual belief revision: semantics and computation. In *Proceedings of the 22nd national conference on Artificial Intelligence (AAAI'07)*, pages 440–445. AAAI Press.

Kourosias, G. and Makinson, D. (2007). Parallel interpolation, splitting, and relevance in belief change. *The Journal of Symbolic Logic*, 72(3):994–1002.

Makinson, D. (2009). Propositional relevance through letter-sharing. *Journal of Applied Logic*, pages 377–387.

Negri, S. and von Plato, J. (2001). *Structural Proof Theory*. Cambridge University Press.

Nykänen, M., Eloranta, S., Niinivaara, O., and Hakli, R. (2011). Cooperative replies to unbelievable assertions: A dialogue protocol based on logical interpolation. Technical Report C-2011-1, Department Of Computer Science, University of Helsinki, Finland. Available at `http://www.cs.helsinki.fi/group/protean/crua/`.

Parikh, R. (1999). Beliefs, belief revision, and splitting languages. In Moss, L., Ginzburg, J., and de Rijke, M., editors, *Logic, Language, and Computation*, volume 2, pages 266–278. CSLI Publications.

Parsons, S., Wooldridge, M., and Amgoud, L. (2003). Properties and complexity of some formal inter-agent dialogues. *Journal of Logic and Computation*, 13:347–376.

Reiter, R. (1988). On integrity constraints. In *2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 97–111. Morgan Kaufmann.

Sperber, D. and Wilson, D. (2004). Relevance theory. In *The Handbook of Pragmatics*, pages 607–632. Blackwell, Oxford.

van Veenen, J. and Prakken, H. (2006). A protocol for arguing about rejections in negotiation. In Parsons, S., Maudet, N., Moraitis, P., and Rahwan, I., editors, *Second International Workshop on Argumentation in Multi-Agent Systems (ArgMAS 2005)*, number 4049 in Lecture Notes in Artificial Intelligence (LNAI), pages 138–153. Springer.

Walton, D. N. and Krabbe, E. C. W. (1995). *Commitment in Dialogue: Basic Commitments in Interpersonal Dialogue*. SUNY series in logic and language. State University of New York Press.