

# Evidence Elimination in Multi-Agent Justification Logic

Bryan Renne  
Faculty of Philosophy  
University of Groningen  
Oude Boteringestraat 52  
9712 GL Groningen  
The Netherlands  
<http://bryan.renne.org/>

## Abstract

This paper presents a logic combining *Dynamic Epistemic Logic*, a framework for reasoning about multi-agent communication, with a new multi-agent version of *Justification Logic*, a framework for reasoning about evidence and justification. This novel combination incorporates a new kind of *multi-agent evidence elimination* that cleanly meshes with the multi-agent communications from Dynamic Epistemic Logic, resulting in a system for reasoning about multi-agent communication and evidence elimination for groups of interacting rational agents.

## 1 Introductory Example

Consider the following email messages sent between mutual friends Anne (from Atlanta), Bryan, and Charlie (from Columbus).

$x_1$  { To: Bryan, Charlie  
From: Anne  
Date: 1 Jan 2009, 15:00 EST  
Subject: Job  
Bryan, Charlie: I may get a job in Columbus! If I get the job, then I will move to Columbus. Charlie: do you have space for a roommate?

---

© ACM, 2009. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Aviad Heifetz, editor, Proceedings of the 12th Conference of Theoretical Aspects of Rationality and Knowledge (TARK XII), pages 227–236, Stanford, California, USA, 2009. ACM. <http://doi.acm.org/10.1145/1562814.1562845>

$x_2$  { To: Anne, Bryan  
From: Charlie  
Date: 1 Jan 2009, 17:00 EST  
Subject: Re: Job  
Anne: it would be great if you were to get the job and then move here. And yes, my spare room is available if you want it.

$x_3$  { To: Bryan, Charlie  
From: Anne  
Date: 5 Jan 2009, 15:00 EST  
Subject: Job!  
Bryan, Charlie: I got the job!

$x_4$  { To: Anne, Bryan  
From: Charlie  
Date: 5 Jan 2009, 17:00 EST  
Subject: Re: Job!  
Anne: it is great you got this job! I will prepare the room for you. How exciting!

After this sequence of messages, Bryan and Charlie are each able to conclude that Anne is moving to Columbus. After all, each of Bryan and Charlie has message  $x_1$  as evidence that Anne will move if she gets the job along with message  $x_3$  as evidence that Anne indeed got the job. So by combining their evidence  $x_1$  and  $x_3$ , they each have evidence that Anne will move.

Furthermore, since Charlie replied to each of Anne's messages  $x_1$  and  $x_3$  with his own messages  $x_2$  and  $x_4$ , respectively indicating that Charlie has received and understood each of Anne's messages, Bryan not only has his combined evidence  $x_1$  and  $x_3$  that Anne is moving but can also see on the basis of Charlie's messages  $x_2$  and  $x_4$  that Charlie too has combined evidence  $x_1$  and  $x_3$  to conclude that Anne is moving.

But now we add a twist to the story with the following private message from Anne to Bryan.

$x_5$  { To: Bryan  
From: Anne  
Date: 7 Jan 2009, 15:00 EST  
Subject: Rescinded  
Bryan: due to the faltering economy, the company has rescinded my job offer! I have not yet told Charlie because he is so excited about my coming to Columbus. Ugh!

Clearly, this message has the effect of causing Bryan to set aside (or *eliminate*) his evidence  $x_3$  relevant to the assertion that Anne got the job. But since  $x_3$  was sent privately to Bryan, Bryan has no reason to believe that Charlie will doubt the evidence  $x_3$  that Anne got the job. Therefore, while message  $x_5$  causes Bryan to eliminate his evidence  $x_3$  relevant

to the assertion that Anne got the job—and so to abandon his combined evidence  $x_1$  and  $x_3$  that she is moving—Bryan will nevertheless maintain his belief based on the combination of evidence  $x_2$  and  $x_4$  that Charlie still believes based on the combination of evidence  $x_1$  and  $x_3$  that Anne is moving to Columbus.

## 2 Introduction

As indicated by our email example, this paper concerns reasoning about multi-agent communication and evidence elimination. Our work here is part of a larger project aimed at joining two areas of study: *Justification Logic*,<sup>1</sup> a family of logics for reasoning about evidence and justification, and *Dynamic Epistemic Logic*,<sup>2</sup> a family of logics for reasoning about multi-agent communication and belief. Before we discuss our rationale for seeking a combination of these theories, let us first say a few words about each of them in isolation.

*Justification Logic* has been promoted as a logic for reasoning about evidence and justification (Artemov 2008; Fitting 2009). The idea here is to remedy a deficiency found in the standard use of modal logic for reasoning about the justifications agents have for the beliefs that they possess. To illustrate this deficiency, consider a valid modal formula of the form  $B_i\varphi \supset B_i\psi$ , where  $B_i$  is a modal operator and each of  $\varphi$  and  $\psi$  are formulas. This formula, read, “if agent  $i$  believes  $\varphi$ , then agent  $i$  believes  $\psi$ ,” stipulates a certain connection between agent  $i$  believing one thing,  $\psi$ , and agent  $i$  believing another thing,  $\varphi$ . But we see that this formula does not provide a *reason* as to why agent  $i$ ’s belief in one thing follows whenever he or she has belief in another. The formula merely asserts that such conditional belief holds without providing any explanation as to why this is the case.

Justification Logic aims to better explicate this situation by first introducing structured syntactic objects called *terms* and then allowing us to form new formulas of the form  $t :_i \varphi$  whenever  $t$  is a term,  $i$  is an agent, and  $\varphi$  is a formula we have already formed. The idea is to identify the structure of the term  $t$  with an abstract description of a derivation within a given theory of Justification Logic in a way that satisfies *Artemov’s Internalization Theorem* (Artemov 2001): if  $p$  is a proof of a theorem  $\chi$  in the logic, then there is a systematic way to construct a term  $u_p$  whose structure reflects the structure of  $p$  such that  $u_p :_i \chi$  is also a theorem of the logic. This provides us with a *proof-based notion of evidence*, in the sense that the appearance of a term  $t$  in a theorem  $t :_i \varphi$  encodes a description as to why it is (according to the theory) that agent  $i$  believes that  $\varphi$  holds. This leads us to read the formula  $t :_i \varphi$  as “ $t$  is agent  $i$ ’s evidence that  $\varphi$  is true.”

Returning to the deficiency of modal logic, Justification Logic enables us to replace the statement  $B_i\varphi \supset B_i\psi$  of conditional (modal) belief with a more nuanced statement of conditional *evidence-based belief*: for a term  $s_t$  built up from  $t$ , the formula  $t :_i \varphi \supset s_t :_i \psi$

---

<sup>1</sup>See: Artemov 2008; Artemov and Nogina 2005; Artemov 2001; Fitting 2009; Fitting 2005; Kuznets 2008; Renne 2008a.

<sup>2</sup>See: Baltag and Moss 2004; Baltag, Moss, and Solecki 1998; Baltag, van Ditmarsch, and Moss 2008; Renne 2008b; van Benthem 2006; van Benthem, van Eijck, and Kooi 2006; van Ditmarsch, van der Hoek, and Kooi 2007.

says, “if  $t$  is agent  $i$ ’s evidence for  $\varphi$ , then  $s_t$  is agent  $i$ ’s evidence for  $\psi$ .” Since  $s_t$  is built up from  $t$ , we see that this tells us that in case agent  $i$  has justification  $t$  for his or her belief in  $\varphi$ , then agent  $i$  may justify his or her belief in  $\psi$  by inserting the argument  $t$  into the appropriate places in the argument  $s_t$ , thereby yielding an argument supporting  $\psi$ . In this way, Justification Logics provide an in-language notion of *proof-based evidence*. To distinguish this notion of evidence from other notions of evidence, it may be helpful to think of proof-based evidence in the following intuitive way:

To have *proof-based evidence* for  $\varphi$   
 is to have conclusive argumentation  
 supporting  $\varphi$ .

One immediate difficulty for the proof-based notion of evidence found in many Justification Logics is that it is *static*: in these theories, either agent  $i$  has an argument supporting  $\varphi$ , in which case he or she has and will always have evidence for  $\varphi$ , or else agent  $i$  does not have an argument supporting  $\varphi$ , in which case he or she does not and will never have evidence for  $\varphi$ . Such Justification Logics do not allow agents to *learn* new evidence, to *forget* old evidence, or to *have a change of mind* about the reliability of existing evidence. This is clearly at odds with everyday experience where, for example, agent  $i$  may believe  $\varphi$  on the basis of an argument  $s_t$  only to find out later that the evidence  $t$  on which  $s_t$  is crucially dependent is completely unreliable, thereby forcing agent  $i$  to throw out both  $t$  and  $s_t$  and hence the belief in  $\varphi$  on the basis of  $s_t$ . What is missing from basic Justification Logics is a certain *evidence dynamics*, whereby certain basic pieces of evidence can change their status from “good” to “bad” (or the other way around).

A further criticism of the proof-based notion of evidence is that it neglects the intuitive *social* role that evidence plays as a vehicle for *persuasion*: if agent  $i$  has an argument supporting  $\varphi$  but agent  $j$  does not, then agent  $i$  ought to be able to tell agent  $j$  of  $i$ ’s argument supporting  $\varphi$ , which might have the effect of convincing agent  $j$  to believe  $\varphi$ . Here what is missing is a means of describing how evidence dynamics can be brought about as a result of *communication*. But note that the why in which this communication takes place must also be taken into account:  $i$ ’s argument might be told to  $j$  in public, which then might also affect the beliefs of other agents who hear of  $i$ ’s argument; alternatively,  $i$ ’s argument may be told to  $j$  in private, which would probably not effect the others’ beliefs (though they might develop suspicions as to what  $i$  and  $j$  are talking about when they see  $i$  and  $j$  depart for a closed-door meeting).

One framework that elegantly describes a wide range of multi-agent communications—whether public or private, with or without deception, having or not having suspicion, and so on—is the framework of *Dynamic Epistemic Logic*.<sup>3</sup> In Dynamic Epistemic Logic, a communication is modeled using what we call an *update frame*, which is a finite Kripke frame<sup>4</sup> whose worlds have been labeled by formulas. The idea is that each world  $w$  in an

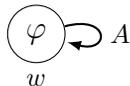
---

<sup>3</sup>See: Baltag and Moss 2004; Baltag, Moss, and Solecki 1998; Baltag, van Ditmarsch, and Moss 2008; Renne 2008b; van Benthem 2006; van Benthem, van Eijck, and Kooi 2006; van Ditmarsch, van der Hoek, and Kooi 2007.

<sup>4</sup>A *finite Kripke frame* is a finite directed graph whose nodes are called “worlds” and whose edges are

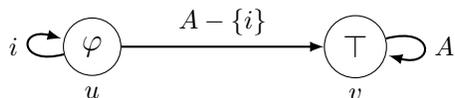
update frame represents a possible communication of the formula  $\chi_w$  that labels  $w$ . The structure of the arrows in the update frame describes the agents' conditional uncertainties as to which communication is actually taking place: in case there is an  $i$ -arrow from world  $w$  to world  $w'$ , then agent  $i$  will think it possible that the formula  $\chi_{w'}$  is communicated if it is the case that the formula  $\chi_w$  is actually communicated. This allows us to use update frames to represent a wide variety of communications.

As an example, for a nonempty finite set  $A$  of agents, the update frame



represents the *public announcement of  $\varphi$* , since for every agent  $i \in A$ , in case  $\chi_w = \varphi$  is the formula that is actually communicated, then, since there is a unique  $i$ -arrow from  $w$  to  $w$ ,<sup>5</sup> agent  $i$  thinks that the only formula that can possibly be communicated is the formula  $\chi_w = \varphi$  itself. Thus to have such a communication occur is to make it the case that every agent will simultaneously come to know  $\varphi$ , which is intuitively what happens when a public announcement of  $\varphi$  occurs. Of course, the update frame above for the public announcement of  $\varphi$  also says something about *higher-order knowledge*: since there is a unique  $i$ -arrow from  $w$  to  $w$  for an agent  $i \in A$  and since there is also a unique  $j$ -arrow from  $w$  to  $w$  for another agent  $j \in A$ ,  $j$  knows that  $i$  knows that  $\varphi$  is to be communicated (after all, at the unique world  $w$  that  $j$  considers possible, agent  $i$  knows that  $\chi_w$  is the formula that is communicated, this because the only possibility that  $i$  considers is the formula  $\chi_w$  itself). It is not too hard to see that the update frame above describes a situation in which it is *common knowledge* that  $\varphi$  is communicated; that is, each agent knows that  $\varphi$  is communicated, each agent knows that each agent knows that  $\varphi$  is communicated, each agent knows that each agent knows that each agent knows that  $\varphi$  is communicated, and so on for any finite number of occurrences of the phrase “each agent knows that” appearing before the phrase “ $\varphi$  is communicated.”

As another example, the update frame



represents the *private announcement of  $\varphi$  to agent  $i$* , since in case  $\chi_u = \varphi$  is the formula that is actually communicated, the unique  $i$ -arrow from  $u$  to  $u$  says that  $i$  thinks that the only formula that can possibly be communicated is the formula  $\chi_u = \varphi$ , whereas the unique  $j$ -arrow from  $u$  to  $v$  for each  $j \in A - \{i\}$  says that any other agent  $j$  thinks that the formula  $\chi_v = \top$  (the propositional constant for truth, which conveys no new information)

---

labeled by agent names. We will call the edges of a finite Kripke frame “arrows” and we will use the phrase “ $X$ -arrow” to refer to an arrow that is labeled by  $X$ .

<sup>5</sup>We adopt the following drawing convention for update frames: for a (possibly empty) set  $S$  of agents, if a drawing of an update frame contains an  $S$ -arrow from a world  $w$  to a world  $w'$ , then what is meant is that for each agent  $a \in S$ , there is to be an  $a$ -arrow from  $w$  to  $w'$ .

was communicated.<sup>6</sup> Thus if  $\varphi$  is actually communicated, then the information about  $\varphi$  is sent to agent  $i$  while all the other agents gain no new information. This is intuitively what it is to have a private announcement of  $\varphi$  to agent  $i$ .

Using symbols to denote update frames, Dynamic Epistemic Logic allows us to form new formulas of the form  $[U, u]\varphi$ , where  $U$  is a symbol for an update frame,  $u$  is a symbol naming a world in the update frame  $U$  (here  $u$  indicates that the formula  $\chi_u$  is what is actually communicated), and  $\varphi$  is a formula that has already been formed. The formula  $[U, u]\varphi$  then says, “after the update  $(U, u)$ , [where the formula  $\chi_u$  is actually communicated,] we have that  $\varphi$  is true.” (Note that we use the words “update” and “communication” interchangeably because very complicated update frames are better thought of as generalized “informational updates” due to the fact that there is not always an intuitive, everyday communicative type corresponding to each complex update frame; see the discussion in Renne (2006).) Since the formula  $\varphi$  occurring in  $[U, u]\varphi$  can be a formula describing various agents’ beliefs, we see that Dynamic Epistemic Logic can be used to describe how communications affect the beliefs of agents; for example,  $[U, u]B_i\psi$  says that agent  $i$  believes  $\psi$  after the occurrence of the update  $(U, u)$ .

While a communication can affect agent belief, the language of Dynamic Epistemic Logic—itsself an extension of the language of modal logic—does not provide us with the means to describe *reasons* for such changes in agent belief. To illustrate, let  $p$  be a propositional letter and then consider a valid formula of the form  $p \supset [U, u]B_i p$  (“if  $p$  is true, then, after update  $(U, u)$ , agent  $i$  believes  $p$ ”). This formula stipulates a connection between the occurrence of the update  $(U, u)$  and agent  $i$  coming to believe the fact  $p$ . In essence, the content of the update  $(U, u)$  brings about this change in agent  $i$ ’s belief, though the reason why this particular change in belief is brought about is left unstated—the formula says only that this communication brings about  $i$ ’s change in belief *for some reason*. (This is similar to the case of modal logic itself, where the formula  $B_i\varphi \supset B_i\psi$  does not say *why* it is that  $i$ ’s belief in  $\psi$  follows whenever he or she believes  $\varphi$ .) Indeed, despite the fact that an explanation based on the structure of  $U$  justifies this change in  $i$ ’s belief, this explanation is not expressible within the language itself. So for example, if  $U$  is the public announcement of  $p$ , then the explanation of why  $(U, u)$  causes  $i$  to believe  $p$  would go something like “since a public announcement of a fact causes all agents to come to learn that fact, agent  $i$  will come to believe the fact  $p$  after it is publicly announced;” however, such an explanation cannot be formulated within the language. Furthermore, if we suppose that another agent  $j$  believes that update  $(U, u)$  will cause agent  $i$  to come to believe the fact  $p$ ,

$$B_j(p \supset [U, u]B_i p) ,$$

then, just as in modal logic, there is no way for the language to describe agent  $j$ ’s justification for this belief. In short, the language is not only lacking the capacity to express the agents’

---

<sup>6</sup>More precisely: the  $A$ -arrow from  $v$  to  $v$  makes it so that in case  $\chi_u = \varphi$  is the formula that is actually communicated, each of the non- $i$  agents mistakenly thinks that it is common knowledge that  $\top$  is communicated (and so no new information is conveyed). So while  $i$  learns the new information  $\varphi$ —this due to the  $i$ -arrow from  $u$  to  $u$ —the other agents have the mistaken common belief that no new information has been communicated.

reasons for *holding* their beliefs (as is the case with modal logic) but it is also unable to express the agents' reasons for *changing* their beliefs.

To summarize what we have said thus far, we have argued that each of Justification Logic and Dynamic Epistemic Logic would be better suited for modeling, describing, and studying rational agency if it had additional features. In the case of Justification Logic, we have argued for the ability to handle *evidence dynamics*, taking into account the *social* role that evidence plays as a vehicle for persuasion in various kinds of communication, be it public or private, with or without deception, having or not having suspicion, and so on. In the case of Dynamic Epistemic Logic, we have argued for an *in-language account of evidence* so that agents can provide justifications not only as to why they hold certain beliefs but also as to why they would change these beliefs. We thus think it natural to look for a combined framework—something that might eventually be called *Dynamic Justification Logic*—that brings together the complementary strengths from each of Dynamic Epistemic Logic and Justification Logic so as to address each area's respective weakness.

In this paper, we begin the project of Dynamic Justification Logic by proposing a theory called  $\text{JLce}^A$  that combines multi-agent communication and belief from Dynamic Epistemic Logic with a new multi-agent Justification Logic for reasoning about evidence and justification.  $\text{JLce}^A$ , the *theory of Justification Logic with multi-agent communication and evidence elimination*, allows us to reason not only about multi-agent belief dynamics arising as the result of Dynamic-Epistemic-Logic-style communications—hereafter called (*multi-agent communications*)—but also about a kind of multi-agent evidence dynamics we call *evidence elimination*, whereby a piece of evidence relevant to a particular claim is to be set aside as part of a multi-agent communication whose content effectively undermines the applicability of the evidence for the claim. Our email example indicates one of the many kinds of multi-agent evidence eliminations that  $\text{JLce}^A$  can reason about.

After defining the language, axiomatics, and semantics of  $\text{JLce}^A$ , we will show how this theory can be used to formalize the evidence elimination from our email example. Due to space constraints, we will omit proofs in this text; the interested reader may find full details in Renne (2009a), an extended version of this paper.

### 3 Syntax

Our language, called  $\text{UL}^A$ , is used to reason about multi-agent communication and evidence elimination for a finite nonempty set  $A$  of rational agents.  $\text{UL}^A$  consists of two kinds of syntactic objects: *terms*, which represent pieces of evidence, and *formulas*, which express propositions. For a language  $L$ , we will write  $\mathcal{T}(L)$  for the set of all terms in  $L$ , and we will write  $L$  both for the name of the language and also for the set of formulas in this language.

$\text{UL}^A$  makes use of what we call *update frames*, which are structures  $(W, R, f, \mathcal{A})$  consisting of a finite, transitive Kripke frame  $(W, R)$  said to be *underlying*  $U$  (note that  $R : A \rightarrow 2^{W \times W}$  is a function assigning a transitive binary relation  $R_i$  to each agent  $i \in A$ ), a *formula-labeling function*  $f : W \rightarrow L$  that assigns a formula  $f(w)$  in a language  $L$  to each world  $w \in W$ , and an *evidence-labeling function*  $\mathcal{A} : A \rightarrow (\mathcal{T}(L) \times L) \rightarrow 2^W$  that assigns to each agent  $i \in A$  a

function  $\mathcal{A}_i$  that itself assigns to each term-formula pair  $(t, \varphi) \in \mathcal{T}(L) \times L$  a set  $\mathcal{A}_i(t, \varphi)$  of worlds at which the term  $t$  is considered by agent  $i$  to be evidence relevant to the assertion that  $\varphi$  is true. We write  $u \in U$  to mean that  $u$  is a world in the Kripke frame underlying  $U$ . A *pointed update frame* is a pair  $(U, u)$  consisting of an update frame  $U$  and a world  $u \in U$ .

To allow us to utilize the components of an update frame  $U = (W, R, f, \mathcal{A})$  without having to name these components, we adopt the following abbreviations: writing the name “ $U$ ” of the update frame  $U$  where one would expect a set of worlds is an abbreviation for the set  $W$ , writing  $U_i$  is an abbreviation for  $R_i$ , writing  $U^l$  is an abbreviation for  $f$ , and writing  $U^e$  is an abbreviation for  $\mathcal{A}$ .

Informally speaking, to say that  $t$  is agent  $i$ 's evidence relevant to the assertion that  $\varphi$  (*is true*) means that  $i$  considers evidence  $t$  to be probative for  $\varphi$ —that is,  $i$  thinks that  $t$  generally tends to demonstrate or prove  $\varphi$ —though  $i$  need not be persuaded to believe  $\varphi$  on the basis of evidence  $t$ . So to have  $i$  take evidence  $t$  as relevant for  $\varphi$  is just to say that  $i$  considers  $t$  to be admissible as evidence for  $\varphi$ , regardless of whether this admissibility also comes with belief. We will write  $t \gg_i \varphi$  to mean that  $t$  is  $i$ 's evidence relevant to the assertion that  $\varphi$  is true.

When agent  $i$  has a piece of evidence  $t$  relevant to the assertion  $\varphi$  and agent  $i$  also believes that  $\varphi$  is true, then we will say that  $t$  is *agent  $i$ 's evidence that  $\varphi$  (is true)* and we will write  $t :_i \varphi$ . So we have two notions of evidence: the weaker notion  $t \gg_i \varphi$  of relevant evidence and the stronger notion  $t :_i \varphi$  of relevant evidence plus belief.

Pointed update frames  $(U, u)$  will be used to describe an *update*, which is a multi-agent communication. Originally devised by Baltag, Moss, and Solecki (BMS) (Baltag and Moss 2004; Baltag, Moss, and Solecki 1998) and now standardized in Dynamic Epistemic Logic (Baltag, van Ditmarsch, and Moss 2008; Renne 2008b; van Ditmarsch, van der Hoek, and Kooi 2007), a pointed update frame  $(U, u)$  describes the communication of the formula  $U^l(u)$  within a network of possible formulas  $U^l(v)$  for  $v \in U$  that might be communicated. So while the update  $(U, u)$  actually communicates the formula  $U^l(u)$ , the agents have uncertainty as to which communication actually occurs according to the structure of the Kripke frame underlying  $U$  and the way in which  $U^l$  labels worlds in  $U$  by formulas. We explained this in greater detail above in the introduction, where we gave examples of two update frames, one for the public announcement of  $\varphi$  and one for the private announcement of  $\varphi$  to agent  $i$ .

Our contribution in this paper is to add a notion of *multi-agent evidence elimination* to the above-described BMS-picture. We thus view each world in an update frame  $U$  as constituting a *possible communication-elimination*: for an update  $(U, u)$ , in case  $u \in U_i^e(x_k, \varphi)$  for a term  $x_k$  and a formula  $\varphi$ , then the communication-elimination possibility simultaneously communicates  $U^l(u)$  and eliminates agent  $i$ 's evidence  $x_k$  relevant to  $\varphi$ , which has the effect of falsifying  $x_k \gg_i \varphi$ . But an elimination of one piece of evidence  $t$  relevant for an assertion can have consequences for other pieces of evidence  $s_i$  that are built up from  $t$ . We saw this in our email example: Anne's final email  $x_5$  eliminated Bryan's evidence  $x_3$  relevant to Anne getting the job, and this elimination itself triggered an elimination of Bryan's combined evidence  $x_1$  and  $x_3$  that Anne is moving to Columbus.

Acknowledging the first multi-agent language for Justification Logic (Yavorskaya (Sidon)

2008), we define a new multi-agent *update language*, written  $UL^A$ , which consists of the *terms*  $t$  and the *formulas*  $\varphi$  formed by the following grammar.

$$\begin{aligned}
t & ::= c_k \mid x_k \mid t_1 \cdot_{\varphi} t_2 \mid t_1 + t_2 \mid !t \mid t^{U,u} \\
\varphi & ::= p_k \mid \perp \mid \top \mid \varphi_1 \star \varphi_2 \mid \neg\varphi \mid B_i\varphi \mid \\
& \quad t \gg_i \varphi \mid t :_i \varphi \mid t \gg_i^{U,u} \varphi \mid [U, u]\varphi \\
k & \in \mathbb{N}, \star \in \{\supset, \vee, \wedge, \equiv\}, i \in A
\end{aligned}$$

Crucial requirement: in the above grammar,  $(U, u)$  is a pointed update frame such that for each world  $v \in U$  and each agent  $i \in A$ , the formula  $U^i(v)$  is formed by the above grammar and, in addition, the domain of  $U_i^e$  is finite and each term-formula pair in the domain of  $U_i^e$  is formed by the above grammar.<sup>7</sup> The terms  $c_k$  are called *constants* and the terms  $x_k$  are called *variables*; the constants and variables are together called the *atomic terms*. We use the symbol  $\star$  as a metavariable ranging over binary Boolean connectives. We assign the following informal readings to the key formulas in  $UL^A$ :  $B_i\varphi$  is read, “agent  $i$  believes that  $\varphi$  is true”;  $t \gg_i \varphi$  is read, “ $t$  is agent  $i$ ’s evidence relevant to the assertion that  $\varphi$  is true”;  $t :_i \varphi$  is read, “ $t$  is agent  $i$ ’s evidence that  $\varphi$  is true”;  $t \gg_i^{U,u} \varphi$  is read, “update  $(U, u)$  eliminates agent  $i$ ’s evidence  $t$  relevant to the assertion that  $\varphi$  is true”; and  $[U, u]\varphi$  is read, “after update  $(U, u)$ , we have that  $\varphi$  is true.”

Each of the term-forming operations in the grammar for  $JLce^A$  represents a kind of evidence or a certain logical combination of evidence; let us discuss this now. The constants  $c_k$  play a special role as evidence relevant to the axioms of our theory, and the variables  $x_k$  are used as contingent relevant evidence that may be directly affected by an evidence elimination. The operation  $\cdot_{\varphi}$  takes evidence  $t$  relevant to an implication  $\varphi \supset \psi$  and evidence  $s$  relevant to the antecedent  $\varphi$  and produces the evidence  $t \cdot_{\varphi} s$  relevant to the consequent  $\psi$ , in accord with the rule of Modus Ponens.<sup>8</sup> The operation  $+$  is the union of pieces of relevant evidence:  $t + s$  is relevant evidence for anything for which one or more of  $t$  and  $s$  is relevant evidence. The operation  $!$  (“bang”) is an evidence checker: in case  $t$  is relevant evidence for  $\varphi$ , then  $!t$  checks that  $t$  is evidence for  $\varphi$ . Finally,  $-^{U,u}$  represents an assumption that agents make about the reliability of their own evidence: in case  $t$  is evidence relevant to  $\varphi$ , then  $t^{U,u}$  is evidence relevant to the assertion that  $\varphi$  is true (even) after the occurrence of update  $(U, u)$ . (So the agents trust their relevant evidence so much that they assume that this evidence is not affected by the occurrence of updates.)

Given an update frame  $U$ , we define a simple axiomatic theory called the *U-calculus* in Figure 1. This theory, inspired by Kuznets’  $\ast$ -calculi (Kuznets 2008), will be used to describe

<sup>7</sup>The reason for assuming finiteness is to ensure that it is possible in principle to write out a finite sequence of symbols that describes the pointed update frame  $(U, u)$ . Convention: when a term-formula pair  $(t, \varphi)$  is not in the domain of  $U_i^e$ , then we take this to mean that  $U_i^e(t, \varphi) = \emptyset$ .

<sup>8</sup>For Justification Logic aficionados: in order to properly formulate the effects of eliminations on terms of the form  $t_1 \cdot_{\varphi} t_2$ , we need to keep track of the antecedent  $\varphi$  that was used in forming relevant evidence  $t_1 \cdot_{\varphi} t_2$  for  $\psi$  given that we have evidence  $t_1$  relevant to  $\varphi \supset \psi$  and evidence  $t_2$  relevant to  $\varphi$ . Including the subscript formula  $\varphi$  within the term  $t_1 \cdot_{\varphi} t_2$  allows us to do just this.

SCHEMES

$$\text{EV. } w \vdash x_k \gg_i \varphi \quad \text{if } w \in U_i^e(x_k, \varphi)$$

RULES

$$\begin{array}{c} \frac{w \vdash t \gg_i (\varphi \supset \psi)}{w \vdash (t \cdot_{\varphi} s) \gg_i \psi} \text{ (EAL)} \qquad \frac{w \vdash s \gg_i \varphi}{w \vdash (t \cdot_{\varphi} s) \gg_i \psi} \text{ (EAR)} \\ \\ \frac{w \vdash t \gg_i \varphi \quad w \vdash s \gg_i \varphi}{w \vdash (t + s) \gg_i \varphi} \text{ (ES)} \qquad \frac{w \vdash t \gg_i \varphi}{w \vdash !t \gg_i (t :_i \varphi)} \text{ (EC)} \\ \\ \frac{w' \vdash t \gg_i \varphi \quad wU_i w'}{w \vdash t \gg_i \varphi} \text{ (EM)} \qquad \frac{w \vdash t \gg_i \varphi}{w \vdash t^{U,u} \gg_i [U, u] \varphi} \text{ (EU)} \end{array}$$

Figure 1: The  $U$ -calculus

the effect evidence eliminations have on evidence relevance in the following way: the update  $(U, u)$  eliminates agent  $i$ 's evidence  $t$  relevant to the assertion that  $\varphi$ —meaning that  $t \gg_i \varphi$  will be false after update  $(U, u)$ —if and only if  $u \vdash t \gg_i \varphi$  is derivable in the  $U$ -calculus. We will write  $U, u \vdash t \gg_i \varphi$  to mean that  $u \vdash t \gg_i \varphi$  is derivable in the  $U$ -calculus, and we will write  $U, u \not\vdash t \gg_i \varphi$  to mean that  $u \vdash t \gg_i \varphi$  is not derivable in the  $U$ -calculus. It is not hard to see that either  $U, u \vdash t \gg_i \varphi$  or else  $U, u \not\vdash t \gg_i \varphi$ , and determining which of these is the case is a decidable question.

To state the axiomatics, we define a notion of *composition* for update frames that allows us to take update frames  $U = (W, R, f, \mathcal{A})$  and  $U' = (W', R', f', \mathcal{A}')$  and build an update frame  $U \circ U'$  such that executing the update  $(U, u)$  and then executing the update  $(U', u')$  has the same combined effect as executing the single update  $(U \circ U', (u, u'))$ . To define the *composition*  $U \circ U' := (W^\circ, R^\circ, f^\circ, \mathcal{A}^\circ)$ , we use standard definitions from Dynamic Epistemic Logic (Baltag, van Ditmarsch, and Moss 2008; Renne 2008b; van Ditmarsch, van der Hoek, and Kooi 2007) for the first three components: set  $W^\circ := W \times W'$ , allow  $(v, v')R_i^\circ(w, w')$  if and only if  $vR_i w$  and  $v'R'_i w'$ , and set  $f^\circ(v, v') := \neg[U, v] \neg U'(v')$ . To define the final component  $\mathcal{A}^\circ$ , we allow  $(v, v') \in \mathcal{A}_i^\circ(t, \varphi)$  if and only if  $v \in \mathcal{A}(t, \varphi)$  or  $v' \in \mathcal{A}'(t, \varphi)$ . It is not hard to see that  $(U \circ U', (u, u'))$  eliminates agent  $i$ 's evidence  $t$  relevant to the assertion that  $\varphi$  if and only if at least one of  $(U, u)$  or  $(U', u')$  eliminates  $i$ 's evidence  $t$  relevant to the assertion that  $\varphi$ ; that is, we have  $U \circ U', (u, u') \vdash t \gg_i \varphi$  if and only if  $U, u \vdash t \gg_i \varphi$  or  $U', u' \vdash t \gg_i \varphi$ .

Finally, we state in two parts the axiomatics of our *theory of Justification Logic with multi-agent communication and evidence elimination*, written  $\text{JLce}^A$ . The first part is the axiomatic theory  $\text{AX}^A$  (Figure 2), and the second part is our theory  $\text{JLce}^A$  itself (Figure 3).<sup>9</sup>

<sup>9</sup>We observe two points. First, in the statement of Axiom E (Figure 2), we have offloaded to the  $U$ -calculus the work of determining whether  $t \gg_i^{U,u} \varphi$  should be provably equivalent to  $\top$  or provably equivalent to  $\perp$ . In this way, the formula  $t \gg_i^{U,u} \varphi$  expresses whether update  $(U, u)$  eliminates  $i$ 's evidence  $t$  relevant to  $\varphi$

Terminology: an  $\text{AX}^A$ -theorem is any formula derivable using only the axioms and rules found in Figure 2; similarly, a  $\text{JLce}^A$ -theorem is any formula derivable using only the rules found in Figure 3, though one should take care to note that Rule AX of  $\text{JLce}^A$  (Figure 3) says that every  $\text{AX}^A$ -theorem is also a  $\text{JLce}^A$ -theorem.<sup>10</sup>

Artemov (Artemov 2001) first identified a key result of Justification Logics called the property of *Internalization*. Internalization describes the way in which a Justification Logic allow us to describe proofs of the theory using terms, which bolsters our reading of terms as pieces of evidence for the formulas that they label. We formulate Internalization for  $\text{JLce}^A$  as follows.

**Theorem 3.1** (Artemov’s Internalization Theorem; Artemov 2001; Renne 2009a). If  $\text{JLce}^A \vdash \varphi$ , then for each  $i \in \mathcal{A}$ , there is a term  $t \in \mathcal{T}(\text{UL}^A)$  such that  $\text{JLce}^A \vdash t :_i \varphi$ .

We observe that our formulation of the  $U$ -calculus in Figure 1 indicates our intention to restrict our notion of evidence elimination in such a way that whenever update  $(U, u)$  eliminates agent  $i$ ’s evidence  $t$  relevant to  $\varphi$ —thereby falsifying  $t \gg_i \varphi$ —this elimination can be traced back to an elimination of the form  $x_k \gg_i \psi$ . We have imposed this restriction so as to guarantee that constants  $c_k$  continue to act as evidence for  $\text{AX}^A$ -axioms (see Rule CN of  $\text{AX}^A$ ), a requirement for our proof of Theorem 3.1 to go through (Renne 2009a). It might be possible to elegantly loosen this restriction so as to allow the elimination of constants  $c_k$  evidencing non- $\text{AX}^A$ -axioms or to allow the direct elimination of non-atomic terms, but we leave these issues to future work.

---

in accord with the  $U$ -calculus, a theory we defined to determine exactly when such an elimination should take place. But note that offloading the work in Axiom E to the  $U$ -calculus is not strictly necessary; in fact, it is possible to embed the  $U$ -calculus into the theory  $\text{JLce}^A$  by writing equivalences  $(t \gg_i^{U,u} \varphi) \equiv \top$  or  $(t \gg_i^{U,u} \varphi) \equiv \perp$  for each of the possible forms for term  $t$  according to whether  $U, u \vdash t \gg_i \varphi$ . See Renne (2009b) for details on how this is done in a theory of simple evidence elimination for Justification Logic. The reason we have offloaded the work in this way in this paper is to save space and to simplify the presentation of the axiomatics.

Second, we have assumed that agent belief is governed by the modal theory K4. While other choices are possible in Justification Logic (Kuznets 2008; Renne 2008a), they would introduce complications here that would distract us from the key issues of this paper: multi-agent communication and evidence elimination.

<sup>10</sup>The role of the theory  $\text{AX}^A$  is to cleanly state the axioms of our theory  $\text{JLce}^A$ . In particular, the axioms of  $\text{JLce}^A$  consist of instances of the axiom schemes from Figure 2 and, in addition, for each instance  $\varphi$  of an axiom scheme from Figure 2, each natural number  $n \in \mathbb{N}$ , each sequence  $\{k_j\}_{j=0}^n$  of natural numbers, and each sequence  $\{i_j\}_{j=0}^n$  of agents in  $A$ , the formula

$$c_{k_0} :_{i_0} (c_{k_1} :_{i_1} (c_{k_2} :_{i_2} (\cdots :_{i_n} (c_{k_n} :_{i_n} \varphi) \cdots)))$$

is also an axiom of  $\text{JLce}^A$ . The reader can easily see that this collection of axioms is just the set of  $\text{AX}^A$ -theorems (Figure 2). So  $\text{AX}^A$  describes the “axioms” of the theory  $\text{JLce}^A$ , which then closes these “axioms” under the rules (from Figure 3) of Modus Ponens (MP), Belief Necessitation (BN), and Update Necessitation (UN). Thus we axiomatize  $\text{JLce}^A$  as in Figure 3: Rule AX includes the  $\text{AX}^A$ -theorems as the “axioms” of  $\text{JLce}^A$ , and then these “axioms” are closed under Rules MP, BN, and UN.

BELIEF/EVIDENCE SCHEMES	ELIMINATION/UPDATE SCHEMES
CL. Classical Propositional Logic	ET. $(t \gg_i^{U,u} \varphi) \equiv \top$ if $U, u \vdash t \gg_i \varphi$ (Figure 1)
BK. $B_i(\varphi \supset \psi) \supset (B_i\varphi \supset B_i\psi)$	E $\perp$ . $(t \gg_i^{U,u} \varphi) \equiv \perp$ if $U, u \not\vdash t \gg_i \varphi$ (Figure 1)
B4. $B_i\varphi \supset B_i(B_i\varphi)$	UA. $[U, u]q \equiv U^l(u) \supset q$ , for $q \in \{p_k, \perp, \top\}$
EA. $t \gg_i(\varphi \supset \psi) \supset$ $(s \gg_i \varphi \supset (t \cdot_\varphi s) \gg_i \psi)$	U $\star$ . $[U, u](\varphi \star \psi) \equiv [U, u]\varphi \star [U, u]\psi$
ES. $t \gg_i \varphi \supset (t + s) \gg_i \varphi$ $s \gg_i \varphi \supset (t + s) \gg_i \varphi$	UN. $[U, u]\neg\varphi \equiv U^l(u) \supset \neg[U, u]\varphi$
EC. $t \gg_i \varphi \supset !t \gg_i(t :_i \varphi)$	UB. $[U, u]B_i\varphi \equiv U^l(u) \supset \bigwedge_{uU_i v} B_i[U, v]\varphi$
EM. $t \gg_i \varphi \supset B_i(t \gg_i \varphi)$	UR. $[U, u](t \gg_i \varphi) \equiv$ $U^l(u) \supset ((t \gg_i \varphi) \wedge \neg(t \gg_i^{U,u} \varphi))$
EU. $t \gg_i \varphi \supset t^{U,u} \gg_i [U, u]\varphi$	UE. $[U, u](t \gg_i^{U',u'} \varphi) \equiv U^l(u) \supset (t \gg_i^{U',u'} \varphi)$
C. $t :_i \varphi \equiv (t \gg_i \varphi \wedge B_i\varphi)$	UU. $[U, u][U', u']\varphi \equiv [U \circ U', (u, u')]\varphi$

RULES

$$\frac{k \in \mathbb{N} \quad i \in A \quad \vdash \varphi \quad (\text{CN})}{\vdash c_k :_i \varphi}$$

Figure 2: The theory  $\text{AX}^A$

RULES

$$\frac{\text{AX}^A \vdash \varphi}{\vdash \varphi} \quad (\text{AX}) \qquad \frac{\vdash \varphi \supset \psi \quad \vdash \varphi}{\vdash \psi} \quad (\text{MP})$$

$$\frac{\vdash \varphi}{\vdash B_i\varphi} \quad (\text{BN}) \qquad \frac{\vdash \varphi}{\vdash [U, u]\varphi} \quad (\text{UN})$$

Figure 3: The theory  $\text{JLce}^A$

## 4 Semantics

The semantics of  $\text{UL}^A$  is our adaptation of a Kripke-style semantics for Justification Logic due to Fitting (Fitting 2005) and Mkrtychev (Mkrtychev 1997). A *Fitting model (for  $\text{UL}^A$ )* is a tuple  $M = (W, R, V, \mathcal{A})$  consisting of a transitive Kripke model  $(W, R, V)$  that is said to be *underlying*  $M$  and an evidence-labeling function  $\mathcal{A} : A \rightarrow (\mathcal{T}(\text{UL}^A) \times \text{UL}^A) \rightarrow 2^W$  satisfying each of the following properties: *Constant Specification*  $S$ : for each  $k \in \mathbb{N}$  and each  $\text{AX}^A$ -theorem  $\varphi$  (Figure 2), we have  $\mathcal{A}_i(c_k, \varphi) = W$ ; *Application*:  $\mathcal{A}_i(t, \varphi \supset \psi) \cap \mathcal{A}_i(s, \varphi) \subseteq \mathcal{A}_i(t \cdot_\varphi s, \psi)$ ; *Sum*:  $\mathcal{A}_i(t, \varphi) \cup \mathcal{A}_i(s, \varphi) \subseteq \mathcal{A}_i(t + s, \varphi)$ ; *Checker*:  $\mathcal{A}_i(t, \varphi) \subseteq \mathcal{A}_i(!t, t :_i \varphi)$ ; *Monotonicity*:  $\Gamma R_i \Delta$  and  $\Gamma \in \mathcal{A}_i(t, \varphi)$  together imply that  $\Delta \in \mathcal{A}_i(t, \varphi)$ ; and *Update*:  $\mathcal{A}_i(t, \varphi) \subseteq \mathcal{A}_i(t^{U,u}, [U, u]\varphi)$ . For a Fitting model  $F$ , we write  $\Gamma \in M$  to mean that  $\Gamma$  is a world in the Kripke model underlying  $M$ . A *pointed Fitting model* for  $\text{UL}^A$  is a pair  $(M, \Gamma)$  consisting of a Fitting model  $M$  and a world  $\Gamma \in M$ . The notion of truth for  $\text{UL}^A$ -formulas at pointed Fitting models is given by an induction on  $\text{UL}^A$ -formula construction in which Boolean cases are handled as is usual in modal logic (Blackburn, de Rijke, and Venema 2001; Renne 2009a); the key inductive cases for  $\text{UL}^A$ -formulas are as follows.

- $M, \Gamma \models B_i \varphi$  means that  $M, \Delta \models \varphi$  for each  $\Delta \in M$  such that  $\Gamma R_i \Delta$ .
- $M, \Gamma \models t \gg_i \varphi$  means that  $\Gamma \in \mathcal{A}_i(t, \varphi)$ .
- $M, \Gamma \models t :_i \varphi$  means that  $\Gamma \in \mathcal{A}_i(t, \varphi)$  and  $M, \Delta \models \varphi$  for each  $\Delta \in M$  such that  $\Gamma R_i \Delta$ .
- $M, \Gamma \models t \gg_i^{U,u} \varphi$  means that  $U, u \vdash t \gg_i \varphi$ . (See Figure 1.)
- $M, \Gamma \models [U, u]\varphi$  means that either  $M, \Gamma \not\models U^l(u)$  or else both  $M, \Gamma \models U^l(u)$  and  $M^U, (\Gamma, u) \models \varphi$ , where the components of the tuple

$$M^U := (W^U, R^U, \mathcal{A}^U, V^U)$$

are given as follows.

$$W^U := \{(\Delta, v) \in (W \times U) : M, \Delta \models U^l(v)\}.$$

$$R_i^U := \{((\Delta, v), (\Delta', v')) : \Delta R_i \Delta' \ \& \ v U_i v'\}.$$

$$\mathcal{A}_i^U(t, \psi) := \{(\Delta, v) : \Delta \in \mathcal{A}_i(t, \psi) \ \& \ U, v \not\models t \gg_i \psi\}.$$

$$V^U(p_k) := \{(\Delta, v) : \Delta \in V(p_k)\}.$$

If  $M$  is a Fitting model, then  $M^U$  is also a Fitting model. To say that a  $\text{UL}^A$ -formula  $\varphi$  is *valid*, written  $\models \varphi$ , means that  $\varphi$  is true at all pointed Fitting models.

The inductive cases for formulas of the form  $t \gg_i^{U,u} \varphi$  and of the form  $[U, u]\varphi$  each delegate all or part of their work to the  $U$ -calculus (Figure 1). This may seem strange because we are admitting the  $U$ -calculus—a syntactic notion—into our semantics. However, since it is our intention for evidence eliminations to respect the intended meanings of the term-forming operations (described earlier in the section on syntax)—this so that the elimination of one or more parts of a combination  $t$  of multiple pieces of evidence shall affect  $t$  itself—some simple theory describing the logical consequences that the elimination of a simple term has on more

complex terms is a necessary part of the semantics. One may also take some comfort in the fact that the  $U$ -calculus is a simple, decidable theory.

**Theorem 4.1** (Renne 2009a).  $\text{JLce}^A \vdash \varphi$  if and only if  $\models \varphi$ .

## 5 Formalized Example

We now use  $\text{JLce}^A$  to formalize our email example from the beginning of this paper. To begin, we let  $J$  and  $M$  denote propositional letters, and we write  $J \supset M$  (“job” implies “moves”) to describe Anne’s conditional moving plan. Email messages  $x_1$  through  $x_4$  provide us with our initial setup  $X$ , a conjunction of the formulas  $x_1 :_B (J \supset M)$  (“ $x_1$  is Bryan’s evidence that  $J \supset M$ ”),  $x_1 :_C (J \supset M)$  (“ $x_1$  is Charlie’s evidence that  $J \supset M$ ”),  $x_2 :_B (x_1 :_C (J \supset M))$  (“ $x_2$  is Bryan’s evidence that  $x_1 :_C (J \supset M)$ ”),  $x_3 :_B J$  (“ $x_3$  is Bryan’s evidence that  $J$ ”),  $x_3 :_C J$  (“ $x_3$  is Charlie’s evidence that  $J$ ”), and  $x_4 :_B (x_3 :_C J)$  (“ $x_4$  is Bryan’s evidence that  $x_3 :_C J$ ”). It is not difficult to see that  $\text{JLce}^A \vdash X \supset (x_1 \cdot_J x_3) :_B M$  and  $\text{JLce}^A \vdash X \supset (x_1 \cdot_J x_3) :_C M$ ; that is, given our setup  $X$ , each of Bryan and Charlie has evidence  $x_1 \cdot_J x_3$  that Anne is moving to Columbus.

Let us now look at Bryan’s evidence about Charlie’s evidence. First, we have that  $\text{JLce}^A \vdash x_1 :_C (J \supset M) \supset (x_3 :_C J \supset (x_1 \cdot_J x_3) :_C M)$ , from which it follows by Artemov’s Internationalization Theorem (Theorem 3.1) that there is a term  $t$  such that

$$t :_B (x_1 :_C (J \supset M) \supset (x_3 :_C J \supset (x_1 \cdot_J x_3) :_C M))$$

is a  $\text{JLce}^A$ -theorem. Reasoning in  $\text{JLce}^A$ , we have  $\text{JLce}^A \vdash X \supset s :_B ((x_1 \cdot_J x_3) :_C M)$ —that is, “given setup  $X$ , Bryan has evidence  $s$  that Charlie has evidence  $x_1 \cdot_J x_3$  that  $M$ ”—where  $s$  is the term

$$(t \cdot_{(x_1 :_C (J \supset M))} x_2) \cdot_{(x_3 :_C J)} x_4 \cdot$$

So from the initial situation  $X$  given by messages  $x_1$  through  $x_4$ , each of Bryan and Charlie has evidence  $x_1 \cdot_J x_3$  that Anne is moving to Columbus, and Bryan has evidence  $s$  that Charlie has evidence  $x_1 \cdot_J x_3$  that Anne is moving to Columbus.

Now let us examine how Anne’s private email  $x_5$  to Bryan affects Bryan’s evidence. Adapting the Dynamic Epistemic Logic definition of the *private announcement to an agent* (described in the introduction; see also: Baltag and Moss (2004), Renne (2008b), van Ditmarsch, van der Hoek, and Kooi (2007)), we define an update frame  $\text{PRI}_i^{(k,\varphi)} := (W, R, f, \mathcal{A})$ , called the *private elimination of agent  $i$ ’s evidence  $x_k$  relevant to  $\varphi$* , according to the following:  $W := \{u, v\}$ ,  $R_i := \{(u, u), (v, v)\}$ ,  $R_j := \{(u, v), (v, v)\}$  if  $j \neq i$ ,  $f(u) := \top$ ,  $f(v) := \top$ ,  $\mathcal{A}_i(x_k, \varphi) := \{u\}$ ,  $\mathcal{A}_j(x_l, \psi) := \emptyset$  if  $j \neq i$  or  $(x_l, \psi) \neq (x_k, \varphi)$ . We picture  $\text{PRI}_i^{(k,\varphi)}$  in Figure 4. We will use the update  $(\text{PRI}_B^{(3,J)}, u)$  to represent the effect of Anne’s final message  $x_5$ .

We have that  $\text{PRI}_B^{(3,J)}, u \vdash x_3 \gg_B J$  by Axiom V of the  $\text{PRI}_B^{(3,J)}$ -calculus (Figure 1), from which it follows that  $\text{PRI}_B^{(3,J)}, u \vdash (x_1 \cdot_J x_3) \gg_B M$  by Rule EAR. We therefore have

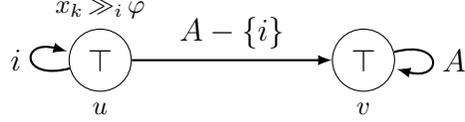


Figure 4: The update frame  $\text{PRI}_i^{(k,\varphi)}$ , the private elimination of agent  $i$ 's evidence  $x_k$  relevant to  $\varphi$

that  $\text{JLce}^A \vdash (x_1 \cdot_J x_3) \gg_B^{\text{PRI}_B^{(3,J)},u} M$  by Axiom  $\text{E}\top$  (Figure 2) and propositional reasoning, from which it follows that  $\text{JLce}^A \vdash [\text{PRI}_B^{(3,J)},u] \neg((x_1 \cdot_J x_3) \gg_B M)$  by Axioms  $\text{UR}$  and  $\text{UN}$  (Figure 2) and propositional reasoning. Applying Axiom  $\text{C}$  (Figure 2) and modal reasoning, it follows that  $\text{JLce}^A \vdash [\text{PRI}_B^{(3,J)},u] \neg((x_1 \cdot_J x_3) :_B M)$ . This is, after Anne's private message to Bryan eliminating Bryan's evidence  $x_3$  relevant to  $J$ , it is not the case that  $x_1 \cdot_J x_3$  is Bryan's evidence that  $J$  (Anne gets the job). And yet we have that  $\text{PRI}_B^{(3,J)},u \not\vdash s \gg_B ((x_1 \cdot_J x_3) :_C M)$  and thus that  $\text{JLce}^A \vdash X \supset [\text{PRI}_B^{(3,J)},u] s :_B ((x_1 \cdot_J x_3) :_C M)$  by Axioms  $\text{E}\perp$  and  $\text{UR}$  (Figure 2) and propositional reasoning. In words: given setup  $X$ , after Anne's private message to Bryan eliminating Bryan's evidence  $x_3$  relevant to  $J$ , it is (still) the case that Bryan has evidence  $s$  that Charlie has evidence  $x_1 \cdot_J x_3$  that Anne is moving, despite the fact that Bryan himself does not consider  $x_1 \cdot_J x_3$  evidence that Anne is moving.

## 6 Conclusion

The work in this paper is part of a larger project, *Dynamic Justification Logic*, whose aim is to combine the frameworks of Dynamic Epistemic Logic and Justification Logic in order to reason about belief and evidence dynamics arising from multi-agent communications. The role of this paper in the project is to introduce *multi-agent evidence elimination*, one kind of dynamic operation on multi-agent evidence that causes an agent to set aside a piece of evidence relevant to a given assertion and then determine how this affects other pieces of evidence as a result of the ways in which evidence may be logically combined in Justification Logic. In earlier work (Renne 2008a), we studied a notion of *evidence introduction*, whereby a piece of evidence may be introduced as relevant for a given assertion; however, the particular operation we defined there lacks the tight integration with the semantics of Dynamic Epistemic Logic that we found for evidence elimination in this paper, and so the techniques developed here might well be adapted to a notion of *multi-agent evidence introduction* as well. Beyond investigating these two operations on evidence, there are a number of natural directions in which we might proceed, including the introduction of *preferences and preference change* over evidence, which would allow us to tie an agent's beliefs (or the relative strength among his or her various beliefs) to his or her preference ordering on evidence. This work would naturally dovetail and complement recent work in Dynamic Epistemic Logic on *preference upgrade* and *Belief Revision* (Baltag and Smets 2007; van Benthem 2004; van Benthem and Liu 2007), furthering our general view that the tools and techniques of Dy-

namic Epistemic Logic and Justification Logic can work together to provide a fuller picture of belief and evidence dynamics for groups of interacting rational individuals.

## References

- Artemov, S. (2008). The logic of justification. *The Review of Symbolic Logic* 1(4), 477–513.
- Artemov, S. and E. Nogina (2005). Introducing justification into epistemic logic. *Journal of Logic and Computation* 15(6), 1059–1073.
- Artemov, S. N. (2001). Explicit provability and constructive semantics. *The Bulletin of Symbolic Logic* 7(1), 1–36.
- Baltag, A. and L. S. Moss (2004). Logics for epistemic programs. *Synthese* 139(2), 165–224.
- Baltag, A., L. S. Moss, and S. Solecki (1998). The logic of common knowledge, public announcements, and private suspicions. In I. Gilboa (Ed.), *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK VII)*, Evanston, IL, USA, pp. 43–56.
- Baltag, A. and S. Smets (2007). A qualitative theory of dynamic interactive belief revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge (Eds.), *Selected papers from LOFT’06*, Texts in Logic and Games. Amsterdam University Press.
- Baltag, A., H. P. van Ditmarsch, and L. S. Moss (2008). Epistemic logic and information update. In P. Adriaans and J. van Benthem (Eds.), *Handbook on the Philosophy of Information*, pp. 369–463. Elsevier.
- Blackburn, P., M. de Rijke, and Y. Venema (2001). *Modal Logic*. Cambridge University Press.
- Fitting, M. (2005). The Logic of Proofs, semantically. *Annals of Pure and Applied Logic* 132(1), 1–25.
- Fitting, M. (2009). Reasoning with justifications. Forthcoming in *Studia Logica*.
- Kuznets, R. (2008). *Complexity Issues in Justification Logic*. Ph. D. thesis, CUNY Ph.D. Program in Computer Science.
- Mkrtychev, A. (1997). Models for the Logic of Proofs. In S. Adian and A. Nerode (Eds.), *Logical Foundations of Computer Science*, Volume 1234 of *Lecture Notes in Computer Science*, pp. 266–275. Springer.
- Renne, B. (2006). Public and private communication are different: Results on relative expressivity. *Synthese* 165(2), 225–245.

- Renne, B. (2008a). *Dynamic Epistemic Logic with Justification*. Ph. D. thesis, The City University of New York.
- Renne, B. (2008b, July). A survey of Dynamic Epistemic Logic. Manuscript.
- Renne, B. (2009a, March). Multi-agent justification logic: Communication and evidence elimination. Manuscript.
- Renne, B. (2009b, March). Simple evidence elimination in justification logic. Manuscript.
- van Benthem, J. (2004). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics* 14(2), 129–155.
- van Benthem, J. (2006). One is a lonely number: Logic and communication. In Z. Chatzidakis, P. Koepke, and W. Pohlers (Eds.), *Logic Colloquium '02*, Volume 27 of *Lecture Notes in Logic*. Association for Symbolic Logic.
- van Benthem, J. and F. Liu (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics* 17(2), 157–182.
- van Benthem, J., J. van Eijck, and B. Kooi (2006). Logics of communication and change. *Information and Computation* 204(11), 1620–1662.
- van Ditmarsch, H., W. van der Hoek, and B. Kooi (2007). *Dynamic Epistemic Logic*. Springer.
- Yavorskaya (Sidon), T. (2008). Interacting explicit evidence systems. *Theory of Computing Systems* 43, 272–293.