

PDF version of the entry
Epistemic Logic
<https://plato.stanford.edu/archives/sum2024/entries/logic-epistemic/>

from the SUMMER 2024 EDITION of the

STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Co-Principal Editors: Edward N. Zalta & Uri Nodelman

Associate Editors: Colin Allen, Hannah Kim, & Paul Oppenheimer

Faculty Sponsors: R. Lanier Anderson & Thomas Icard

Editorial Board: <https://plato.stanford.edu/board.html>

Library of Congress ISSN: 1095-5054

Notice: This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

Stanford Encyclopedia of Philosophy
Copyright © 2024 by the publisher
The Metaphysics Research Lab
Department of Philosophy
Stanford University, Stanford, CA 94305

Epistemic Logic
Copyright © 2024 by the authors
Rasmus Rendsvig, John Symons, and Yanjing Wang

All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

Epistemic Logic

First published Fri Jun 7, 2019; substantive revision Fri Dec 1, 2023

Epistemic logic is a subfield of philosophical logic concerned with logical approaches to knowledge, belief, and related notions. Though any logic with an epistemic interpretation may be called an *epistemic logic*, the most widespread type of epistemic logics in use at present are modal logics. Knowledge and belief are represented via the modal operators K and B , often with a subscript indicating the agent that holds the attitude. Formulas $K_a\varphi$ and $B_a\varphi$ are then read “agent a knows that φ ” and “agent a believes that φ ”, respectively. Epistemic logic allows the formal exploration of the implications of epistemic principles. For example, the formula $K_a\varphi \rightarrow \varphi$ states that what is known is true, while $K_a\varphi \rightarrow K_aK_a\varphi$ states that what is known is known to be known. The semantics of epistemic logic are typically given in terms of possible worlds *via* Kripke models such that the formula $K_a\varphi$ is read to assert that φ is true in all worlds agent a considers epistemically possible relative to its current information. The central problems that have concerned epistemic logicians include, for example, determining which epistemic principles are most appropriate for characterizing knowledge and belief, the logical relations between different conceptions of knowledge and belief, and the epistemic features of groups of agents. Beyond philosophy proper, epistemic logic flourishes in theoretical computer science, AI, economics, and related fields.

- 1. Introduction
- 2. The Modal Approach to Knowledge
 - 2.1 The Formal Language of Epistemic Logic
 - 2.2 Higher-Order Attitudes
 - 2.3 The Partition Principle and Modal Semantics

- 2.4 Kripke Models and The Indistinguishability Interpretation of Knowledge
- 2.5 Epistemological Principles in Epistemic Logic
- 2.6 Principles of Knowledge and Belief
- 3. Knowledge in Groups
 - 3.1 Multi-Agent Languages and Models
 - 3.2 Notions of Group Knowledge
- 4. Beyond Knowing That
- 5. Logical Omniscience
- Bibliography
- Academic Tools
- Other Internet Resources
- Related Entries

1. Introduction

Aristotelian texts set the groundwork for discussions of the logic of knowledge and belief, particularly *De Sophisticis Elenchis* as well as the *Prior* and *Posterior Analytics*. While Aristotle addressed the four alethic modes of possibility, necessity, impossibility, and contingency, Buridan, Pseudo Scotus, Ockham, and Ralph Strode, helped to extend Aristotle's insights to epistemic themes and problems (Boh 1993; Knuuttila 1993). During this period, Pseudo Scotus and William of Ockham supplemented Aristotle's study of mental acts of cognition and volition (see Boh 1993: 130). Ivan Boh's studies of the history of fourteenth and fifteenth century investigations into epistemic logic provide an excellent coverage of the topic, especially his *Epistemic Logic in the Later Middle Ages* (1993).

According to Boh, the English philosopher Ralph Strode formulated a fully general system of propositional epistemic rules in his influential 1387 book *Consequences* (Boh 1993: 135). Strode's presentation built on the

earlier logical treatises of Ockham and Burley. Problems of epistemic logic were also discussed between the 1330s and 1360s by the so-called Oxford Calculators, most prominently by William Heytesbury and Richard Kilvington. By the fifteenth century, Paul of Venice and other Italian philosophers also engaged in sophisticated reflection on the relationship between knowledge, truth, and ontology.

Discussions of epistemic logic during the medieval period share a similar set of foundational assumptions with contemporary discussions. Most importantly, medieval philosophers explored the connection between knowledge and veracity: If I know p , then p is true. Furthermore, many medieval discussions begin with an assumption similar to G.E. Moore's observation that an epistemic agent cannot coherently assert " p but I do not believe (know) p ". Sentences of this form are generally referred to as *Moore sentences*.

Modern treatments of the logic of knowledge and belief grew out of the work of philosophers and logicians writing from 1948 through the 1950s. Rudolf Carnap, Jerzy Łoś, Arthur Prior, Nicholas Rescher, G.H. von Wright, and others recognized that our discourse concerning knowledge and belief admits of an axiomatic-deductive treatment. Among the many important papers that appeared in the 1950s, von Wright's seminal work (1951) is widely acknowledged as having initiated the formal study of epistemic logic as we know it today. Von Wright's insights were extended by Jaakko Hintikka in his book *Knowledge and Belief: An Introduction to the Logic of the Two Notions* (1962). Hintikka provided a way of interpreting epistemic concepts in terms of possible world semantics and as such it has served as the foundational text for the study of epistemic logic ever since.

In the 1980s and 1990s, epistemic logicians focused on the logical properties of systems containing groups of knowers and later still on the

epistemic features of so-called “multi-modal” contexts. Since the 1990s work in dynamic epistemic logic has extended traditional epistemic logic by modeling the dynamic process of knowledge acquisition and belief revision. In the past two decades, epistemic logic has come to comprise a broad set of formal approaches to the interdisciplinary study of knowledge and belief.

Interest in epistemic logic extends well beyond philosophers. Recent decades have seen a great deal of interdisciplinary attention to epistemic logic with economists and computer scientists actively developing the field together with logicians and philosophers. In 1995 two important books signaled the fertile interplay between computer science and epistemic logic: Fagin, Halpern, Moses, and Vardi (1995) and Meyer and van der Hoek (1995). Work by computer scientists has become increasingly central to epistemic logic in the intervening years.

Among philosophers, there is increased attention to the interplay between these formal approaches and traditional epistemological problems (see for example, van Benthem 2006; Hendricks & Symons 2006; Stalnaker 2006; Artemov 2008; Holliday 2018; Baltag, Bezhanishvili, et al. 2019).

Several introductory texts on epistemic logic exist, e.g., Fagin et al. (1995), van Benthem (2011); van Ditmarsch, van der Hoek, and Kooi (2007); van Ditmarsch, Halpern, et al. (2015); Gochet and Gribomont (2006); and Meyer (2001) with Lenzen (1980) providing an overview of early developments.

2. The Modal Approach to Knowledge

Until relatively recently, epistemic logic focused almost exclusively on propositional knowledge. In cases of propositional knowledge, an agent or a group of agents bears the propositional attitude of knowing towards

some proposition. For example, when one says: “Zoe knows that there is a hen in the yard”, one asserts that Zoe is the agent who bears the propositional attitude *knowing* towards the proposition expressed by the English sentence “there is a hen in the yard”. Now imagine that Zoe does not know whether there is a hen in the yard. For example, it might be the case that she has no access to information about whether there is or is not a hen in the yard. In this case her lack of information means that she will consider two scenarios as being possible, one in which there is a hen in the yard and one in which there is not.

Perhaps she has some practical decision that involves not only hens but also the presence of frightening dogs in the yard. She might wish to feed the hens but will only do so if there is no dog in the yard. If she were ignorant of whether there is a dog in the yard, the number of scenarios she must consider in her deliberations grows to four. Clearly, one needs to consider epistemic alternatives when one does not have complete information concerning the situations that are relevant to one’s decisions. As we shall see below, possible worlds semantics has provided a useful framework for understanding the manner in which agents can reason about epistemic alternatives.

While epistemic logicians had traditionally focused on *knowing that*, one finds a range of other uses of knowledge in natural language. As Y. Wang (2018b) points out, the expressions *knowing how*, *knowing what*, *knowing why* are very common, appearing almost just as frequently (sometimes more frequently) in spoken and written language as *knowing that*. Recently non-standard epistemic logics of such know-wh expressions have been developed, to which we will come back in Section 4. In the remainder of this section, we will introduce the basics of the epistemic logic of knowing that.

2.1 The Formal Language of Epistemic Logic

Recent work in epistemic logic relies on a modal conception of knowledge. In order to be clear about the role of modality in epistemic logic it is helpful to introduce the basic elements of the modern formalism. For the sake of simplicity we begin with the case of knowledge and belief for a single agent, postponing consideration of multiple agents to Section 3,

A prototypical epistemic logic language is given by first fixing a set of *propositional variables* p_1, p_2, \dots . In applications of epistemic logic, propositional variables are given specific interpretations: For example, p_1 could be taken to represent the proposition “there is a hen in the yard” and p_2 the proposition “there is a dog in the yard”, etc. The propositional variables represent propositions which are represented in no finer detail in the formal language. As such, they are therefore often referred to as *atomic propositions* or simply *atoms*. Let $Atom$ denote the set of atomic propositions.

Apart from the atomic propositions, epistemic logic supplements the language of propositional logic with a modal operator, K_a , for knowledge and B_a , for belief.

$K_a\varphi$ reads “Agent a knows that φ ”

and similarly

$B_a\varphi$ reads “Agent a believes that φ ”.

In many recent publications on epistemic logic, the full set of formulas in the language is given using a so-called *Backus-Naur Form* (BNF). This is simply a notational technique derived from computer science that provides

a recursive definition of the formulas deemed grammatically “correct”, i.e., the set of *well-formed formulas*:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid B_a\varphi, \text{ for } p \in Atom.$$

It should be noted that the Greek letter φ stands for the *syntactic category* of formula. So this definition says: an atom p is a formula; $\neg\varphi$ is a formula if φ is a formula (read \neg as ‘it is not the case that’); $(\varphi \wedge \varphi)$ is a formula whenever any two formulas are connected by the \wedge symbol (read \wedge as ‘and’); and $K_a\varphi$ and $B_a\varphi$ are formulas whenever φ is a formula (the readings were indicated above). Note that in a non-BNF recursive specification of the language, the Greek variable φ would be used a *metavariable* ranging over formulas, and one would normally state the clause for conjunctions as: $(\varphi \wedge \psi)$ is a formula whenever φ and ψ are formulas. But the BNF let’s us get away with just using φ to get the same effect.

We will call this basic language that includes both a Knowledge and a Belief operator, \mathcal{L}_{KB} . As in propositional logic, additional connectives are defined from \neg and \wedge : Typical notation is ‘ \vee ’ for ‘or’, ‘ \rightarrow ’ for ‘if..., then ...’ and ‘ \leftrightarrow ’ for ‘... if, and only if, ...’. Also typically \top (‘top’) and \perp (‘bottom’) is used to denote the constantly true proposition and the constantly false proposition, respectively.

As we shall see below, $K_a\varphi$ is read as stating that φ holds in *all* of the worlds accessible to a . In this sense, K can be regarded as behaving similarly to the ‘box’ operator, \Box , often used to denote necessity. In evaluating $K_a\varphi$ at a possible world w , one is in effect evaluating a *universal quantification* over all the worlds accessible from w . The universal quantifier \forall in first-order logic has the existential quantifier \exists as its *dual*: This means that the quantifiers are mutually definable by taking either \forall as primitive and defining $\exists x\varphi$ as short for $\neg\forall x\neg\varphi$ or by taking \exists as primitive and defining $\forall x\varphi$ as $\neg\exists x\neg\varphi$. In the case of K_a , it may be

seen that the formula $\neg K_a \neg \varphi$ makes an *existential quantification*: It says that there *exists* an accessible world that satisfies φ . In the literature, a dual operator for K_a is often introduced. The typical notation for $\neg K_a \neg$ includes $\langle K_a \rangle$ and \hat{K}_a . This notation mimics the diamond-shape \diamond , which is the standard dual operator to the box \Box , which in turn is standard notation for the universally quantifying modal operator (see the entry on modal logic).

More expressive languages in epistemic logic involve the addition of operators for various notions of group knowledge (see Section 3). For example, as we discuss below, the *common knowledge* operator and so-called *dynamic* operators are important additions to the language of epistemic logic. Dynamic operators can indicate for example the *truthful public announcement* of φ : $[\varphi!]$. A formula $[\varphi!]\psi$ is read “if φ is truthfully announced to everybody, then after the announcement, ψ is the case”. The question of what kinds of expressive power is added with the addition of operators is a research topic that is actively being investigated in dynamic epistemic logic. So, for example, adding $[\varphi!]$ by itself to \mathcal{L}_{KB} does *not* add expressive power, but in a language that also includes common knowledge, it does (e.g., van Ditmarsch, van der Hoek, & Kooi 2007).

2.2 Higher-Order Attitudes

Notice that, for example, $K_a K_a p$ is a formula in the language we introduced above. It states that agent a knows that agent a knows that p is the case. Formula with *nested* epistemic operators of this kind express a *higher-order* attitude: an attitude concerning the attitude of some agent.

Higher-order attitudes is a recurring theme in epistemic logic. The aforementioned Moore sentences, e.g., $B_a(p \wedge B_a \neg p)$ express a higher-order attitude. So do many of the epistemic principles discussed in the literature and below. Consider the following prominent epistemic principle

involving higher-order knowledge: $K_a \varphi \rightarrow K_a K_a \varphi$. Is it reasonable to require that knowledge satisfies this scheme, i.e., that if somebody knows φ , then they know that they know φ ? In part, we might hesitate before accepting this principle in virtue of the higher-order attitude involved. This is a matter of ongoing discussion in epistemic logic and epistemology.

2.3 The Partition Principle and Modal Semantics

The semantics of the formal language introduced above is generally presented in terms of so-called possible worlds. In epistemic logic possible worlds are interpreted as epistemic alternatives. Hintikka was the first to explicitly articulate such an approach (1962). This is another central feature of his approach to epistemology which continues to inform developments today. It may be stated, simplified,^[1] as follows:

Partition Principle: Any propositional attitude partitions the set of possible worlds into those that are in accordance with the attitude and those that are not.

The partition principle may be used to provide a semantics for the knowledge operator. Informally,

$K_a \varphi$ is true in world w if, and only if, φ is true in every world w' compatible with the information a has at w .

Here, agent a knows that φ just in case the agent has information that rules out every possibility of error and rules out every case where $\neg \varphi$.

2.4 Kripke Models and The Indistinguishability Interpretation of Knowledge

Since the 1960s *Kripke models*, defined below, have served as the basis of the most widely used semantics for all varieties of modal logic. The use of Kripke models in the representation of epistemic concepts involves taking a philosophical stance with respect to those concepts. One widespread interpretation, especially in theoretical economics and theoretical computer science, understands knowledge in terms of informational indistinguishability between possible worlds. What we will refer to here as the *indistinguishability interpretation* goes back to Lehmann (1984), and Aumann (1976) has an equivalent partition-based framework.

As the indistinguishability interpretation concerns knowledge, but not belief, we will be working with a language without belief operators. Therefore, let the language \mathcal{L}_K be given by the Backus-Naur form

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \text{ for } p \in Atom.$$

As we shall see, the indistinguishability interpretation involves very stringent requirements in order for something to qualify as knowledge. We introduce it here for pedagogical purposes, putting the formal details of the interpretation in place so as to introduce and explain relatively less extreme positions thereafter.

Consider again the case of Zoe, the hen, and the dog. The example involves two propositions, which we will identify with the formal atoms:

p read as “there is a hen in the yard”.

and

q read as “there is a dog in the yard”.

It is worth emphasizing that for the purposes of our formalization of this scenario, these two are the *only* propositions of interest. We are restricting our attention to $Atom = \{p, q\}$. In early presentations of epistemic logic and in much of standard epistemic logic at present, *all* the atoms of interest are included from the outset. Obviously, this is an idealized scenario. It is important to notice what this approach leaves out. Considerations that are not captured in this way include the appearance of novel atoms; the idea that other atomic propositions might be introduced at some future state via some process of learning for example, or the question of an agent’s awareness of propositions; the scenario in which an agent might be temporarily *unaware* of some atom due to some psychological or other factor (see Section 5 for references to so-called *awareness logic*). For now, the main point is that standard epistemic logic begins with the assumption that the set $Atom$ exhausts the space of propositions for the agent.

With two atoms, there are four different ways a world could consistently be. We can depict each by a box:



The four boxes may be formally represented by a set $W = \{w_1, w_2, w_3, w_4\}$, typically called a set of *possible worlds*. Each world is further labeled with the atoms true at that world. They are labeled by a function V , the *valuation*. The valuation specifies which atoms are true at each world in the following way: Given an atom p , $V(p)$ is the subset of worlds at which p is true.^[2] That w_1 is labeled with p and q thus means that $w_1 \in V(p)$ and $w_1 \in V(q)$. In the illustration, $V(p) = \{w_1, w_2\}$ and $V(q) = \{w_1, w_3\}$.

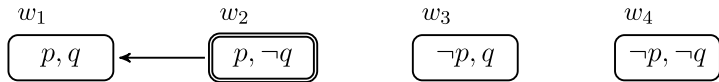
For presentational purposes, assume that there really is a hen in the yard, but no dog. Then w_2 would represent the *actual world* of the model. In

illustrations, the actual world is commonly highlighted:



Now, assume that the hen is always clucking, but that the dog never barks, and that although Zoe has acute hearing, she cannot see the yard. Then there are certain possible worlds that Zoe cannot *distinguish*: possible ways things may be which she cannot tell apart. For example, being in the world with only a hen ($p, \neg q$), Zoe cannot tell if she is in the world with both hen and dog (p, q): her situation is such that Zoe is aware of two ways things could be but her information does not allow her to eliminate either.

To illustrate that one possible world cannot be distinguished from another, an arrow is typically drawn from the former to the latter:



Here, arrows represent a *binary relation* on possible worlds. In modal logic in general, it is referred to as the **accessibility relation**. Under the indistinguishability interpretation of epistemic logic, it is sometimes called the **indistinguishability relation**. Formally, denote the relation R_a , with the subscript showing the relation belongs to agent a . The relation is a subset of the set of *ordered pairs* of possible worlds, $\{(w, w') : w, w' \in W\}$. One world w “points” to another w' if $(w, w') \in R_a$. In this case, w' is said to be *accessible* (*indistinguishable*) from w . In the literature, this is often written $wR_a w'$ or $R_a w w'$. The notation ‘ $w' \in R_a(w)$ ’ is also common: the set $R_a(w)$ is then the worlds accessible from w , i.e.,

$$R_a(w) := \{w' \in W : (w, w') \in R_a\}.$$

A final note: the set $\{(w, w') : w, w' \in W\}$ is often written $W \times W$, the *Cartesian product* of W with itself.

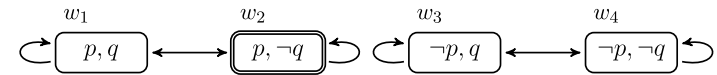
For R_a to faithfully represent a relation of indistinguishability, what worlds should it relate? If Zoe was plunged in w_1 for example, could she tell that she is not in w_2 ? No: the relation of indistinguishability is *symmetric*—if one cannot tell a from b , neither can one tell b from a . That a relation is symmetric is typically drawn by omitting arrow-heads altogether or by putting them in both directions:



Which of the remaining worlds are indistinguishable? Given that the hen is always clucking, Zoe has information that allows her to distinguish w_1 and w_2 from w_3 and w_4 and *vice versa*, cf. symmetry. Hence, no arrows between these. The worlds w_3 and w_4 are indistinguishable. This brings us to the following representation:



Since no information will ever allow Zoe to distinguish something from itself, any possible world is thus related to itself and the indistinguishability relation is *reflexive*:



The standard interpretation of the Zoe example in terms of a possible worlds model is now complete. Before turning to a general presentation of the indistinguishability interpretation, let us look at what Zoe knows.

Recall the informal modal semantics of the knowledge operator from above:

$K_a\varphi$ is true in world w if, and only if, φ is true in every world w' compatible with the information a has at w .

To approach a formal definition, take ' $w \models \varphi$ ' to mean that φ is true in world w . Thus we can, define truth of $K_a\varphi$ in w by

$w \models K_a\varphi$ iff $w' \models \varphi$ for all w' such that wR_aw' .

This definition states that a knows φ in world w if, and only if, φ is the case in all the worlds w' which a cannot distinguish from w .

So, where does that leave Zoe? First off, the definition allows us to evaluate her knowledge in each of the worlds, but seeing as w_2 is the actual world, it is the world of interest. Let a denote Zoe, here are some examples of what we can say about Zoe's knowledge in w_2 :

1. $w_2 \models K_ap$. Zoe knows that the hen is in the yard as all the worlds indistinguishable from w_2 that would be w_1 and w_2 make p true.
2. $w_2 \models \neg K_aq$. Zoe does not know the dog is in the yard, as one of the indistinguishable worlds in fact w_2 itself makes q false.
3. $w_2 \models K_aK_ap$. Zoe knows that she knows p because
 - a. $w_2 \models K_ap$ (cf. 1.) and
 - b. $w_1 \models K_ap$.
4. $w_2 \models K_a\neg K_aq$. Zoe knows that she does not know q because
 - a. $w_2 \models \neg K_aq$ (cf. 2.) and
 - b. $w_1 \models \neg K_aq$.

We could say a lot more about Zoe's knowledge: every formula of the epistemic language without belief operators may be evaluated in the

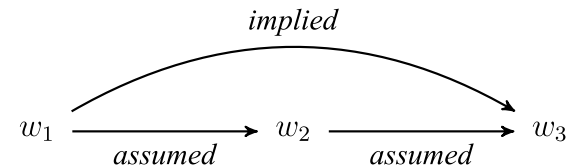
model. It thus represents all Zoe's higher-order information about her own knowledge of which points 3. and 4. are the first examples.

One last ingredient is required before we can state the indistinguishability interpretation in its full generality. In the example above, it was shown that the indistinguishability relation was both *symmetric* and *reflexive*. Formally, these properties may be defined as follows:

Definition: A binary relation $R \subseteq W \times W$ is

- a. **reflexive** iff for all $w \in W$, wRw ,
- b. **symmetric** iff for all $w, w' \in W$, if wRw' , then $w'Rw$.

The missing ingredient is then the relational property of *transitivity*. 'Shorter than' is an example of a transitive relation: Let x be shorter than y , and let y be shorter than z . Then x must be shorter than z . So, given w_1, w_2 and w_3 , if the relation R holds between w_1 and w_2 and between w_2 and w_3 , then the arrow between w_1 and w_3 is the consequence of requiring the relation to be transitive:



Formally, transitivity is defined as follows:

Definition: A binary relation $R \subseteq W \times W$ is **transitive** iff for all $w, w', w'' \in W$, if wRw' and $w'Rw''$, then wRw''

A relation that is both reflexive, symmetric and transitive is called an **equivalence relation**.

With all the components in place, let us now define the Kripke model:

Definition: A *Kripke model* for \mathcal{L}_K is a tuple $M = (W, R, V)$ where

- W is a non-empty set of possible worlds,
- R is a binary relation on W , and
- $V: Atom \rightarrow \mathcal{P}(W)$ is a valuation.

In the definition, ‘ $\mathcal{P}(W)$ ’ denotes the *powerset* of W : It consists of all the subsets of W . Hence $V(p)$, the valuation of atom p in the model M , is some subset of the possible worlds: Those where p is true. In this general definition, R can be any relation on W .

To specify which world is actual, one last parameter is added to the model. When the actual world is specified a Kripke model is commonly called *pointed*:

Definition: A *pointed Kripke model* for \mathcal{L}_K is a pair (M, w) where

- $M = (W, R, V)$ is a Kripke model, and
- $w \in W$.

Finally, we may formally define the semantics that was somewhat loosely expressed above. This is done by defining a relation between pointed Kripke models and the formulas of the formal language. The relation is denoted ‘ \models ’ and is often called the *satisfaction relation*.

The definition then goes as follows:

Definition: Let $M = (W, R_a, V)$ be a Kripke model for \mathcal{L}_K and let (M, w) be a pointed Kripke model. Then for all $p \in Atom$ and all $\varphi, \psi \in \mathcal{L}_K$

$(M, w) \models p$	iff $w \in V(p)$
$(M, w) \models \neg\varphi$	iff not $(M, w) \models \varphi$
$(M, w) \models (\varphi \wedge \psi)$	iff $(M, w) \models \varphi$ and $(M, w) \models \psi$
$(M, w) \models K_a\varphi$	iff $(M, w') \models \varphi$ for all $w' \in W$ such that wR_aw' .

The formula φ is *satisfied* in the pointed model (M, w) iff $(M, w) \models \varphi$.

In full generality, the indistinguishability interpretation holds that for K_a to capture knowledge, the relation R_a must be an equivalence relation. A pointed Kripke model for which this is satisfied is often referred to as an *epistemic state*. In epistemic states, the relation is denoted by a tilde with a subscript: \sim_a .

Given pointed Kripke models and the indistinguishability interpretation, we have a semantic specification of one concept of knowledge. With this approach, we can build models of situations involving knowledge as we did with the toy example of Zoe and the hens. We can use these models to determine what the agent does or does not know. We also have the formal foundations in place to begin asking questions concerning how the agent’s knowledge or uncertainty develops when it receives *new information*, a topic studied in dynamic epistemic logic.

We may also ask more general questions concerning the concept of knowledge modeled using pointed Kripke models with indistinguishability relations: Instead of looking at a particular model at the time and asking which formulas the model makes true, we can ask which general principles all such models agree on.

2.5 Epistemological Principles in Epistemic Logic

Settling on the correct formal representation of knowledge involves reflecting carefully on the epistemological principles to which one is committed. An uncontroversial example of such a principle which most philosophers will accept is veridicality:

If a proposition is known, then it is true.

$$K_a\varphi \rightarrow \varphi.$$

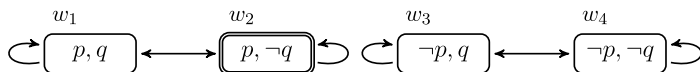
In a formal context this principle can be understood to say that if φ is known then it should always be satisfied in one's models. If it turns out that some of one's chosen models falsify the veridicality principle, then most philosophers would simply deem those models unacceptable.

Returning to pointed Kripke models, we can now ask which principles these models commit one to. In order to begin answering this question, we need to understand the most general features of our formalism. The strategy in modal logic in general (see Blackburn, de Rijke, & Venema 2001) is to abstract away from any given model's *contingent* features. Contingent features would include, for example, the specific number of worlds under consideration, the specific valuation of the atoms, and the choice of an actual world. In this case, the only features that are not contingent are those required by the general definition of a pointed Kripke model.

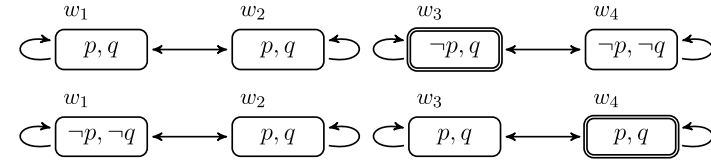
To abstract suitably, take a pointed Kripke model $(M, w) = (W, R, V, w)$. To determine whether the relation of this model is an equivalence relation we only need to consider the worlds and the relation. The pair of these elements constitute the fundamental level of the model and is called the *frame* of the model:

Definition: Let $(M, w) = (W, R, V, w)$ be a pointed Kripke model. Then the pair (W, R) is called the **frame** of (M, w) . Any model (M', w') which shares the frame (W, R) is said to be **built on** (W, R) .

Consider again the epistemic state for Zoe from above:



Several other models may be built on the same frame. The following are two examples:



With the notion of a frame, we may define the notion of validity of interest. It is the second term defined in the following:

Definition: A formula φ is said to be **valid in the frame** $F = (W, R)$ iff every pointed Kripke model built on F satisfies φ , i.e., iff for every $(M, w) = (F, V, w) = (W, R, V, w)$, $(M, w) \models \varphi$. A formula φ is **valid on the class of frames** \mathbf{F} (written $\mathbf{F} \models \varphi$) iff φ is valid in every frame F in \mathbf{F} .

The set of formulas valid on a class of frames \mathbf{F} is called the **logic** of \mathbf{F} . Denote this logic that is, the set $\{\varphi \in \mathcal{L}_K : \mathbf{F} \models \varphi\}$ by $\Lambda_{\mathbf{F}}$. This is a *semantic* approach to defining logics, each just a set of formulas. One may also define logics *proof-theoretically* by defining a logic as the set of formulas provable in some system. With logics as just sets of formulas, *soundness* and *completeness* results may then be expressed using set inclusion. To exemplify, let \mathbf{A} be a set of axioms and write $\mathbf{A} \vdash \varphi$ when φ is provable from \mathbf{A} using some given set of deduction rules. Let the resulting logic the set of theorems be denoted $\Lambda_{\mathbf{A}}$. It is the set of formulas from \mathcal{L}_K provable from \mathbf{A} , i.e., the set $\{\varphi \in \mathcal{L}_K : \mathbf{A} \vdash \varphi\}$. The logic $\Lambda_{\mathbf{A}}$ is sound with respect to \mathbf{F} iff $\Lambda_{\mathbf{A}} \subseteq \Lambda_{\mathbf{F}}$ and complete with respect to \mathbf{F} iff $\Lambda_{\mathbf{F}} \subseteq \Lambda_{\mathbf{A}}$.^[3]

Returning to the indistinguishability interpretation of knowledge, we may then seek to find the epistemological principles which the interpretation is committed to. There is a trivial answer of little direct interest: Let EQ be

the class of frames with equivalence relations. Then the logic of the indistinguishability interpretation is the set of formulas of \mathcal{L}_K which are valid over **EQ**, i.e., the set $\Lambda_{\text{EQ}} := \{\varphi \in \mathcal{L}_K : \text{EQ} \models \varphi\}$. Not very informative.

Taking an *axiomatic* approach to specifying the logic, however, yields a presentation in terms of easy to grasp principles. To start with the simplest, then the principle **T** states that knowledge is *factual*: If the agent knows φ , then φ must be true. The more cumbersome **K** states that if the agent knows an implication, then if the agent knows the antecedent, it also knows the consequent. I.e., if we include the derivation rule *modus ponens* (from $\varphi \rightarrow \psi$ and φ , conclude ψ) as rule of our logic of knowledge, **K** states that knowledge is *closed under implication*. The principle **B** states that if φ is true, then the agent knows that it considers φ possible. Finally, **4** states that if the agent knows φ , then it knows that it knows φ . **T**, **B** and **4** in the table below (the names are historical and not all meaningful).

K	$K_a(\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$
T	$K_a\varphi \rightarrow \varphi$
B	$\varphi \rightarrow K_a\widehat{K}_a\varphi$
4	$K_a\varphi \rightarrow K_aK_a\varphi$

In lieu of epistemological intuitions, we could discuss a concept of knowledge by discussing these and other principles. Should we accept **T** as a principle that knowledge follows? What about the others? Before we proceed, let us first make clear how the four above principles relate to the indistinguishability interpretation. To do so, we need the notion of a *normal modal logic*. In the below definition, as in the above principles, we are technically using *formula schemas*. For example, in $K_a\varphi \rightarrow \varphi$, the φ is a variable ranging over formulas in \mathcal{L}_K . Thus, strictly speaking, $K_a\varphi \rightarrow \varphi$ is not a formula, but a *scheme* for obtaining a formula. A *modal instance* of $K_a\varphi \rightarrow \varphi$ is then the formula obtained by letting φ be some

concrete formula from \mathcal{L}_K . For example, $K_ap \rightarrow p$ and $K_a(p \wedge K_aq) \rightarrow (p \wedge K_aq)$ are both modal instances of **T**.

Definition: Let $\Lambda \subseteq \mathcal{L}_K$ be a set of modal formulas. Then Λ is a **normal modal logic** iff Λ satisfies all of the following:

1. Λ contains all modal instances of the classical propositional tautologies.
2. Λ contains all modal instances of **K**.
3. Λ is closed under *modus ponens*: If $\varphi \in \Lambda$ and $\varphi \rightarrow \psi \in \Lambda$, then $\psi \in \Lambda$.
4. Λ is closed under *generalization* (a.k.a. *necessitation*): If $\varphi \in \Lambda$, then $K_a\varphi \in \Lambda$.

There is a unique *smallest* normal modal logic (given the set *Atom*) that which contains exactly what is required by the definition and *nothing more*. It is often called the **minimal normal modal logic** and is denoted by the boldface **K** (not to be confused with the non-boldface **K** denoting the schema).

The logic **K** is just a set of formulas from \mathcal{L}_K . I.e., $\mathbf{K} \subseteq \mathcal{L}_K$. Points 1.4. gives a perspective on this set: They provide an *axiomatization*. Often, as below, the schema **K** is referred to as an axiom, though really the instantiations of **K** are axioms.

To **K**, we can add additional principles as axioms (axiom schemes) to obtain stronger logics (logics that have additional theorems: Logics Λ for which $\mathbf{K} \subseteq \Lambda$). Of immediate interest is the logic called **S5**:

Definition: The logic **S5** is the smallest normal modal logic containing all modal instances of **T**, **B**, and **4**.

Here, then, is the relationship between the above four principles and the indistinguishability interpretation:

Theorem 1: The logic **S5** is the logic of the class of pointed Kripke models build on frames with equivalence relations. I.e., $\mathbf{S5} = \Lambda_{\text{EQ}}$.

What does this theorem tell us with respect to the principles of knowledge, then? In one direction it tells us that if one accepts the indistinguishability interpretation, then one has implicitly accepted the principles K, T, B and 4 as reasonable for knowledge. In the other direction, it tells us that if one finds that **S5** is the appropriate logic of knowledge *and* one finds that pointed Kripke models are the right way to semantically represent knowledge, then one must use an equivalence relation. Whether one should interpret this relation in terms of indistinguishability, though, is a matter on which logic is silent.

In discussing principles for knowledge, it may be that some of the four above seem acceptable, while others do not: One may disagree with the acceptability of B and 4, say, while accepting K and T. In understanding the relationship between **S5** and equivalence relations, a more fine-grained perspective is beneficial: Theorem 1 may be chopped into smaller pieces reflecting the contribution of the individual principles K, T, 4 and B to the equivalence requirement i.e., that the relation should be at the same time reflexive, symmetric and transitive.

Theorem 2: Let $F = (W, R)$ be a frame. Then:

- All modal instances of K are valid in F .
- All modal instances of T are valid in F iff R is reflexive.
- All modal instances of B are valid in F iff R is symmetric.
- All modal instances of 4 are valid in F iff R is transitive.

There are a number of insights to gain from Theorem 2. First, if one wants to use *any* type of Kripke model to capture knowledge, then one must accept K. Skipping some details, one must in fact accept the full logic **K** as this is the logic of the class of *all* Kripke models (see, e.g., Blackburn, de Rijke, & Venema 2001).

Second, the theorem shows that there is an intimate relationship between the individual epistemic principles and the properties on the relation. This, in turn, means that one, in general, may approach the “logic” in epistemic logic from two sides from intuitions about the accessibility relation or from intuitions about epistemic principles.

Several normal modal logical systems weaker than **S5** have been suggested in the literature. Here, we specify the logics by the set of their modal axioms. For example, the logic **K** is given by $\{K\}$, while **S5** is given by $\{K, T, B, 4\}$. To establish nomenclature, the following table contains a selection of principles from the literature with the frame properties they characterize, cf. Aucher (2014) and Blackburn, de Rijke, & Venema (2001), on the line below them. The frame conditions are not all straightforward.

In Table 1, the subscript on R_a is omitted to ease readability, and so is the domain of quantification W over which the worlds variables x, y, z range.

K	$K_a(\varphi \rightarrow \psi) \rightarrow (K_a\varphi \rightarrow K_a\psi)$ None: <i>Not applicable</i>
D	$K_a\varphi \rightarrow \widehat{K}_a\varphi$ Serial: $\forall x\exists y, xRy$.
T	$K_a\varphi \rightarrow \varphi$ Reflexive: $\forall x, xRx$.
4	$K_a\varphi \rightarrow K_aK_a\varphi$ Transitive: $\forall x, y, z$, if xRy and yRz , then xRz .
B	$\varphi \rightarrow K_a\widehat{K}_a\varphi$ Symmetric: $\forall x, y$, if xRy , then yRx .
5	$\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$ Euclidean: $\forall x, y, z$, if $xR_a y$ and $xR_a z$, then yRz .
.2	$\widehat{K}_aK_a\varphi \rightarrow K_a\widehat{K}_a\varphi$ Confluent: $\forall x, y$, if xRy and xRy' , then $\exists z, yRz$ and $y'Rz$.
.3	$(\widehat{K}_a\varphi \wedge \widehat{K}_a\psi) \rightarrow (\widehat{K}_a(\varphi \wedge \widehat{K}_a\psi) \vee \widehat{K}_a(\varphi \wedge \psi) \vee \widehat{K}_a(\psi \wedge \widehat{K}_a\varphi)$ No branching to the right: $\forall x, y, z$, if xRy and xRz , then yRz or $y = z$ or zRy
.3.2	$(\widehat{K}_a\varphi \wedge \widehat{K}_aK_a\psi) \rightarrow K_a(\widehat{K}_a\varphi \vee \psi)$ Semi-Euclidean: $\forall x, y, z$, if xRy and xRz , then zRx or yRz .
.4	$(\varphi \wedge \widehat{K}_aK_a\varphi) \rightarrow K_a\varphi$ Name unknown to authors: $\forall x, y$, if xRy , then $\forall z$, if xRz , then $x = z$ or yRz .

TABLE 1. Epistemic Principles and their frame conditions.

Adding epistemic principles as axioms to the basic minimal normal modal logic **K** yields new, normal modal logics. A selection is:

K	{K}
T	{K, T}
D	{K, D}
KD4	{K, D, 4}

KD45	{K, D, 4, 5}
S4	{K, T, 4}
S4.2	{K, T, 4, .2}
S4.3	{K, T, 4, .3}
S4.4	{K, T, 4, .4}
S5	{K, T, 5}

TABLE 2. Logic names and axioms

Different axiomatic specifications may produce the same logic. Notice, e.g., that the table's axiomatic specification $\{K, T, 5\}$ of **S5** does not match that given in the definition preceding Theorem 1, $\{K, T, B, 4\}$. Note also, there is more than one axiomatization of **S5**: the axioms $\{K, T, 5\}$, $\{K, T, B, 4\}$, $\{K, D, B, 4\}$ and $\{K, D, B, 5\}$ all give the **S5** logic (e.g., Chellas 1980). An often seen variant is $\{K, T, 4, 5\}$. However, it is redundant to add it as all its instances can be proven from K, T and 5. But as both 4 and 5 capture important epistemic principles (see Section 2.6), 4 is often sometimes included for the sake of philosophical transparency. For more equivalences between modal logics, see, e.g., the entry on modal logic or Chellas (1980) or Blackburn, de Rijke, and Venema (2001).

Logics may be stronger or weaker than each other, and knowing the frame properties of their axioms may help us to understand their relationship. For example, as 4 is derivable from $\{K, T, 5\}$, all the theorems of **S4** are derivable in **S5**. **S5** is thus *at least as strong* as **S4**. In fact, **S5** is also *strictly stronger*: It can prove things which **S4** cannot.

That **S5** may be axiomatized both by $\{K, T, B, 4\}$ and $\{K, T, 5\}$ may be seen through the frame properties of the axioms: every reflexive and euclidean relation (T and 5) is an equivalence relation (T, B, and 4). This also shows the redundancy of 4: If one has assumed a relation reflexive and euclidean, then it adds nothing new to additionally assume it to be

transitive. In general, having an understanding of the interplay between relational properties is of great aid in seeing relationships between modal logics. For example, noticing that every reflexive relation is also serial means that all formulas valid on the class of serial models are also valid on the class of reflexive models. Hence, every theorem of **D** is thus a theorem of **T**. Hence **T** is at least as strong as **D** (i.e., $\mathbf{D} \subseteq \mathbf{T}$). That **T** is also strictly stronger (not $\mathbf{T} \subseteq \mathbf{D}$) can be shown by finding a serial, non-reflexive model which does not satisfy some theorem of **T** (for example $K_a p \rightarrow p$).

2.6 Principles of Knowledge and Belief

With the formal background of epistemic logic in place, it is straightforward to slightly vary the framework in order to accommodate the concept of belief. Return to the language \mathcal{L}_{KB} of both knowledge and belief:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\psi \mid B_a\psi, \text{ for } p \in \text{Atom}.$$

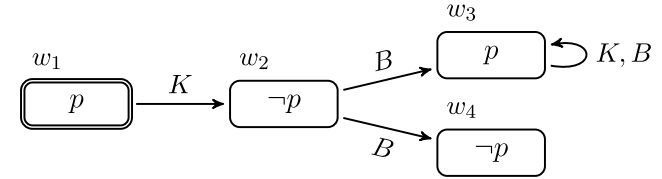
To interpret knowledge and belief formulas together in pointed Kripke models, all that is needed is an additional relation between possible worlds:

Definition: A *pointed Kripke model* for \mathcal{L}_{KB} is a tuple $(M, w) = (W, R_K, R_B, V, w)$ where

- W is a non-empty set of possible worlds,
- R_K and R_B are a binary relations on W ,
- $V: \text{Atom} \rightarrow \mathcal{P}(W)$ is a valuation, and
- $w \in W$.

R_K is the relation for the knowledge operator and R_B the relation for the belief operator. The definition makes no further assumptions about their

properties. In the figure below we provide an illustration, where the arrows are labeled in accordance with the relation they correspond to. The reflexive loop at w_3 is a label indicating that it belongs to both relations, i.e., $(w_3, w_3) \in R_K$ and $(w_3, w_3) \in R_B$.



The satisfaction relation is defined as above, but with the obvious changes for knowledge and belief:

$$(M, w) \models K_a\varphi \text{ iff } (M, w') \models \varphi \text{ for all } w' \in W \text{ such that } wR_K w'.$$

$$(M, w) \models B_a\varphi \text{ iff } (M, w') \models \varphi \text{ for all } w' \in W \text{ such that } wR_B w'.$$

The indistinguishability interpretation puts very strong requirements on the accessibility relation for knowledge. These have now been stripped away and so has any commitment to the principles T, B, D, 4 and 5. Taking Kripke models as basic semantics, we are still committed to K, though this principle is not unproblematic as we shall see below in our discussion of the problem of logical omniscience.

Of the principles from Table 1, T, D, B, 4 and 5 have been discussed most extensively in the literature on epistemic logic, both as principles for knowledge and as principles for belief. The principle T for knowledge

$$K_a\varphi \rightarrow \varphi$$

is broadly accepted. Knowledge is commonly taken to be *veridical*—only true propositions can be known. For, e.g., Hintikka (1962) and Fagin et al.

(1995), the failure of T for belief is the defining difference between the two notions.

Though belief is not commonly taken to be veridical, beliefs are typically taken to be *consistent*. I.e., agents are taken to *never* believe the contradiction that is, any formula equivalent with $(p \wedge \neg p)$ or \perp , for short. That beliefs should be consistent is then captured by the principle

$$\neg B_a \perp.$$

The principle $\neg B_a \perp$ is, on Kripke models, equivalent with the principle D, $B_a \varphi \rightarrow \widehat{B}_a \varphi$. Hence the validity of $\neg B_a \perp$ requires serial frames. Witness, e.g., its failure in w_1 above: As there are no worlds accessible through R_B , all accessible worlds satisfy \perp . Hence w_1 satisfies $B_a \perp$, violating consistency. Notice also that $\neg B_a \perp$ may be re-written to $\widehat{B}_a \top$, which is true at a world just in case some world is accessible through R_B . Its validity thus ensures seriality.

Notice that the veridicality of knowledge ensures its consistency: Any reflexive frame is automatically serial. Hence accepting $K_a \varphi \rightarrow \varphi$ implies accepting $\neg K_a \perp$.

Of the principles D, 4 and 5, the two latter have received so far the most attention, both for knowledge and for belief. They are commonly interpreted as governing of *principled access* to own mental states. The 4 principles

$$\begin{aligned} K_a \varphi &\rightarrow K_a K_a \varphi \\ B_a \varphi &\rightarrow B_a B_a \varphi \end{aligned}$$

are often referred to as *principles of positive introspection*, or for knowledge the '*KK*' principle. Both principles are deemed acceptable by, e.g., Hintikka (1962) on grounds *different* from introspection. He argues based on an autoepistemic analysis of knowledge, using a non-Kripkean

possible worlds semantics called *model systems*. Hintikka holds that when an agent commits to knowing φ , the agent commits to holding the same attitude no matter what new information the agent will encounter in the future. This entails that in all the agent's epistemic alternatives—for Hintikka, all the model sets (partial descriptions of possible worlds) where the agent knows at least as much they now do—the agent still knows φ . As $K_a \varphi$ thus holds in all the agent's epistemic alternatives, Hintikka concludes that $K_a K_a \varphi$. Likewise, Hintikka endorses 4 for belief, but Lenzen raises objections (Lenzen 1978: ch. 4).

Williamson argues against the general acceptability of the principle (Williamson 2000: ch. 5) for a concept of knowledge based on slightly inexact observations, a so-called *margin of error principle* (see, e.g., Aucher 2014 for a short summary).

The 5 principles

$$\begin{aligned} \neg K_a \varphi &\rightarrow K_a \neg K_a \varphi \\ \neg B_a \varphi &\rightarrow B_a \neg B_a \varphi \end{aligned}$$

are often referred to as *principles of negative introspection*. Negative introspection is quite controversial as it poses very high demands on knowledge and belief. The schema 5 may be seen as a *closed world assumption* (Hendricks 2005): The agent has a complete overview of all the possible worlds and own information. If $\neg \psi$ is considered possible ($\widehat{K}_a \neg \psi$, i.e., $\neg K_a \psi$), then the agent knows it is considered possible ($K_a \neg K_a \psi$). Such a closed world assumption is natural when constructing hyper-rational agents in, e.g., computer science or game theory, where the agents are assumed to reason as hard as logically possible about their own information when making decisions.

Arguing against 5 is Hintikka (1962), using his conception of epistemic alternatives. Having accepted T and 4 for knowledge, 5 stands or falls with

the assumption of a symmetric accessibility relation. But, Hintikka argues, the accessibility relation is not symmetric: If the agent possesses some amount of information at model set s_1 , then the model set s_2 where the agent has learned something more will be an epistemic alternative to s_1 . But s_1 will not be an epistemic alternative to s_2 , because in s_1 , the agent does by hypothesis not know as much as it does in s_2 . Hence the relation is not symmetric, so 5 is not a principle of knowledge, on Hintikka's account.

Given Hintikka's non-standard semantics, it is a bit difficult to pin down whether he would accept a normal modal logic as the logic of knowledge and belief, but if so, then **S4** and **KD4** would be the closest candidates (see Hendricks & Rendsvig 2018 for this point). By contrast, for knowledge von Kutschera argued for **S4.4** (1976), Lenzen suggested **S4.2** (1978), van der Hoek argued for **S4.3** (1993), and Fagin, Halpern, Moses, and Vardi (1995) and many others use **S5** for knowledge and **KD45** for belief.

Beyond principles governing knowledge and principles governing belief, one may also consider principles governing the interplay between knowledge and belief. Three principles of interest are

KB1	$K_a\varphi \rightarrow B_a\varphi$
KB2	$B_a\varphi \rightarrow K_aB_a\varphi$
KB3	$B_a\varphi \rightarrow B_aK_a\varphi$

The principles KB1 and KB2 were introduced by Hintikka (1962), who endorses only the former (see §3.7), noting that Plato is also committed to KB1 in *Theatetus*. The first principle, KB1, captures the intuition that knowledge is a stronger notion than belief. The second—like 4 and 5—captures the idea that one has privileged access to one's own beliefs. The third, stemming from Lenzen (1978), captures the notion that beliefs are held with some kind of conviction: if something is believed, it is believed to be known.

Though the interaction principles KB1—KB3 may look innocent on their own, they may lead to counterintuitive conclusions when combined with specific logics of knowledge and belief. First, Voorbraak (1993) shows that combining 5 for knowledge and D for belief with KB1, implies that

$$B_aK_a\varphi \rightarrow K_a\varphi$$

is a theorem of the resulting logic. Assuming that knowledge is truthful, this theorem entails that agents cannot believe to know something which happens to be false.

If additionally KB3 is added, the notions of knowledge and belief *collapse*. I.e., it may be proven that $B_a\varphi \rightarrow K_a\varphi$, which, in combination with KB1 entails that

$$B_a\varphi \leftrightarrow K_a\varphi.$$

Hence, the two notions have collapsed to one. This was stated in 1986, by Kraus and Lehmann.

If one is not interested in knowledge and belief collapsing, one must thus give something up: One cannot have both 5 for knowledge, D for belief and KB1 and KB3 governing their interaction. Again, results concerning correspondence between principles and relational properties may assist: In 1993, van der Hoek showed, based on a semantic analysis that where the four principles are jointly sufficient for collapse, *no subset of them is*. Giving up any one principle will thus eliminate the collapse. Weakening KB1 to hold only for non-modal formulas is also sufficient to avoid collapse (cf. Halpern 1996).

For more on epistemic interaction principles, the principles .2, .3, .3.2. and .4, and relations to so-called *conditional beliefs*, see Aucher (2014). For an introduction to conditional beliefs and relations to several other types of knowledge from the philosophical literature, see Baltag and Smets (2008).

The latter also includes discussion concerning the interdefinability of various notions, as does Halpern, Samet, and Segev (2009) for knowledge and (non-conditional) belief.

Note that although we mainly focus on the Kripke semantics in this entry, there are other epistemic models for potentially weaker logic, such as neighborhood models (e.g., van Benthem et al. 2014) and topological models (e.g., Baltag, Bezhanishvili, et al. 2019).

3. Knowledge in Groups

We human beings are preoccupied with the epistemic states of other agents. In ordinary life, we reason with varying degrees of success about what others know. We are especially concerned with what others know about us, and often specifically about what they know about what we know.

Does she know that I know where she buried the treasure?

Does she know that I know that she knows?

And so on.

Epistemic logic can reveal interesting epistemic features of systems involving groups of agents. In some cases, for example, emergent social phenomena depend on agents reasoning in particular ways about the knowledge and beliefs of other agents. As we have seen, traditional systems of epistemic logic applied only to single-agent cases. However, they can be extended to groups or multi-agent systems in a relatively straightforward manner.

As David Lewis noted in his book *Convention* (1969) many prominent features of social life depend on agents assuming that the rules of some

practice are matters of *common knowledge*. For example, drivers know that a red traffic light indicates that they should stop at an intersection. However, for the convention of traffic lights to be in place at all, it is first necessary that drivers must also know that other drivers know that *red* means *stop*. In addition, drivers must also know that everyone knows that everyone knows that The conventional role of traffic lights relies on all drivers knowing that all drivers know the rule, that the rule is a piece of common knowledge.

A variety of norms, social and linguistic practices, agent interactions and games presuppose common knowledge, first formalized by Aumann (1976) and with earliest epistemic logical treatments by Lehmann (1984) and by Halpern and Moses (1984). In order to see how epistemic logic sheds light on these phenomena, it is necessary to introduce a little more formalism. Following the standard treatment (see, e.g., Fagin et al. 1995), we can syntactically augment the language of propositional logic with n knowledge operators, one for each agent involved in the group of agents under consideration. The primary difference between the semantics given for a mono-agent and a multi-agent semantics is roughly that n accessibility relations are introduced. A modal system for n agents is obtained by joining together n modal logics where for simplicity it may be assumed that the agents are homogenous in the sense that they may all be described by the same logical system. The epistemic logical system for n agents consists of n copies of a certain modal system. In such an extended epistemic logic it is possible to express that some agent in the group knows a certain fact that an agent knows that another agent knows a fact etc. It is possible to develop the logic even further: Not only may an agent know that another agent knows a fact, but they may all know this fact simultaneously.

3.1 Multi-Agent Languages and Models

To represent knowledge for a set \mathcal{A} of n agents, first let's stipulate a language. Let \mathcal{L}_{Kn} be given by the *Backus-Naur form*

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \text{ for } p \in \text{Atom}, i \in \mathcal{A}.$$

To represent knowledge for all n agents jointly in pointed Kripke models, all that is needed is to add suitably many relations:

Definition: A *pointed Kripke model* for \mathcal{L}_{Kn} is a tuple $(M, w) = (W, \{R_i\}_{i \in \mathcal{A}}, V, w)$ where

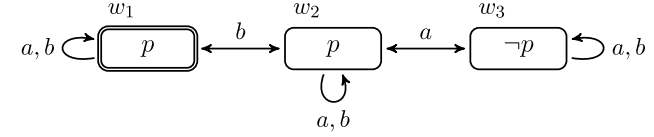
- W is a non-empty set of possible worlds,
- For every $i \in \mathcal{A}$, R_i is a binary relation on W ,
- $V: \text{Atom} \rightarrow \mathcal{P}(W)$ is a valuation, and
- $w \in W$.

To also incorporate beliefs, simply apply the same move as in the single agent case: augment the language and let there be two relations for each agent.

The definition uses a family of relations $\{R_i\}_{i \in \mathcal{A}}$. In the literature, the same is denoted $(W, R_i, V, w)_{i \in \mathcal{A}}$. Alternatively, R is taken to be a function sending agents to relations, i.e., $R: \mathcal{A} \rightarrow \mathcal{P}(W \times W)$. Then for each $i \in \mathcal{A}$, $R(i)$ is a relation on W , often denoted R_i . These are stylistic choices.

When considering only a single agent, it is typically not relevant to include more worlds in W than there are possible valuations of atoms. In multi-agent cases, this is not the case: to express the different forms of available higher-order knowledge, many copies of “the same” world are needed. Let us exemplify for $\mathcal{A} = \{a, b\}$, $\text{Atom} = \{p\}$ and each $R_i, i \in \mathcal{A}$, an equivalence relation. Let us represent that both a and b

know p , but b does not know that a knows p , i.e., $K_ap \wedge K_bp \wedge \neg K_bK_ap$. Then we need three worlds:



If we delete w_1 and let w_2 be the actual world, then a would lose knowledge in p : both p worlds are needed. In general, if W is assumed to have any fixed, finite size, there will be some higher-order information formula that cannot be satisfied in it.

3.2 Notions of Group Knowledge

Multi-agent systems are interesting for other reasons than to represent higher-order information. The individual agents' information may also be pooled to capture what the agents know jointly, as group knowledge (see Baltag, Boddy, & Smets 2018 for a recent discussion). A standard notion in this style is *distributed knowledge*: The knowledge the group *would have* if the agents share all their individual knowledge. To represent it, augment the language \mathcal{L}_{Kn} with operators

$$D_G \text{ for } G \subseteq \mathcal{A},$$

to make $D_G\varphi$ a well-formed formula. Where $G \subseteq \mathcal{A}$ is a group of agents, the formula $D_G\varphi$ reads that it is *distributed knowledge in the group G that φ* .

To evaluate $D_G\varphi$, we define a new relation from those already present in the model. The idea behind the definition is that if some one agent has eliminated a world as an epistemic alternative, then so will the group. Define the relation as the intersection of the individual agents' relations:

$$R_G^D = \bigcap_{i \in G} R_i$$

In the three state model, R_G^D contains only the three loops. To evaluate a distributed knowledge formula, use the same form as for other modal operators:

$$(M, w) \models D_G \varphi \text{ iff } (M, w') \models \varphi \text{ for all } w' \in W \text{ such that } wR_G^D w'.$$

It may be the case that some very knowing agent knows all that is distributed knowledge in G , but it is not guaranteed. To capture that all the agents know φ , we could use the conjunction of the formulas $K_i \varphi$ for $i \in \mathcal{A}$, i.e., $\bigwedge_{i \in \mathcal{A}} K_i \varphi$. This is a well-defined formula if \mathcal{A} is finite (which it typically is). If \mathcal{A} is not finite, then $\bigwedge_{i \in \mathcal{A}} K_i \varphi$ is not a formula in \mathcal{L}_{Kn} , as it has only finite conjunctions. As a shorthand for $\bigwedge_{i \in \mathcal{A}} K_i \varphi$, it is standard to introduce the *everybody knows* operator, E_G :

$$E_G \varphi := \bigwedge_{i \in \mathcal{A}} K_i \varphi.$$

In the three world model, $K_a p \wedge K_b p$, so $E_{\{a,b\}} p$.

That everybody knows something does not mean that this knowledge is shared between the members of the group. The three world model exemplifies this: Though $E_{\{a,b\}} p$, it also the case that $\neg K_b E_{\{a,b\}} p$.

To capture that there is no uncertainty in the group about φ nor *any higher-order* uncertainty about φ being known by all agents, no formula in the language \mathcal{L}_{Kn} is enough. Consider the formula

$$E_G^k \varphi$$

where E_G^k is short for k iterations of the E_G operator. Then for no natural number k will the formula $E_G^k \varphi$ be enough: it could be the case that b doesn't know it! To rectify this situation, one could try

$$\bigwedge_{k \in \mathbb{N}} E_G^k \varphi$$

but this is not a formula as \mathcal{L}_{Kn} only contains finite conjunctions.

Hence, though the E_G operator is definable in the language \mathcal{L}_{Kn} , a suitable notion of *common knowledge* is not. For that, we again need to define a new relation on our model. This time, we are interested in capturing that nobody considers φ epistemically possible *anywhere*. To build the relation, we therefore first take the union the relations of all the agents in G , but this is not quite enough: to use the standard modal semantic clause, we must also be able to reach all of the worlds in this relation *in a single step*. Hence, let

$$R_G^C := \left(\bigcup_{i \in G} R_i \right)^*$$

where $(\cdot)^*$ is the operation of taking the *reflexive transitive closure*. If R is a relation, then $(R)^*$ is R plus the reflexive arrow and all the pairs missing to make R a transitive relation. Consider the three world model: With the relation $\bigcup_{i \in \{a,b\}} R_i$, we can reach w_3 from w_1 in two steps, stopping over at w_2 . With $(\bigcup_{i \in \{a,b\}} R_i)^*$, w_3 is reachable in one step: By the newly added transitive link from w_1 to w_3 .

To represent common knowledge, augment the *Backus-Naur form* of \mathcal{L}_{Kn} with operators

$$C_G \text{ for } G \subseteq \mathcal{A},$$

to make $C_G \varphi$ a well-formed formula. Evaluate such formulas by the semantic clause

$$(M, w) \models C_G \varphi \text{ iff } (M, w') \models \varphi \text{ for all } w' \in W \text{ such that } wR_G^C w'.$$

Varying the properties of the accessibility relations R_1, R_2, \dots, R_n , as described above results in different epistemic logics. For instance, system **K** with common knowledge is determined by all frames, while system **S4** with common knowledge is determined by all reflexive and transitive frames. Similar results can be obtained for the remaining epistemic logics (Fagin et al. 1995, van Ditmarsch et al. 2007). For an informal discussion, consult the entry on common knowledge.

4. Beyond Knowing That

So far, we have focused on *know-that*, the standard notion analyzed in epistemic logic. However, as mentioned earlier, in natural language, we also express knowledge in terms of know-wh, i.e., the verb “know” followed by an embedded question, such as:

- Alice *knows whether* the claim is true.
- Bob *knows what* the password is.
- Charlie *knows how* to prove the theorem.
- Dave *knows why* Charlie *knows how* to prove the theorem.

Logical research on know-wh also dates back to Hintikka (1962), who dedicated a chapter to the formal account of *know-who*, and proposed to formalize other *know-wh* alike. According to Hintikka, to know “who Mary is” is to know the answer to the corresponding question “Who is Mary?”. Essentially, one must be able to identify the person denoted by the name *Mary* (supposing it is a proper name). Therefore, it is a piece of *de re* knowledge, i.e., the knowledge of an *object*. To render it formally, Hintikka treats “Bob knows who Mary is” as “there is a person such that Bob *knows that* this person is Mary”, which connects know-that with know-wh by using quantification over objects. To express Hintikka’s

formalization, we need to extend our basic language of propositional modal logic to *first-order modal logic*, the modal logic with quantifiers and predicates. As another example, “Bob knows who murdered Dave” can be formalized in first-order epistemic logic as $\exists x K_{\text{Bob}} \text{Murder}(x, \text{Dave})$. Note that the order of the quantifier and the know-that modality is crucial to capture *de re* knowledge. The swapped *de dicto* version $K_{\text{Bob}} \exists x \text{Murder}(x, \text{Dave})$ merely says “Bob knows that Dave is murdered”. For Hintikka, it is also important how the agent is identified, cf. (Hintikka & Symons 2003).

Such treatments of know-wh are also supported by a large body of research on the semantics of (embedded) questions in linguistics (e.g., Groenendijk & Stokhof 1982; Harrah 2002), and epistemology of knowledge-wh (e.g., Stanley & Williamson 2001). Linguists have discussed various readings of epistemic expressions with embedded questions, in particular the so-called *mention-some* and *mention-all* interpretations. The mention-some interpretation requires that one knows at least one (correct) answer in order to possess know-wh knowledge, while the mention-all interpretation requires that one knows all the (correct) answers (cf. Groenendijk & Stokhof 1982). For example, a (strong) exhaustive mention-all reading for “Bob knows who came to the party” is formalized by $\forall x (K_{\text{Bob}} \text{Came}(x) \vee K_{\text{Bob}} \neg \text{Came}(x))$ where \forall replaces \exists as the leading quantifier, contra the previous examples. The interpretation of know-wh is also context-dependent in general, which introduces further complications (e.g., Aloni 2001, 2018). Coming back to logic, Hintikka (1962) discussed the notion of consistency in such a logic of know-who and various philosophical issues, such as trans-world identity. In his latter work, Hintikka (2003) also made use of his *Independent Friendly Logic* to go beyond the first-order quantification, in order to capture wh-knowledge about higher-order entities, such as “I know whom every young mother should trust” with the intention of “trusting *her own mother*”.

Although Hintikka gave a basic semantics of know-wh, the systematical study of the epistemic logics of know-wh was not developed until recently. This is partly due to the underdevelopment of first-order modal logic, compared to its propositional brother (cf., Gochet & Gribomont 2006 for a survey on first-order epistemic logic). Various well-known systems of first-order modal logic do not enjoy the technically desirable properties of propositional modal logic, but suffers incompleteness, undecidability and the failure of interpolation (cf., Braüner & Ghilardi 2007). In particular, it is very hard to find decidable fragments of first-order modal logic to make it computationally appealing (e.g., Hodkinson, Wolter, & Zakharyashev 2000).

Nevertheless, there are works on clarifying the *de re* and *de dicto* distinctions in the setting of quantified epistemic logic (e.g., Grove 1995; Corsi & Orlandelli 2013; Holliday & Perry 2014; Occhipinti Liberman & Rendsvig 2022), and a number of application-driven frameworks based on first-order epistemic logic (e.g., Mika Cohen & Dam 2007; Kaneko & Nagashima 1996; Sturm, Wolter, & Zakharyashev 2000; Belardinelli & Lomuscio 2011), although they are not directly about know-wh. As one of a few exceptions directly related to know-wh based on quantified modal logic, Rendsvig (2012) discusses Frege’s Puzzle about Identity in a framework where agents may have knowledge-who about the referents of proper names.

Another line of research directly related to know-who is based on *term-modal logic* proposed by (Fitting, Thalmann, & Voronkov 2001), a variant of quantified modal logic where modalities are subscripted by *terms* instead of indexes. This allows formulas such as $\exists y \exists x K_y \text{Murder}(x, \text{Dave})$ (“there is someone that knows who murdered Dave”), or $\exists x K_{\text{Bob}}(\text{Murder}(x, \text{Dave}) \wedge x = \text{Adam})$ (“Bob knows who murdered Dave and that is Adam”). Starting from (Kooi 2007), various versions of (dynamic) term-modal logic were proposed to discuss know-

who, such as Occhipinti Liberman, Achen, & Rendsvig 2020; and Y. Wang, Wei, & Seligman 2022. These are closer to the original systems suggested by Hintikka (1962), where the names of agents are also treated as terms. Term-modal logic formalization of know-who in social networks, and the dynamics that influence it, is a main theme in Occhipinti Liberman & Rendsvig 2022. A few decidable fragments of term-modal logics were discovered in Orlandelli & Corsi 2017; Padmanabha & Ramanujam 2019a, 2019b; and Occhipinti Liberman, Achen, and Rendsvig 2020, where a recent review of the term-modal logic literature can be found.

Concerning a different type of know-wh, some formal discussions on know-how have appeared in the setting of *strategy logics*. It is observed that in the setting of *Epistemic Alternating Temporal Logic*, just combining “know” and “can” does not capture know-how (cf., Jamroga & van der Hoek 2004; Herzig 2015) but *de dicto* knowledge: I know there is some way to guarantee φ . Instead of explicitly using the quantifiers, people introduced alternative semantics for the coalition operators in epistemic variants of ATL (e.g., Jamroga & Ågotnes 2007; Maubert, Pinchinat, Schwarzentruher, & Stranieri 2020; Maubert, Murano, Pinchinat, Schwarzentruher, & Stranieri 2020).

Recently, a *know-wh first* approach has emerged that treats know-wh as an independent modality, just as the modality of know-that (cf. the survey by Y. Wang 2018b). The general motivation of this line of research is to focus on the logical behavior of each know-wh as a *primitive concept*, rather than breaking it down in the syntax into quantifiers, modalities, and predicates in the full language of predicate modal logic. This approach is inspired by the early work of Plaza 1989 and Ma & Guo 1983, where a Kv operator is introduced to capture the notion of “knowing the value of a constant”. Semantically, Kvd can be viewed as a “bundle” of $\exists x K(d = x)$, which says the agent *knows what* the value of d is. In this

line of research, various languages were proposed with both the know-that operator and the know-wh operator of certain types, whose semantics is given by the first-order modal interpretation, as we described above for the Kv operator.

As an example of a *know-wh first* framework, consider the following epistemic language of (goal-directed) know-how and know-that discussed by Li & Wang (2021a), where $Kh_i\varphi$ says that agent i knows how to achieve the goal φ .

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid Kh_i\varphi$$

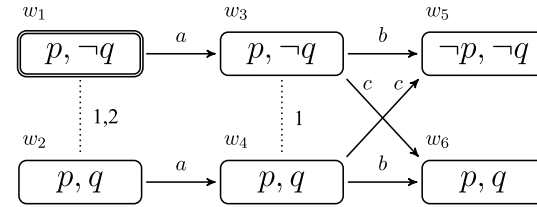
The semantics is given over a type of Kripke models with both epistemic relations $\sim_i, i \in \mathcal{A}$ and labeled action relations $\xrightarrow{a}, a \in Act$, where $Act = \cup_{i \in \mathcal{A}} Act_i$ is the union of the sets of actions of each agent, which do not explicitly described in the formal language. Such a model takes the form $\mathcal{M} = (W, \{\sim_i\}_{i \in \mathcal{A}}, \{\xrightarrow{a}\}_{a \in Act}, V)$. The satisfaction relation between a pointed Kripke model \mathcal{M}, w and a $Kh_i\varphi$ formula is given by the following $\exists xK$ -style schema:

$\mathcal{M}, w \models Kh_i\varphi \iff \text{there is a plan } \pi \text{ such that for all } w' \sim_i w:$ <ol style="list-style-type: none"> 1. π is <i>strongly executable</i> by i on w'; 2. $\mathcal{M}, v \models \varphi$ for each final state v reachable by executing π.
--

A plan π can have different forms, such as a finite linear sequence of actions, a conditional strategy given the knowledge of agents, or even a program with loops (cf., Li & Wang 2021b). A plan is *strongly executable* intuitively means the plan will never get stuck and always terminates successfully.

To exemplify the framework, suppose that a patient is experiencing some rare symptom p . To know the cause (q or $\neg q$), Doctor 1 suggests the

patient first take a special test (a), which is only available at Doctor 1's hospital, and send the results to Doctor 2 in another city, who is more experienced in examining the results. Doctor 2 will then know the cause, depending on which, different medicines (b or c) should be subscribed to cure the patient. The situation is depicted below as a model with both epistemic relations (dotted lines labeled by 1, 2; reflexive ones omitted) and action relations (solid lines labeled by a, b, c). Note that only Doctor 1 can execute action a , and only Doctor 2 can perform b or c . After action a , Doctor 2 knows whether she is at w_3 or at w_4 , but Doctor 1 is still uncertain about it.



Intuitively, Doctor 1 knows how to let Doctor 2 know how to cure the patient ($\neg p$), although neither Doctor 1 nor Doctor 2 knows how to cure the patient alone. This can be expressed by $\neg Kh_1\neg p \wedge \neg Kh_2\neg p \wedge Kh_1((K_2q \vee K_2\neg q) \wedge Kh_2\neg p)$, which is true at w_1 and w_2 , given a proper notion of plans.

Finally, logics of know-how may be obtained for this framework. The axioms depend on the specific assumptions made of the Kripke models, as in standard epistemic logic. For example, if we allow *conditional knowledge-based plans* and consider agents with *perfect recall* (as in Fervari, Herzig, Li, & Wang 2017), then the following axioms and rules on top of the S5 proof system for K_i are complete:

$$\text{EMP} \quad K_i p \rightarrow Kh_i p \quad \text{KhK} \quad Kh_i p \rightarrow Kh_i K_i p$$

$$\begin{array}{ll}
\text{KKh} & Kh_i p \rightarrow K_i Kh_i p \quad \text{KhKh} \quad Kh_i Kh_i p \rightarrow Kh_i p \\
\text{Khbot} & \neg Kh_i \perp \quad \text{MONOKh} \quad \frac{\vdash \varphi \rightarrow \psi}{Kh_i \varphi \rightarrow Kh_i \psi}
\end{array}$$

Note that, in such a *know-wh first* approach, the axioms directly capture the characteristic properties of the specific type of know-wh in concern rather than specifying the behavior of quantifiers, know-that modalities, and predicates in first-order modal logic. For example, in the above proof system for a specific goal-directed know-how based on conditional plans, EMP proposes that the knowledge of p entails the knowledge of how to achieve p (simply by maintaining the status quo); KKh postulates that know-how is introspective, which may distinguish know-how from mere *ability to*, that the agent may not know it has; KhKh asserts that if one knows how to know how, then one already knows how, which reflects the compositionality of the conditional plans witnessing the knowledge-how; KhK, meanwhile, expresses that knowledge of how to achieve a goal implies the ability to knowingly achieve it.

As in standard epistemic logic, these axioms, although arguably contentious, provide an engaging forum for philosophical investigations. For instance, the interactive axioms about both K and Kh may help us sharpen our understanding of knowledge-how and knowledge-that. KKh may raise philosophical interests, given the heated debate about its counterpart, the KK-principle, in standard epistemic logic. Moreover, KhK is not intuitively valid if we allow one to forget what was known (no perfect recall): you may know how to get drunk, but by the time you are really drunk, you may not realize it. As another example without K , if we allow *activity* know-how alongside goal-directed know-how, then KhKh does not hold intuitively: knowing how to let myself know how to swim (by hiring a private coach) does not mean I know how to swim right now. There are of course other ways to make these axioms valid or invalid. The

rich landscape of possible axioms encourages thorough and systematic philosophical exploration.

Within the *know-wh first* approach, other know-wh modalities, semantics and logics have been studied. Here we list a few existing works, not meant to be exhaustive, and their connections to other well-known logics.

The logic of “**know-whether**” (and ignorance) has been explored in van der Hoek & Lomuscio 2003; Fan, Wang, & Ditmarsch 2015; and Fine 2018. It is closely connected to the *non-contingency logics* in the literature. Variants of such logics are under extensive study, e.g., Fan 2019, 2021. In addition, “know-whether” has proven useful in simplifying the axiomatization of common knowledge, as shown by Herzig & Perrotin (2020). The logic of “**know-what**” is another area that received extensive attention (Y. Wang & Fan 2013; Baltag 2016), extending upon the logic of “know-value” by Plaza (1989). It also has surprising connections with *weakly aggregative modal logics*. The logic of “**know-how**” has been studied in planning-based approaches such as Y. Wang 2018a; Fervari, Herzig, et al. 2017; Li & Wang 2021b), and coalition-based approaches such as Naumov & Tao 2017, 2018. These works are connected to epistemic strategy logics such as *Alternating Epistemic Temporal Logic*. The logic of “**know-why**” was explored by Xu, Wang, & Studer (2021) as a fragment of a quantified version of *justification logic*. The logic of “**know-who**” has been investigated in Wang, Wei & Seligman 2022; and Epstein & Naumov 2021), and has connections to *term-modal logic*. Such know-wh logics are applied to AI, e.g., for planning and reasoning, as demonstrated by Li & Wang 2021a; Naumov & Tao 2020, 2019; and Jiang & Naumov 2022.

These know-wh logics share some common features. First, they are mostly *non-normal* (in the technical sense). For example, knowing how to achieve φ and knowing how to achieve ψ together does not entail knowing how to

achieve $\varphi \wedge \psi$, contra the normal modal logical aggregation principle $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$ viewing know-how as a \Box modality. Moreover, since the know-wh modalities semantically correspond to complicated constructions expressible in first-order modal logic, the models tend to be rich in information. However, the languages of know-wh logics are simple, causing a mismatch between the syntax and semantics. This leads to difficulties in axiomatizing these logics, requiring new techniques (e.g., Gu & Wang 2016). Finally, these logics are often *decidable*, contra full first-order modal language approaches to know-wh. Many know-wh logics can be viewed as one-variable fragments of first-order modal logic. Extending the idea of know-wh logics, Y. Wang (2017) proposed to study the fragments of first-order modal logic where a quantifier always occurs with a modality without any restriction on the number of variables or the arity of the predicates. This leads to the so-called *bundled fragments*, many of which are decidable too (cf. the survey by Liu, Padmanabha, Ramanujam, & Wang 2023).

Besides the first-order epistemic logic and the know-wh first approaches to the logic of know-wh, there is a linguistically motivated approach to know-wh based on *inquisitive logic* (e.g., Ciardelli & Roelofsen 2011). Departing from the usual logical frameworks, inquisitive logic formalizes reasoning patterns with both statements and questions in a unified picture based on *inquisitive semantics*. This provides the opportunity to handle know-wh as it appears in the natural language: the verb know followed by an (embedded) question. Therefore, it is possible to use modal inquisitive logic to express know-wh as well, as demonstrated in Ciardelli 2014, 2016, 2023; and Ciardelli & Roelofsen 2015. For example, consider the following language of *inquisitive modal logic*:

$$\varphi ::= p \mid \perp \mid (\varphi \wedge \varphi) \mid (\varphi \rightarrow \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \backslash \vee \varphi) \mid \Box\varphi$$

One can use $\Box(\varphi \backslash \vee \neg\varphi)$ to express know-whether φ , where $(\varphi \backslash \vee \neg\varphi)$ is an inquisitive disjunction representing a *polar question* about whether φ .

In a first-order inquisitive setting, one can use the combination of the modality and the inquisitive existential quantifier to express mention-some interpretation for know-wh statements (cf., Ciardelli 2023). Moreover, there are intimate connections between inquisitive logic and the logics of know-wh. In particular, inquisitive logic and many other intermediate logics can be interpreted as epistemic logics of know-how (to prove/resolve/solve) (e.g., H. Wang, Wang, & Wang 2022). For example, in the propositional setting, a state (a set of possible worlds) that supports $(\varphi \backslash \vee \neg\varphi)$ according to inquisitive semantics can be viewed as an S5 epistemic model for an agent that knows how to answer the question $(\varphi \backslash \vee \neg\varphi)$.

The logic of know-wh is an active area of research with many open questions. In addition to numerous specific technical questions about existing epistemic logics of know-wh in various approaches, there are also many conceptual questions about the nature of know-wh that encourage systematic studies. For example, what are the corresponding dynamics that update knowledge-wh? There are some initial attempts to model specific dynamics in the setting of know-value (van Eijck et al. 2017; and M. Cohen, Tang, & Wang 2021), know-how (Areces, Fervari, Saravia, & Velázquez-Quesada 2022), and know-who (Occhipinti Liberman & Rendsvig 2022). Moreover, what are the corresponding group knowledge of know-wh? Besides the coalition-based know-how logics such as Naumov & Tao 2018, some attempts on distributed/commonly know-whether have been explored by Su 2017; and Fan, Grossi, et al. 2020 [Other Internet Resources]). Further, how different types of know-wh interact? An attempt to blend know-how and know-value was pioneered in Jiang & Naumov 2022, in a setting with strategic games with data. Moreover, linguistic discussions on the non-exhaustive interpretations of embedded question suggests more refined structures of know-wh (e.g., Xiang 2016). For example, does someone know how to prove a theorem if they know one correct proof, but believe falsely that another proof also

works? Some initial logical discussions can be found in Yang 2023. Finally, although there is a large body of philosophical research regarding the logical principles of know-that (such as introspection and logical omniscience), the corresponding know-wh principles are seldom discussed philosophically. In general, it is a promising direction to forge links with the contemporary epistemology of knowledge-wh (see, for example, Pavese 2021 [2022]).

5. Logical Omniscience

The principal complaint against the approach taken by epistemic logicians is that it is committed to an excessively idealized picture of human reasoning. Critics have worried that the relational semantics of epistemic logic commits one to a closure property for an agent's knowledge that is implausibly strong, given actual human reasoning abilities. The closure properties give rise to what has come to be called the problem of logical omniscience:

Whenever an agent c knows all of the formulas in a set Γ and A follows logically from Γ , then c also knows A .

In particular,

- c knows all theorems of their epistemic logic Λ : if φ is a theorem in Λ (i.e., a logical consequence of $\Gamma = \emptyset$), then so is $K_i\varphi$ by the inference rule *generalization*, and
- c knows all the logical consequences of any formula that the agent knows (letting Γ consist of a single formula): Assume the agent knows φ , i.e. $K_c\varphi$, and that $\varphi \rightarrow \psi$ is a theorem. By the latter and *generalization*, also $K_c(\varphi \rightarrow \psi)$ is a theorem. Hence we have $K_c(\varphi \rightarrow \psi)$ and $K_c\varphi$. By the axiom $K((K_c(\varphi \rightarrow \psi) \rightarrow (K_c\varphi \rightarrow K_c\psi))$, we thus also have $K_c\psi$.

The concern here is that finite agents are constrained by limits on their cognitive capacities and reasoning abilities. The account of knowledge and belief that epistemic logic seems committed to involves superhuman abilities like knowing all the tautologies. Thus, the concern is that epistemic logic is simply unsuited to capturing actual knowledge and belief as these notions figure in ordinary human life.

Hintikka recognized a discrepancy between the rules of epistemic logic and the way the verb “to know” is ordinarily used already in the early pages of *Knowledge and Belief*. He pointed out that

it is clearly inadmissible to infer “he knows that q ” from “he knows that p ” solely on the basis that q follows logically from p , for the person in question may fail to see that p entails q , particularly if p and q are relatively complicated statements. (1962: 30–31)

Hintikka's first reaction to what came to be called the problem of logical omniscience was to see the discrepancy between ordinary usage of terms like “consistency” and formal treatments of knowledge as indicating a problem with our ordinary terminology. If a person knows the axioms of a mathematical theory but is unable to state the distant consequences of the theory, Hintikka denied that it is appropriate to call that person inconsistent. In ordinary human affairs, Hintikka claimed, the charge of inconsistency when directed towards an agent has the connotation of being irrational or dishonest. Thus, from Hintikka's perspective we should choose some other term to capture the situation of someone who is rational and amenable to persuasion or correction but not logically omniscient. Non-omniscient, rational agents can be in a position to say that “I know that p but I don't know whether q ” even in case q can be shown to be entailed by p . He then suggests that q should be regarded as *defensible* given the agent's knowledge and the denial of q should be regarded as

indefensible. This choice of terminology was criticized insofar as it attaches the pejorative *indefensible* to some set of proposition, even though the fault actually lies in the agent's cognitive capacities (Chisholm 1963; Hocutt 1972; Jago 2007).

Hintikka's early epistemic logic can be understood as a way of reasoning about what is implicit in an agent's knowledge even in cases where the agent itself is unable to determine what is implicit. Such an approach risks being excessively idealized and its relevance for understanding human epistemic circumstances can be challenged on these grounds.

Few philosophers were satisfied with Hintikka's attempt to revise our ordinary use of the term "consistent" as he presented it in *Knowledge and Belief*. However, he and others soon provided more popular ways of dealing with logical omniscience. In the 1970s responses to the problem of logical omniscience introduced semantical entities that explain why the agent appears to be, but in fact is not really guilty of logical omniscience. Hintikka (1978) called these entities "impossible possible worlds" (see also the entry on impossible worlds and Jago 2014). The basic idea is that an agent may mistakenly count among the worlds consistent with its knowledge, some worlds containing logical contradictions. The mistake is simply a product of the agent's limited resources; the agent may not be in a position to detect the contradiction and may erroneously count them as genuine possibilities. In some respects, this approach can be understood as an extension of the aforementioned response to logical omniscience that Hintikka had already outlined in *Knowledge and Belief*.

In the same spirit, entities called "seemingly possible" worlds are introduced by Rantala (1975) in his urn-model analysis of logical omniscience. Allowing impossible possible worlds or seemingly possible worlds in which the semantic valuation of the formulas is arbitrary to a certain extent provides a way of making the appearance of logical

omniscience less threatening. After all, on any realistic account of epistemic agency, the agent is likely to consider (albeit inadvertently) worlds in which the laws of logic do not hold. Since no real epistemic principles hold broadly enough to encompass impossible and seemingly possible worlds, some conditions must be applied to epistemic models such that they cohere with epistemic principles (for criticism of this approach see Jago 2007: 336–337).

Alternatively to designing logics in which the knowledge operators do not exhibit logical omniscience, *awareness logic* offers an alternative: Change the interpretation of $K_a\varphi$ from "*a* knows that φ " to "*a* implicitly knows that φ " and take *explicit* knowledge that φ to be implicit knowledge that φ and awareness of φ . With awareness not closed under logical consequence, such a move allows for notion of explicit knowledge not logically omniscient. As agents neither have to compute their implicit knowledge nor can they be held responsible for answering queries based on it, logical omniscience is problematic only for explicit knowledge, the *problem* of logical omniscience is thus averted. Though logical omniscience is an epistemological condition for implicit knowledge, the agent itself may actually fail to realize this condition. For more on awareness logic, see, e.g., the seminal Fagin & Halpern (1987), Halpern & Pucella (2011), or Velazquez-Quesada (2011) and Schipper (2015) for overviews.

Debates about the various kinds of idealization involved in epistemic logic are ongoing in both philosophical and interdisciplinary contexts.

Bibliography

Aloni, Maria, 2001, *Quantification Under Conceptual Covers*, Ph.D. Thesis, University of Amsterdam. [Aloni 2001 available online]

- , 2018, “Knowing-Who in Quantified Epistemic Logic”, in van Ditmarsch and Sandu 2018: 109–129. doi:10.1007/978-3-319-62864-6_4
- Areces, Carlos, Raul Fervari, Andrés R. Saravia, and Fernando R. Velázquez-Quesada, 2023, “First Steps in Updating Knowing How”, in *DaLi 2022: Dynamic Logic. New Trends and Applications*, Carlos Areces and Diana Costa (eds.), (Lecture Notes in Computer Science 13780), Cham: Springer International Publishing, 1–16. doi:10.1007/978-3-031-26622-5_1
- Arló-Costa, Horacio, Vincent F. Hendricks, and Johan van Benthem (eds.), 2016, *Readings in Formal Epistemology*, Cham: Springer International Publishing. doi:10.1007/978-3-319-20451-2
- Artemov, Sergei, 2008, “The Logic of Justification”, *The Review of Symbolic Logic*, 1(4): 477–513. doi:10.1017/S1755020308090060
- Aucher, Guillaume, 2014, “Principles of Knowledge, Belief and Conditional Belief”, in *Interdisciplinary Works in Logic, Epistemology, Psychology and Linguistics: Dialogue, Rationality, and Formalism*, Manuel Rebuschi, Martine Batt, Gerhard Heinzmann, Franck Lihoreau, Michel Musiol, and Alain Trognon (eds.), Cham: Springer International Publishing, 97–134. doi:10.1007/978-3-319-03044-9_5
- Aumann, Robert J., 1976, “Agreeing to Disagree”, *The Annals of Statistics*, 4(6): 1236–1239. Reprinted in Arló-Costa, Hendricks, and van Benthem 2016: 859–862. doi:10.1214/aos/1176343654, doi:10.1007/978-3-319-20451-2_40
- Baltag, Alexandru, 2016, “To Know Is to Know the Value of a Variable”, in *Advances in Modal Logic* (Volume 11), Lev Beklemishev, Stéphane Demri, and András Máté (eds.), Rickmansworth, UK: College Publications, 135–155. [Baltag 2016 available online]
- Baltag, Alexandru, Nick Bezhanishvili, Aybüke Özgün, and Sonja Smets, 2019, “A Topological Approach to Full Belief”, *Journal of Philosophical Logic*, 48(2): 205–244. doi:10.1007/s10992-018-9463-4
- Baltag, Alexandru, Rachel Boddy, and Sonja Smets, 2018, “Group Knowledge in Interrogative Epistemology”, in van Ditmarsch and Sandu 2018: 131–164. doi:10.1007/978-3-319-62864-6_5
- Baltag, Alexandru and Sonja Smets, 2008, “A Qualitative Theory of Dynamic Interactive Belief Revision”, in *Logic and the Foundations of Game and Decision Theory* (LOFT 7), Texts in Logic and Games: Volume 3, G. Bonanno, W. van der Hoek, and M. Wooldridge (eds.), Amsterdam: Amsterdam University Press, 9–58.
- Belardinelli, Francesco and Alessio Lomuscio, 2011, “First-Order Linear-Time Epistemic Logic with Group Knowledge: An Axiomatisation of the Monodic Fragment”, *Fundamenta Informaticae*, 106(2–4): 175–190.
- van Benthem, Johan, 2006, “Epistemic Logic and Epistemology: The State of Their Affairs”, *Philosophical Studies*, 128(1): 49–76. doi:10.1007/s11098-005-4052-0
- , 2011, *Logical Dynamics of Information and Interaction*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511974533
- van Benthem, Johan, David Fernández-Duque, and Eric Pacuit, 2014, “Evidence and Plausibility in Neighborhood Structures”, *Annals of Pure and Applied Logic*, 165(1): 106–133. doi:10.1016/j.apal.2013.07.007
- Blackburn, Patrick, Maarten de Rijke, and Yde Venema, 2001, *Modal Logic*, Cambridge: Cambridge University Press. doi:10.1017/CBO9781107050884
- Boh, Ivan, 1993, *Epistemic Logic in the Later Middle Ages*, (Topics in Medieval Philosophy), London/New York: Routledge.
- Braüner, Torben and Silvio Ghilardi, 2007, “First-Order Modal Logic”, in *Handbook of Modal Logic*, Patrick Blackburn, Johan van Benthem, and Frank Wolter (eds.), (Studies in Logic and Practical Reasoning

- 3), Amsterdam, Boston: Elsevier, 549–620. doi:10.1016/S1570-2464(07)80012-7
- Chellas, Brian F., 1980, *Modal Logic: An Introduction*, Cambridge: Cambridge University Press.
- Chisholm, Roderick M., 1963, “The Logic of Knowing”, *The Journal of Philosophy*, 60(25): 773–795. doi:10.2307/2022834
- Ciardelli, Ivano, 2014, “Modalities in the Realm of Questions: Axiomatizing Inquisitive Epistemic Logic”, in *Advances in Modal Logic* (Volume 10), Rajeev Goré, Barteld Kooi, and Agi Kurucz (eds.), Rickmansworth, UK: College Publications, 94–113. [Ciardelli 2014 available online]
- , 2016, *Questions in Logic*, Ph.D. Thesis, University of Amsterdam.
- , 2023, *Inquisitive Logic: Consequence and Inference in the Realm of Questions*, (Trends in Logic 60), Cham: Springer International Publishing. doi:10.1007/978-3-031-09706-5
- Ciardelli, Ivano and Floris Roelofsen, 2011, “Inquisitive Logic”, *Journal of Philosophical Logic*, 40(1): 55–94. doi:10.1007/s10992-010-9142-6
- , 2015, “Inquisitive Dynamic Epistemic Logic”, *Synthese*, 192(6): 1643–1687. doi:10.1007/s11229-014-0404-7
- Cohen, Michael, Wen Tang, and Yanjing Wang, 2021, “De Re Updates”, in *Proceedings Eighteenth Conference on Theoretical Aspects of Rationality and Knowledge* (TARK 2021), Joseph Halpern and Andrés Perea (eds.), Electronic Proceedings in Theoretical Computer Science 335, 103–117. doi:10.4204/EPTCS.335.9
- Cohen, Mika and Mads Dam, 2007, “A Complete Axiomatization of Knowledge and Cryptography”, in *22nd Annual IEEE Symposium on Logic in Computer Science* (LICS 2007), IEEE, 77–88. doi:10.1109/LICS.2007.4
- Corsi, Giovanna and Eugenio Orlandelli, 2013, “Free Quantified Epistemic Logics”, *Studia Logica*, 101(6): 1159–1183. doi:10.1007/s11225-013-9528-x
- van Ditmarsch, Hans, Joseph Y. Halpern, Wiebe van der Hoek, and Barteld Kooi (eds.), 2015, *Handbook of Epistemic Logic*, London: College Publications.
- van Ditmarsch, Hans, Wiebe van der Hoek, and Barteld Kooi, 2007, *Dynamic Epistemic Logic*, Dordrecht: Springer Netherlands. doi:10.1007/978-1-4020-5839-4
- van Ditmarsch, Hans and Gabriel Sandu (eds.), 2018, *Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, (Outstanding Contributions to Logic, 12), Cham: Springer International Publishing. doi:10.1007/978-3-319-62864-6
- van Eijck, Jan, Malvin Gattinger, and Yanjing Wang, 2017, “Knowing Values and Public Inspection”, in *ICLA 2017: Logic and Its Applications* (Lecture Notes in Computer Science: Volume 10119), Sujata Ghosh and Sanjiva Prasad (eds.), Berlin, Heidelberg: Springer, 77–90. doi:10.1007/978-3-662-54069-5_7
- Epstein, Sophia and Pavel Naumov, 2021, “Epistemic Logic of Know-Who”, *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(13): 11479–11486. doi:10.1609/aaai.v35i13.17367
- Fagin, Ronald and Joseph Y. Halpern, 1987, “Belief, Awareness, and Limited Reasoning”, *Artificial Intelligence*, 34(1): 39–76. doi:10.1016/0004-3702(87)90003-8
- Fagin, Ronald, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi, 1995, *Reasoning About Knowledge*, Cambridge, MA: The MIT Press.
- Fan, Jie, 2019, “Bimodal Logics with Contingency and Accident”, *Journal of Philosophical Logic*, 48(2): 425–445. doi:10.1007/s10992-018-9470-5
- , 2021, “A Logic for Disjunctive Ignorance”, *Journal of Philosophical Logic*, 50(6): 1293–1312. doi:10.1007/s10992-021-09599-4
- Fan, Jie, Yanjing Wang, and Hans van Ditmarsch, 2015, “Contingency and Knowing Whether”, *The Review of Symbolic Logic*, 8(1): 75–107.

- doi:10.1017/S1755020314000343
- Fervari, Raul, Andreas Herzig, Yanjun Li, and Yanjing Wang, 2017, “Strategically Knowing How”, in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence IJCAI 2017*, Carles Sierra (ed.), 1031–1038. doi:10.24963/ijcai.2017/143
- Fine, Kit, 2018, “Ignorance of Ignorance”, *Synthese*, 195(9): 4031–4045. doi:10.1007/s11229-017-1406-z
- Fitting, Melvin, Lars Thalmann, and Andrei Voronkov, 2001, “Term-Modal Logics”, *Studia Logica*, 69(1): 133–169. doi:10.1023/A:1013842612702
- Gochet, Paul and Pascal Gribomont, 2006, “Epistemic Logic”, in Dov M. Gabbay and John Woods (eds.), *Handbook of the History of Logic 7: Logic and the Modalities in the Twentieth Century* (Studies in Logic and Practical Reasoning), Amsterdam: North-Holland, 99–105.
- Groenendijk, Jeroen and Martin Stokhof, 1982, “Semantic Analysis of WH-Complements”, *Linguistics and Philosophy*, 5(2): 175–233. doi:10.1007/BF00351052
- Grove, Adam J., 1995, “Naming and Identity in Epistemic Logic Part II: A First-Order Logic for Naming”, *Artificial Intelligence*, 74(2): 311–350. doi:10.1016/0004-3702(95)98593-D
- Gu, Tao and Yanjing Wang, 2016, “‘Knowing Value’ Logic as a Normal Modal Logic”, in *Advances in Modal Logic*, (Volume 11), Lev Beklemishev, Stéphane Demri, and András Máté (eds.), Rickmansworth, UK: College Publications, 362–381. [Gu and Wang 2016 available online]
- Halpern, Joseph Y., 1996, “Should Knowledge Entail Belief?”, *Journal of Philosophical Logic*, 25(5): 483–494. doi:10.1007/BF00257382
- Halpern, Joseph Y., Dov Samet, and Ella Segev, 2009, “Defining Knowledge in Terms of Belief: The Modal Logic Perspective”, *The Review of Symbolic Logic*, 2(3): 469–487. doi:10.1017/S1755020309990141
- Halpern, Joseph Y. and Riccardo Pucella, 2011, “Dealing with Logical Omniscience: Expressiveness and Pragmatics”, *Artificial Intelligence*, 175(1): 220–235. doi:10.1016/j.artint.2010.04.009
- Halpern, Joseph Y. and Yoram Moses, 1984, “Knowledge and Common Knowledge in a Distributed Environment”, in *Proceedings of the Third Annual ACM Symposium on Principles of Distributed Computing* (PODC ’84), Vancouver: ACM Press, 50–61. doi:10.1145/800222.806735
- Harrah, David, 2002, “The Logic of Questions”, in *Handbook of Philosophical Logic* (Volume 8), Dov M. Gabbay and F. Guenther (eds.), second edition, Dordrecht: Springer Netherlands, 1–60. doi:10.1007/978-94-010-0387-2_1
- Hendricks, Vincent F., 2005, *Mainstream and Formal Epistemology*, Cambridge: Cambridge University Press. doi:10.1017/CBO9780511616150
- Hendricks, Vincent F. and Rasmus K. Rendsvig, 2018, “Hintikka’s Knowledge and Belief in Flux”, in van Ditmarsch and Sandu 2018: 317–337. doi:10.1007/978-3-319-62864-6_13
- Hendricks, Vincent F. and John Symons, 2006, “Where’s the Bridge? Epistemology and Epistemic Logic”, *Philosophical Studies*, 128(1): 137–167. doi:10.1007/s11098-005-4060-0
- Herzig, Andreas, 2015, “Logics of Knowledge and Action: Critical Analysis and Challenges”, *Autonomous Agents and Multi-Agent Systems*, 29(5): 719–753. doi:10.1007/s10458-014-9267-z
- Herzig, Andreas and Elise Perrotin, 2020, “On the Axiomatisation of Common Knowledge”, in *Advances in Modal Logic* (Volume 13), Nicola Olivetti, Rineke Verbrugge, Sara Negri, and Gabriel Sandu (eds.), Rickmansworth, UK: College Publications, 309–328. [Herzig and Perrotin 2020 available online]
- Hintikka, Jaakko, 1962 [2005], *Knowledge and Belief: An Introduction to the Logic of the Two Notions* (Texts in Philosophy: 1), second edition,

- Vincent F. Hendriks and John Symons (eds.), London: College Publications.
- , 1969, “Semantics for Propositional Attitudes”, in *Philosophical Logic*, J. W. Davis, D. J. Hockney, and W. K. Wilson (eds.), Dordrecht: Springer Netherlands, 21–45. doi:10.1007/978-94-010-9614-0_2
- , 1978, “Impossible Possible Worlds Vindicated”, in *Game-Theoretical Semantics*, Esa Saarinen (ed.) (SLAP 5), Dordrecht: Springer Netherlands, 367–379. doi:10.1007/978-1-4020-4108-2_13
- , 2003, “A Second Generation Epistemic Logic and Its General Significance”, in *Knowledge Contributors*, Vincent F. Hendricks, Klaus Frovin Jørgensen, and Stig Andur Pedersen (eds.), Dordrecht: Springer Netherlands, 33–55. doi:10.1007/978-94-007-1001-6_3
- , 2007, “Epistemology without Knowledge and without Belief”, in *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning*, Cambridge: Cambridge University Press, 11–37. doi:10.1017/CBO9780511619298.002
- Hintikka, Jaakko and John Symons, 2003, “Systems of Visual Identification in Neuroscience: Lessons from Epistemic Logic”, *Philosophy of Science*, 70(1): 89–104. doi:10.1086/367871
- Hocutt, Max O., 1972, “Is Epistemic Logic Possible?”, *Notre Dame Journal of Formal Logic*, 13(4): 433–453. doi:10.1305/ndjfl/1093890705
- Hodkinson, Ian, Frank Wolter, and Michael Zakharyashev, 2000, “Decidable Fragments of First-Order Temporal Logics”, *Annals of Pure and Applied Logic*, 106(1–3): 85–134. doi:10.1016/S0168-0072(00)00018-X
- van der Hoek, Wiebe, 1993, “Systems for Knowledge and Belief”, *Journal of Logic and Computation*, 3(2): 173–195. doi:10.1093/logcom/3.2.173
- van der Hoek, Wiebe and Alessio Lomuscio, 2004, “A Logic for Ignorance”, in *DALT 2003: Declarative Agent Languages and Technologies* (Lecture Notes in Computer Science 2990), João Leite, Andrea Omicini, Leon Sterling, and Paolo Torroni (eds.), Berlin/Heidelberg: Springer, 97–108. doi:10.1007/978-3-540-25932-9_6
- Holliday, Wesley H., 2018, “Epistemic Logic and Epistemology”, in *Introduction to Formal Philosophy*, Sven Ove Hansson and Vincent F. Hendricks (eds.), Cham: Springer International Publishing, 351–369. doi:10.1007/978-3-319-77434-3_17
- Holliday, Wesley H. and John Perry, 2014, “Roles, Rigidity, and Quantification in Epistemic Logic”, in *Johan van Benthem on Logic and Information Dynamics* (Outstanding Contributions to Logic 5), Alexandru Baltag and Sonja Smets (eds.), Cham: Springer International Publishing, 591–629. doi:10.1007/978-3-319-06025-5_22
- Jago, Mark, 2007, “Hintikka and Cresswell on Logical Omniscience”, *Logic and Logical Philosophy*, 15(4): 325–354. doi:10.12775/LLP.2006.019
- , 2014, *The Impossible: An Essay on Hyperintensionality*, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198709008.001.0001
- Jamroga, Wojciech and Thomas Ågotnes, 2007, “Constructive Knowledge: What Agents Can Achieve under Imperfect Information”, *Journal of Applied Non-Classical Logics*, 17(4): 423–475. doi:10.3166/jancl.17.423-475
- Jamroga, Wojciech and Wiebe van der Hoek, 2004, “Agents That Know How to Play”, *Fundamenta Informaticae*, 63(2–3): 185–219.
- Jiang, Junli and Pavel Naumov, 2022, “Data-Informed Knowledge and Strategies”, *Artificial Intelligence*, 309: article 103727. doi:10.1016/j.artint.2022.103727

- Kaneko, Mamoru and Takashi Nagashima, 1996, “Game Logic and Its Applications I”, *Studia Logica*, 57(2–3): 325–354. doi:10.1007/BF00370838
- Knuuttila, Simo, 1993, *Modalities in Medieval Philosophy*, (Topics in Medieval Philosophy), New York: Routledge.
- Kooi, Barteld P., 2007, “Dynamic Term-Modal Logic”, in *A Meeting of the Minds, Proceedings of the Workshop on Logic, Rationality and Interaction (LORI-I)*, Johan van Benthem, Shier Ju, and Frank Veltman (eds.), London: College Publications, 173–185.
- Kraus, Sarit and Daniel Lehmann, 1986, “Knowledge, Belief and Time”, in *Automata, Languages and Programming*, Laurent Kott (ed.), Berlin, Heidelberg: Springer Berlin Heidelberg, 186–195.
- Kutschera, Franz von, 1976, *Einführung in Die Intensionale Semantik*, (De Gruyter Studienbuch: Grundlagen Der Kommunikation), Berlin/New York: De Gruyter.
- Lehmann, Daniel, 1984, “Knowledge, Common Knowledge and Related Puzzles (Extended Summary)”, *Proceedings of the Third Annual ACM Symposium on Principles of Distributed Computing (PODC '84)*, 62–67. doi:10.1145/800222.806736
- Lenzen, Wolfgang, 1978, *Recent Work in Epistemic Logic*, (Acta Philosophica Fennica, 30), Amsterdam: North Holland Publishing Company.
- , 1980, *Glauben, Wissen Und Wahrscheinlichkeit: Systeme Der Epistemischen Logik*, (Library of Exact Philosophy, 12), Wien: Springer.
- Lewis, David K., 1969, *Convention: A Philosophical Study*, Cambridge, MA: Harvard University Press.
- Li, Yanjun and Yanjing Wang, 2021a, “Knowing How to Plan”, in *Proceedings Eighteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2021)*, Joseph Halpern and Andrés Perea (eds.), Electronic Proceedings in Theoretical Computer Science 335, 233–247. doi:10.4204/EPTCS.335.22
- , 2021b, “Planning-Based Knowing How: A Unified Approach”, *Artificial Intelligence*, 296: article 103487. doi:10.1016/j.artint.2021.103487
- Liu, Mo, Anantha Padmanabha, R. Ramanujam, and Yanjing Wang, 2023, “Are Bundles Good Deals for First-Order Modal Logic?”, *Information and Computation*, 293: article 105062. doi:10.1016/j.ic.2023.105062
- Meyer, John-Jules Ch, 2001, “Epistemic Logic”, in *The Blackwell Guide to Philosophical Logic*, Lou Goble (ed.), Oxford: John Wiley & Sons, 183–202.
- Meyer, John-Jules Ch. and Wiebe van der Hoek, 1995, *Epistemic Logic for AI and Computer Science*, (Cambridge Tracts in Theoretical Computer Science, 41), Cambridge: Cambridge University Press.
- Ma Xiwen and Guo Weide, 1983, “W-JS: A Modal Logic of Knowledge”, in *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-83)*, Volume 1: 398–401. [Ma and Guo 1983 available online]
- Maubert, Bastien, Aniello Murano, Sophie Pinchinat, Francois Schwarzenruber, and Silvia Stranieri, 2020, “Dynamic Epistemic Logic Games with Epistemic Temporal Goals”, in *ECAI 2020* (Frontiers in Artificial Intelligence and Applications 325), Amsterdam: IOS Press, 155–162. [Maubert, Murano et al. 2020 available online]
- Maubert, Bastien, Sophie Pinchinat, Francois Schwarzenruber, and Silvia Stranieri, 2020, “Concurrent Games in Dynamic Epistemic Logic”, in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI 2020)*, Christian Bessiere (ed.), 1877–1883. doi:10.24963/ijcai.2020/260

- Naumov, Pavel and Jia Tao, 2017, “Together We Know How to Achieve: An Epistemic Logic of Know-How (Extended Abstract)”, in *Proceedings Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge* (TARK 2017), Jérôme Lang (ed.), Electronic Proceedings in Theoretical Computer Science 251, 441–453. doi:10.4204/EPTCS.251.32
- , 2018, “Together We Know How to Achieve: An Epistemic Logic of Know-How”, *Artificial Intelligence*, 262: 279–300. doi:10.1016/j.artint.2018.06.007
- , 2019, “Knowing-How under Uncertainty”, *Artificial Intelligence*, 276: 41–56. doi:10.1016/j.artint.2019.06.007
- , 2020, “An Epistemic Logic of Blameworthiness”, *Artificial Intelligence*, 283: article 103269. doi:10.1016/j.artint.2020.103269
- Occhipinti Liberman, Andrés, Andreas Achen, and Rasmus Kræmmer Rendsvig, 2020, “Dynamic Term-Modal Logics for First-Order Epistemic Planning”, *Artificial Intelligence*, 286: article 103305. doi:10.1016/j.artint.2020.103305
- Occhipinti Liberman, Andrés and Rasmus K. Rendsvig, 2022, “Reasoning about Epistemic Social Network Dynamics Using Dynamic Term-Modal Logic”, *Journal of Logic and Computation*, 32(6): 1067–1087. doi:10.1093/logcom/exac019
- Orlandelli, Eugenio and Giovanna Corsi, 2018, “Decidable Term-Modal Logics”, in *Multi-Agent Systems and Agreement Technologies: 15th European Conference, EUMAS 2017, and 5th International Conference, AT 2017* (Lecture Notes in Computer Science 10767), Francesco Belardinelli and Estefanía Argente (eds.), Cham: Springer International Publishing, 147–162. doi:10.1007/978-3-030-01713-2_11
- Padmanabha, Anantha and R. Ramanujam, 2019a, “The Monodic Fragment of Propositional Term Modal Logic”, *Studia Logica*, 107(3): 533–557. doi:10.1007/s11225-018-9784-x

- , 2019b, “Two Variable Fragment of Term Modal Logic”. in *44th International Symposium on Mathematical Foundations of Computer Science* (MFCS 2019), Peter Rossmanith, Pinar Heggernes, and Joost-Pieter Katoen (eds.), Leibniz International Proceedings in Informatics 138, article 30. doi:10.4230/LIPIcs.MFCS.2019.30
- Pavese, Carlotta, 2021 [2022], “Knowledge How”, *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), Edward N. Zalta and Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2022/entries/knowledge-how/>.
- Plaza, Jan A., 1989, “Logics of Public Communications”, in *Proceedings of the Fourth International Symposium on Methodologies for Intelligent Systems: Poster Session Program*, M. L. Emrich, S. Pfeifer, M. Hadzikadic, and Z. W. Ras (eds.), Oak Ridge National Laboratory, 201–216; reprinted, 2007, “Logics of Public Communications”, *Synthese*, 158(2): 165–179. doi:10.1007/s11229-007-9168-7
- Rantala, Veikko, 1975, “Urn Models: A New Kind of Non-Standard Model for First-Order Logic”, *Journal of Philosophical Logic*, 4(4): 455–474. doi:10.1007/BF00558760
- Rendsvig, Rasmus K., 2012, “Modeling Semantic Competence: A Critical Review of Frege’s Puzzle about Identity”, in *New Directions in Logic, Language and Computation: ESSLLI 2010 and ESSLLI 2011 Student Sessions, Selected Papers* (Lecture Notes in Computer Science 7415), Daniel Lassiter and Marija Slavkovik (eds.), Berlin/Heidelberg: Springer, 140–157. doi:10.1007/978-3-642-31467-4_10
- Schipper, Burkhard C., 2015, “Awareness”, in van Ditmarsch et al. 2015: 77–146.
- Stalnaker, Robert, 2006, “On Logics of Knowledge and Belief”, *Philosophical Studies*, 128(1): 169–199. doi:10.1007/s11098-005-

- 4062-y
- Stanley, Jason and Timothy Williamson, 2001, “Knowing How”, *Journal of Philosophy*, 98(8): 411–444. doi:10.2307/2678403
- Sturm, Holger, Frank Wolter, and Michael Zakharyashev, 2000, “Monodic Epistemic Predicate Logic”, in *Logics in Artificial Intelligence: European Workshop JELIA 2000* (Lecture Notes in Computer Science 1919), Manuel Ojeda-Aciego, Inma P. De Guzmán, Gerhard Brewka, and Luís Moniz Pereira (eds.), Berlin/Heidelberg: Springer, 329–344. doi:10.1007/3-540-40006-0_23
- Su, Xingchi, 2017, “Distributed Knowing Whether (Extended Abstract)”, in *LORI 2017: Logic, Rationality, and Interaction* (Lecture Notes in Computer Science 10455), Alexandru Baltag, Jeremy Seligman, and Tomoyuki Yamada (eds.), Berlin, Heidelberg: Springer, 653–657. doi:10.1007/978-3-662-55665-8_47
- Velazquez-Quesada Fernando Raymundo, 2011, “Small Steps in Dynamics of Information”, Ph.D. Thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Voorbraak, Franciscus Petrus Johannes Maria, 1993, “As Far as I Know: Epistemic Logic and Uncertainty”, Ph.D. Thesis, Department of Philosophy, Utrecht University.
- Wang, Haoyu, Yanjing Wang, and Yunsong Wang, 2022, “Inquisitive Logic as an Epistemic Logic of Knowing How”, *Annals of Pure and Applied Logic*, 173(10): article 103145. doi:10.1016/j.apal.2022.103145
- Wang, Yanjing, 2017, “A New Modal Framework for Epistemic Logic”, in *Proceedings Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge* (TARK 2017), Jérôme Lang (ed.), Electronic Proceedings in Theoretical Computer Science 251, 515–534. doi:10.4204/EPTCS.251.38
- , 2018a, “A Logic of Goal-Directed Knowing How”, *Synthese*, 195(10): 4419–4439. doi:10.1007/s11229-016-1272-0
- , 2018b, “Beyond Knowing That: A New Generation of Epistemic Logics”, in van Ditmarsch and Sandu 2018: 499–533. doi:10.1007/978-3-319-62864-6_21
- Wang, Yanjing and Jie Fan, 2013, “Knowing That, Knowing What, and Public Communication: Public Announcement Logic with Kv Operators”, in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI 2013)*, Francesca Rossi (ed.), 1147–1154. [Wang and Fan 2013 available online]
- Wang, Yanjing, Yu Wei, and Jeremy Seligman, 2022, “Quantifier-Free Epistemic Term-Modal Logic with Assignment Operator”, *Annals of Pure and Applied Logic*, 173(3): article 103071. doi:10.1016/j.apal.2021.103071
- Williamson, Timothy, 2000, *Knowledge and Its Limits*, Oxford: Oxford University Press. doi:10.1093/019925656X.001.0001
- Wright, Georg Henrik von, 1951, *An Essay in Modal Logic*, (Studies in Logic and the Foundations of Mathematics), Amsterdam: North-Holland Publishing Company.
- Xiang, Yimei, 2016, *Interpreting Questions with Non-Exhaustive Answers*, Ph.D. Thesis, Harvard University.
- Xu, Chao, Yanjing Wang, and Thomas Studer, 2021, “A Logic of Knowing Why”, *Synthese*, 198(2): 1259–1285. doi:10.1007/s11229-019-02104-0
- Yang, Yuanzhe, 2023, “Knowledge-wh and False Belief Sensitivity: A Logical Study (An Extended Abstract)”, in *Proceedings Nineteenth Conference on Theoretical Aspects of Rationality and Knowledge* (TARK 2023), Rineke Verbrugge (ed.), Electronic Proceedings in Theoretical Computer Science 379, 527–544. doi:10.4204/EPTCS.379.40

Academic Tools

- 🔗 How to cite this entry.
- 🔗 Preview the PDF version of this entry at the Friends of the SEP Society.
- 📖 Look up topics and thinkers related to this entry at the Internet Philosophy Ontology Project (InPhO).
- 📄 Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

- Fan, Jie, Davide Grossi, Barteld Kooi, Xingchi Su, and Rineke Verbrugge, 2020, “Commonly Knowingly Whether”, manuscript.
- Hintikka’s World, a graphical, pedagogical tool for learning about epistemic logic, higher-order reasoning and knowledge dynamics.
- Modal Logic Playground, a graphical interface for drawing and evaluating formulas of modal propositional logic.

Related Entries

common knowledge | Heytesbury, William | impossible worlds | logic: and information | logic: dynamic epistemic | logic: justification | logic: modal | logic: of belief revision | multi-modal logic, philosophical aspects of

Notes to Epistemic Logic

1. In his 1969, Hintikka writes

My basic assumption (slightly oversimplified) is that an attribution of any propositional attitude to the person in question involves a division of all possible worlds (more precisely, all the worlds which we can distinguish in the part of language we use in making

the attribution) into two classes: into those possible worlds which are in accordance with the attitude in question and into those worlds which are incompatible with it. (p. 25)

A similar view was held throughout his authorship, echoed in 2007 by

What the concept of knowledge involves in a purely logical perspective is thus a dichotomy of the space of all possible scenarios into those that are compatible with what I know and those that are incompatible with my knowledge. This observation is all we need for most of epistemic logic. (p. 15)

2. In the literature, the valuation is sometimes defined the other way around, i.e., defined not by it assigning to each atom a set of worlds, but by it assigning to each world a set of atoms. In the example, we would thus look at $V(w_1)$ the valuation of the world rather than $V(p)$ the valuation of the atom. $V(w_1)$ would be $\{p, q\}$. The two approaches are equivalent.

A second alternative used, e.g., in the entry on modal logic is to let the valuation be a function from the set of pairs of worlds and atoms $W \times Atom$ to the set truth values for *true* and *false*, $\{T, F\}$. Then $V(w_1, p) = T$ while $V(w_2, q) = F$. Again, this approach is equivalent to the one used here.

3. This paragraph refers to *weak* completeness. For the difference between weak and *strong completeness*, and for general meta-theoretical results for modal logic, see, e.g., Blackburn, de Rijke, and Venema 2001.

Copyright © 2024 by the authors
Rasmus Rendsvig, John Symons, and Yanjing Wang