Dynamic Epistemic Logic I: Modeling Knowledge and Belief

Eric Pacuit*
University of Maryland, College Park

Abstract

Dynamic epistemic logic, broadly conceived, is the study of logics of *information change*. This is the first paper in a two–part series introducing this research area. In this paper, I introduce the basic logical systems for reasoning about the knowledge and beliefs of a group of agents.

1. Introduction

The first to propose a (modal) logic of knowledge and belief was Jaako Hintikka in his seminal book *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, published in 1962. However, the general study of formal semantics for knowledge and belief (and their logic) really began to flourish in the 1990s with fundamental contributions from computer scientists (Fagin et al. 1995; Meyer and van der Hoek 1995) and game theorists (Aumann 1999; Bonanno and Battigalli 1999). As a result, the field of Epistemic Logic developed into an interdisciplinary area focused on explicating epistemic issues in, for example, game theory (Brandenburger 2007), economics (Samuelson 2004), computer security (Halpern and Pucella 2003; Ramanujam and Suresh 2005), distributed and multiagent systems (Halpern and Moses 1990; van der Hoek and Wooldridge 2003), and the social sciences (Parikh 2002; Gintis 2009). Nonetheless, the field has not loss touch with its philosophical roots: See (Holliday 2012; Egré 2011; Stalnaker 2006; van Benthem 2006a; and Sorensen 2002) for logical analyses aimed at mainstream epistemology.

Inspired, in part, by issues in these different 'application' areas, a rich repertoire of epistemic and doxastic attitudes have been identified and analyzed in the epistemic logic literature. The challenge for a logician is not to argue that one particular account of belief or knowledge is *primary*, but, rather, to explore the logical space of definitions and identify interesting relationships between the different notions. I do not have the space for a comprehensive overview of the many different logics of knowledge and belief. Instead, I will confine the discussion to three logical frameworks needed for the survey of *dynamic* logics of knowledge and belief found in the second paper.

The formal models introduced in this article can be broadly described as 'possible worlds models', familiar in much of the philosophical logic literature. Let $\mathcal{A}$ be a finite set of agents and At a (finite or countable) set of atomic sentences. Elements $p \in$ At are intended to describe *basic properties* of the situation being modeled, such as 'it is raining' or 'the red card is on the table'. Setting aside any conceptual difficulties surrounding the use of 'possible worlds',[1] a nonempty set W of *states*, or *possible worlds*, will be used to represent different possible *scenarios* or *states of affairs*. A *valuation function* associates with each atomic proposition p a set of states where p is true: $V : $ At $\rightarrow \wp(W)$.

In this article, I will introduce different logical systems that have been used to reason about the knowledge and beliefs of a group of agents. My focus is on the underlying intuitions

and a number of illustrative examples rather than the general logical theory of these systems (e.g., issues of decidability, complexity, completeness and definability). Although I will include pointers to this more technical material throughout this article, interested readers are encouraged to consult two textbooks on modal logic (Blackburn et al. 2002; van Benthem 2010) for a systematic presentation.

2. Knowledge

The model in this section is based on a very simple (and ubiquitous) idea: An agent is *informed* that φ is true when φ is true throughout the agent's current range of possibilities. In an epistemic model, the agents' 'range of possibilities' are described in terms of *epistemic accessibility relations* R_i on the set of states W (i.e., $R_i \subseteq W \times W$). That is, given a relation R_i on W, the set $R_i(w) = \{v \mid wR_iv\}$ is agent i's 'range of possibilities' at state w.

Definition 2.1 (Epistemic Model) Let $\mathcal{A}$ be a finite set of agents and At a (finite or countable) set of atomic propositions. An *epistemic model* is a tuple $\langle W, \{R_i\}_{i \in \mathcal{A}}, V \rangle$ where $W \neq \varnothing$, for each $i \in \mathcal{A}$, $R_i \subseteq W \times W$ is a relation, and $V : \text{At} \to \wp(W)$ is a valuation function.

Epistemic models are used to describe what the agents *know* about the situation being modeled. A simple propositional modal language is often used to make this precise: Let $\mathcal{L}_K$ be the set of sentences generated by the following grammar:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_i\varphi$$

where $p \in \text{At}$ (the set of atomic propositions). The additional propositional connectives $(\to, \leftrightarrow, \vee)$ are defined as usual, and the dual of K_i, denoted L_i, is $\neg K_i \neg \varphi$. Following the standard usage in the epistemic logic and game theory literature, the intended interpretation of $K_i\varphi$ is 'agent i knows that φ'. An alternative interpretation (which is more natural in many situations) is 'agent i is informed that φ is true'.

Each state of an epistemic model represents a possible scenario that can be described in the formal language given above: If $\varphi \in \mathcal{L}_K$, then $\mathcal{M}, w \vDash \varphi$ means φ is *true* in the situation represented by w. This can be made precise as follows:

Definition 2.2 (Truth for $\mathcal{L}_K$) Let $\mathcal{M} = \langle W, \{R_i\}_{i \in \mathcal{A}}, V \rangle$ be an epistemic model. For each $w \in W$, φ *is true at state* w, denoted $\mathcal{M}, w \vDash \varphi$, is defined by induction on the structure of φ:

- $\mathcal{M}, w \vDash p$ iff $w \in V(p)$
- $\mathcal{M}, w \vDash \neg\varphi$ iff $\mathcal{M}, w \nvDash \varphi$
- $\mathcal{M}, w \vDash \varphi \wedge \psi$ iff $\mathcal{M}, w \vDash \varphi$ and $\mathcal{M}, w \vDash \psi$
- $\mathcal{M}, w \vDash K_i\varphi$ iff for all $v \in W$, if wR_iv then $\mathcal{M}, v \vDash \varphi$

It is important to recognize that $K_i\varphi$ *defines* a set of states in which it is true that agent i 'knows' that φ (or agent i is 'informed' that φ). It does not *explain* why or how agent i came to the conclusion that φ is true. Indeed, there are different ways to understand exactly how these models represent the agents' *knowledge* (and later, beliefs). The crucial

interpretative step is to explain what it means for a state v to be accessible for agent i from state w.

The first, and most neutral, interpretation of wR_iv is *everything that agent i knows in state w is true in state v*. Under this interpretation, the agents' knowledge is *not* defined in terms of more primitive notions. Instead, an epistemic model represents the 'implicit consequences' of what the agents are assumed to know in the situation being modeled.

A second use of epistemic models is to formalize a substantive theory of knowledge. In this case, the agents' knowledge is *defined* in terms of more primitive concepts (which are built into the definition of the accessibility relation). For instance, suppose that wR_iv means that *agent i has the same experiences and memories in both w and v* (Lewis 1996). Another example is wR_iv means that *at state w, agent i cannot rule out state v (according to i's current observations and any other evidence she has obtained)*. In both cases, what the agents *know* is defined in terms of more primitive notions (in terms of 'sameness of experience and memory' or a 'ruling–out' operation).

Finally, epistemic models have been used to make precise informal notions of knowledge found in both the computer science and game theory literature. For example, consider a system of processors executing some distributed program. When designing such a distributed program, it is natural to use statements of the form 'if processor i *knows* that the message was delivered, then . . .'. The notion of knowledge being used here can be given a concrete interpretation as follows: A possible world, or global state, is a complete description of the system at a fixed moment. At each possible world, processors are associated with a *local view*, defined in terms of the values of the variables that the processor has access to. Then, a global state s is accessible for processor i from another global state s' provided i has the same local view in both s and s'. A similar approach can be used to analyze game-theoretic situations. For example, in a poker game, the states are the different distribution of cards. Then, a state v is accessible for player i from state w provided i has the same *information* in both states (eg., i was dealt the same cards in both states). In both cases, the goal is to develop a useful model of knowledge based on a concrete definition of possible worlds and the epistemic accessibility relation rather than a general analysis of 'knowledge'.

The issues raised above are important conceptual and methodological considerations, and they help us understand the scope of an epistemic analysis using the logical system introduced in this section. However, the general distinctions should not be overstated as they tend to fade when analyzing specific examples. Consider the following running example: Suppose that there are two coins each sitting in different drawers and two agents, Ann (a) and Bob (b). We are interested in describing what Ann and Bob know about the two coins. Of course, there are many facts about the coins that we may want to represent (eg., the exact position of the coins in the drawer, what type of coins are in the drawers: Are they nickels? Are they quarters?, etc.), but to keep things simple consider four atomic propositions: For $i = 1, 2$, let H_i denote 'the coin in drawer i is facing heads up' and T_i denote 'the coin in drawer i is facing tails up'. Suppose that Ann looked at the coin in drawer 1 and Bob looked at the coin in drawer 2. Given what Ann and Bob have observed, we can represent their knowledge about the coins in the diagram below: Suppose that both coins are facing heads up (so w_1 is the actual state). We draw an edge labeled with a (*respectively b*) between two states if a (*respectively b*) cannot distinguish them (based on their observations). Of course, each state cannot be distinguished from itself; however, in order to minimize the clutter in the diagram, I do not draw these reflexive edges.
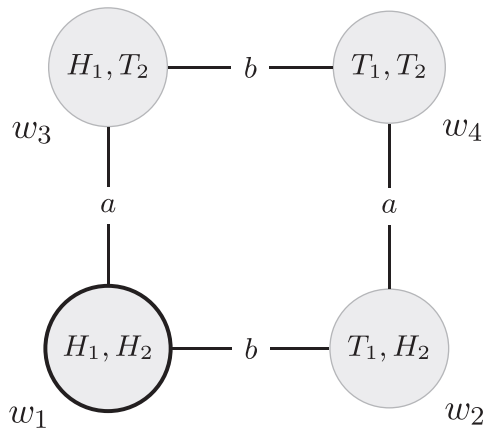
The reader is invited to verify that $K_a H_1 \wedge K_b H_2$ is true at state w_1, as expected. However, much more is true at state w_1. In particular, while Ann does not know which way the coin is facing in the second drawer, she does know that Bob knows whether the coin is facing heads up or tails up (i.e., $K_a(K_b H_2 \vee K_b T_2)$ is true at w_1). Thus, states in an epistemic model also describe the agents' *higher order knowledge*: the information about what the other agents know about each other. The implicit assumption underlying the above model is that the agents correctly observe the face of the coin when they look in the drawer *and* they take it for granted that the other agent correctly perceived the other coin. This is a substantive assumption about what the agents know about each other and can be dropped by adding states to the model:
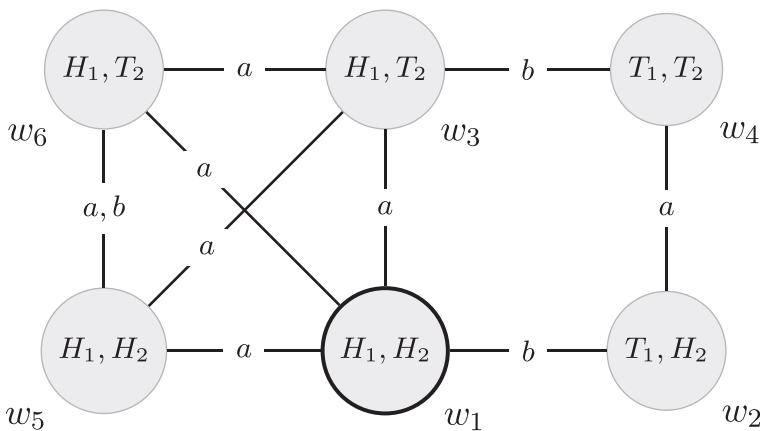
Now, at state w_1 in the above model, Ann considers it possible that Bob does not know the position of the coin in the second drawer (so, $K_a(K_b H_2 \vee K_b T_2)$ is not true at w_1). Of course, there are other implicit assumptions built into the modified model. Indeed, it is not hard to see that one always finds implicit assumptions in a finite epistemic model. This raises an intriguing question: Are there epistemic models that make no, or at least as few as possible, implicit assumptions about what each agent knows about the other agents' knowledge? This question has been extensively discussed in the literature on the epistemic foundations of game theory – see the discussion and references in Samuelson (2004) and Roy and Pacuit (2011) for more on this topic.

In the above example, the agents' accessibility relation satisfied a number of additional properties. In particular, the relations are reflexive (for each w, wR_iw), transitive (for each w, v, s if wR_iv and vR_is then wR_is) and Euclidean (for each w, v, s if wR_iv and wR_is then vR_is). It is not hard to see that relations that are reflexive and Euclidean must also be symmetric (for all w, v if wR_iv then vR_iw). Relations that are reflexive, transitive and symmetric are called *equivalence relations*. Assuming that the agents' accessibility relations are equivalence relations (in the remainder of this article and in part II, we use the notation $\sim_i$ instead of R_i when the accessibility relations are equivalence relations) is a substantive assumption about the nature of the informational attitude being represented in these models. In fact, by using standard techniques from the mathematical theory of modal logic, I can be much more precise about what properties of knowledge are being assumed. In particular, *modal correspondence theory* rigorously relates properties of the relation in an epistemic model with modal formulas (cf. Blackburn et al. 2002, Chapter 3).[2] Table 1 lists some key formulas in the language $\mathcal{L}_K$ with their corresponding property.

Viewed as a description, even an idealized one, of *knowledge*, the properties in Table 1 have drawn many criticisms. The underlying assumption of *logical omniscience*, witnessed by the validity of the first axiom (each agent's knowledge is closed under known logical consequences) and the validity of the inference rule, called *necessitation, from φ infer $K_i\varphi$* (each agent knows all validities),[3] has generated the most extensive criticisms (Stalnaker 1991) and responses (Halpern and Pucella 2011). Furthermore, the two introspection principles have also been the subject of intense discussion (cf. Williamson 2000; Egré and Bonnay 2009). These discussions are fundamental to the theory of knowledge and its formalization, but here I choose to bracket them, and, instead, take epistemic models for what they are: precise descriptions of what the (modeler takes the) agents (to) know about the situation being modeled.

3. Adding Belief

A characteristic feature of a rational agent is her *beliefs*, and the ability to correct them when they turn out to be wrong. There are many different formal representations of beliefs – see (Huber 2011) and (Halpern 2003) for overviews.

A simple modification of the above epistemic models allows us to represent both the agents' knowledge and 'beliefs': An epistemic doxastic model is a tuple $\langle W, \{R_i^K\}_{i\in\mathcal{A}}, \{R_i^B\}_{i\in\mathcal{A}}, V \rangle$ where both R_i^K and R_i^B are relations on W. Truth of a belief operator $B_i\varphi$ is defined precisely as in Definition 2.2, replacing R_i with R_i^B. This points to a logical analysis of both informational attitudes with various 'bridge principles' relating knowledge and belief (such as knowing something implies believing it or if an agent believes φ then the agent knows that she believes it). However, we do not discuss this line of research[4] here since these models are not our preferred ways of representing the agents' beliefs (see, for example, Halpern 1996; Stalnaker 2006 for a discussion).

Table 1. Formulas of $\mathcal{L}_K$ and their corresponding properties.

Name	Formula	Property
K axiom	$K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi)$	—
Truth axiom	$K_i\varphi \rightarrow \varphi$	Reflexive
Positive introspection	$K_i\varphi \rightarrow K_iK_i\varphi$	Transitive
Negative introspection	$\neg K_i\varphi \rightarrow K_i\neg K_i\varphi$	Euclidean

3.1. MODELS OF BELIEF VIA PLAUSIBILITY

A key aspect of beliefs not yet represented in the epistemic doxastic models sketched above is that they are *revisable* in the presence of new information. In this section, I introduce models that describe both the knowledge and beliefs of a group of agents together with the key modal languages that have been used to reason about these structures. Building on the extensive literature on the theory of belief revision initiated with the classic paper (Alchourrón et al. 1985), these models will be used in part II to study *dynamic* logics of belief. Before we get started, it is worth stressing that, as in the previous section, the models introduced below are intended to describe the knowledge and beliefs of a group of agents without committing to any specific psychological or philosophical theory of knowledge and/or beliefs.

Taking a cue from Adam Grove's classic paper (Grove 1988), the key idea is to endow epistemic ranges with a *plausibility ordering* for each agent: a pre-order (reflexive and transitive) $w \preceq_i v$ that says 'agent i considers world w at least as plausible as v'. As a convenient notation, for $X \subseteq W$, we set $Min_{\preceq_i}(X) = \{v \in X | v \preceq_i w \text{ for all } w \in X \}$, the set of minimal elements of X according to $\preceq_i$. This is the subset of X that agent i considers the 'most plausible'. Thus, while the $\sim_i$ partitions the set of possible worlds according to i's 'hard information', the plausibility ordering $\preceq_i$ represents i's 'soft information' about which of the possible worlds agent i considers more 'plausible'. The models defined below have been extensively studied by logicians (van Benthem 2007; van Ditmarsch 2005; Baltag and Smets 2006), game theorists (Board 2004) and computer scientists (Boutilier 1992; Lamarre and Shoham 1994).

Definition 3.1 (Epistemic Plausability Model) Let $\mathcal{A}$ be a finite set of agents and At a (finite or countable) set of atomic propositions. An *epistemic plausability model* is a tuple $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, \{\preceq_i\}_{i \in \mathcal{A}}, V \rangle$ where $\langle W, \{\sim_i\}_{i \in \mathcal{A}}, V \rangle$ is an epistemic model and, for each $i \in \mathcal{A}$, $\preceq_i$ is a well-founded[5], reflexive and transitive relation on W satisfying, for all $w, v \in W$:

1. *plausibility implies possibility*: if $w \preceq_i v$ then $w \sim_i v$.
2. *Locally connected*: if $w \sim_i v$ then either $w \preceq_i v$ or $v \preceq_i w$.

Remark 3.2. Note that if $w \not\sim_i v$ then, since $\sim_i$ is symmetric, we also have $v \not\sim_i w$, and so by property 1, $w \not\preceq_i v$ and $v \not\preceq_i w$. Thus, we have the following equivalence: $w \sim_i v$ iff $w \preceq_i v$ or $v \preceq_i w$. In what follows, unless otherwise stated, I assume that $\sim_i$ is *defined* as follows: $w \sim_i v$ iff $w \preceq_i v$ or $v \preceq_i w$.

In order to reason about these structures, extend the basic epistemic language $\mathcal{L}_K$ with a conditional belief operator: Let $\mathcal{L}_{KB}$ be the set of sentences generated by the following grammar:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid B_i^{\varphi}\psi \mid K_i\varphi$$

where $p \in$ At (the set of atomic propositions) and $i \in \mathcal{A}$. The same conventions apply as above with the additional convention that we write $B_i\varphi$ for $B_i^{\top}\varphi$.

Let $[w]_i = \{v | w \sim_i v\}$ be the equivalence class of w under $\sim_i$, called i's *information cell* at w. Then, local connectedness implies that $\preceq_i$ totally orders $[w]_i$ and well-foundedness implies that $Min_{\preceq_i}([w]_i \cap X)$ is nonempty if $[w]_i \cap X \neq \emptyset$.

Definition 3.3 (Truth for $\mathcal{L}_{KB}$) Suppose that $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, \{\preceq_i\}_{i \in \mathcal{A}}, V \rangle$ is an episte-mic plausibility model. The definition of truth for formulas from $\mathcal{L}_K$ is given in Definition 2.2. The conditional belief operator is defined as follows:

- $\mathcal{M}, w \vDash B_i^{\varphi} \psi$ iff for all $v \in Min_{\preceq_i}([w]_i \cap [\![\varphi]\!]_{\mathcal{M}})$, $\mathcal{M}, v \vDash \psi$

where $[\![\varphi]\!]_{\mathcal{M}} = \{w \mid \mathcal{M}, w \vDash \varphi\}$.

Thus, ψ is believed conditional on φ, if i's most plausible φ-worlds (i.e., the states satisfying φ that i has not ruled out and considers most plausible) all satisfy ψ. Then, the definition of plain belief (which is defined to be $B^{\top}$) is:

$$\mathcal{M}, w \vDash B_i \varphi \quad \text{iff} \quad \text{for each } v \in Min_{\preceq_i}([w]_i), \mathcal{M}, v \vDash \varphi$$

Recall the example of Ann and Bob, and the two coins in separate drawers. The following epistemic plausibility model describes a possible configuration of beliefs before the agents observe their respective coins: I draw an arrow from v to w if $w \preceq v$ (to keep the clutter down, I do not include all arrows. The remaining arrows can be inferred by transitivity and reflexivity).
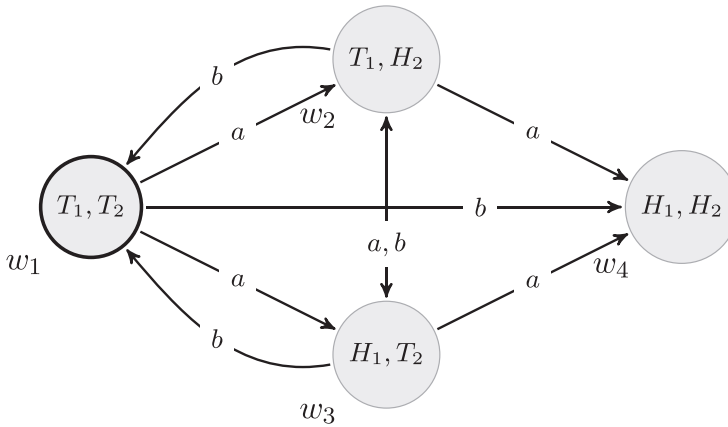
Suppose that both coins are facing tails up, so w_1 is the actual state. Following the conven-tion from Remark 3.2, we have $[w_1]_a = [w_1]_b = \{w_1, w_2, w_3, w_4\}$, and so, neither Ann nor Bob knows this fact. Furthermore, both Ann and Bob believe that both coins are facing heads up (i.e., $w_1 \vDash B_a(H_1 \wedge H_2) \wedge B_b(H_1 \wedge H_2)$ since $Min_{\preceq_a}([w_1]_a) = Min_{\preceq_b}([w_1]_b) = \{w_4\}$). How-ever, Ann and Bob do have different *conditional* beliefs. Ann believes that the position of the coins in the two drawers are independent; and so, she believes that H_2 is true even under the supposition that T_1 is true (and vice versa for the other coin: $w_1 \vDash B_a^{T_1} H_2 \wedge B_a^{T_2} H_1$). On the other hand, Bob believes that the coins are somehow correlated; and so, under the supposi-tion that T_1 is true, Bob believes that the coin in the second drawer must also be facing tails up (and vice versa for the other coin: $w_1 \vDash B_b^{T_1} T_2 \wedge B_b^{T_2} T_1$).

So, conditional beliefs describe an agent's *disposition* to change his or her beliefs in the presence of (perhaps surprising) evidence.[6] This is reflected in the logic of the conditional belief operator. An immediate observation is that $B_i(\varphi \rightarrow \psi)$ does not imply $B_i^{\varphi} \psi$.[7] Another, more fundamental, observation is that conditioning is nonmonotonic in the sense that $B_i^{\varphi} \alpha \rightarrow B_i^{\varphi \wedge \psi} \alpha$ is not valid. Nonetheless, weaker forms of monotonicity do hold. These and other key logical principles of conditional belief are shown in Table 2.

Table 2. Key principles of conditional belief.

Success:	$B_i^\varphi \varphi$		
Knowledge entails belief	$K_i \varphi \rightarrow B_i^\psi \varphi$		
Full introspection:	$B_i^\varphi \psi \rightarrow K_i B_i^\varphi \psi$	and	$\neg B_i^\varphi \psi \rightarrow K_i \neg B_i^\varphi \psi$
Cautious monotonicity:	$(B_i^\varphi \alpha \wedge B_i^\varphi \beta) \rightarrow B_i^{\varphi \wedge \beta} \alpha$		
Rational monotonicity:	$(B_i^\varphi \alpha \wedge \neg B_i^\varphi \neg \beta) \rightarrow B_i^{\varphi \wedge \beta} \alpha$		

The reader is invited to check that all of the formulas in Table 2 are valid in any epistemic plausibility model. The full introspection principles are valid since we assume that each agent's plausibility ordering is uniform. That is, each agent i has a single plausibility order over the set of all possible worlds, which totally orders each of i's information cells.[8] The intuition is that all the worlds that an agent i has ruled out at state w (based on her observations and/or background knowledge) are considered the least plausible overall. A more general semantics is needed if one wants to drop the assumption of full introspection and that knowledge entails belief. Specifically, the key idea is to define the agent's plausibility ordering as a *ternary* relation where $x \preceq_i^w y$ means i considers x at least as plausible as y in state w. So, agents may have different plausibility orderings at different states. The logic of these more general structures has been studied by Oliver Board (2004) (see also an earlier paper by John Burgess 1981).

The other principles highlight the close connection with the AGM postulates of rational belief revision (Alchourrón et al. 1985).[9] While success and cautious monotonicity are often taken to be constitutive of believing something *under the supposition that φ is true*, rational monotonicity has generated quite a bit of discussion. The most well-known criticism comes from Robert Stalnaker (1994a) who suggests the following counterexample: Suppose that Ann initially believes that the composer Verdi is Italian $I(v)$ and Bizet and Satie are French $(F(b) \wedge F(s))$. Conditioning on the fact that Verdi and Bizet are compatriots $(C(v,b))$,[10] Ann still believes that Satie is French $(B_a^{C(v,b)} F(s)$ is true since supposing that Verdi and Bizet are compatriots does not conflict with the belief that Satie is French). Under the supposition that Verdi and Bizet are compatriots, Ann thinks it is (doxastically) possible that Verdi and Satie are compatriots $(\neg B^{C(v,b)} \neg C(v,s)$ is true since $C(v,b)$ is consistent with all three being French). Rational monotonicity gives us that Ann believes Satie is French under the supposition that $C(v,b) \wedge C(v,s)$ (i.e., $B^{C(v,b) \wedge C(v,s)} F(s)$ must be true). However, supposing that $C(v,b) \wedge C(v,s)$ is true implies all three composers are compatriots, and this could be because they are all Italian. Much has been written in response to this counterexample. However, a detailed analysis of this literature would take us too far away from the main objective of this section: a concise introduction to the main formal models of knowledge and belief.

I conclude this section by discussing two additional notions of belief. To that end, we need some additional notation: The plausibility relation $\preceq_i$ can be lifted to subsets of W as follows[11]:

$$X \preceq_i Y \text{ iff } x \preceq_i y \text{ for all } x \in X \text{ and } y \in Y$$

Exend the language $\mathcal{L}_{KB}$ with two belief operators: B_i^r ('robust belief') and B_i^s ('strong belief'). Let $\mathcal{L}_{KB}^+$ be the resulting language. Semantics for this language is given by adding the following clauses to Definition 3.3.

- *Robust belief* [12]: $\mathcal{M}, w \models B_i^r \varphi$ iff for all v if $v \preceq_i w$ then $\mathcal{M}, v \models \varphi$. Thus, φ is robustly believed if φ is true in *all* states the agent considers at least as plausible as the current state.

This stronger notion of belief has also been called *certainty* by some authors (Shoham and Leyton-Brown (2009), Section 13.7).

- *Strong belief:* $\mathcal{M}, w \vDash B_i^s \varphi$ iff there is a v with $w \sim_i v$ and $\mathcal{M}, v \vDash \varphi$ and $\{x \mid \mathcal{M}, x \vDash \varphi\} \cap [w]_i \preceq_i \{x \mid \mathcal{M}, x \vDash \neg\varphi\} \cap [w]_i$ (recall that $[w]_i$ is the equivalence class of w under $\sim_i$). This notion has been studied by (Stalnaker 1994b; Battigalli and Siniscalchi 2002).

The following example illustrates the logical relationships between the various notions of belief and knowledge we have discussed. Consider the following plausibility model with four states for a single agent (since there is only one agent, I do not use subscripts):

Note that the set of minimal states is $\{v_2\}$, so if $v_2 \in V(p)$, then the agent believes p (Bp is true at all states). Suppose that w is the 'actual world' and consider the following truth assignments of an atomic proposition p.

- $V(p) = \{v_0, w, v_2\}$. Then $\mathcal{M}, w \vDash B^r p$, but $\mathcal{M}, w \nvDash B^s p$, so robust belief need not imply strong belief.
- $V(p) = \{v_2\}$. Then $\mathcal{M}, w \vDash B^s p$, but $\mathcal{M}, w \nvDash B^r p$, so strong belief need not imply robust belief.
- $V(p) = \{v_0, v_1, w, v_2\}$. Then $\mathcal{M}, w \vDash Kp \wedge B^s p \wedge B^r p \wedge Bp$ (in fact, it is easy to see that knowledge implies belief, robust belief and strong belief).

Note that, unlike beliefs, conditional beliefs may be inconsistent (i.e., $B_i^{\varphi} \bot$ may be true at some state). In such a case, agent i cannot (on pain of inconsistency) revise by φ, but this will happen only if the agent has hard information that φ is false. Indeed, $K_i \neg \varphi$ is logically equivalent to $B_i^{\varphi} \bot$ (over the class of epistemic plausibility models). This suggests the following (dynamic) characterization of an agent's hard information as unrevisable beliefs:

- $\mathcal{M}, w \vDash K_i \varphi$ iff $\mathcal{M}, w \vDash B_i^{\psi} \varphi$ for all ψ.

Robust and strong belief can be similarly characterized by restricting the set of formulas that an agent can condition on:

- $\mathcal{M}, w \vDash B_i^r \varphi$ iff $\mathcal{M}, w \vDash B_i^{\psi} \varphi$ for all ψ with $\mathcal{M}, w \vDash \psi$: That is, the agent robustly believes φ iff she continues to believe φ given any true formula.
- $\mathcal{M}, w \vDash B_i^s \varphi$ iff $\mathcal{M}, w \vDash B_i \varphi$ and $\mathcal{M}, w \vDash B_i^{\psi} \varphi$ for all ψ with $\mathcal{M}, w \vDash \neg K_i(\psi \to \neg\varphi)$: That is, the agent strongly believes φ iff she believes φ and continues to believe φ given any evidence (truthful or not) that is not known to contradict φ.

Finally, there is an elegant axiomatization of epistemic plausibility models in a modal language containing knowledge operators and robust belief operators using the following characterizations of conditional and strong belief (Baltag and Smets 2009):

- $$B_i^{\varphi} \psi := L_i \varphi \to L_i \big(\varphi \wedge B_i^r (\varphi \to \psi) \big)$$

- $$B_i^s \varphi := L_i \varphi \wedge K_i \big(\varphi \to B_i^r \varphi \big)$$

As discussed above, each K_i satisfies logical omniscience, veracity and both positive and negative introspection. Robust belief, B_i^r, shares all of these properties except negative

introspection. Modal correspondence theory can again be used to characterize the remaining properties (see Table 3).

3.2. MODELS OF BELIEF VIA PROBABILITY

The logical frameworks introduced in the previous section represent an agent's 'all-out' or 'full' beliefs. However, there is a large body of literature within (formal) epistemology and game theory that works with a *quantitative* conception of belief. Graded beliefs have also been subjected to sophisticated logical analyses (see, for example, Fagin et al. 1990; Fagin and Halpern 1994; Heifetz and Mongin 2001; Zhou 2009; Goldblatt 2010).

The dominant approach to formalizing graded beliefs is (subjective) probability theory. A *probability measure* on a set W is a function assigning a positive real number to (some) subsets of W such that $\pi(W) = 1$ and for disjoint subsets $E, F \subseteq W$ ($E \cap F = \varnothing$) $\pi(E \cup F) = \pi(E) + \pi(F)$. For simplicity, I assume in this section that W is finite. Then, the definition of a probability measure can be simplified: a probability measure on a finite set W is a function $\pi : W \to [0,1]$ such that for each $E \subseteq W$, $\pi(E) = \sum_{w \in E} \pi(w)$ and $\pi(W) = 1$. Nothing that follows hinges on the assumption that W is finite, but if W is infinite, then there are a number of important mathematical details that add some complexity to the forthcoming definitions.[13] *Conditional probability* is defined in the usual way: $\pi(E|F) = \frac{\pi(E \cap F)}{\pi(F)}$ if $\pi(F) > 0$ ($\pi(E|F)$ is undefined when $\pi(F) = 0$).

The model we study in this section is very close to the epistemic plausibility of Definition 3.1 with probability measures in place of plausibility orderings:

Definition 3.4 (Epistemic Probability Model) Suppose that $\mathcal{A}$ is a finite set of agents, At is a (finite or countable) set of atomic propositions and W is a finite set of states. An *epistemic probability model* is a tuple $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, \{\pi_i\}_{i \in \mathcal{A}}, V \rangle$ where $\langle W, \{\sim_i\}_{i \in \mathcal{A}}, V \rangle$ is an epistemic model and, for each $i \in \mathcal{A}$, π_i is a probability measure on W. We also assume that each π_i is *weakly regular*[14] in the sense that for each $w \in W$, $\pi_i([w]_i) > 0$.

The probability measures π_i represent agent i's prior beliefs about the likelihood of each element of W. Agents then receive private information, represented by the equivalence relations $\sim_i$ and update their initial beliefs with that information. A variety of modal languages have been proposed to reason about graded beliefs. In this section, we focus on a very simple language containing a knowledge modality $K_i \varphi$ ('i is informed that φ is true') and $B_i^q \varphi$ ('i believes φ is true to degree at least q' or 'i's degree of belief in φ is at least q') where q is a rational number. More formally, let $\mathcal{L}_{KB}^{prob}$ be the smallest set of formulas generated by the following grammar:

$$ p \mid \neg\varphi \mid \varphi \wedge \psi \mid B_i^q \varphi \mid K_i \varphi $$

where $q \in \mathbb{Q}$ (the set of rational numbers), $i \in \mathcal{A}$ and $p \in \mathsf{At}$.

Table 3. Properties of robust belief and knowledge.

Knowledge entails robust belief	$K_i \varphi \to B_i^r \varphi$
Local connectedness	$\left(K_i(\varphi \vee B_i^r \psi) \wedge K_i(\psi \vee B_i^r \varphi) \right) \to K_i \varphi \vee K_i \psi$

Definition 3.5 (Truth for $\mathcal{L}_{KB}^{prob}$) Suppose that $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, \{\pi_i\}_{i \in \mathcal{A}}, V \rangle$ is an epistemic probability model. The definition of truth for formulas from $\mathcal{L}_K$ is given in Definition 2.2. The belief operator is defined as follows:

- $\mathcal{M}, w \vDash B_i^q \varphi$ iff $\pi_i(\llbracket \varphi \rrbracket_{\mathcal{M}} \mid [w]_i) \geq q$

where $\llbracket \varphi \rrbracket_{\mathcal{M}} = \{w \mid \mathcal{M}, w \vDash \varphi\}$.

Note that since we assume for each $w \in W$, $\pi_i([w]_i) > 0$, the above definition is always well defined. The logic of these models is very similar to the qualitative version presented in the previous section. In particular, knowledge entails belief $(K_i \varphi \rightarrow B_i^q \varphi)$, and the full introspection principles ($B_i^q \varphi \rightarrow K_i B_i^q \varphi$ and $\neg B_i^q \varphi \rightarrow K_i \neg B_i^q \varphi$) are both valid. As before, full introspection can be dropped by assuming agents have different probability measures at different states. [15]

The graded notion of belief has much in common with its qualitative version. In particular, B_i^q satisfies both positive and negative introspection (these both follow from full introspection, the fact that knowledge entails belief and logical omniscience: if $\varphi \rightarrow \psi$ is valid then so is $B_i^q \varphi \rightarrow B_i^q \psi$). There is also an analog to the success axiom (although we cannot state it in our language since we do not have conditional belief operators in $\mathcal{L}_{KB}^{prob}$):

$$\pi_i(\llbracket \varphi \rrbracket_{\mathcal{M}} \mid \llbracket B_i^q \varphi \rrbracket_{\mathcal{M}}) \geq q$$

It is a simple (and instructive!) exercise to verify that the above property is true in any epistemic probability model. In addition to principles describing what the agents know and believe about their own beliefs, there are principles that ensure the different B_i^q operators fit together in the right way (see Table 4).

A distinctive feature of the complete logic of epistemic probability models is the following inference rule reflecting the completeness (in the topological sense) of the real numbers:

Archimedian Rule: If $\psi \rightarrow B_i^p \varphi$ is valid for each $p < q$, then $\psi \rightarrow B_i^q \varphi$ is valid.
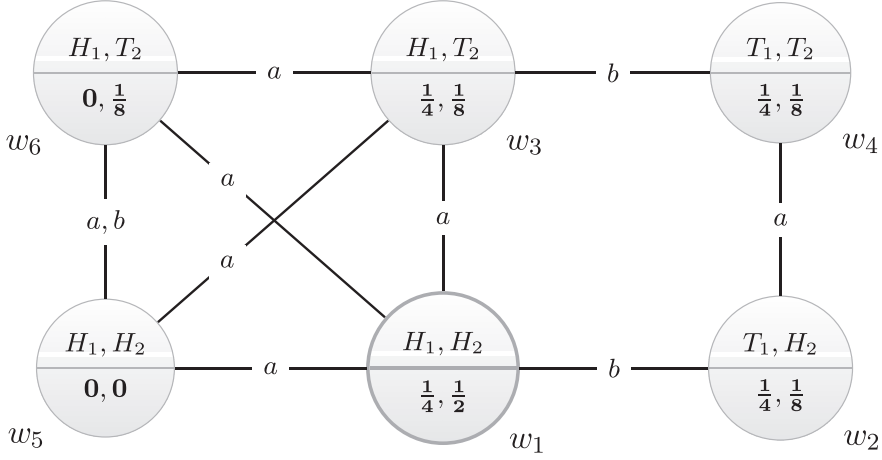
I conclude this section with some brief comments about probability zero events and the relationship between knowledge $(K_i \varphi)$ and belief with probability one $(B_i^1 \varphi)$. Note that it is possible that in an epistemic probability model, there are states w and v with $w \sim_i v$ and $\pi_i(v) = 0$. In such a case, agent i is sure that state v is not the actual state, but does not consider v impossible (i.e., i cannot 'rule out' v according to her information). In particular, this means that $B_i^1 \varphi \rightarrow K_i \varphi$ is not valid. The following example illustrates these issues: The numbers in the lower half of the circle indicate Ann and Bob's initial probability for that state (Ann's probability is on the left and Bob's is on the right, so, for example, $\pi_a(w_1) = \frac{1}{4}$ and $\pi_b(w_1) = \frac{1}{2}$).

Table 4. Principles of probalistic belief.

$B_i^0 \varphi$

$B_i^1 \top$

$B_i^q(\varphi \wedge \psi) \wedge B_i^p(\varphi \wedge \neg \psi) \rightarrow B_i^{q+p} \varphi, \quad q + p \leq 1$

$\neg B_i^q(\varphi \wedge \psi) \wedge \neg B_i^p(\varphi \wedge \neg \psi) \rightarrow \neg B_i^{q+p} \varphi, \quad q + p \leq 1$

$B_i^q \varphi \rightarrow \neg B_i^p \neg \varphi, \quad q + p > 1$

H_1, T_2
$0, \frac{1}{8}$
w_6

H_1, T_2
$\frac{1}{4}, \frac{1}{8}$
w_3

T_1, T_2
$\frac{1}{4}, \frac{1}{8}$
w_4

a b

a

a, b a a

a

H_1, H_2
$0, 0$
w_5

H_1, H_2
$\frac{1}{4}, \frac{1}{2}$
w_1

T_1, H_2
$\frac{1}{4}, \frac{1}{8}$
w_2

a b

The following observations illustrate the notions introduced above:

- Ann does not know the direction the coin is facing in the second drawer and believes that H_2 and T_2 are equally likely: $\mathcal{M}, w_1 \models \neg K_a H_2 \wedge \neg K_a T_2 \wedge B_a^{\frac{1}{2}} H_2 \wedge B_a^{\frac{1}{2}} T_2$.
- Bob does not know the direction the coin is facing in the first drawer, but believes it is more likely facing heads up: $\mathcal{M}, w_1 \models \neg K_b H_1 \wedge \neg K_b T_1 \wedge B_b^{\frac{4}{5}} H_1 \wedge B_b^{\frac{1}{5}} T_1$. [16]
- Ann does not know that Bob knows whether H_2, but she is certain that he knows whether H_2 (in the sense that she assigns probability one to him knowing whether): $\mathcal{M}, w_1 \models \neg K_a(K_b H_2 \vee K_b T_2) \wedge B_a^1(K_b H_2 \vee K_b T_2)$.

Of course, epistemic probability models provide a more fine-grained representation of the agents' beliefs than their qualitative counterparts. However, the relationship between the two models is subtle and touches on many issues beyond the scope of this article.[17]

4. Group Notions

Both game theorists and logicians have extensively discussed different notions of knowledge and belief for a group, such as *common* knowledge and belief. These notions have played a fundamental role in the analysis of distributed algorithms (Halpern and Moses 1990) and social interactions (Chwe 2001). In this section, I introduce and formally define various group informational attitudes. I can only scratch the surface of the extensive literature discussing the numerous logical and philosophical issues that arise here (see Vanderschraaf and Sillari (2009) for an in-depth discussion of this literature[18]).

Consider the statement 'everyone in group G knows that φ'. With finitely many agents, this can be easily defined in the epistemic language $\mathcal{L}_{KB}$:

$$K_G \varphi := \bigwedge_{i \in G} K_i \varphi$$

where $G \subseteq \mathcal{A}$. The first nontrivial informational attitude for a group that we study is *common knowledge*. If φ is common knowledge for the group G, then not only does everyone in the group know that φ is true but also this fact is completely transparent to all members of the group. There are different ways to make precise what it means for something to be 'completely transparent' to a group of agents.[19] The approach I follow here is to iterate the everyone knows operator:

$$\varphi \wedge K_G\varphi \wedge K_GK_G\varphi \wedge K_GK_GK_G\varphi \wedge \cdots$$

The above formula is an *infinite* conjunction and, so, is not a formula in our epistemic language $\mathcal{L}_{KB}$ (by definition, there can be at most finitely many conjunctions in any formula). In fact, using standard modal logic techniques, one can show that there is no formula of $\mathcal{L}_{KB}$ that is logically equivalent to the above infinite conjunction. Thus, we must extend our basic epistemic language with a modal operator $C_G\varphi$ with the intended meaning 'φ is common knowledge among the group G'. The idea is to *define* $C_G\varphi$ to be true precisely when φ is true, everyone in G knows that φ is true, everyone in G knows that everyone in G knows that φ is true, and so on *ad infinitum*.[20]

Before giving the details of this definition, consider $K_GK_GK_G\varphi$. This formula says that 'everyone from group G knows that everyone from group G knows that everyone from group G knows that φ'. When will this be true at a state w in an epistemic model? First some notation: *a path of length n for G* in an epistemic model is a sequence of states $(w_0, w_1, \ldots, w_n)$ where for each $l = 0, \ldots, n-1$, we have $w_l \sim_i w_{l+1}$ for some $i \in G$ (for example $w_0 \sim_1 w_1 \sim_2 w_2 \sim_1 w_3$ is a path of length 3 for $\{1,2\}$). Thus, $K_GK_GK_G\varphi$ is true at state w iff every path of length 3 for G starting at w leads to a state where φ is true. This suggests the following definition:

Definition 4.1 (Interpretation of C_G) Let $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, V \rangle$ be an epistemic model[21] and $w \in W$. The truth of formulas of the form $C_G\varphi$ is

$$\mathcal{M}, w \vDash C_G\varphi \text{ iff } \text{ for all } v \in W, \text{ if } wR_G^C v \text{ then } \mathcal{M}, v \vDash \varphi$$

where $R_G^C := (\cup_{i \in G} \sim_i)^*$ is the reflexive transitive closure[22] of $\cup_{i \in G} \sim_i$.

It is well known that for any relation R on W, if wR^*v then there is a finite R-path starting at w ending in v. Thus, we have $\mathcal{M}, w \vDash C_G\varphi$ iff every finite path for G from w ends with a state satisfying φ. The logical analysis is more complicated in languages with a common knowledge operator; however, the formulas given in Table 5 can be said to characterize[23] common knowledge.

The first formula captures the 'self-evident' nature of common knowledge: if φ is common knowledge then everyone in the group knows this.

The approach to defining common knowledge outlined above can be viewed as a recipe for defining common (robust/strong) belief. For example, suppose $wR_i^B v$ iff $v \in Min_{\preceq_i}([w]_i)$ and define R_G^B to be the transitive closure[24] of $\cup_{i \in G} R_i^B$. Then, *common belief of φ*, denoted $C_G^B\varphi$, is defined in the usual way:

$$\mathcal{M}, w \vDash C_G^B\varphi \text{ iff for each } v \in W, \text{ if } wR_G^B v \text{ then } \mathcal{M}, v \vDash \varphi$$

While common belief also validates the fixed-point and induction axiom, its logic does differ from the logic of common knowledge. The most salient difference is that common

Table 5. Principles of common knowledge.

Fixed point	$C_G\varphi \to K_GC_G\varphi$
Induction	$(\varphi \wedge C_G(\varphi \to K_G\varphi)) \to C_G\varphi$

knowledge satisfies negative introspection ($\neg C_G \varphi \rightarrow C_G \neg C_G \varphi$ is valid) while common belief does not ($\neg C_G^B \varphi \rightarrow C_G^B \neg C_G^B \varphi$ is not valid). See Bonanno (1996) and Lismont and Mongin (1994, 2003) for more information about the logic of common belief. A probabilistic variant of common belief was introduced by Monderer and Samet (1989).

I conclude by briefly discussing another notion of 'group knowledge': *distributed knowledge*. Intuitively, φ is distributed knowledge among a group of agents if φ would be known if all the agents in the group put all their information together. Formally, given an epistemic model (beliefs do not play a role here) $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, V \rangle$, let $R_G^D = \cap_{i \in G} \sim_i$, then define

$$\mathcal{M}, w \vDash D_G \varphi \text{ iff for all } v \in W, \text{ if } w R_G^D v \text{ then } \mathcal{M}, v \vDash \varphi.$$

Note that $D_G \varphi$ is *not* simply equivalent to $\bigwedge_{i \in G} K_i \varphi$ as illustrated by the following example (as usual reflexive arrows are not depicted in the diagram):
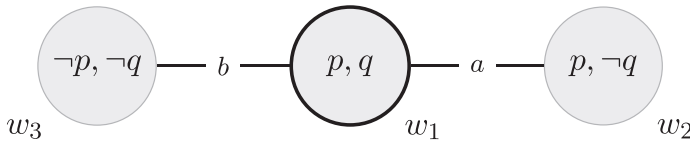
There is distributed knowledge at w_1 that q is true (agent a knows that p is true and agent b knows that $p \rightarrow q$ is true); however, neither agent individually knows that q is true. The logical analysis of distributed knowledge has raised a number of interesting technical and conceptual issues (see Fagin et al. 1995; Roelofsen 2007; van Benthem 2011; Gerbrandy 1999; Baltag and Smets 2010).

Short Biography

Eric Pacuit is an assistant professor in the Department of Philosophy at the University of Maryland, College Park and a research fellow in the Tilburg Institute for Logic and Philosophy of Science. Eric received his PhD in computer science from the Graduate Center of the City University of New York. He was a postdoctoral researcher at the Institute for Logic, Language and Computation at the University of Amsterdam and in the Departments of Philosophy and Computer Science at Stanford University. His postdoctoral research at the ILLC was funded by the Natural Science Foundation. Eric's primary research interests are in logic, game theory and formal epistemology. Currently, his research is focused on logics of rational agency and foundational issues in game theory and social choice theory. This research is supported by a Vidi grant from NWO called 'A Formal Analysis of Social Procedures'.

Notes

★ Correspondence: Department of Philosophy, University of Maryland, College Park. Email: epacuit@umd.edu.

[1] Note that there is nothing 'metaphysical' attached to the term 'possible world' here. Indeed, the notion of a 'possible state' (i.e., a 'possible world') plays a key role in many computer science applications, such as in the semantics of programming languages and the specification of (multiagent) systems. Nonetheless, the basic modeling choices are not without controversy. The 'conceptual difficulties' are related to issues raised by Jon Barwise and John Perry in their development of *situation semantics* (Barwise and Perry 1983) and issues surrounding the underlying assumption of logical omniscience (Stalnaker 1991; Parikh 2005).

[2]   To be more precise, the key notion here is *frame definability*: A frame is a pair $\langle W,R \rangle$ where W is a nonempty set and R a relation on W. A modal formula is valid on a frame if it is valid in every model based on that frame. It can be shown that some modal formulas have first-order *correspondents* P where for any frame $\langle W,R \rangle$, the relation R has property P iff φ is valid on $\langle W,R \rangle$.

[3]   It is important to distinguish the valid inference rule if φ is valid then so is $K_i\varphi$ from the formula $\varphi \to K_i\varphi$. The latter is certainly not valid as it trivializes the notion of knowledge (anything that is true is known) while necessitation is a rule in every normal modal logic.

[4]   One key issue runs as follows: Suppose that p is something agent i certain of, but is false: i.e., $\neg p \wedge B_i p$ is true. Then, since K_i satisfies the veracity axiom, $\neg K_i p$ must be true. By negative introspection, this implies $K_i\neg K_i p$ is true. Since knowledge implies belief, we have $B_i\neg K_i p$ is true. Since $B_i p$ is true and belief implies believing it is known, we have $B_i K_i p$ is true. But this means $B_i(K_i p \wedge \neg K_i p)$ is true, which contradicts the assumption that B_i is consistent (i.e., $\neg B_i \bot$ is valid). Thus, there is a conflict between standard assumptions of knowledge and belief, and certain bridge principles relating knowledge and belief. Of course, one may question the validity of these bridge principles (especially the assumption that believing something implies believing that one knows it). See Halpern (1996) and van der Hoek (1993) for extended analyses.

[5]   Well-foundedness is a standard assumption in the literature. It is only needed to ensure that for any nonempty set X, $Min_{\preceq_i}(X)$ is nonempty. This is important only when W is infinite – and there are ways around this in current logics of belief revision. Moreover, the condition of connectedness can also be lifted, but it is used here for convenience.

[6]   See Leitgeb (2007) for an extensive discussion about a number of philosophical issues surrounding how to interpret conditional beliefs *qua* belief.

[7]   Suppose that there are two states w and v with p true only at w and q true only at v. Then $p \to q$ is true at v and q is false at w. Suppose that $v \preceq_i w$. Then, $B_i(p \to q)$ is true at w, but $B^p q$ is not true at w. This example shows that conditional belief does not reduce to belief in a *material conditional*. This is, of course, not surprising. A much more interesting question is about the relationship between a conditional belief and a belief in an (*indicative*) *conditional*. There is a broad literature on this issue. A good entrée into the field is (Leitgeb 2010a).

[8]   In general, each agent's plausibility ordering will not be a total relation as states in different information cells cannot be compared by the plausibility ordering.

[9]   This close connection should not come as a surprise, since epistemic plausibility models are essentially the sphere models of Grove (1988).

[10]   Note that in this example, $I(v)$, $F(b)$ and $F(s)$ are all atomic propositions and $C(i,j)$ is defined to be $(I(i) \wedge I(j)) \vee (F(i) \wedge F(j))$.

[11]   This is only one of many possible choices here (cf. Liu 2011, Chapter 4), but it is the most natural in this setting.

[12]   In the dynamic doxastic logic literature, this notion is often called *safe belief*. However, this terminology conflicts with Timothy Williamson's notion of 'safety' (Williamson 2000).

[13]   See Halpern (2003), Chapter 1 and Billingsley (1995) for details.

[14]   A probability measure is *regular* provided $\pi(E) > 0$ for each (measurable) subset E.

[15]   Formally, replace each probability measure π_i with functions $P_i : W \to \Delta(W)$ where $\Delta(W)$ is the class of probability measures on W.

[16]   To see, for example, that $\mathcal{M}, w_1 \vDash B_b^{\frac{4}{5}} H_1$, we need to calculate the following conditional probability: $\pi_b\left([\![H_1]\!]_{\mathcal{M}} \mid [w_1]_b \right) = \pi_b\left(\{w_1, , w_3, , w_5, , w_6\} \middle| \{w_1, , w_2\} \right) = \frac{\pi_b(\{w_1\})}{\pi_b(\{w_1, , w_2\})} = \frac{1/2}{1/2 + 1/8} = \frac{4}{5}$.

[17]   It is tempting to identify 'more plausible' with 'more probable'. There are a number of reasons why this is not a good idea. The difficulty here boils down to a deep foundational problem: Can a rational agent's full beliefs be *defined* in terms of her graded beliefs and/or vice versa? A second, more conceptual, observation is that the two models represent different aspects of the agents' states of belief. To illustrate the difference, suppose that Ann is flipping a biased coin. It may be much more likely to land heads, but landing heads and tails are both *plausible* outcomes, whereas landing on its side is not a plausible outcome. Furthermore, landing on its side is more plausible than hovering in the air for 10 minutes. So, the plausibility ordering describes the agents' all-out judgements about the priority between the states, which is not directly represented in a probability measure. A good overview of the main philosophical issues is found in Christensen (2007). See Hawthorne and Bovens (1999), Leitgeb (2010b) and Arló-Costa and Pedersen (2012) for discussions that are particularly relevant to the logical frameworks introduced in this paper.

[18]   The textbooks Fagin et al. (1995) and van Benthem (2011) also provide illuminating discussions of key logical issues.

[19]   Barwise (1988) discusses three main approaches: (i) the iterated view; (ii) the fixed-point view; and (iii) the shared situation view. In this paper, I only discuss the first approach.

[20]   It is worth pointing out that this is not what David Lewis had in mind when he first formalized common knowledge (Lewis 1973). For Lewis, the infinite conjunction is a necessary but not a sufficient condition for common knowledge. See Cubitt and Sugden (2003) for an illuminating discussion and a reconstruction of Lewis' notion of common knowledge. Nonetheless, following Aumann (1976), the definition given in this section has become standard in the game theory and epistemic logic literature.

[21]   The same definition will of course hold for epistemic plausibility and epistemic probability models.

[22] The reflexive transitive closure of a relation R is the smallest relations R^* containing R that is reflexive and transitive.
[23] Techniques similar to the previously mentioned *correspondence theory* can be applied here to make this precise: see van Benthem (2006b) for a discussion.
[24] Since beliefs need not be factive, we do not force R_G^B to be reflexive.

Works Cited

Alchourrón, C. E., P. Gärdenfors, and D. Makinson. 'On the Logic of Theory Change: Partial Meet Contraction and Revision Functions.' *Journal of Symbolic Logic* 50 (1985): 510–530.

Arló-Costa, H. and A. P. Pedersen. 'Belief and Probability: A General Theory of Probability Cores.' *Journal of Approximate Reasoning* 53(3) (2012): 293–315.

Aumann, R. 'Agreeing to Disagree.' *The Annals of Statistics* 4 (1976): 1236–1239.

———. 'Interactive Epistemology I: Knowledge.' *International Journal of Game Theory* 28 (1999): 263–300.

Baltag, A. and S. Smets. 'Conditional Doxastic Models: A Qualitative Approach to Dynamic Belief Revision.' *Proceedings of WOLLIC 2006, Electronic Notes in Theoretical Computer Science*. Eds. G. Mints and R. de Queiroz. Elsevier, Volume 165. 2006. 5–21.

———. 'ESSLLI 2009 Course: Dynamic Logics for Interactive Belief Revision.' (2009). Slides available at http://alexandru.tiddlyspot.com.

———. 'Correlated Knowledge, an Epistemic-Logic View on Quantum Entanglement.' *International Journal of Theoretical Physics* 49(12) (2010): 3005–3021.

Barwise, J. 'Three Views of Common Knowledge.' *TARK '88: Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning About Knowledge*. San Francisco, CA, USA: Morgan Kaufmann Publishers, 1988. 365–379.

Barwise, J. and J. Perry. *Situations and Attitudes*. CSLI Publications, Stanford, 1983.

Battigalli, P. and M. Siniscalchi. 'Strong Belief and Forward Induction Reasoning.' *Journal of Economic Theory* 105 (2002): 356–391.

van Benthem, J. 'Epistemic Logic and Epistemology: The State of Their Affairs.' *Philosophical Studies* 128 (2006a): 49–76.

———. 'Modal Frame Correspondences and Fixed-Points.' *Studia Logica* 83(1-3) (2006b): 133–155.

———. 'Dynamic Logic for Belief Revision.' *Journal of Applied Non-classical Logics* 17(2) (2007): 129–155.

———. *Logical Dynamics of Information and Interaction*. Cambridge University Press, Cambridge, 2011.

———. *Modal Logic for Open Minds*. CSLI Publications, Stanford, 2010.

Billingsley, P. *Probability and Measure*. Wiley, Hoboken, 1995.

Blackburn, P., M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, Cambridge, 2002.

Board, O. 'Dynamic Interactive Epistemology.' *Games and Economic Behavior* 49 (2004): 49–80.

Bonanno, G. 'On the Logic of Common Belief.' *Mathematical Logical Quarterly* 42 (1996): 305–311.

Bonanno, G. and P. Battigalli. 'Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory.' *Research in Economics* 53(2) (1999): 149–225.

Boutilier, C. 'Conditional Logics for Default Reasoning and Belief Revision.' Ph. D. thesis, University of Toronto, 1992.

Brandenburger, A. 'The Power of Paradox: Some Recent Developments in Interactive Epistemology.' *International Journal of Game Theory* 35 (2007): 465–492.

Burgess, J. 'Quick Completeness Proofs for Some Logics of Conditionals.' *Notre Dame Journal of Formal Logic* 22 (1981): 76–84.

Christensen, D. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford University Press, 2007.

Chwe, M. S.-Y. *Rational Ritual*. Princeton University Press, 2001.

Cubitt, R. P. and R. Sugden. 'Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory.' *Economics and Philosophy* 19(2) (2003): 175–210.

van Ditmarsch, H. 'Prolegomena to Dynamic Logic for Belief Revision.' *Synthese: Knowledge, Rationality, and Action* 147 (2005): 229–275.

Egré, P. 'Epistemic Logic.' *The Continuum Companion to Philosophical Logic*. Eds. L. Horsten and R. Pettigrew. Continuum, London 2011.

Egré, P. and D. Bonnay. 'Inexact Knowledge with Introspection.' *Journal of Philosophical Logic* 38(2) (2009): 179–228.

Fagin, R. and J. Halpern. 'Reasoning about Knowledge and Probability.' *Journal of the ACM* 41(2) (1994): 340–367.

Fagin, R., J. Halpern, and N. Megiddo. 'A Logic for Reasoning About Probabilities.' *Information and Computation* 87(1) (1990): 78–128.

Fagin, R., J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. The MIT Press, Boston, 1995.

Gerbrandy, J. 'Bisimulations on Planet Kripke.' Ph. D. thesis, University of Amsterdam, 1999.

Gintis, H. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press, Princeton, 2009.

Goldblatt, R. 'Deductive Systems for Coalgebras over Measurable Spaces.' *Journal of Logic and Computation* 20(5) (2010): 1069–1100.

Grove, A. 'Two Modelings for Theory of Change.' *Journal of Philosophical Logic* 17 (1988): 157–179.

Halpern, J. 'Should Knowledge Entail Belief?' *Journal of Philosophical Logic* 25(5) (1996): 483–494.

———. *Reasoning about Uncertainty*. The MIT Press, Boston, 2003.

Halpern, J. and Y. Moses. 'Knowledge and Common Knowledge in a Distributed Environment.' *Journal of the ACM* 37(3) (1990): 549–587.

Halpern, J. and R. Pucella. 'Modeling Adversaries in a Logic for Security Protocol Analysis.' *Formal Aspects of Security*. 2003.

———. 'Dealing with Logical Omniscience: Expressiveness and Pragmatics.' *Artificial Intelligence* 175(1) (2011): 220–235.

Hawthorne, J. and L. Bovens. 'The Preface, the Lottery, and the Logic of Beleif.' *Mind* 108 (1999): 241–264.

Heifetz, A. and P. Mongin. 'Probability Logic for Type Spaces.' *Games and Economic Behavior* 35 (2001): 31–53.

Hintikka, J. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, 1962.

van der Hoek, W. and M. Wooldridge. 'Cooperation, Knowledge, and Time: Alternating-Time Temporal Epistemic Logic and its Applications.' *Studia Logica* 75(1) (2003): 125–157.

van der Hoek, W. 'Systems for Knowledge and Belief.' *Journal of Logic and Computation* 3(2) (1993): 173–195.

Holliday, W. H. 'Epistemic Logic and Epistemology.' *Handbook of Formal Philosophy*. Eds. V. Hendricks and S. O. Hansson. Springer, 2012.

Huber, F. 'Formal Representations of Belief'. *The Stanford Encyclopedia of Philosophy*. Winter 2011 ed. Ed. E. N. Zalta. 2011.

Lamarre, P. and Y. Shoham. 'Knowledge, Certainty, Belief and Conditionalisation.' *Proceedings of the International Conference on Knowledge Representation and Reasoning*, 1994. 415–424.

Leitgeb, H. 'Beliefs in Conditionals vs. Conditional Beliefs.' *Topoi* 26(1) (2007): 115–132.

———. 'On on the Ramsey Test without Triviality.' *Notre Dame Journal of Formal Logic* 51(1) (2010a): 21–54.

———. 'Reducing Belief Simpliciter to Degrees of Belief.' Technical report, Bristol. Manuscript, 2010b.

Lewis, D. *Counterfactuals*. Oxford: Blackwell Publishers, 1973.

———. 'Elusive Knowledge.' *Australasian Journal of Philosophy* 74(4) (1996): 549–567.

Lismont, L. and P. Mongin. 'On the Logic of Common Belief and Common Knowledge.' *Theory and Decision* 37(1) (1994): 75–106.

———. 'Strong Completeness Theorems for Weak Logics of Common Belief.' *Journal of Philosophical Logic* 32(2) (2003): 115–137.

Liu, F. *Reasoning about Preference Dynamics*. Springer-Verlag, 2011.

Meyer, J.-J. and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, Cambridge, 1995.

Monderer, D. and D. Samet. 'Approximating Common Knowledge with Common Beliefs.' *Games and Economic Behavior* 1(2) (1989): 170–190.

Parikh, R. 'Social Software.' *Synthese* 132 (2002): 187–211.

———. 'WHAT Do We Know, and What Do WE Know?' *Proceedings of Theoretical Aspects of Rationality and Knowledge*. Ed. R. van der Meyden. National University of Singapore, 2005.

Ramanujam, R. and S. Suresh. 'Deciding Knowledge Properties of Security Protocols.' *Proceedings of Theoretical Aspects of Rationality and Knowledge*, 2005. 219–235.

Roelofsen, F. 'Distributed Knowledge.' *Journal of Applied Non-Classical Logics* 17(2) (2007): 255–273.

Roy, O. and E. Pacuit. 'Substantive Assumptions in Interaction: A Logical Perspective.' *Synthese* 190(5) (2011): 891–908.

Samuelson, L. 'Modeling Knowledge in Economic Analysis.' *Journal of Economic Literature* 57 (2004): 367–403.

Shoham, Y. and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, Cambridge, 2009.

Sorensen, R. 'Formal Problems about Knowledge.' *Oxford Handbook of Epistemology*. Ed. P. K. Moser. 2002.

Stalnaker, R. 'The Problem of Logical Omniscience I.' *Synthese* 89 (1991): 425–440.

———. 'Nonmonotonic Consequence Relations.' *Fundamenta Informaticae* 21 (1994a): 7–21.

———. 'On the Evaluation of Solution Concepts.' *Theory and Decision* 37(42) (1994b).

———. 'On Logics of Knowledge and Belief.' *Philosophical Studies* 128 (2006): 169–199.

Vanderschraaf, P. and G. Sillari. 'Common Knowledge.' *The Stanford Encyclopedia of Philosophy*. Spring 2009 ed. Ed. E. N. Zalta. 2009.

Williamson, T. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.

Zhou, C. 'A Complete Deductive System for Probability Logic.' *Journal of Logic and Computation* 19(6) (2009): 1427–1454.