Epistemic Logic and Logical Omniscience: A Survey

Kwang Mong Sim*
Department of Computer Science and Engineering, Ho Sin-Hang Building, The Chinese University of Hong Kong, Shatin, NT, Hong Kong

This survey brings together a collection of epistemic logics and discusses their approaches in alleviating the logical omniscience problem. Of particular note is the logic of implicit and explicit belief. Explicit belief refers to information actively held by an agent, while implicit belief refers to the logical consequence of explicit belief. Ramifications of Levesque’s logic include nonstandard epistemic logic and the logics of awareness and local reasoning. Models of nonstandard epistemic logic are defined with respect to nonstandard proportional logic to weaken its semantics. In the logic of awareness, an agent can only believe a concept that it is aware of. Closely related to awareness are S-1 and S-3 epistemic operators which can be used to model skeptical and credulous agents. The logic of local reasoning provides a semantics for representing the fact that agents can have different clusters of beliefs which may contradict each other. Other variations include epistemic structures which are generalizations of the logic of local reasoning and fusion epistemic models which provide an account that agents can combine information conjunctively or disjunctively. Another closely related approach is the logic of explicit propositions which captures the insight that agents can hold beliefs independently without putting them together. © 1997 John Wiley & Sons, Inc.

I. INTRODUCTION

Reasoning about knowledge and belief were first studied in philosophy under the topics of epistemic (knowledge) and doxastic (belief) logic.1 Epistemic logic and doxastic logic are ramifications of modal logic.2 They are primarily logics that involve notions such as “knowing that” and “believing that”1 and the formal logical analysis of reasoning about these notions. The concepts of knowledge and belief are related but different; an agent cannot know a fact that is false but the agent can believe it.3 Since the logics discussed in this article address issues of representing information in the knowledge bases of intelligent agents which is typically not required to be true, the term belief seems to be

*E-mail: kmsim@cs.unimaas.nl; Tel: (31) 43 388-2023; Fax: (31) 43 325-2392.
more appropriate here. However, for the ease of exposition, this article prefers to use the two terms quite interchangeably throughout the discussion.

In addition, it is noteworthy to mention that aside from its long standing tradition in philosophy, reasoning about knowledge has also been investigated and applied in other disciplines such as game theory, distributed systems, and artificial intelligence. In game theory, reasoning about knowledge is crucial in decision making because each person (or player) affects everybody else. Therefore, one cannot make decisions without some knowledge or beliefs about what others will act or do. Reasoning about knowledge may provide useful insights in understanding and designing protocols in distributed systems. For instance, in a distributed system, one can view communication as a transformation of the system’s state of knowledge. The analysis and design of distributed protocols may be carried out by reasoning about the state of knowledge of groups of processors when the system goes through an execution.

A. Reasoning About Knowledge in Artificial Intelligence

In artificial intelligence, research in reasoning about knowledge focuses on designing intelligent agents that can reason about the current state of the world, beliefs of other agent’s beliefs and its own beliefs. The motivation for reasoning about knowledge in artificial intelligence, stems from areas such as planning, nonmonotonic reasoning, and natural language processing. For example, in multiagent planning an agent may need to predict how other agents will react in order to construct any complex plan. In multiagent nonmonotonic reasoning, an agent may need to reason about other agent’s default beliefs. In computational linguistics, the system needs to take into account the mental state of the person it is communicating with when interpreting and generating utterances. All these require the agent to know about the beliefs of other agents.

The literature in reasoning about knowledge in artificial intelligence forms a rich collection and expositions of (some representative examples of) these logics occupy the main bulk of this article. However, to set the stage for later discussions it seems prudent to review the classical approach for epistemic reasoning based on possible worlds.

II. CLASSICAL EPISTEMIC LOGIC

The idea of applying the possible-worlds semantics to model knowledge and belief was originally due to Hintikka. The intuitive idea of the possible-worlds paradigm is to acknowledge in the semantics that things might have happened differently from the way they did in fact happen. Thus, besides the true state of affairs (actual world), there are other possible states of affairs. Under this interpretation, an agent is said to know (or believe) a fact if it is true in all the states that it considers possible. The language that is employed to formalize these ideas is typically some variations of propositional modal logic. Such a language uses a set \( P \) of primitive letters (atomic sentences \( . . . , p, q, r, . . . \)) to represent knowledge in a world. Formulas from \( \text{PL} \) are formed by letters in \( P \)
Table I. Characteristic axioms and rules of inference of knowledge and belief.

<table>
<thead>
<tr>
<th>Name</th>
<th>Axiom Schema</th>
<th>Restriction on $R_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$</td>
<td>$K_i \alpha \land K_i(\alpha \supset \beta) \supset K_i\beta$</td>
<td>none</td>
</tr>
<tr>
<td>$T$</td>
<td>$K_i\alpha \supset \alpha$</td>
<td>reflexive</td>
</tr>
<tr>
<td>$D$</td>
<td>$K_i\alpha \supset \neg K_i \neg \alpha$</td>
<td>serial</td>
</tr>
<tr>
<td>$4$</td>
<td>$K_i\alpha \supset K_iK_i\alpha$</td>
<td>transitive</td>
</tr>
<tr>
<td>$5$</td>
<td>$\neg K_i\alpha \supset K_i \neg K_i\alpha$</td>
<td>euclidean</td>
</tr>
</tbody>
</table>

Rules of Inference

<table>
<thead>
<tr>
<th>Name</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP (Modus Ponen)</td>
<td>From $\alpha$ and $\alpha \supset \beta$ infer $\beta$</td>
</tr>
<tr>
<td>RN (Necessitation)</td>
<td>From $\alpha$ infer $K_i\alpha$</td>
</tr>
</tbody>
</table>

and logical connectives such as $\neg$ and $\land$, together with the modal operators $K_1, \ldots, K_n$. Thus, it follows that if $\alpha$ and $\beta$ are formulas of $\text{PL}$, then so are $\neg \alpha$, $\land \beta$, and $K_i\alpha$ (for $i = 1, \ldots, n$ where $K_i\alpha$ reads “agent $i$ knows $\alpha$”).

To give semantics to sentences in $\text{PL}$, a Kripke structure is usually employed. A Kripke Structure $M$ for a system of $n$ agents is a tuple $<S, \pi, R_1, \ldots, R_n>$ where $S$ is a set of possible worlds or states and $\pi$ associates each world in $S$ with a truth assignment. That is, for each state $w \in S$, $\pi(w)(p)$ is a mapping from a primitive letter $p \in P$ to {true, false}. Each $R_i$ is a binary (accessibility) relation on $S$ that captures the possibility relation according to agent $i$. Thus, $wR_iw_1$ holds if agent $i$ considers world $w_1$ possible when in world $w$. In addition, a relation $|=\,$ can be defined, where $M, w |= \alpha$ means “$\alpha$ is true at $M, w$” or “$\alpha$ holds at $M, w$”. Some of the clauses defined using $|=\,$ are as follows:\cite{4,14,15}

1. $M, w |= p$ iff $\pi(w)(p) =$ true.
2. $M, w = \neg \alpha$ iff $M, w \not|= \alpha$.
3. $M, w = \alpha \land \beta$ if both $M, w |= \alpha$ and $M, w |= \beta$.
4. $M, w |= K_i\alpha$ iff $M, w_1 |= \alpha$ for all $w_1$ such that $wR_iw_1$.

Clause (1) simply states that $\pi$ is used to define the semantics of primitive sentences and (2) and (3) are just recursive clauses for $\neg$ and $\land$. $M, w |= K_i\alpha$ means that agent $i$ knows a formula $\alpha$ at world $w$ in a structure $M$, if $\alpha$ is true in all worlds $w_i$ that agent $i$ considers possible in world $w$. Notions of validity and satisfiability are defined in the usual sense as in Ref. 2.

A. Axiom Systems for Knowledge and Belief

It has been noted\cite{4,14,15} that by putting different constraints on the accessibility relation $R_i$, different properties of knowledge can be captured. Some of the common characteristics of $R_i$ are best summarized in Table I. The axiom schemas $K, T, 4,$ and $5$ (shown in Table I) are sometimes called the distribution axiom, the
knowledge axiom, the positive introspection axiom, and the negative introspection axiom, respectively. While the distribution axiom states that knowledge is closed under implication, the knowledge axiom says that an agent can only know true facts (it is interesting to note that the knowledge axiom is sometime used to differentiate knowledge and belief). The positive and negative introspection axiom state that an agent knows (believes) what it knows (believes) and what it does not know (believe), respectively. An agent capable of both positive and negative introspection is said to be fully introspective. The standard logic of knowledge S5 usually contains the schema $K\alpha$, $T\alpha$, $4\alpha$, and $5\alpha$ together with all instances of propositional tautologies while the standard logic of belief is obtained by replacing the axiom $T$ with the axiom $D$. In addition, most logics of knowledge and belief usually have the two rules of inference: modus ponen (MP) and knowledge generalization or rule of necessitation (RN) (see Table I).

Although the possible-worlds approach can be easily altered to represent different properties of knowledge and belief, it suffers from what Hintikka termed the logical omniscience problem.

III. THE LOGICAL OMNISCIENCE PROBLEM

Logical omniscience requires an agent to know all logical consequences of its beliefs (that is, the set of beliefs held by the agent is closed under implication) and all valid sentences (including tautologies). The various problems associated with logical omniscience are described below.

Consequential closure. If an agent knows a set of formulas $\Gamma$ and if $\Gamma$ logically implies a formula $\alpha$ then the agent will also know $\alpha$ (that is, if $\Gamma$ and $\Gamma \supset \alpha$ are both true in every world then $\alpha$ must also be true in every world). The problem with closure under implication is that it does not consider what an agent believes directly but what the world would be like if what it believed were true.

Irrelevant beliefs. Under the possible-worlds interpretation, a valid sentence is one that is true in every world that the agent considers possible. This will also mean that the agent will need to know all tautologies in addition to its active beliefs. For instance, aside from what it already knows, the agent should also know tautologies such as “it is raining in Boston at this instant or it is not raining in Boston at this instant.” The problem is that the possible-worlds approach does not distinguish between an agent’s active beliefs and tautologies that are irrelevant to the agent’s beliefs. One way to block irrelevant beliefs is to only consider beliefs that are actively held by the agent. This may be achieved by not assigning truth values to sentences that are not relevant to the agent’s belief set.

Inconsistent beliefs. Additionally, an agent cannot believe both a sentence $\alpha$ and its negation $\neg\alpha$ without believing every sentence. This follows from the fact that none of the possible worlds is consistent with $(\alpha \land \neg\alpha)$ [since $(\alpha \land \neg\alpha)$ is false in every world]. One of the possible solutions is to allow worlds that support both the truth and the falsity of a sentence and thus allowing inconsistent beliefs to be satisfiable.
Computational intractability. Computationally, these problems result in intractability because the agent is required to compute all logical consequences of its beliefs. In practice, resource-bounded agents usually do not have enough time or memory to derive an explicit representation of each fact. Hintikka\textsuperscript{18} remarked that the possible-worlds semantics is unsuitable for modeling human reasoning because humans do not seem to be logically omniscient.

A. Approaches to Mitigate Logical Omniscience

It has been noted that logical omniscience is perhaps one of the most hotly contested topics in reasoning about knowledge.\textsuperscript{12} In the last decade, there has been a resurgence of work (including some best-article prizes at IJCAI and AAAI conferences) to mitigate the logical omniscience problem. They are roughly classified in Ref. 3 under the syntactic approach, the semantic (or Montague–Scott) approach, the impossible-worlds approach and the nonstandard propositional logic approach.

Syntactic approach. One method for distinguishing beliefs is by examining their syntactic shapes. Levesque\textsuperscript{19} called this the syntactic approach. In this approach, an epistemic model is a structure consisting of an explicit set of sentences (that is not necessarily closed under logical consequence); each formula that an agent believes is a member of this set. For instance, under this interpretation, an agent may believe \( p \lor q \) without believing \( p \land (r \lor \neg r) \). This approach was pursued by Eberle,\textsuperscript{20} Moore and Hendrix.\textsuperscript{21} Another variant of this approach is considered by Konolige\textsuperscript{22} who represents an agent’s beliefs by a set of base facts that is closed under a set of deduction rules. Under this approach, logical omniscience is avoided by allowing the set of deduction rules to be incomplete.

Montague–Scott approach. Another approach, due to Montague\textsuperscript{23,24} and Scott\textsuperscript{25} represents an agent’s beliefs by a set of sets (or clusters) of possible worlds. A Montague–Scott model is sometimes referred to as a minimal (or neighborhood) model.\textsuperscript{2} A brief exposition of the Montague–Scott model is given in Section IV-B. In this approach every proposition is identified with a set of possible worlds; an agent knows a proposition \( \alpha \) if \( \alpha \) is true in at least one of the sets of worlds that it considers possible. Some of the problems of logical omniscience are avoided because each cluster is not closed under its superset. Vardi’s\textsuperscript{26} epistemic structures pursues such an approach. Other variants includes cluster models (the logic of local reasoning\textsuperscript{17,27}) and Vardi’s fusion epistemic models.\textsuperscript{28} A description of epistemic structures, the logic of local reasoning, and fusion epistemic models will be given in Sections IV-B.1, IV-B.2, and IV-B.3, respectively.

Impossible-worlds approach. Still another approach called the impossible-worlds approach,\textsuperscript{29-33} augments possible worlds with impossible worlds. An impossible world is a world in which valid formulas are not necessarily true or in which inconsistent formulas may be true\textsuperscript{33} (where valid and inconsistent refer to logical validity and inconsistency in the sense of classical logic). However, impossible worlds are worlds that are admitted only as epistemic alternatives but are not logically possible.\textsuperscript{18} Thus, in this approach an agent may not know all tautologies.
of classical logic because there may be impossible worlds that it considers possible where these tautologies fail to be true. A variant of this approach is Levesque’s\textsuperscript{19} logic of implicit and explicit belief (a detailed description will be given in Sec. IV-A.1).

**Nonstandard logic approach.** Finally, another approach which is based on a nonstandard logic is proposed by Fagin, Halpern, and Vardi.\textsuperscript{34} In this approach, possible worlds are models of a nonstandard propositional logic (NPL). A slightly more detailed description of NPL will be given in Section IV-A.5.

**IV. EPISTEMIC LOGICS IN ARTIFICIAL INTELLIGENCE**

In the design of intelligent systems, a logic that is used to model finite agents operating in a real-time environment should take into account their limited reasoning capabilities. This section describes some of the research in artificial intelligence that deals with the knowledge representation and reasoning of non-omniscient (resource-bounded) agents. It elaborates some of the approaches for alleviating the problem of logical omniscience (mentioned in Sec. III-A). As mentioned earlier, the literature in reasoning about knowledge form a very large collection and space limitation precludes introducing all of them here. Thus, the logics that are presented in this article clearly reflects my own bias. However, it serves the emphasis to mention that for ease of exposition, this article adopts the position that classes of epistemic logics are identified by the structure of their semantic models (that is either the Kripke structure or the Montague–Scott structure).

**A. Epistemic Logic Based on Kripke Structure**

Although Levesque’s\textsuperscript{19} logic of implicit and explicit belief is a variant of the impossible-worlds approach, the semantic model is essentially a Kripke structure. This is also the case for Fagin, Halpern, and Vardi’s\textsuperscript{34} nonstandard epistemic logic (Sec. IV-A.5), where a Kripke structure is used to give semantics to a nonstandard modal propositional logic rather than a model propositional logic. In addition, there were several outgrowths of Levesque’s logic that employed some variant of the Kripke structure. Among them were the logic of awareness (Sec. IV-A.3), the Cadoli–Schaerf epistemic model (Sec. IV-A.4) and Lakemeyer’s\textsuperscript{35} extension to Levesque’s model (Sec. IV-A.2).

1. Levesque’s Logic of Implicit and Explicit Belief

In his model, Levesque defined explicit belief as those sentences actively held by an agent and implicit belief as logical consequence of its explicit belief. His idea of avoiding some aspects of logical omniscience is to relax the condition that worlds have to be complete and consistent. This is achieved by decoupling the semantics of a formula with two independent notions of truth support and falsity support. In this approach, a sentence in a given situation can be true supported, false supported, both true and false supported, and neither true nor false supported.
Thus, a situation semantics\textsuperscript{36} rather than a possible-worlds semantics is employed. The main deviation from a possible-worlds semantics is that incomplete situations (where a sentence can be neither true or false) and incoherent situations (where a sentence can be both true and false) are allowed. A complete situation (or possible world) is one that is not incoherent and not incomplete.

An epistemic model for implicit and explicit belief is a tuple $M_L = \langle S, \Sigma, T, F \rangle$ where $S$ is a set of all situations and $\Sigma$ is any subset of $S$. $T$ and $F$ are functions from $P$ (an enumerably infinite set of atomic sentences) to $P(S)$, such that for any sentence $p \in P$, $T(p)$ is those situations that support the truth of $p$ and $F(p)$ is those that support the falsity of $p$. Thus, a situation is incoherent if it is in the set $T(p) \cap F(p)$. Consequently, there are two support relations $|=_{\top}$ and $|=_{\bot}$ between situations and formulas from a propositional language $PL$. $PL$ is a language that is formed using $P$ and the standard connectives $\wedge, \neg, \lor, = =$ together with two epistemic operators $L$ and $B$ (representing implicit belief and explicit belief, respectively); the only restriction is that sentences containing $B$ or $L$ cannot occur within the scope of the two epistemic operators. Thus, for all $\alpha \in PL, M_L, s |=_{\top} \alpha$ means that $s$ supports the truth of $\alpha$ and $M_L, s |=_{\bot} \alpha$ means that $s$ supports the falsity of $\alpha$. In Ref. 19, the following semantics is defined:

\begin{align*}
(L\text{-}p\top) & & M_L, s |=_{\top} p \text{ iff } s \in T(p) \\
(L\text{-}p\bot) & & M_L, s |=_{\bot} p \text{ iff } s \in F(p) \\
(L\text{-}\neg\top) & & M_L, s |=_{\top} \neg \alpha \text{ iff } M_L, s |=_{\bot} \alpha \\
(L\text{-}\neg\bot) & & M_L, s |=_{\bot} \neg \alpha \text{ iff } M_L, s |=_{\top} \alpha \\
(L\text{-}\land\top) & & M_L, s |=_{\top} \alpha \land \beta \text{ iff } M_L, s |=_{\top} \alpha \text{ and } M_L, s |=_{\top} \beta \\
(L\text{-}\land\bot) & & M_L, s |=_{\bot} \alpha \land \beta \text{ iff } M_L, s |=_{\bot} \alpha \text{ or } M_L, s |=_{\bot} \beta \\
(L\text{-}B\top) & & M_L, s |=_{\top} B\alpha \text{ iff } M_L, t |=_{\top} \alpha \text{ for all } t \in \Sigma \\
(L\text{-}B\bot) & & M_L, s |=_{\bot} B\alpha \text{ iff } M_L, s \neq t B\alpha. \\
(L\text{-}L\top) & & M_L, s |=_{\top} L\alpha \text{ iff } M_L, t |=_{\top} \alpha \text{ for all } t \in W(\Sigma) \\
(L\text{-}L\bot) & & M_L, s |=_{\bot} L\alpha \text{ iff } M_L, s \neq t L\alpha.
\end{align*}

In $(L\text{-}LT)$, $W(\Sigma)$ is the set of complete situations where each possible world $w \in W(\Sigma)$ is compatible with some situation $s$ in $\Sigma$. A possible world $w$ is compatible with another situation $s$ if every sentence that is true (respectively, false) in $s$ is also true (respectively, false) in $w$. More formally, for each $s \in S$ the function $W$ is defined as follows:

$W(s) = \{ w \in S \mid$ for every $p \in P,$

| (i) $w$ is a member of exactly one of $T(p)$ and $F(p)$ |
| (ii) if $s \in T(p)$ then $w \in T(p)$ |
| (iii) if $s \in F(p)$ then $w \in F(p)$ |
Also, for any subset $\Sigma$ of $S$, $W(\Sigma) = \cup_{s \in \Sigma} W(s)$. Thus, $W(\Sigma)$ is simply a set of possible worlds. Since implicit belief is defined with respect to a set of possible worlds, it is clear that the $L$ operator is closed under implication. Also, all tautologies and valid sentences are implicitly believed. In Levesque’s logic, a formula $\alpha$ is valid (written $\models \alpha$) if it is true in every complete situation (possible world) $s$ in every model $M_L$. A sentence $\alpha$ is true or is satisfied at a situation $s$ if $s \models \alpha$ holds.

While implicit belief suffers from logical omniscience, explicit belief defined in (L-BT) and (L-BF) seems to have nonomniscient properties. In particular, the following sets of sentences given in (LO1), (LO2), (LO3), and (LO4) are satisfiable.

\[(LO1) \quad B\alpha \land B(\alpha \supset \beta) \land \neg B\beta\]
\[(LO2) \quad B\alpha \land B(\neg \alpha) \land \neg B\beta\]
\[(LO3) \quad \neg B(\alpha \lor \neg \alpha)\]
\[(LO4) \quad B\alpha \land \neg B(\alpha \land (\beta \lor \neg \beta))\]

**Avoiding consequential closure.** In the classical possible-worlds semantics, the assumption that $(\alpha \supset \beta)$ is logically valid means that $\beta$ is true in every interpretation (possible worlds) in which $\alpha$ is true. Hintikka’s original idea to solve the logical omniscience problem is to allow epistemic alternatives (or worlds) in which $\beta$ is false when both $(\alpha \supset \beta)$ and $\alpha$ are true. Levesque’s notion of explicit belief admits such epistemic alternatives in that (LO1) is satisfiable (Fig. 1). In Figure 1, $s \models \alpha$ and $s \models B(\alpha \supset \beta)$ but $s \not\models B\beta$. Thus, explicit belief is not closed under implication. It is noted that the situation $s_1$ (shown in Fig.
1) is an incoherent situation. The fact that (LO1) is satisfiable is due to the presence of some incoherent situations.

**Admitting inconsistent beliefs.** Satisfiability of (LO2) means that an agent can explicitly believe inconsistent beliefs such as $B\alpha \land B\neg\alpha$ and without believing every sentence (this is illustrated in Fig. 2). In Figure 2, $s \models B\alpha$ and $s \models B(\neg\alpha)$ but $s \not\models B\beta$. However, inconsistent beliefs are only possible if every visible situation is incoherent.

**Blocking tautologies and closure under valid implications.** (LO3) states that tautologies are not explicitly believed, while (LO4) says explicit belief is not closed under valid implication, thus, even though $\alpha \supset (\beta \lor \neg\beta)$ is valid, $B\alpha \supset B(\alpha \land (\beta \lor \neg\beta))$ is not. (LO3) is illustrated in Figure 3. The satisfiability of (LO3) and (LO4) stems from the presence of incomplete situations. In Figure 3 for instance, neither the truth nor the falsity of $\alpha$ is supported in $s_3$. Thus, it follows that $S \not\models B(\alpha \lor \neg\alpha)$.

**Proof theory for explicit belief.** Beside having nonomniscient properties, a proof theory is required to reason about explicit belief. The proof theory for explicit belief rests heavily on the tautological entailment in relevance logic. In fact, the idea of a situation bears strong similarities to the notion of set-ups in relevance logic. Levesque’s contribution was to make the connection between explicit belief and tautological entailment by proving that $\models (B\alpha \supset B\beta)$ iff $\alpha$ entails $\beta$. This means that the logic of explicit belief contains relevance logic as a subpart and thus the constraints of explicit belief can be characterized by a set of axioms that characterize tautological entailment.

**Comments.** In this connection, agents in Levesque’s logic are committed to believe all sentences that are tautologically entailed by its knowledge base. This
means that they are perfect reasoners in relevance logic, as Vardi\textsuperscript{28} has pointed out. On this account, it seems that the logic of implicit and explicit belief is psychologically implausible for modeling human reasoning since it is no more plausible to expect people to reason perfectly in relevance logic than in classical logic. In addition, Levesque’s logic is rather restrictive because it can only be applied to a single agent environment. This issue has been taken up by Fagin and Halpern\textsuperscript{17,27} in their logic of awareness and logic of local reasoning.

2. Logic of (Nested) Implicit and Explicit Belief

In his model, Lakemeyer\textsuperscript{35} had employed not one but two accessibility relations. A Lakemeyer model for implicit and explicit belief is a tuple \( \mathbf{M}_{\mathbf{IL}} = (S, T, F, R^+, R^-) \) where \( S, T, \) and \( F \) are defined as in Section IV-A.1 and \( R^+ \) and \( R^- \) are binary relations on \( S \). \( R^+ \) is used when an agent is confirming a belief while \( R^- \) is used to determine a disbelief. These two accessibility relations capture the intuition that in a given situation, an agent uses different sets of possible situations to confirm or disconfirm a belief. However, in a possible world, \( R^+ \) and \( R^- \) coincide.

**Definition 1.** For all \( w \in W(\Sigma) \) and for all situations \( s \in S \): \( wR^+s \iff wR^-s \).

The language \( \mathbf{PL} \) that is employed is similar to those in Section IV-A.1 except that nesting of modal operators is allowed. In this case sentences that do not contain \( \mathbf{B} \) and \( \mathbf{L} \) are called \textit{objective} sentences and sentences that occur in the scope of a \( \mathbf{B} \) or \( \mathbf{L} \) are called \textit{subjective} sentences. However, there is a constraint...
stating that no \( L \) may occur within the scope of a \( B \). Thus, for all \( \psi \in PL \) the following semantics was defined:\(^{35}\)

\[
\begin{align*}
(\text{LL-T}p) & \quad M_{LL}, s =^t p \iff s \in T(p) \\
(\text{LL-F}p) & \quad M_{LL}, s =^f p \iff s \in F(p) \\
(\text{LL-T}¬) & \quad M_{LL}, s =^t \neg \alpha \iff M_{LL}, s =^f \alpha \\
(\text{LL-F}¬) & \quad M_{LL}, s =^f \neg \alpha \iff M_{LL}, s =^t \alpha \\
(\text{LL-T}Λ) & \quad M_{LL}, s =^t \alpha \land \beta \iff M_{LL}, s =^t \alpha \text{ and } M_{LL}, s =^t \beta \\
(\text{LL-F}Λ) & \quad M_{LL}, s =^f \alpha \land \beta \iff M_{LL}, s =^f \alpha \text{ or } M_{LL}, s =^f \beta \\
(\text{LL-TB}) & \quad M_{LL}, s =^t B_\alpha \iff \text{for all } t \in \Sigma \text{ sR} + t \Rightarrow M_{LL}, t =^t \alpha \\
(\text{LL-FB}) & \quad M_{LL}, s =^f B_\alpha \iff \text{for some } t \in \Sigma \text{ sR} - t \land M_{LL}, t =^f \alpha \\
(\text{LL-TL}) & \quad M_{LL}, s =^t L_\alpha \iff \text{for all } t \in W(\Sigma) \text{ sR} + t \Rightarrow M_{LL}, t =^t \alpha \\
(\text{LL-FL}) & \quad M_{LL}, s =^f L_\alpha \iff M_{LL}, s =^f \neg \alpha.
\end{align*}
\]

The definitions of (LL-Tp), (LL-Fp), (LL-T¬), (LL-F¬), (LL-TΛ), and (LL-FΛ) are similar to those in Section IV-A.1. The only difference is that in (LL-TB), \( B_\alpha \) is true supported at \( s \) if all situations accessible from \( s \) via \( R^+ \) support the truth of \( \alpha \). For (LL-FB), the situation \( s \) supports the falsity of \( B_\alpha \) just in case if there is a situation that is accessible from \( s \) via \( R^- \) which does not support the truth of \( \alpha \).

**Nested beliefs.** For explicit belief, the properties (LO1), (LO2), (LO3), and (LO4) (presented in Section IV-A.1) are still retained. Furthermore, these properties also hold for nested sentences within the limit of the language above. For instance, \( \neg B(\neg B_\alpha \lor \neg B_\alpha) \) is satisfiable.

### 3. Logic of Awareness

Other researchers beside Lakemeyer, such as Fagin and Halpern\(^{17,27}\) have also extended Levesque’s logic. In their logic of awareness, Fagin and Halpern deal with both the issues of nested beliefs and multiagent reasoning. The language that is employed is a propositional language \( PL \), built from a countably infinite set \( P \) of atomic letters \( \{ p_1, p_2, p_3, \ldots \} \) and the usual logical connectives \( \neg, \land, \lor, \implies, \iff \), together with two sets of modal operators \( L_1, \ldots, L_n \) and \( B_1, \ldots, B_n \). Since the logic of awareness deals with multiagent reasoning, a set \( A \) of \( n \) agents is assumed and \( i \) is an index referring to agents in \( A \). Thus, for \( i = 1, \ldots, n \), \( B_i \alpha \) reads “agent \( i \) explicitly believes \( \alpha \)” and \( L_i \alpha \) means “agent \( i \) implicitly believes \( \alpha \)” Additionally, arbitrary nesting of the \( L_i \) and \( B_i \) in formulas are also permitted since nested beliefs among different agents are allowed in the logic of awareness.

A Kripke structure for awareness\(^{17,27}\) is a tuple \( M_A = (S, A_1, \ldots, A_n, R_1, \ldots, R_n, \pi) \) where \( S \) is a set of possible worlds, \( \pi \) is a truth assignment \((\pi: P \rightarrow \{true, false\})\), \( R_1, \ldots, R_n \) are serial, transitive, and euclidean relations
on \( S \) and \( A_i \) is a function associating each possible world with a set of primitive formulas \( P \subseteq \Psi \), where \( P \) is the set of all primitive sentences. That is, \( A_i(s) \) is the set of primitive formulas that agent \( i \) is aware of at state \( s \). Two support relations \( \models_A \) and \( \models_T \) are defined for each set of formula \( \Psi \). This places a restriction such that in every state only the formulas in \( \Psi \) are defined with either \( \text{true} \) or \( \text{false} \). \( \models_A \) and \( \models_T \) together with the \( A_i(s) \) function provide some effect of partial situations from the perspective of agent \( i \). For instance, a sentence \( \alpha \) may be true at state \( s \), but agent \( i \) may not be aware of \( \alpha \) at \( s \) (that is \( \alpha \not\in A_i(s) \)). Thus, for all \( \psi \in PL \) the following semantics for the logic of awareness is defined:\textsuperscript{27,29}

\[
\begin{align*}
(A-pT) & \quad M_\alpha, s \models_T p \iff \pi(p, s) = \text{true} \; \text{and} \; p \in \Psi \\
(A-pF) & \quad M_\alpha, s \models_F p \iff \pi(p, s) = \text{false} \; \text{and} \; p \in \Psi \\
(A-p) & \quad M_\alpha, s \models p \iff \pi(p, s) = \text{true} \\
(A-\lnot T) & \quad M_\alpha, s \models \lnot \alpha \iff M_\alpha, s \models \lnot \alpha \\
(A-\lnot F) & \quad M_\alpha, s \models \lnot \alpha \iff M_\alpha, s \models \lnot \alpha \\
(A-\lnot) & \quad M_\alpha, s \models \lnot \alpha \iff M_\alpha, s \not\models \alpha \\
(A-\land T) & \quad M_\alpha, s \models \alpha \land \beta \iff M_\alpha, s \models \alpha \; \text{and} \\ & \quad M_\alpha, s \models \beta \\
(A-\land F) & \quad M_\alpha, s \models \alpha \land \beta \iff M_\alpha, s \models \alpha \; \text{or} \; M_\alpha, s \models \beta \\
(A-\land) & \quad M_\alpha, s \models \alpha \land \beta \iff M_\alpha, s \models \alpha \; \text{and} \; M_\alpha, s \models \beta \\
(A-B, T) & \quad M_\alpha, s \models T B \alpha \iff M_\alpha, t \models T \alpha \; \text{for all} \; t \; \text{such that} \; (s, t) \in R, \\
(A-B, F) & \quad M_\alpha, s \models F B \alpha \iff M_\alpha, t \models F \alpha \; \text{for some} \; t \; \text{such that} \; (s, t) \in R, \\
(A-B, c) & \quad M_\alpha, s \models B \alpha \iff M_\alpha, t \models T B \alpha, \; (P \; \text{is the set of all primitive letters.}) \\
(A-L, T) & \quad M_\alpha, s \models T L \alpha \iff M_\alpha, t \models T \alpha \; \text{for all} \; t \; \text{such that} \; (s, t) \in R, \\
(A-L, F) & \quad M_\alpha, s \models F L \alpha \iff M_\alpha, t \models F \alpha \; \text{for some} \; t \; \text{such that} \; (s, t) \in R, \\
(A-L, c) & \quad M_\alpha, s \models L \alpha \iff M_\alpha, t \models \alpha \; \text{for all} \; t \; \text{such that} \; (s, t) \in R, 
\end{align*}
\]

Clauses \((A-pT)\) and \((A-pF)\) give the property that for each set of primitive sentences \( \Psi \):

\[
\begin{align*}
(i) & \quad M_\alpha, s \models T \alpha \Rightarrow M_\alpha, s \models \alpha \\
(ii) & \quad M_\alpha, s \models F \alpha \Rightarrow M_\alpha, s \models \lnot \alpha 
\end{align*}
\]

Both \((i)\) and \((ii)\) are proven in Ref. 27. Clauses \((A-\lnot)\) and \((A-\land)\) are just recursive definitions of \((A-p)\). In clause \((A-B, T)\), \( M_\alpha, s \models T B \alpha \) means that agent \( i \) at state \( s \) explicitly believes \( \alpha \) relative to \( \Psi \), if \( \alpha \) is true relative to \( \Psi \cap A_i(s) \) at all the states accessible from \( s \). It also implies that \( B \alpha \) is true at \( s \). This is because \( M_\alpha, s \models T B \alpha \Rightarrow M_\alpha, s \models B \alpha \) follows from property \((i)\) above. In the case of \( M_\alpha, s \models B \alpha \), agent \( i \) explicitly believes \( \alpha \) relative to \( P \), where as for \( M_\alpha, s \models T B \alpha \)
Figure 4. Awareness.

agent $i$ explicitly believes $\alpha$ relative to some subset of $\mathbf{P}$. *Implicit* belief is defined in clause (A-L) does not take the awareness function into account. Thus, $L_i \alpha$ is true at $s$ if $\alpha$ is true at all states accessible from $s$, regardless of whether agent $i$ is aware of $\alpha$. In addition, a sentence $\alpha$ is valid if $M_A, s \models \alpha$ for all models $M_A$ and all states $s$; $\alpha$ is satisfiable if $\neg \alpha$ is not valid.

**Example.** An example to illustrate implicit and explicit belief in the context of awareness is shown in Figure 4. Let $\Psi = \{\alpha, \beta\}$ be a subset of $\mathbf{P}$. That is, for every state, letters in $\Psi$ are defined with either true or false. A further restriction is provided by $A_i(s)$, which can be thought of as a syntactic filter. In this example, $A_i(s) = \{\alpha\}$. Thus, in Figure 4, agent $i$ explicitly believes $\alpha$ relative to $\Psi$ at state $s$ because in every state accessible from $s$ (that is $s_1$, $s_2$, and $s_3$) $\alpha$ is true relative to $\Psi \cap A_i(s) = \{\alpha\}$. Although $\beta$ is true in $s_1$, $s_2$, and $s_3$, agent $i$ does not explicitly believe $\beta$ because it is not aware of $\beta$ [that is, $\beta$ is not in $\Psi \cap A_i(s)$]. Thus, in Figure 4, it follows that $M_A, s \not\models B_i \beta$ and $M_A, s \not\models B_i \beta$ and, consequently, $M_A, s \not\models B_i \beta \lor \neg B_i \beta$. However, agent $i$ implicitly believes $\beta$ since it is true in all accessible states and the awareness function is not taken into account. Similarly, agent $i$ also believes $\alpha$ implicitly.

**Avoiding logical omniscience.** Some of properties of Levesque’s explicit belief are present in the logic of awareness. For instance (LO3) and (LO4) are still satisfiable. Further, the example in Figure 4 shows that formula such as $\neg B_i (B_i \beta \lor \neg B_i \beta)$ (which is in the form of (LO3) but with arbitrary nesting of
beliefs) is satisfiable. However, since Fagin and Halpern dispense the use of incoherent situations, an agent cannot hold inconsistent beliefs and explicit belief is closed under implication in this logic. Thus, both (LO1) and (LO2) are not satisfiable in the logic of awareness.

4. Cadoli–Schaerf Episemic Model

In the previous section, the awareness function was used as a syntactic filter. The approach adopted by Cadoli and Schaerf is somewhat similar. A set of countably infinite atomic letters $P = \{p_1, p_2, p_3, \ldots \}$ is assumed but with the following interpretation:

**DEFINITION 2.**

(1) A 3-interpretation of $P$ is a truth assignment $I_3$ which maps every letter $p \in P$ to \{true, false, T\} (2) A 2-interpretation of $P$ is a truth assignment $I_2$ which maps every letter $p \in P$ to \{true, false\} (3) A 1-interpretation of $P$ is a truth assignment $I_1$ which maps every letter $p \in P$ to \{false\}

A generalized notion of 1- and 3- interpretation can be defined by restricting the possibility of mapping some subsets of primitive letters $S$ to $\bot$ and $T$, where $P \supset S$, giving the following definition:

**DEFINITION 3.**

(1) An $S$-3 interpretation of $P$ is a truth assignment $I_3$ which maps every letter of $S$ to \{true, false\} and $P \setminus S$ to \{true, false, T\} where $P \supset S$. Thus, for every letter $p \in P$ there are three possibilities:
   (i) $I_3(p) = 1$, $I_3(\neg p) = 0$
   (ii) $I_3(p) = 0$, $I_3(\neg p) = 1$
   (iii) $I_3(p) = 1$, $I_3(\neg p) = 1$ iff $p \in P \setminus S$

(2) An $S$-1 interpretation of $P$ is a truth assignment $I_1$ which maps every letter of $S$ to \{true, false\} and $P \setminus S$ to \{\bot\} where $P \supset S$. Thus, for every letter $p \in P$ there are three possibilities:
   (i) $I_1(p) = 1$, $I_1(\neg p) = 0$ iff $p \in S$
   (ii) $I_1(p) = 0$, $I_1(\neg p) = 1$ iff $p \in S$
   (iii) $I_1(p) = 0$, $I_1(\neg p) = 0$ iff $p \in P \setminus S$

Modal Extensions to Definition 3 are given as follows:

**S-3-Kripke interpretation.** An S-3-Kripke model is a tuple $M = (Sit, R, I_3)$, where $Sit$ is a set of situations, $R$ is an accessibility relations on $Sit$, and $I_3$ is an S-3-interpretation for every situation in $Sit$.

**S-1-Kripke interpretation.** An S-1-Kripke model is a tuple $M = (Sit, R, I_1)$ where $Sit$ is a set of situations, $R$ is an accessibility relations on $Sit$, and $I_1$ is an S-1-interpretation for every situation in $Sit$.

Cadoli and Schaerf proposed that the ideas of S-3 and S-1 Kripke interpretation can also be used to model nonomniscient agents.
**Modeling Resource-bounded agents.** A Cadoli–Schaerf model for epistemic reasoning is a tuple $M_{CS} = \langle \text{Sit}, V, R \rangle$ where $\text{Sit}$ is a set of situations and $R$ is a reflexive, transitive, and euclidean relation on $\text{Sit}$. $V$ is a valuation mapping situations to truth assignments ($V: \text{Sit} \rightarrow \{T, I, F\}$). A situation with an S-3 interpretation (respectively, S-1 interpretation) is called an S-3 (respectively, S-1) situation while a possible world has a 2-interpretation. The set of S-1 situations, S-3 situations, and possible worlds are denoted $S-1(\text{Sit})$, $S-3(\text{Sit})$, and $W(\text{Sit})$, respectively. In addition, a propositional language $PL$ (built from $P$ together with the usual connectives $\wedge$ and $\vee$, and two modal operators $K^1$ and $K^2$) is assumed. Thus, for all $\psi \in PL$ the following semantics was defined:

\[
\begin{align*}
\text{(CS-p)} & \quad M_{CS}, s \models p \iff V(s)(p) = 1, \text{ where } p \text{ is atomic} \\
\text{(CS-\neg)} & \quad M_{CS}, s \models \neg \alpha \iff V(s)(\neg \alpha) = 1
\end{align*}
\]

$\text{(CS-p)}$ and $\text{(CS-\neg)}$ can be expanded as follows (sentences of the form $\alpha \wedge \beta$ may also be inductively defined in the usual fashion):

\[
\begin{align*}
\text{(CS-I)} & \quad M_{CS}, s \models p \iff s \in S-3(\text{Sit}) \text{ and } I^1(s)(p) = 1 \\
\text{(CS-\neg I)} & \quad M_{CS}, s \models \neg \alpha \iff s \in S-3(\text{Sit}) \text{ and } I^1(s)(\neg \alpha) = 1 \\
\text{(CS-\neg I)} & \quad M_{CS}, s \models \neg \alpha \iff s \in S-3(\text{Sit}) \text{ and } I^1(s)(\neg \alpha) = 1 \\
\text{(CS-F)} & \quad M_{CS}, s \models p \iff s \in W(\text{Sit}) \text{ and } F(s)(p) = 1 \\
\text{(CS-F)} & \quad M_{CS}, s \models \neg \alpha \iff s \in W(\text{Sit}) \text{ and } F(s)(\neg \alpha) = 1
\end{align*}
\]

The modal operator $K^2$ is defined over a set of S-3 situations, while the $K^1$ modal operator is taken over a set of S-1 situations as stated below.

\[
\begin{align*}
\text{(CS-K)} & \quad M_{CS}, s \models K^2 \alpha \iff \forall t \in S-3(\text{Sit}) \ sRt \Rightarrow M_{CS}, t \models \alpha \\
\text{(CS-\neg K)} & \quad M_{CS}, s \models \neg K^2 \alpha \iff \exists t \in S-3(\text{Sit}) \ sRt \wedge M_{CS}, t \models \neg \alpha \\
\text{(CS-K)} & \quad M_{CS}, s \models K^1 \alpha \iff \forall t \in S-1(\text{Sit}) \ sRt \Rightarrow M_{CS}, t \models \alpha \\
\text{(CS-\neg K)} & \quad M_{CS}, s \models \neg K^1 \alpha \iff \exists t \in S-1(\text{Sit}) \ sRt \wedge M_{CS}, t \models \neg \alpha
\end{align*}
\]

$K^1$ was designed as a filter for selecting complete but not necessarily coherent situations while $K^2$ was meant to select situations that are coherent but not necessarily complete. In this regard, the Cadoli–Schaerf model can be viewed as an extension of Levesque’ model. For $K^1$, it is easily checked that both $B\alpha \wedge B(\alpha \supset \beta) \wedge \neg B\beta$ and $B\alpha \wedge B(\neg \alpha \vee \beta)$ are satisfiable. However, $\neg B(\alpha \vee \neg \alpha)$ and $B\alpha \wedge B(\alpha \wedge \neg \beta)$ are both unsatisfiable because for every sentence $\alpha \in PL$ in an S-3 situation, at least one of $\alpha$ and $\neg \alpha$ must be true. The converse holds for the $K^2$ operator. Additionally, it is noteworthy to mention that the $K^2$ and the $K^1$ operator can be used to model skeptical agents and credulous agents, respectively.
5. Nonstandard Epistemic Logic

Levesque’s notion of an incoherent situation has been criticized as unintuitive.28 Fagin, Halpern, and Vardi34 argue that in the real world formulas can only either be true or false. In fact, their nonstandard epistemic logic (NE) was proposed as an alternative to Levesque’s model to achieve the same results but without using incoherent and incomplete situations. Fagin et al.’s approach is based on a nonstandard propositional logic (NPL). The semantics of NPL is defined such that each world \( w \) is associated with a sister world \( w^* \). While the truth of a formula \( a \) is determined by \( w \), the semantics of negated formulas \( \neg a \) is given by \( w^* \). Thus, \( \neg a \) is true in \( w \) if \( a \) is not true in \( w^* \). However, if \( w = w^* \), then the definition of negation is as the usual sense and for each world \( w \), \( w^{**} = w \). This approach which was due to Routley and Routley39 has a similar effect of decoupling the truth values of a formula and its negation (which was one of the novel features of Levesque’s model). The difference is that in this approach possible worlds are used as models of nonstandard propositional logic and this leads to the idea of a nonstandard Kripke structure.

A nonstandard Kripke structure for a system of \( n \) agents is a tuple \( M_{NE} = (S, \pi, R_1, \ldots, R_n, *) \) where \( S \) is a set of possible states or worlds, \( \pi \) is a truth assignment to the primitive sentences for each world \( s \in S \) and \( R_1, \ldots, R_n \), are binary relations on \( S \) and * is defined as above. A propositional language PL as defined in Section IV-A.3 is assumed. Thus, for all \( \psi \in PL \) the following semantics for an NE model is defined:

- **(NE-p)** \( M_{NE}, s \models p \) iff \( \pi(s)(p) = \text{true} \)
- **(NE-\neg)** \( M_{NE}, s \models \neg \alpha \) iff \( M_{NE}, s^* \not\models \alpha \)
- **(NE-\land)** \( M_{NE}, s \models \alpha \land \beta \) iff \( M_{NE}, s \models \alpha \) and \( M_{NE}, s \models \beta \)
- **(NE-K)** \( M_{NE}, s \models K\alpha \) iff \( M_{NE}, t \models \alpha \) for all \( t \) such that \( (s, t) \in R_i \)

The definitions of (NE-p), (NE-\neg), and (NE-K) are similar to those in the standard Kripke structure. The main deviation is the definition of negation (NE-\neg). For instance, if \( \neg K\alpha \) holds at \( s \) then \( M_{NE}, s^* \not\models K\alpha \) which means that there is a world \( t \) that is accessible from \( s^* \) (instead of \( s \)) such that \( M_{NE}, t \not\models \alpha \).

B. Epistemic Logic Based on Minimal (Montague–Scott) Models

Throughout the literature in epistemic logic, there were several epistemic models that were based on minimal models. A minimal (Montague–Scott) model is a structure \( M = (W, N, P) \) where \( W \) is a set of possible worlds, \( N \) is a function assigning to each possible world \( s \in W \) a collection of sets of possible world (\( N: W \rightarrow \mathcal{P}(\mathcal{P}(W)) \)). \( I \) is an intensional assignment which associates each atomic sentence \( p \in P \) with a set of possible worlds (\( I: P \rightarrow \mathcal{P}(\mathcal{P}(W)) \)). In this case of an arbitrary sentence \( \alpha \), its truth set is expressed as \( \|\alpha\|^M = \{w \mid w \models \alpha\} \). Thus, the following semantics can be defined:
Clause (MS-\(p\)) states that \(p\) is true at a world \(w\) if \(w\) is in the set \(I(p)\). (MS-\(\neg\)) and (MS-\(\wedge\)) are recursive clauses for \(\neg\) and \(\wedge\). The definition of necessity is given in (MS-\(K\)) which says that \(\alpha\) is necessary if its truth set \(\|\alpha\|^M\) is among the collection of subsets of \(W\) at \(w\).

**Reasoning about knowledge.** There are various interpretations for epistemic models that are based on minimal models. Vardi’s\(^{26}\) epistemic structures (Sec. IV-B.1) are closest to the definitions given above. Other variants include Fagin and Halpern’s\(^{17,27}\) logic of local reasoning (Sec. IV-B.2) which employs the notion of “belief cells” and Vardi’s\(^{28}\) fusion epistemic models (Sec. IV-B.3) constructed from world fusion\(^{30}\). However, it may be interesting to note that Delgrande’s\(^{40}\) logic of explicit propositions (Sec. IV-B.4) adopts two different interpretations; the first is close to the original Montague–Scott semantics and the second is based on the “belief cell” approach.

### 1. Epistemic Structures

Vardi’s\(^{26}\) Epistemic Structures (ES) are based on the Montague–Scott model where epistemic notions are modeled by *epistemic sets*. An epistemic set is a set of sentences that an agent believes or knows. In this approach, each sentence \(\alpha\) is associated with a proposition (or an intension) which is a set of worlds in which \(\alpha\) is satisfied.

An epistemic structure for a system of \(n\) agents is a tuple \(M_{ES} = (S, N_1, \ldots, N_n, I)\) where \(S\) is a set of possible states or worlds, \(I\) is an intensional assignment which associates each atomic sentence \(p \in P\) with a set of possible worlds. Intensional assignments can also be extended to any arbitrary sentence \(\alpha \in PL\) (a propositional language), that is \(I(\alpha) = \{t | M_{ES}, t |= \alpha\}\). \(N_1, \ldots, N_n\) are epistemic assignments, where \(N_i\) assigns to agent \(i\) an epistemic set (a set of sets of possible worlds). The following semantics for an epistemic structure is defined:\(^{26}\)

\[
\begin{align*}
(ES-p) & \quad M_{ES}, s |= p \text{ where } p \in P, \text{ if } s \in I(p) \\
(ES-\neg) & \quad M_{ES}, s |= \neg \alpha, \text{ if } M_{ES}, s \not|= \alpha \\
(ES-\wedge) & \quad M_{ES}, s |= \alpha \wedge \beta \iff M_{ES}, s |= \alpha \text{ and } M_{ES}, s |= \beta \\
(ES-E) & \quad M_{ES}, s |= E_i \alpha \iff I(\alpha) \in N_i(s)
\end{align*}
\]

The definitions in (ES-\(p\)), (ES-\(\neg\)), and (ES-\(\wedge\)) follow directly from those in (MS-\(p\)), (MS-\(\neg\)) and (MS-\(\wedge\)), respectively. In the case of (ES-\(E\)) agent \(i\) believes (or knows) a sentence \(\alpha\) just in case if \(I(\alpha)\) is in \(N_i(s)\) (that is, there is a set \(T \in N_i(s)\) such that \(\forall t \in T \implies M_{ES}, t |= \alpha\)).
Avoiding logical omniscience. (LO1), (LO2), and (LO3) are satisfiable in ES but (LO4) is not. The fact that $B\alpha \land B(\alpha \supset \beta) \land \neg B\beta$ is satisfiable is because the two beliefs $B\alpha$ and $B(\alpha \supset \beta)$ may be expressed by different sets of worlds say $\{w_1\}$ and $\{w_2\}$. Thus, the two propositions $\{w_1\}$ and $\{w_2\}$ may be necessary at a world $w$ (where $N(w) = \{\{w_1\}, \{w_2\}\}$), but their intersection $\{w_1\} \cap \{w_2\} = \emptyset$ may not be necessary at $w$. The same reason holds for the satisfiability of $B\alpha \land B(\neg \alpha) \land \neg B\beta$ because $B\alpha$ and $B(\neg \alpha)$ may be characterized by different sets of worlds. To see that $\neg B(\alpha \lor \neg \alpha)$ is satisfiable, consider an ES structure with just one world $w$ and $N(w) = \emptyset$. Thus, at $w$, there is no necessary proposition, so that valid sentences such as $B(\alpha \lor \neg \alpha)$ do not hold. However, $B\alpha \land \neg B(\alpha \lor (\beta \lor \neg \beta))$ is not satisfiable because of the fundamental inference rule from $\alpha = \beta$ infer $B\alpha = B\beta$ in the neighborhood semantics. Since $\alpha = \alpha \land (\beta \lor \neg \beta)$ is valid, it follows that $B\alpha = B(\alpha \land (\beta \lor \neg \beta))$.

Constraints on $N_i(s)$. In addition, Vardi also mentions that conditions can be placed on $N_i(s)$ to express different modes of reasoning. For instance, by imposing that $N_i(s)$ is nonempty, one obtains the property that an agent does not believe sentences such as $(\alpha \land \neg \alpha)$. Since $(\alpha \land \neg \alpha)$ is false in every world, it corresponds to the empty set. Another possible restriction on $N_i(s)$ is to require that sets in $N_i(s)$ are closed under supersets [that is if $U \subseteq N_i(s)$ and $V \supseteq U$ then $V \subseteq N_i(s)$]. This means that if an agent believes a sentence $(\alpha \land \beta)$, then it also believes any sentence that is more general than $(\alpha \land \beta)$ (for example $B(\alpha \land \beta) \supset B\alpha$). In fact, by imposing on $N_i(s)$ the two conditions mentioned above together with another condition $S \subseteq N_i(s)$, one obtains the logic of local reasoning.

2. Logic of Local Reasoning

The motivation for local reasoning is that humans do not focus their attention on all issues simultaneously. Thus, if an agent believes a sentence $\alpha$ then it is said that it believes $\alpha$ in a certain frame of mind. In the logic for local reasoning, an agent is viewed as perceiving different frames of mind (or “society of mind”) where each frame (or cluster or belief cell) is modeled with a different set of possible worlds.

The epistemic model for local reasoning can be thought of as a special case of the minimal model described in chapter 7 of ref. 2. A model for local reasoning in a system of $n$ agents is a tuple $M_{\text{LR}} = \langle S, C_1, \ldots , C_n, \pi \rangle$ where $S$ is a set of possible states or worlds, $\pi$ is a truth assignment to the primitive sentences $p \in \text{P}$ for each world $s \in S$, and $C_i(s)$ is a nonempty set of nonempty subsets of $S$. Thus, $C_i(s) = \{F_1, \ldots , F_i\}$ means that in a world $s$, agent $i$ can believe different sets of possible worlds (frames of mind) represented by $F_1, \ldots , F_i$. A propositional language $\text{PL}$ similar to that defined in Section IV-A.3 is assumed. Thus, for all $\psi \in \text{PL}$ the following semantics for local reasoning is defined:

$$\text{(LR-p)} \quad M_{\text{LR}}, s \models p \quad (\text{where } p \text{ is a primitive letter}) \iff \pi(s, p) = \text{true}$$

$$\text{(LR-\neg)} \quad M_{\text{LR}}, s \models \neg \alpha \quad \text{if } M_{\text{LR}}, s \models \neg \alpha$$
(LR-\land) \ M_{LR}, s \models \alpha \land \beta \iff \ M_{LR}, s \models \alpha \ \text{and} \ M_{LR}, s \models \beta$

(LR-B) \ M_{LR}, s \models B_\alpha \iff \exists F \in C_i(s) \text{ such that } \forall t \in F \Rightarrow M_{LR}, t \models \alpha.$

(LR-L) \ M_{LR}, s \models L_\alpha \iff \ M_{LR}, t \models \alpha \ \text{for all} \ t \in \bigcap_{F \in C_i(s)} F.$

Two modal operators $B$ and $L$ are defined, respectively, for local (or explicit) belief and implicit belief. $B_\alpha$ is interpreted as agent $i$ believes $\alpha$ in some frame of mind. This kind of belief is called weak belief.\textsuperscript{17} $L_\alpha$ is interpreted as agent $i$ implicitly believes $\alpha$ as a result of putting together the information of its various frames of mind. This is illustrated in Figure 5, in which frame 2 is a frame where the agent combines the information from frame 1 and frame 3. The shaded circles represent the set of worlds that are common in both frame 1 and frame 3. The example in Figure 5 also shows that by intersecting sets of possible worlds, information from different frames can be used to eliminate possibilities. It is interesting to note that implicit belief in the logic of local reasoning corresponds to the smallest proposition [consisting of the set of worlds that are members of every proposition in $N(w)$] in the original neighborhood semantics mentioned in Section IV-B. An even stronger notion called strong belief ($S$) is defined in Ref. 17; a strong belief is one that is true in every frame of mind:

\textit{Avoiding Logical Omnipotence.} In this approach, local belief is not closed under implication. Thus, $B_\alpha \land B(\alpha \supset \beta) \land \neg B_\beta$ is satisfiable because an agent may believe $\alpha$ in one frame of mind and $\alpha \supset \beta$ in another, but it may not be in a frame of mind to bring these facts together. An agent can also hold inconsistent beliefs such as $B_\alpha \land \neg B_\alpha$. This follows from the fact that an agent may believe $\alpha$ in one frame of mind and believes $\neg \alpha$ in another as illustrated in Figure 5. However, $\neg B(\alpha \lor \neg \alpha)$ and $B_\alpha \land \neg B_\alpha \land (\beta \lor \neg \beta)$ are both not satisfiable. In the logic of local reasoning, agents are perfect reasoners within each frame of mind. As such, an agent’s set of beliefs is closed under valid implication and all valid formulas are believed.

### 3. Fusion Epistemic Models

A Nonstandard (NS) fusion epistemic model\textsuperscript{28} for a system of $n$ agents is a tuple $M_{NS} = (S, N_1, \ldots, N_n, I)$ where $S$ is a set of nonstandard worlds.\textsuperscript{30} $I$ is an intensional assignment which associates every atomic sentence $p \in P$ with a set of nonstandard worlds (that is $I: P \rightarrow 2^S$) and a propositional language $PL$ (formed from the usual connectives together with $B_i$) is assumed. $N_1, \ldots, N_n$ are epistemic assignments, where $N_i$ assigns to agent $i$ a nonstandard world expression (that is $N_i: A \times S \rightarrow E$, where $A$ is a set of agents and $E$ is a class of nonstandard world expressions). These nonstandard world expressions can be constructed by the two world-fusion operations: schematization $\cap$ and superposition $\cup$\textsuperscript{30} and $E$ is closed under the following conditions:
Figure 5. Transforming local (explicit) belief to implicit belief.

(1) $E \supset S$
(2) If $s_j \in S$ for all $j$ in an index set $I$, then $\cap s_j \in E$
(3) If $s_j \in S$ for all $j$ in an index set $I$, then $\cup s_j \in E$

**Schematized and superposed belief.** Schematization combines worlds conjunctively, while superposition fuses worlds together disjunctively. For instance,
obtains in the schematization of two worlds \( s_1 \) and \( s_2 \) (that is \( s_1 \lor s_2 \)) if \( p \) holds in both \( s_1 \) and \( s_2 \). It is noted that schematized worlds may be underdetermined and thus neither \( p \) nor \( \neg p \) may hold in \( s_1 \lor s_2 \), while superposed worlds may be overdetermined and both \( p \) and \( \neg p \) may hold in \( s_1 \lor s_2 \). Based on these ideas, Vardi defined two notions of belief; one with respect to schematized worlds (called schematized belief) and the other is taken over superposed worlds (called superposed belief). Formal definitions for schematized belief and superposed belief are given in (NS-Sch) and (NS-Sup), respectively. Thus, for all \( \psi \in \mathbf{PL} \) the following semantics for a fusion epistemic model is defined:

\[
\begin{align*}
\text{(NS-\(p\))} & \quad M_{\text{NS}}, s \models p \text{ where } p \in P, \text{ if } s \in \Pi(p) \\
\text{(NS-\(\neg\))} & \quad M_{\text{NS}}, s \models \neg \alpha \text{ if } M_{\text{NS}}, s \not\models \alpha \\
\text{(NS-\(\land\))} & \quad M_{\text{NS}}, s \models \alpha \land \beta \text{ iff } M_{\text{NS}}, s \models \alpha \text{ and } M_{\text{NS}}, s \models \beta \\
\text{(NS-Sch)} & \quad \text{If } N(i, s) = \bigcap s \text{ then } M_{\text{NS}}, s \models B_\alpha \text{ if } M_{\text{NS}}, s \models \alpha \text{ for all } j \in J \\
\text{(NS-Sup)} & \quad \text{If } N(i, s) = \bigcap s \text{ then } M_{\text{NS}}, s \models B_\alpha \text{ if } M_{\text{NS}}, s_j \models \alpha \text{ for some } j \in J
\end{align*}
\]

Since a schematized belief is considered over a cluster of schematized worlds, it may be viewed as a semantic counterpart of the \( K_{\text{S}} \) epistemic operator. Similarly, superposed belief may be considered as a semantic counterpart of the \( K_{\text{C}} \) operator. In addition, it is noted that both fusion epistemic models and epistemic structures can be viewed as outgrowths of the logic of local reasoning. While Vardi’s epistemic structures are generalizations of the logic of local reasoning, fusion epistemic models provide an account that agents can combine information conjunctively or disjunctively.

\section*{4. Logic of Explicit Propositions}

The treatment of explicit belief in Lakemeyer and Levesque’s model was being criticized for being inadequate because of the validity of the collective conjunction axiom (B12, in Sec. IV-A.1). (B12) says that believing two sentences \( \alpha \) and \( \beta \), is equivalent to believing their conjunction \( \alpha \land \beta \). That is \( B\alpha \land B\beta \supset B(\alpha \land \beta) \) and \( B(\alpha \land \beta) \supset B\alpha \land B\beta \) are valid. Although \( B(\alpha \land \beta) \supset B\alpha \land B\beta \) seems plausible, \( B\alpha \land B\beta \supset B(\alpha \land \beta) \) seems unintuitive. For instance, the fact that an agent believes that “3 is a prime number” and “penguins cannot fly” does not necessarily mean that it should also believe “3 is a prime number and penguins cannot fly” since it may not connect the two sentences together. Furthermore, Delgrande\(^40\) remarked that (B12) is valid because there is no way of imposing that \( B\alpha \land B\beta \) does not imply \( B(\alpha \land \beta) \) by using the standard notion of an accessibility relation. This is because the truth of \( B\alpha \) and \( B\beta \) are determined by the same set of situations. He suggested that explicit belief should be individually characterized by a set of (possibly different) situations by defining their semantics relative to a minimal model.

A minimal model for explicit propositions is a tuple \( M_{\text{EP}} = \langle S, N_T, N_F, T, F \rangle \) where \( S \) is a set of all situations. \( T \) and \( F \) are defined as in Section IV-A.1.
The deviation from the Levesque–Lakemeyer model is the use of $N_T$ and $N_F$ in place of $R^+$ and $R^-$. $N_T$ expresses those propositions that an agent believes while $N_F$ are used to specify those propositions which are not believed. A propositional language $PL$ as defined in Sections IV-A.1 and IV-A.2 is assumed. Thus, for all $\psi \in PL$ the following semantics for explicit propositions is defined:

\[
\begin{align*}
(EP-T \land) \quad & M_{EP}, s \models \alpha \land \beta \text{ iff } M_{EP}, s \models \alpha \text{ and } M_{EP}, s \models \beta \\
(EP-F \land) \quad & M_{EP}, s \models \alpha \land \beta \text{ iff } M_{EP}, s \models \alpha \text{ or } M_{EP}, s \models \beta \\
(EP-T B) \quad & M_{EP}, s \models B \alpha \text{ iff } \|\alpha\|^M \in N_T(s), \text{ where } \|\alpha\|^M = \{t \in M_{EP}, t \models \alpha\} \\
(EP-F B) \quad & M_{EP}, s \models B \alpha \text{ iff } \|\alpha\|^M \not\in N_T(s) \\
(EP-T B') \quad & M_{EP}, s \models B \alpha \text{ iff there is some } C \in N_T(s) \text{ such that } M_{EP}, t \models \alpha \text{ for all } t \in C \\
(EP-F B') \quad & M_{EP}, s \models B \alpha \text{ iff there do not exist } C \in N_T(s), \text{ such that } M_{EP}, t \models \alpha \text{ for all } t \in C
\end{align*}
\]

Avoiding logical omniscience: Although, $B \alpha \land B(\alpha \supset \beta) \land \neg B \beta$, $B \alpha \land B(\neg \alpha) \land \neg B \beta$, and $B(\alpha \lor \neg \alpha)$ are satisfiable, the fundamental inference rule from $\alpha = \beta$ infer $B \alpha = B \beta$ in the neighborhood semantics still presents a problem. If two propositions $\alpha$ and $\beta$ have the same set of supporting situations, then an agent who believes $\alpha$ must also believe $\beta$.

V. DISCUSSION

Levesque’s model is an important contribution to epistemic logic because it is one of the first attempts in knowledge representation to solve the logical omniscience problem. Several models were built based on Levesque’s model to improve some aspects of it. This has taken several directions. The issues which are of interest include: multiagent reasoning, nested beliefs, avoiding the use of incoherent (and incomplete) situations, and disallowing the collective conjunctive axiom.
Multiagent reasoning and nested beliefs. Lakemeyer has extended Levesque’s model by allowing nesting of beliefs. An interesting feature of Lakemeyer’s logic is that belief and disbelief are considered over different sets of situations via two separate accessibility relations. Additionally, Fagin and Halpern’s logic of awareness was proposed to address both the issues of introspection and multiagent reasoning. However, the logic of awareness does not retain all the properties of explicit belief because agents cannot hold inconsistent beliefs since incoherent situations are not allowed.

Avoiding the use of incoherent (and incomplete) situations. Several researchers such as Fagin, Halpern, and Vardi have argued strongly against the use of incoherent situations. They maintain that in the real world formulas can only either be true or false. In fact, in most of their epistemic logics, such as the logic of awareness, the logic of local reasoning and nonstandard epistemic logic, incoherent and incomplete situations were not allowed. The semantics of nonstandard epistemic logic is still defined in terms of classical worlds, although each world \( w \) is associated with a dual \( w^* \). This approach seems to provide the effect of decoupling the truth and falsity of a formula. The logic of awareness dispenses with the use of both incoherent and incomplete situations but the awareness function at each state seems to provide an effect that is similar to an incomplete situation. In the logic of local reasoning an agent can also hold inconsistent beliefs such as \( B \alpha \land \neg \alpha \). This is because an agent may believe a sentence in one frame of mind and believe its negation in another. However, a belief such as \( B(\alpha \land \neg \alpha) \) is not satisfiable.

Disallowing the collective conjunctive axiom. Degrande remarked that the treatment of explicit belief in Lakemeyer and Levesque’s model is inadequate because it does not capture the insight that agents can hold beliefs independently without putting them together. In particular the validity of the collective conjunction axiom is viewed as problematic for the reasons that were given in Section IV-B.4. Additionally, it is noted that validity of the collective conjunction axiom will also mean that \( B(\alpha \land \beta) \land B \alpha \subseteq B(\beta \lor \alpha \land \neg \alpha) \) is valid. Thus, either an agent’s beliefs are closed under logical implication or every situation that it considers possible is incoherent. The issues of the collective conjunctive axiom have been taken up by Delgrande who discards the notion of an accessibility relation by defining explicit belief relative to a minimal model. Since different explicit propositions may be characterized by different sets of situations, the collective conjunctive axiom is not valid in Delgrande’s model.

Despite the above criticisms, it seems noteworthy to mention that Levesque’s notion of explicit belief does account for the fact that some sentences lie outside the scope of an agent’s beliefs. Further, he seems to have anticipated criticisms about incoherent situations, and argued in Ref. 19 that even though incoherent situations cannot be realized, they can be imagined by an agent. In other words, inconsistency can still arise because of an agent’s epistemic shortcomings even though the world is inherently consistent from an ontological perspective. If situations (or worlds) are to be thought of as epistemic alternatives, then there is no reason to assume that an agent will not receive inconsistent information.
VI. CONCLUSION

In this survey, several logics for reasoning about knowledge and belief were reviewed. Although these different models appear to have surface dissimilarities, a closer examination shows that they have strong resemblance. Further, given the current proliferation of epistemic logics, it seems prudent to establish comparisons among the various approaches and to consider the issue of unifications among them. One of the issues that is of theoretical interest is to formulate a unified framework that harmonizes the ideas in these logics. This issue is addressed in Refs. 43 and 44.

References


