

# Hintikka on Epistemological Axiomatizations

Vincent F. Hendricks

Department of Philosophy and Science Studies

Roskilde University, Denmark

vincent@ruc.dk    <http://www.ruc.dk/~vincent>

## Abstract

Jaakko Hintikka's development of, and subsequent contributions to, epistemic logic is one of the great accomplishments in formal philosophy. The epistemological significance of Hintikka's 'logical epistemology' has often been largely neglected by philosophy and its broader intellectual, often technical, environment. The purpose of this paper celebrating the 75th Birthday of Jaakko Hintikka is to discuss his ideas for 'the logic of epistemology or the epistemology of logic' from a *forcing* perspective.

Forthcoming in *Hintikka on Everything — Festschrift in Honor of Jaakko Hintikka's 75th Birthday*, Symons, J. (ed.), Kluwer Academic Publishers, 2004.

## 1 Introduction

Among the many intellectual accomplishments which Jaakko Hintikka is recognized for is his pioneering work in epistemic logic. Although epistemic logic was studied somewhat in the Middle Ages the real break-throughs are to be found in the work of von Wright [59] and most notably Hintikka's seminal book *Knowledge and Belief: An Introduction to the Logic of the Two Notions* from 1962 [24]. There has hardly been an article or book published on the logic of knowledge and belief since not making reference to this exquisite treatise.

For the past 40 years epistemic and doxastic logics have developed into fields of research with wide ranges of application. They are of immanent importance to theoretical computer science, artificial intelligence, linguistics, game theory, economics and social software. Be that as it may, epistemic and

doxastic logics are still in an awkward philosophical position today. Computer scientists, linguistics and other formally minded researchers utilizing the means and methods do not necessarily have an epistemological ambition with their use of epistemic logic. At the same time it is a discipline devoted to the logic of knowledge and belief but alien to epistemologists and philosophers interested in the theory of knowledge.

Hintikka from the very beginning had a strong epistemological ambition with his development of epistemic logic however. It was not to be another technical spin-off of advances in modal and other intensional logics. Its purpose was, and still remains, to elucidate various epistemic notions and reason about knowledge and belief. Epistemic logic is to serve as a *logical epistemology* for mainstream and formal epistemological approaches alike.

Despite Hintikka's original intentions, ambitions and own work the epistemological significance of epistemic logic has in general been neglected and perhaps even sometimes intentionally ignored by both formal and mainstream epistemologists. Epistemology is in the business of dealing with skepticism and the possibility of error—logical epistemology may actually be viewed as being much in the same business. Modal concepts of knowledge quantify over other possible worlds to secure the robustness and steadfastness of knowledge. But the classical conception of infallibilism is taken to require, that for an agent to have knowledge of some hypothesis or proposition,<sup>1</sup> he must be able to eliminate *all* the possibilities of error associated with the hypothesis in question. The set of *all* worlds is considered. This set of possible worlds is too big for knowledge to have scope over. The set includes some rather bizarre worlds inhabited by odd beasts from demons to mad and malicious scientists who have decided to stick your brain in a tank of nutritious fluids to systematically fool you. Or worlds in which contradictions are true. If these worlds were to be considered relevant all the time skepticism would have the upper hand all the time. There may not be a way for an agent to determine that he is not in the world of the beast or the brain. If infallibilism is to be a viable reply to the skeptic, then infallibilism cannot be defined with respect to all possible worlds. Hintikka may be read as saying something similar when it comes to epistemic logic:

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. [31], p. 2.

---

<sup>1</sup>'Hypothesis' and 'proposition' will be used interchangeably.

This way of battling the skeptic by limiting the set of citable possible worlds carrying potential error has been referred to as ‘forcing’ in Hendricks [17], [18] and in particular [19]:

Whenever knowledge claims are challenged by alleged possibilities of error, the strategy is to show that the possibilities of error fail to be genuine in the relevant sense

Logical epistemology or epistemic logic pays homage to the forcing strategy as the partitioning of the space of possible worlds compatible with knowledge attitude determines a certain set over which the epistemic operator is to have scope. Contemporary mainstram epistemologists choose to speak of the *relevant* possible worlds as a subset of the set of all possible worlds.<sup>2</sup> The epistemic logician considers an *accessibility* relation between worlds in a designated class out of the entire universe of possible worlds. There is no principled difference between relevance and accessibility. Informal epistemologies differ by the way in which relevance is forced given, say, perceptual equivalence conditions, counterfactual proximities or conversational contexts circumscribing the possible worlds. Formal epistemologies differ by the way in which the accessibility relation is defined over possible worlds.

Epistemic logicians obtain different epistemic modal systems valid for a knowledge operator by varying (adding, dropping or relativizing) the properties of the accessibility relation from, say, reflexive and transitive to a reflexive, symmetric and transitive relation. Algebraic constraints on the accessibility relation are the forcing foundation for a formal approach to the theory of knowledge like logical epistemology. Constraints on accessibility relations between possible worlds is a way of demonstrating some of the epistemological significance of Hintikka’s philosophical program in epistemic logic already present in *Knowledge and Belief* and of course beyond.

## 2 Epistemic Logic and Skepticism

For a proper syntactic augmentation of the language of the propositional logic with two unary operators  $K_{\Xi}$  and  $B_{\Xi}$  such that

$K_{\Xi}A$  reads ‘Agent  $\Xi$  knows  $A$ ’ and  $B_{\Xi}A$  reads ‘Agent  $\Xi$  believes  $A$ ’

---

<sup>2</sup>Explicit forcing proposals in the epistemological literature are sometimes referred to as ‘*relevant alternatives proposals*’. Cf. Bernecker and Dretske [1].

for some arbitrary proposition  $A$ , Hintikka came up with the following semantic interpretations of the epistemic and doxastic operators [24], [25]:

$K_{\Xi}A \approx$  in all possible worlds compatible with what  $\Xi$  knows it is the case that  $A$

$B_{\Xi}A \approx$  in all possible worlds compatible with what  $\Xi$  knows it is the case that  $A$

The basic assumption is that any ascription of propositional attitudes like knowledge and belief, requires *partitioning* of the set of possible worlds into two compartments: The compartment consisting of possible worlds compatible with the attitude in question and the compartment of worlds incompatible with it. Based on the partition the agent is capable of constructing different ‘world-models’ using the epistemic modal language. He is not necessarily required to know which one of the world-models constructed is the real world-model. All the same, the agent does not consider all these world-models equally possible or accessible from his current point of view. Some world-models may be incommensurable with his current information state or other background assumptions. These incompatible world-models are excluded from the compatibility partition. This is a variation of the forcing strategy. In logical epistemology, as in many mainstream epistemologies, it is typically stipulated that the smaller the set of worlds an agent considers possible, the smaller his uncertainty, at the cost of stronger forcing assumptions.

The set of worlds considered accessible by an agent depends on the actual world, or the agent’s actual state of information. It is possible to capture the forcing dependency by introducing a relation of accessibility,  $R$ , on the set of compatible possible worlds. To express the idea that for agent  $\Xi$ , the world  $w'$  is compatible with his information state, or accessible from the possible world  $w$  which  $\Xi$  is currently in, it is required that  $R$  holds between  $w$  and  $w'$ . This relation is written  $Rww'$  and read ‘world  $w'$  is accessible from  $w$ ’. The world  $w'$  is said to be an *epistemic alternative* to world  $w$  for agent  $\Xi$ . Given the above semantical interpretation, if a proposition  $A$  is true in all worlds which agent  $\Xi$  considers possible then  $\Xi$  knows  $A$ .

Formally, a *frame*  $\mathcal{F}$  for an epistemic system is a pair  $(W, R)$  for which  $W$  is a non-empty set of possible worlds and  $R$  is a binary accessibility relation over  $W$ . A *model*  $\mathbb{M}$  for an epistemic system consists of a frame and a denotation function  $\varphi$  assigning sets of worlds to atomic propositional formulae. Propositions are taken to be sets of possible worlds; namely the set of possible worlds in which they are true. Let *atom* be the set of atomic

propositional formulae, then  $\varphi : atom \longrightarrow P(W)$  where  $P$  denotes the powerset operation. The model  $\mathbb{M} = \langle W, R, \varphi \rangle$  is called a Kripke-model and the resulting semantics Kripke-semantics [34]: An atomic propositional formulae,  $\mathbf{a}$ , is said to be true in a world  $w$  (in  $\mathbb{M}$ ), written  $\mathbb{M}, w \models \mathbf{a}$ , iff  $w$  is in the set of possible worlds assigned to  $\mathbf{a}$ , i. e.  $\mathbb{M}, w \models \mathbf{a}$  iff  $w \in \varphi(\mathbf{a})$  for all  $\mathbf{a} \in atom$ . The formula  $K_{\Xi}A$  is true in a world  $w$ , i.e.  $\mathbb{M}, w \models K_{\Xi}A$ , iff  $\forall w' \in W : \text{if } Rww', \text{ then } \mathbb{M}, w' \models A$ . The semantics for the Boolean connectives are given in the usual recursive way. A modal formula is said to be *valid* in a frame iff the formula is true for all possible assignments in all worlds admitted by the frame.

A nice feature of possible world semantics is that many common epistemic axioms correspond to certain algebraic properties of the frame in the following sense: A modal axiom is valid in a frame if and only if the accessibility relation satisfies some algebraic condition. For an example, the axiom

$$K_{\Xi}A \rightarrow A \tag{1}$$

is valid in all frames in which the accessibility relation is *reflexive* in the sense that every possible world is accessible from itself. (1) is called axiom  $\top$ ,<sup>3</sup> or the *axiom of truth* or *axiom of veridicality*, and says that if  $A$  is known by  $\Xi$ , then  $A$  is true in accordance with the standard tripartite definition of knowledge as true justified belief.

Similarly if the accessibility relation satisfies the condition that

$$\forall w, w', w'' \in W : Rww' \wedge Rw'w'' \rightarrow Rww''$$

then the axiom

$$K_{\Xi}A \rightarrow K_{\Xi}K_{\Xi}A \tag{2}$$

is valid in all *transitive* frames. (2) is called axiom 4 and is also known as the *axiom of self-awareness*, *positive introspection* or *KK-thesis*. The labels all refer to the idea that an agent has knowledge of his knowledge of  $A$  if he has knowledge of  $A$ . Other axioms require yet other relational properties to be met in order to be valid in all frames: If the accessibility relation is reflexive, symmetric and transitive, then

$$\neg K_{\Xi}A \rightarrow K_{\Xi}\neg K_{\Xi}A \tag{3}$$

---

<sup>3</sup>This nomenclature due to Lemmon [36] and later refined by Bull and Segerberg [4] is helpful while cataloguing the axioms typically considered interesting for epistemic logic.

is valid. (3) is called axiom 5 also better known as the *axiom of wisdom*. It is the much stronger thesis that an agent has knowledge of his own ignorance: If  $\Xi$  does not know  $A$ , he knows that he doesn't know  $A$ . The axiom is sometimes referred to as the *axiom of negative introspection*.

As opposed to (1)–(3) there is a formula or axiom which is valid in all possible frames

$$K_{\Xi}(A \rightarrow A') \rightarrow (K_{\Xi}A \rightarrow K_{\Xi}A'). \quad (4)$$

The axiom amounts to the contentious closure condition for knowledge and is also referred to as axiom **K**, or the *axiom of deductive cogency*: If the agent  $\Xi$  knows  $A \rightarrow A'$ , then if  $\Xi$  knows  $A$ ,  $\Xi$  also knows  $A'$ . One rule of inference which is valid in all possible frames is the rule of *necessitation or epistemization* (**N**)

$$\frac{A}{K_{\Xi}A} \quad (5)$$

which says that if  $A$  is true in all worlds of the frame, then so is  $K_{\Xi}A$ .

Logical epistemology unproblematically accepts (4)–(5) but for formal reasons. Neither (4) nor (5) require any assumptions to be made pertaining to the accessibility relation between the possible worlds considered compatible with the knowledge attitude. It actually turns out that (4) together with (5) comprise the characterizing axiom and rule for possible world semantics with binary accessibility relations. All modal logics in which (4) and (5) are valid are called *normal* modal logics.

These axioms in proper combinations make up epistemic modal systems of varying strength depending on the modal formulae valid in the respective systems given the algebraic properties assumed for the accessibility relation. The weakest system of epistemic interest is usually considered to being system **T**. The system includes **T** and **K** as valid axioms. Additional modal strength may be obtained by extending **T** with other axioms drawn from the above pool altering the frame semantics to validate the additional axioms. Reflexivity is the characteristic frame property of system **T**, transitivity is the characteristic frame property of system **S4**, equivalence the characteristic frame property of **S5**, etc. From an epistemological point of view, the algebraic properties of the accessibility relation are really forcing conditions.

The cognitive rationale of logical epistemology must be something like this: The more properties the accessibility relation is endowed with, the more access the agent has to his epistemic universe, and in consequence the more epistemic strength he will obtain. The stronger knowledge, the stronger forcing clauses.

Modal epistemic axioms and systems may be viewed as measures of infallibility and replies to skepticism. For instance, knowing your own knowledge is a way of blocking the skeptic, but knowledge of your own ignorance in terms of axiom 5 is better still. One motivation for the plausibility of axiom 5 is in data-base applications: An agent examining his own knowledge base will be let to conclude that whatever is not in the knowledge base he does not know and hence he will know that he does not.

The axiom of wisdom or negative introspection is a sort of closed world assumption. A closed world assumption is a forcing assumption if anything is, ‘shutting the world down’ with the agent, leaving the skeptic nowhere to go. To know the truth, to know of your knowledge, and to know of your own ignorance as in **S5** requires ‘full’ epistemic access which is exactly why the accessibility relation must be an equivalence relation. A theorem of **S5** is the following

$$\neg A \rightarrow K_{\Xi} \neg K_{\Xi} A \quad (6)$$

which states that if  $A$  is not the case, then  $\Xi$  knows that he does not know  $A$ —the ‘truly Socratic person’ as Girle explains ([13], p. 157) knowing exactly how ignorant he is.

A bit more ignorance, a bit more skepticism and accordingly a bit more fallibility is allowed in **S4**. Since axiom 5 is dropped and (6) is no longer a theorem,  $\{\neg A, \neg K_{\Xi} \neg K_{\Xi} A\}$  and  $\{\neg K_{\Xi} A, \neg K_{\Xi} \neg K_{\Xi} A\}$  are not inconsistent in **S4**. It is possible for an agent to be ignorant of the fact that he does not know when actually he does know. Put differently, the agent is allowed false beliefs about what is known. Yet more ignorance and skepticism are allowed in system **T** because while  $\{K_{\Xi} A, \neg K_{\Xi} K_{\Xi} A\}$  is inconsistent in **S4**, this set of epistemic statements is not inconsistent in **T**. The agent may thus know something without knowing that he does.<sup>4</sup>

What Hintikka recently dubbed ‘first generation epistemic logic’ in [30] is characterized by the ambition that cataloguing the possible complete systems of such logics would allow for choosing the most ‘appropriate’ or ‘intuitive’ ones(s).<sup>5</sup> Hintikka himself settled for **S4** in *Knowledge and Belief*, but he had very strong epistemological arguments for doing so.

---

<sup>4</sup>All the same, a restricted kind of positive introspection is still prevalent in system **T**. Given the rule of necessitation (5),  $\Xi$  knows all the theorems of the epistemic logic. By iteration,  $K_{\Xi} K_{\Xi} A$  is also known. Thus if  $A$  is a theorem,  $\Xi$  knows that he knows  $A$ .

<sup>5</sup>Hintikka’s ‘second generation epistemic logic’ is discussed under the rubric ‘active agenthood’ in Hendricks [18], [19], and [23]. For excellent surveys of epistemic logic and its contemporary themes see also van Benthem [2] and Gochet and Gribomont [14].

## 2.1 The Logic of Autoepistemology

Hintikka stipulated that the axioms or principles of epistemic logic are conditions descriptive of a special kind of general (strong) *rationality* from a first person perspective.<sup>6</sup> The statements which may be proved false by application of the epistemic axioms are not inconsistent meaning that their truth is logically impossible. They are rather rationally ‘indefensible’. Indefensibility is fleshed out as the agent’s epistemic laziness, sloppiness or perhaps cognitive incapacity whenever to realize the implications of what he in fact knows. Defensibility then means not falling victim of ‘epistemic negligence’ as Chisholm calls it [5], [6]. The notion of indefensibility gives away the status of the epistemic axioms and logics. Some epistemic statement for which its negation is indefensible is called ‘self-sustaining’. The notion of self-sustenance actually corresponds to the concept of validity. Corresponding to a self-sustaining statement is a logically valid statement. But this will again be a statement which is rationally indefensible to deny. So in conclusion, epistemic axioms are descriptions of rationality.

There is an argument to the effect that Hintikka early on was influenced by the autoepistemology of G.E. Moore [47] and especially Malcolm [46] and took, at least in part, their autoepistemology to provide a philosophical motivation for epistemic logic. Moore’s common-sense considerations on which autoepistemology is founded deflates the skeptical possibilities of error from various dialectic angles of which one is particularly pertinent to the current discussion. It is called the argument from *incoherence*. The idea is to demonstrate that skepticism has severe difficulties in formulating its own position coherently. As with any argument, a skeptical conclusion presupposes knowledge of a set of premisses. Moore then points to the fact that merely *asserting* these premisses imply at least a doxastic commitment, but most likely an epistemic commitment. The skeptics cannot be retreating to a statement like

*‘There are 9 planets in our solar system  
but it is not the case that I believe it.’* (7)

The statement in (7) is an instance of what later has become known as the *Moore-paradox*. Let it be granted that (7) only involves an error of omission. All the same it still sounds self-contradictory simply given mere assertion. No formulation of skepticism without incoherence, or in Hintikkian terms, skepticism is an irrational or indefensible epistemological position.

---

<sup>6</sup>For a systematic discussion of logical epistemology from first and third person perspectives refer to Hendricks [19].



The argument from incoherence is a first person point argument. Skepticism is thus rejected along these lines. A first person perspective is one of the very characteristics of autoepistemology. This is also suggested in the label ‘autoepistemology’ attaching the Moore-paradox to it: Whatever an agent may know or believe is partly fixed by the concern whether the epistemic or doxastic claim advocated by the inquiring agent fall victim of a Moore-paradox or not. As long as a thesis concerning epistemic commitments does not pan out in a Moore-paradox the inquiring agent is free to adopt it. As an autoepistemologist one may, by way of example, say

$$\text{‘If I believe that } A, \text{ then I believe that I know that } A\text{’} \quad (8)$$

which has later been called the *Moore-principle* and sometimes the *principle of positive certainty*.<sup>7</sup> Formalized (8) amounts to:

$$B_{\Xi}A \rightarrow B_{\Xi}K_{\Xi}A. \quad (9)$$

According to Moore’s theory, there is nothing self-contradictory or incoherent about asserting the principle. No more Moore paradox to the Moore principle than to the widely adopted principle that one knows that one knows if one does the plausibility of which Malcolm argues for below and elsewhere [46].

From Moore’s first person autoepistemological perspective a statement like

$$\text{‘} A \text{ is the case, but I don’t believe whether } A\text{’} \quad (10)$$

is a paradoxical Moorean statement. There is however nothing paradoxical about

$$\text{‘} A \text{ is the case, but } \Xi \text{ doesn’t believe whether } A\text{’} \quad (11)$$

from a third person perspective. In consequence, what for sure may sound quite implausible from the first person perspective, may sound very plausible from the third person perspective on inquiry and vice versa.

The epistemic and doxastic commitments that an agent may hold in the course of inquiry are sensitive the epistemic environment and what the agent in these local circumstances is both willing to and capable of defending or maximizing. He does not necessarily have an over-all skepticism defeating

---

<sup>7</sup>Lamarre and Shoham explain: ‘To the agent, the facts of which he is certain appear to be knowledge’, [35].

method at his disposal: You are doing the best you can, so is the skeptic, but he is probably not doing as well as you are due to incoherence. Forcing in autoepistemology then means:

*Whenever knowledge claims are challenged by alleged possibilities of error, the strategy is to show that on an individual basis one can do no better than what is being done in the current epistemic environment and attempt to show that the skeptic is doing at least as bad as you are but probably even worse*

Epistemic axioms may be interpreted as principles describing a certain strong rationality congruent with autoepistemology. First of all, neither Malcolm nor Moore would object to the idea that knowledge validates axiom T (1). Secondly, in Hintikka's logical system knowledge is closed in the sense of (4), and the argument cited by Hintikka in favor of closure has the flavor of autoepistemology:

In order to see this, suppose that a man says to you, 'I know that  $p$  but I don't know whether  $q$ ' and suppose that  $p$  can be shown to entail logically  $q$  by means of some argument which he would be willing to accept. Then you can point out to him that what he says he does not know is already implicit in what he claims he knows. If your argument is valid, it is irrational for our man to persist in saying that he does not know whether  $q$  is the case. [24], p. 31.

Not accepting (4) is irrational, but the acceptance of (4) does not entail that the agent in question has to be immediately aware of his own rationality, let alone able to immediately compute it from Hintikka's first person perspective on inquiry.

The autoepistemological inspiration is vindicated while Hintikka argues for the plausibility of the *KK*-thesis as a governing axiom of his logic of knowledge. Approximately a decade after the publication of *Knowledge and Belief*, the *KK*-thesis came under heavy attack. *Synthese* dedicated an issue to the matter where especially Ginet and Castenada were on the offensive, while Hintikka and Hilpinen defended.<sup>8</sup> And while defending, Hintikka refers to Malcolm:<sup>9</sup>

Many of the things Malcolm says fall flat if it is not the case that I in fact know what I claim to know. For instance, if I am the victim

---

<sup>8</sup> *Synthese* **21**, 1970.

<sup>9</sup> For a thorough discussion of Hintikka's conception of the *KK*-thesis, refer to Hendricks [17], pp. 253.

of a clever optical trick when I believe that there is an ink-bottle in front of me—and even believe that I know it in the strong sense—then exposing the trick will provide conclusive evidence against claiming that the ink-bottle is there ... More generally, we might perhaps say that if one knows in the strong sense that  $p$ , then it is the case that one will refuse (if acting rationally) to consider any experience compatible with what he in fact knows as evidence against one’s knowing that  $p$ . ([26]), p. 153.

From this Hintikka concludes that Malcolm’s position is sufficiently close to Hintikka’s own for a behavioral identity between the strong knowledge á la Malcolm á la Hintikka:

This is especially interesting in view of the fact that Malcolm himself uses his strong sense of knowing to explain in what sense it might be true that whenever one knows, one knows that one knows. In this respect, too, Malcolm’s strong sense behaves like mine. [26], p. 154.

Besides the requirement of closure and the validity of the  $KK$ -thesis, axiom T is also valid so the suggestion is that a logic of autoepistemology is philosophically congruent with Hintikka’s suggestion for an **S4** axiomatization describing strong rationality.

Although the epistemic logic of autoepistemology may be **S4**, the doxastic logic is another matter, and the affinities with autoepistemology end. Moore’s principle above (8) is a kind of introspection axiom for rational belief or *subjective certainty*. In a combined epistemic and doxastic logical system in which knowledge and belief are approximately equally strong (save for a truth-condition) the agent will (while subjectively reflecting upon his own state of mind with respect to what he believes) be led to believe that he knows the proposition in question if he certainly believes it. Some contemporary logical epistemologists embrace Moore’s principle (e.g. Halpern [15]). Hintikka denies Moore’s principle in *Knowledge and Belief*:

Hence ... and (C.BK) [Moore’s principle] are acceptable only when an unrealistically high standard of defensibility is imposed on one’s beliefs. The conditions would make it (logically) indefensible to suppose that anyone would have given up any of his present beliefs if he had more information than he now has. And this is clearly too stringent a requirement. [24], p. 52.

To Hintikka belief is a significantly weaker commitment than knowledge. For good reason too it turns out: Consider a combined epistemic and doxastic

logic in which belief is understood as subjective certainty such that (9) holds. Assume also that positive doxastic introspection

$$B_{\Xi}A \rightarrow K_{\Xi}B_{\Xi}A \quad (12)$$

holds for belief together with negative doxastic introspection

$$\neg B_{\Xi}A \rightarrow K_{\Xi}\neg B_{\Xi}A. \quad (13)$$

Even subjective certainty, as strong as it may seem in this system, implies a margin of error: The fact that  $\Xi$  is subjectively certain of  $A$  does not necessarily imply that  $A$  is true. Accordingly axiom T will be dropped for subjective certainty and replaced by the consistency axiom D

$$B_{\Xi}A \rightarrow \neg B_{\Xi}\neg A. \quad (14)$$

On the standard definition of knowledge, knowledge implies belief

$$K_{\Xi}A \rightarrow B_{\Xi}A \quad (15)$$

which is also an uncontroversially accepted assumption for knowledge and subjective certainty. The logic of subjective certainty is **KD45**. Knowledge will obviously have to be stronger than subjective certainty, so it must validate **S5**. On assumptions (9), (12)–(15) Lenzen was able to show that  $B_{\Xi}A$  in the end is equivalent to  $K_{\Xi}A$  [37]. So knowledge and belief collapse into each other!<sup>10</sup>

Many contemporary epistemic logics do nevertheless consider strong belief, rational belief or subjective certainty to be approximately as strong as knowledge. Assuming belief is taken to be approximately as strong as **S5** knowledge with the equivalence relation over worlds implies some attractive formal features like readily epistemic and doxastic partitions. This does not by itself make up for the result that the logic of knowledge and belief coincide.

Hintikka denies the axiom of wisdom because introspection alone should not license agents to ascertain whether some proposition in question is known. Other objections to (3) include the following: Under special circumstances axiom 5 suggests that agents can even decide intractable problems as Binmore reveals in [3], and Shin in [53]. Williamson has launched two objections to models of knowledge and belief validating axiom 5. For **S5** knowledge Williamson disagrees with the ones interpreting knowledge in a

---

<sup>10</sup>Stalnaker also discusses this issue in [56].

data-base like fashion to justify the closed world assumption of axiom 5. Even under the closed world assumption it does not follow in general that an agent can ‘survey the totality of its knowledge’.<sup>11</sup> Secondly, Williamson recently noted that the result to the effect that knowledge and belief collapse under the strong understanding of belief in a combined system points to the untenability of axiom 5, not to the unacceptable nature of subjective certainty *per se*. Moore’s principle is not too extravagant an assumption for rational belief, neither are axioms (12), D, (15) nor axioms T, 4 for knowledge. That leaves axiom 5 as the culprit responsible for collapsing the two notions and besides entails the infallibility of the agent’s beliefs: Whatever  $\Xi$  believes is true. On these grounds, Williamson abandons axiom 5 rather than any of the other principles used in the derivation [61]. Voorbraak makes the unusual move of sacrificing (15) accordingly challenging the intuitions of philosophers since antiquity [58]. In Hendricks [17] it is shown how limiting convergent knowledge and (3) conflict, and in Hendricks [19] it is demonstrated how the axiom of wisdom gives rise to both conceptual and technical problems in multi-agent systems.

## 2.2 ‘Epistemologics’

If **S5** assumptions about knowledge and belief are dropped ideal rationality descriptions and autoepistemological considerations may supply a philosophical foundation and motivation for logical epistemology.<sup>12</sup> The treatment of logical epistemology as a branch of modal logic is still quite costly also for much less ambitious logics than **S5**. The principle of closure (4) is enough to generate problems, and worse, skeptical problems. Nozick for instance emphatically denies closure for epistemic operators given his subjunctive definition of knowledge, and a whole range of other epistemic axioms likewise have to go [48].

### 2.2.1 Counterfactuality

According to Nozick, epistemology is not going to get off the ground before the skeptical challenge is met. It must be demonstrated that knowledge is at least possible. The often cited premiss in favor of the skeptical conclusion

---

<sup>11</sup>See [60], p. 317.

<sup>12</sup>From the point of view of autoepistemology, one also suspects that Moore himself would be disinclined to advocate the axiom of negative introspection (axiom 5). Either because it could amount to a Moorean sentence or because it imposes too much rationality on the part of the singular agent—there is a difference between doing the best you can, and then outdoing yourself.

that agents do not know much of anything is this: If the agent cannot be guaranteed to be able to know the *denials* of skeptical hypotheses, then the agent cannot be ascribed knowledge on any other issues. The traditional understanding of infallibilism counting every possible world as relevant supports the pessimistic premiss presented. Some arbitrary skeptical hypothesis is a possibility of error the falsity of which must be known to the agent for him to acquire knowledge of some other common hypothesis in question. The inability to know the denials of skeptical hypotheses suffice for lacking knowledge of the ordinary hypotheses.

The classical thesis of infallibilism supports the skeptical premiss by the demand that  $\Xi$  should be capable of knowing the denials of all the possibilities of error. The closure condition (4) demands that  $\Xi$  only is knowledgable of the denials of those possibilities of error which in effect are known logical consequences of  $\Xi$ 's knowledge.<sup>13</sup> Suppose  $\Xi$  knows the hypothesis that he is currently sitting reading this article on forcing epistemology. Let it also be the case that  $\Xi$  knows that if he is sitting reading this paper, then he is not being fooled by the Cartesian demon. Then  $\Xi$  must also know that he is not being fooled by the demon. If  $\Xi$  does not know that he is not being deceived by the demon then, given  $\Xi$  knows the implication,  $\Xi$  in turn lacks knowledge of the hypothesis that he is sitting reading forcing epistemology. Now this is exactly what the pessimistic premiss pushes for. But  $\Xi$  can know that he is sitting reading this article without knowing that there is no demon of deception seducing him into the false belief that he is sitting reading this paper. Being seated reading this paper implies that no Cartesian demon is leading  $\Xi$  to falsely believe that he is reading this very article.

Two things follow from this reasoning: (1) Everyday knowledge is secured, but (2) knowledge is not closed in the sense of (4) according to Nozick's counterfactual epistemology. If knowledge was to be closed it could fly far away into skepticism.

Having denied the condition of closure the epistemological mission is still not completed. An explanation must still be provided describing how knowledge of common hypotheses is possible joined with an explanation of the failure to know the denials of skeptical hypotheses. This also goes for the situations in which it is known that the common hypothesis at issue implies relevantly rejecting the skeptical hypothesis.

Dretske's solution is to install a modal condition for knowledge imposing

---

<sup>13</sup>... or perhaps rather known logical consequences of  $\Xi$ 's knowledge – including denials of all possibilities of error (the so-called contrast consequences, Dretske [9]).

truth-conduciveness by *sensitivity* [9]:

$$\text{If } A \text{ were not true, } \Xi \text{ would not believe } A. \quad (16)$$

A belief qualifying as knowledge is a belief which is sensitive to the truth: The proposition  $A$  is true in accordance with the standard definition of knowledge. Had  $A$  which is believed been false, the agent would not be led to the belief that  $A$ .

Condition (16) readily explains why closure fails. Proximity relations between possible worlds are introduced due to the semantics for the inserted subjunctive conditional. One may know both antecedents  $A$  and  $(A \rightarrow A')$  relative to one set of relevant worlds accessible from the actual world, and yet fail to know the consequent  $A'$  relative to a different set of possible worlds. Now relative to a set of possible worlds with proximity ‘close’ to the actual world one knows  $A$  and simultaneously knows that  $A$  implies the denial of the skeptical hypothesis, say  $A'$ . But one may all the same fail to know the consequential denial of the skeptical hypothesis itself for knowledge of the skeptical hypothesis is relative to possible worlds with a ‘way-off’ proximity to the actual world. These possible worlds are radically different from the actual world by all means. ‘Way-off’ worlds are accordingly forced out, skepticism far away because closure fails, but the possibility of knowledge prevails.

In the monumental monograph on knowledge, skepticism, free will and other pertinent philosophical issues [48], Nozick completes a definition of counterfactual knowledge along the Dretsian lines:<sup>14</sup>

$\Xi$  knows  $A$  iff

1.  $A$  is true,
2.  $\Xi$  believes that  $A$ ,
3.  $\neg A \mapsto \neg(\Xi \text{ believes that } A)$ ,
4.  $A \mapsto (\Xi \text{ believes that } A)$ .

To see how the definition works, the possible world semantics provides the following account of the truth-conditions for the subjunctive conditional: A subjunctive  $A \mapsto B$  for arbitrary statements  $A$  and  $B$ , is true, insofar, in all those worlds in which  $A$  is true that are in proximity ‘closest’ to the

---

<sup>14</sup>‘ $\mapsto$ ’ denotes the subjunctive conditional.

actual world,  $B$  is also true in these ‘closest’ worlds. More specifically of three worlds  $w, w', w''$  if  $w'$  is closer to  $w$  than  $w''$ , then  $A \multimap B$  will be true in  $w$  iff  $A$  is not true in any world or there exist a world  $w'$  in which  $A$  and  $B$  are true which is closer to  $w$  than any world  $w''$  in which  $A$  is true but  $B$  is false.<sup>15</sup>

For knowledge possession, one does not have to consult all possible worlds as the skeptic would insist: Given the standard semantical analysis of the subjunctives it is enough that the consequent  $B$  holds in those possible worlds which are closest to the actual world such that the antecedent  $A$  holds. Speaking in terms of forcing a subjunctive conditional is true just in case the consequent is forced among the closest worlds to the actual world in which the antecedent holds.

The third condition of the definition above is there to avoid error. The fourth is there to gain truth. The two conditions are collapsible into one condition:  *$\Xi$ 's belief tracks the truth of  $A$ :*

To know is to have a belief that tracks the truth. Knowledge is a particular way of being connected to the world, having a specific real factual connection to the world: tracking it. [48], p. 178.

The idea of introducing the proximity relation is that the agent’s local epistemic environment normally suffices for the truth witnessing Nozick’s first person stance. Although everyday knowledge is possible in many contexts, some contexts are just beyond reach: It is impossible for  $\Xi$  to know that he is not this brain in a vat. Assuming the brain receives the same sensory patterns as it would was it not dumped in the vat, there would not be anything in the input revealing to  $\Xi$  that he was not a brain in a vat. In this devious scenario  $\Xi$  is also barred from knowing that he is sitting reading this paper on forcing. If  $\Xi$  claims to know that he is sitting reading this article, it must follow that he as a prerequisite tacitly approves of the hypothesis that he is not a brain in a vat. Given this prerequisite and *modus tollens* as  $\Xi$  does not know that he is not sunk into the vat he does not know that he is sitting reading this paper either.

Now the possible world in which  $\Xi$  is a brain in a vat is *ceteris paribus* very distant from the actual world. Failure of knowledge in these cases is not devastating to counterfactual epistemology. It hinges on the relevant possibilities of error. True beliefs are only required in possibilities closer

---

<sup>15</sup>This semantic account of the subjunctive follows rather closely Lewis in [42]. Nozick is however not committed to a particular understanding of the semantics and also discusses Stalnaker’s subjunctive semantics from [54]. See furthermore [48], p. 680, footnote 8.



to actuality that any  $\neg A$ -possibilities: Picture a physicist measuring the voltage drop over some LRC-circuit. A student from epistemology class comes to him and asks whether a relevant possibility of error could be that the voltmeter is calibrated incorrectly. The physicist would probably answer ‘yes’ as calibration problems could lead to a measurement error. Then asking the scientist whether being a brain in a vat is a relevant possibility of error would likely result in the physicist asking the student to go back to his course and stop bothering him with silliness.

By his definition of counterfactual knowledge, Nozick accepts the axiom of veridicality (1), and the rule of necessitation (5) also seems to hold:  $A$  is true,  $\Xi$  believes  $A$ , and since  $A$  is true in all possible worlds,  $A$  is also true in close worlds so  $\Xi$  knows  $A$ .<sup>16</sup> But he rejects both closure and the  $KK$ -thesis (2) for counterfactual knowledge:

Some writers have put forth the view that whenever one knows, one knows that one knows. There is an immediate stumbling block to this, however. One may know yet not believe one knows; with no existing belief that one knows to do the tracking of the fact that one knows, one certainly does not know that one knows. [48], p. 246.

An agent may be tracking the truth of  $A$  without tracking the fact that he is tracking the truth of  $A$ . For much the same reason chances are also that Nozick would dismiss the axiom of wisdom (3) because if an agent is not tracking the truth of  $A$  it does not follow that he will be tracking the fact that he is not tracking  $A$ . The first person logic of counterfactual epistemology is thus very weak and *not* normal in contrast to Hintikka’s logical epistemology.

The counterfactual epistemology in general accommodates elements of the contextualistic epistemology of the next section. Dretske’s view of the closure lets knowledge transfer work across known implications insofar as the implications in question are close or relevant. Knowing that one is sitting down reading this article transfers immediately through the known implication to the ‘close’ hypothesis that one is not standing on a street corner doing the same. This knowledge will at the same time not run through the known implication to the ‘way-off’ hypothesis that one is not being fooled by a malicious demon. Dretske’s point seems to be that knowledge acquisition of a hypothesis in some common *context* assumes by default the very falsity of particular ‘way-off’ and irrelevant possibilities of error [9]. These possibilities of error are skirted, or their falsity presupposed in many

---

<sup>16</sup>I’m indebted to Robert Stalnaker for bringing this to my attention.

everyday knowledge acquisition contexts. Lewis strongly subscribes to this contextualistic forcing feature in his modal epistemology – so does Hintikka.

### 2.2.2 Contextuality

Lewis’ new ‘modal epistemology’ [45] is an elegant variation of contextualism which has many (forcing) features in common with Hintikka’s formal theory of knowledge.<sup>17</sup>

Contextualistic epistemology starts much closer to home than counterfactual epistemology. Agents in their local epistemic environments have knowledge—and plenty of it in a variety of (conversational) contexts. Knowledge is not only possible as counterfactual epistemology will have it, it is real human condition. The general contextualistic template for a theory of knowledge is crisply summarized in DeRose’s description of the attribution of knowledge. The description also embodies many of the epistemological themes central to the contextualistic forcing strategy:

Suppose a speaker *A* says, ‘*S* knows that *P*’, of a subject *S*’s true belief that *P*. According to contextualist theories of knowledge attributions, how strong an epistemic position *S* must be in with respect to *P* for *A*’s assertion to be true can vary according to features of *A*’s conversational context. [7], p. 4.

The incentive to take skeptical arguments to knowledge claims seriously is based on an exploitation of the way in which otherwise operational epistemic concepts, notably knowledge, can be gravely disturbed by sudden changes of the linguistic context in which they figure.

The standards for the possession of knowledge vary from context to context depending on what is at stake. In a course on epistemology the standards for knowledge possession fixed by the interlocutors (teacher and students) are usually very high. The conclusions that we know very little, if anything at all, may by the end of class be true. In a discussion after class a fellow student says ‘I know that *Matrix Revolutions* plays in the Park & 86th Street Theater on 125 E. 86th St.’. The circumstances have now changed

---

<sup>17</sup> Admittedly, Lewis does not refer directly to his theory as ‘modal epistemology’. The term figures in reference to Stalnaker’s distinction between modal and hyperintensional epistemology [55] (according to which all ‘forcing’ epistemology is modal epistemology). To differentiate Lewis’ build of contextualism from other contextualistic models, modal epistemology is reserved for Lewis’ version.

For excellent surveys of contextualism refer to Pritchard [50], [51].

and the standards for knowledge possession in this new, presumably, non-skeptical conversational context are lower. The relatively lower standards put us in the comfortable position of maintaining that we know most of what we think we know. It is admittedly to this low epistemic standard but it surely suffices for going to the movies.

Not only may knowledge attributions fluctuate with contexts, they may also be sensitive to who ascribes knowledge to whom. As indicated by DeRose there is a delicate issue to be addressed pertaining to the strength of the position an agent has to be in order for the epistemic commitment to truthfully pan out. This position is context-sensitive, not only to the agent in the environment, but also to possible ascribers of knowledge to the very agent in question. The first-third person dichotomy is immanent in contextualistic epistemologies.

Finally, the strength of the epistemic position is responsible for turning the contextualistic theory of knowledge into a modal account according to DeRose. For every local environmental ‘time-slice’ the epistemic position of the agent remains constant. The epistemic position the agent however *were* to be in to warrant possession of knowledge is a subjunctively defined spatio-temporal function of the context. A strong epistemic position with respect to some hypothesis  $A$  is to have belief as to whether  $A$  is the case and tracking this fact not only through the actual world but through close worlds as well. Maintaining that one’s belief still tracks the truth at long distances increases the strength of the epistemic position with respect to the hypothesis in question. For belief to become knowledge it should be ‘non-accidentally’ true in the actual world and in close ones as well.<sup>18</sup> This way of realizing the forcing relation resembles the construction advanced by the counterfactual epistemology of the previous section using sensitivity or tracking.

Lewis’ modal epistemology as a contextualistic theory of knowledge is particularly engaging as it balances elegantly between mainstream and formal *modi operandi*. This is not too surprising since Lewis through his career was concerned with modal logic, in particular the logic of counterfactuals [42], modal ontology [44] and almost consequently modal epistemology [45]. Modal logics, epistemic logics in particular, are much about partitioning the set of all possible worlds into classes that are in close proximity, similar, relevant or accessible from the actual world and into classes which are not.

As humans we force for knowledge on a daily basis and obtain it. This means partitioning the set of all possible worlds into relevant, irrelevant

---

<sup>18</sup>See further [7], p. 34.

and extravagant possibilities of error determined by the current context. To obtain knowledge eliminate the relevant possibilities of error, ignore the extravagant ones, and succeed over the remaining possible worlds where the hypothesis in question is true. Everything dictated by the current context. There are rules for elimination, ignoring and success. On a new definition of knowledge yet to be formulated, these rules are what Lewis' modal epistemology is about. Only a selected few of them will be discussed here.<sup>19</sup>

Taking infallibility as a basic epistemological condition, for an agent to know a hypothesis, all possibilities of error must be eliminated given the agent's available information. That is, all the possible worlds in which the negation of the hypothesis is the case must be eliminated. This forcing relation is given by different measures. One measure is simply to ignore possibilities extravaganza, another is to use the available evidence to force such that the uneliminated possible worlds are determined by perceptual equivalences over these alternatives with the actual world as the fix-point. The perceptual experience (and memory) the agent has in the actual world fixes the set of uneliminated possible worlds insofar the agent's cognitive apparatus functions the same in these worlds. Suppose that a perceptual experience has the propositional content  $A$ . The perceptual experience with content  $A$  (memory included) eliminates a certain world  $w'$  if and only if the content of the experience the agent has in  $w'$  differs from  $A$ .

Quantifiers are usually restricted to run over some domain of interest. This also goes for the universal quantifier over possible worlds that would lead to error. Every uneliminated world in which the hypothesis holds is restricted to a sub-domain of properly all uneliminated worlds. Saying that the surface is 'clean' in a certain conversational context is to properly ignore the microscopic dust particles laying on the surface. If somebody was to disagree it would have to be because the new interlocutor in the conversational context means clean in a more restrictive sense. The microscopic dust balls in this case suffice for making the assertion about the clean surface false. Words like 'flat' or 'round' behave in the same way, as does the word 'knowledge'. They are context-sensitive.<sup>20</sup>

Alterations of the conversational context occur when a new hypothesis is introduced which for its part is more demanding than any of the other hypotheses currently explicit in the particular context. Such a *non-uniform* introduction implies an increase in the range of possible worlds to be con-

---

<sup>19</sup>See Hendricks [19] for a complete exposition of the rules from a forcing perspective.

<sup>20</sup>The context-sensitivity of various words including 'knowledge' was noted by Lewis much earlier in [43].

sidered for attribution of knowledge. The strength of the required epistemic position mentioned above is increased accordingly. In a context where the usage of ‘knowledge’ remains *uniform* throughout the conversation, the range of possible worlds to be considered remains stable. Given the context-sensitive nature of knowledge, in every context where knowledge attribution is at stake some uneliminated possible worlds are not rendered relevant by the *current* context. The universal quantifier is restricted accordingly. This restriction is very similar to the quantifier restriction on knowledge in logical epistemology. In epistemic logic, knowledge claims are circumscribed by the compartment of possible worlds in accordance with the epistemic attitude, not the incompatible compartment and not the set of all possible worlds.

These considerations essentially pave the way for the colloquially stated but forceful knowledge definition of modal epistemology:

*S* knows that *P* iff *S*’s evidence eliminates every possibility in which not-*P*—Psst!—except for those possibilities that *we* are properly ignoring. [45], p. 378.

During the individual knowledge attribution process, the possible world which the agent takes to be the actual state of affairs is never ignored. Actuality is by reflexivity always a relevant possible world alternative although indexical. It follows that falsity may not properly be supposed. If falsity is never to be presupposed whatever in the end will turn up knowledge must be true, so the classical condition of truth for knowledge is derived. Never ignoring the actual world is referred to as the *rule of actuality*.

Turn next to the ascription of knowledge to others. The way in which the modal knowledge definition is stated italicizes ‘*we*’. What *we* may properly ignore is going to be dependent on whose actuality is being referred to in the context in question. Assuming that there is only one actual world-index in play in non-modal contexts one should expect that the world considered actual by the agent coincides with the world indexed ‘actual’ by the ascribers.

In counterfactual situations referring for instance to what an agent would have known today had he read the paper yesterday, or whether an agent knew yesterday who he was then, fixing the index of actuality is trickier. Had the agent read the paper yesterday he would presumably have known more than he in fact knows today. The agent is ascribing knowledge and ignorance to himself now as the one not having read the paper last night. The ascriber, say  $\Xi'$ , of knowledge to agent  $\Xi$  has an index of actuality demonstratively different from  $\Xi$ ’s index. The index on actuality for  $\Xi'$  is what  $\Xi'$  would have been like knowledge-wise had he read the paper yesterday. Actuality

indices differ for  $\Xi$  and  $\Xi'$  in this situation. Similarly for the attribution of knowledge to  $\Xi$  knowing yesterday who he was. For  $\Xi$ 's reality is defined for his spatio-temporal history up until yesterday; for  $\Xi'$  reality is defined for his spatio-temporal history up to today when the question is popped whether  $\Xi$  knew yesterday who he was. The two world stories are different.  $\Xi$ 's actuality yesterday is different from  $\Xi'$ 's actuality today. Similarly for a host of other situations involving say iterated modal constructions like knowledge of knowledge etc.

The rule of actuality applies both to the ascriber and the ascribed. What may not be properly ignored is the local agent's actuality. Epistemologists considering what  $\Xi$  knows from a third person perspective will attend to whatever possible worlds that  $\Xi$  himself attends to as possible and then some. The set of possible worlds ignored by a third person knowledge attributor for  $\Xi$  will properly be a superset of the possible worlds  $\Xi$  ignores. An agent may know more than what may be ascribed to him because his actuality in some cases differs from the ascribers and his range of viable worlds does as well. Applying the principle of 'epistemic' charity means that while attributing knowledge to an agent in his local epistemic environment, the third person ascriber may ignore fewer possibilities than  $\Xi$ .

Next, a world  $w'$  which 'salient resembles' another world  $w$  enforces a kind of symmetry. If  $w$  may not be properly ignored in virtue of the other rules neither may  $w'$  and vice versa. This accessibility clause is referred to as the *rule of resemblance*. The rule is dangerous and powerful at the same time.

The rule of resemblance is dangerous because not applied carefully invites skepticism and global underdetermination back in. The actual world is left uneliminated by the agent's available evidence. It follows that any other uneliminated world resembles the agent's actual world in one important respect, namely, with respect to the agent's evidence. This will continue to hold even in worlds which otherwise are radically different from the agent's actual world including the demon world. By application of the rule of actuality together with the rule of resemblance leads to the conclusion that these worlds are relevant alternative worlds as well!

There is no obvious remedy to this problem and it reappears with respect to knowledge closure. Agreeing with counterfactual epistemology that closure over arbitrary contexts amounts to a fallacy driving skeptical arguments, modal epistemology holds that closure is possible locally without skepticism. Knowledge is closed for a fixed context. Knowing that you are reading this paper implies that you are reading this paper and not being deceived (by a demon or a mad scientist) in this particular uniform con-

text  $c_1$ . If the context is non-uniformly changed right after the antecedent conditions obtain to a new context  $c_2$ , ‘all bets are off’ [45], p. 382:

$$\underbrace{K_{\Xi}h \rightarrow K_{\Xi}(h \rightarrow h')}_{c_1} \quad \overset{\text{change!}}{\uparrow} \quad \underbrace{\nrightarrow K_{\Xi}h'}_{c_2} \quad (17)$$

Closure fails because the strength of the epistemic position now required in  $c_2$  to attribute knowledge has been increased way beyond  $c_1$  by the increase in possible worlds at issue dictated by  $c_2$ . The range of possible worlds may now include the demon world which is a whole different context. Knowledge is closed under implication because implication preserves truth in a fixed context not over arbitrary contexts.

If knowledge is closed in uniform contexts, then this seems to be exactly what Hintikka could say when presented with the closure challenge and the skeptical invitation. The argument for closure so far rests on autoepistemological and rationality considerations but does not necessarily escape Nozick’s argument against closure. Since *Knowledge and Belief* Hintikka has emphasized the importance of partitioning the set of worlds into the two distinct compartments consisting of the worlds in accordance with the attitude and the ones not. The worlds in accordance with the epistemic attitude may be read in accordance with Lewis’ context-sensitive quantifier restriction on knowledge above. Then, the demon world, brain-in-a-vat world and other derivatives of global underdetermination are simply excluded from the compatibility partition; these extravagant worlds are not in accordance with the epistemic attitude.<sup>21</sup> Thus, these error-possibilities will not disturb the context, or in Hintikkian terms, will not pass over into the compatibility partition, so knowledge is closed for a given compatible partition, i.e. uniform context.<sup>22</sup>

One of Lewis’ rules seem trivial, and yet it furnishes insight as to Lewis’ view of the situation in epistemology today. Knowledge attribution is partly a socially determined process forced by conventional means to be taken seriously. This seriousness is reflected in the *rule of attention*. Which worlds

<sup>21</sup>Global underdetermination amounts to the impossibility of reliable knowledge acquisition anyway as Kelly has argued in [33].

<sup>22</sup>There is however not any obvious way to ensure that such a contextual change is not taking place in Lewis’ modal theory of knowledge. The rules of actuality and resemblance combined immediately permit for such a change to occur. The demon world resembles saliently the actual world with respect to agent’s evidence and should accordingly not be ignored. Lewis readily admits to an *ad hoc* modification of the rule as to exclude this resemblance. Observe that this does not immediately apply to Hintikka’s logical epistemology.

are ignored is context-dependent. When ignored in a specific context these worlds are *really*, not only counterfactually so, tossed out and not to be considered. Attending to even far-fetched possible worlds in a different context make them relevant possibilities again. Relevant possibilities of error undercut infallible knowledge claims and knowledge flies away—becomes elusive.<sup>23</sup>

Buying into too many uneliminated possibilities of error often makes epistemologists end up with buyers regret. Potential counterexamples to knowledge ascriptions are waiting everywhere in the wings of rich domains making the required epistemic position impossible to reach for anybody. No first persons have knowledge in these particularly demanding contexts, no third persons either. Unfortunately, as a discipline epistemology is one such demanding context. The foe of epistemology is not really skepticism but epistemology itself:

That is how epistemology destroys knowledge. But it does so only temporarily. The pastime of epistemology does not plunge us forevermore into its special context. We can still do a lot of proper ignoring, a lot of knowing, and a lot of true ascribing of knowledge to ourselves and others the rest of the time. [45], p. 377.

Modal epistemology concedes to skepticism the high epistemic standards on which the skeptical position operates. These epistemic standards are exceedingly harder to meet than those required for everyday attributions of knowledge. Admitting this much to skepticism licences the concern that these elevated standards are in fact the correct standards to be met for genuine knowledge ascriptions and acquisitions. When push comes to shove, the everyday knowledge attributions do not stand up to these standards, so knowledge attributions on a daily basis are bogus as discussed by Pritchard [49]. Skepticism can never be dodged. The rules may conflict in such a way that skeptical possibilities like hallucinations become relevant. Applying the prohibitive rule of resemblance merely escapes skepticism by *ad hoc* qualifications. This leaves us again ‘caught between the rock of fallibilism, and the whirlpool of skepticism’ as Lewis puts it [45], p. 367. Modal epistemology was supposed to come to the rescue.

As bogus as these ascriptions may seem, they may also be as good as it gets. A similar response to skepticism following ‘smooth’ lines may be

---

<sup>23</sup>Ignoring worlds may from this perspective be seen as a necessary last resort because the available evidence may always be insufficient to block global underdetermination. Ignoring is a precondition for knowledge—love it or leave it.



found in Levi's formal epistemology [39], [40]. To gain truth and avoid error beliefs should be chosen carrying the highest 'epistemic utility'. The epistemic utility embodies truth as well as content. Significant possibilities of error are forgivable just the agent settles for the belief with the highest epistemic utility in the particular context. This may not exactly add up to real knowledge but it is good enough for decision and action. The elevation of the skeptical standards for knowledge is immaterial for common epistemic practice. Infallibilism with respect to all worlds cannot be reached anyway and agents are doing the best they can quantifying over less reaching at least a workable impasse with skepticism. That is the epistemic balance; Agents can act on their 'discount' infallible knowledge, but skeptics can do very little with their high standards. Turning the tables, skeptics are the *real* epistemologists.

Denials of skeptical hypotheses cannot be known on the modal conception of knowledge trans-contextually. So an objection would be that knowledge is not even possible, much less real. A defense would be to simply admit that the logics of knowledge are rather weak at least for the third person knowledge operator and in case of contextual changes. As opposed to counterfactual epistemology's denial of closure it holds for a first person operator in a uniform context in Lewis' modal epistemology. Closure may fail from the third person perspective because the set of worlds to be considered is strictly a superset of the set of worlds the first person operator has to consider leaving room for radical context change, and a failure. There is support to be found for such a defense.

Levi's epistemological program is a version of a first person perspective emphasizing a distinction between *the logic of truth* and *the logic of consistency* and not the first and third person perspectives [41]. Even though related the two distinctions are not exactly the same. Levi denies the validity of various epistemic axioms as axioms of an epistemic logic of truth. This crudely means to reject these axioms as axioms for a third person knowledge operative. An axiom like the *KK*-thesis found to be invalid in counterfactual epistemology is here valid as an axiom serving regulative purposes of maintaining consistency for a rational epistemic agent. The logic of truth for an epistemic agent on the other hand is not necessarily regulated by a principle like the *KK*-thesis. Lewis seems to follow suit because knowledge of knowledge introduces a discrepancy of actualities for the first and the third person operator. Because of the subject-based contextualism enforced by the rule of actuality, the third person operator is to ignore fewer worlds leaving more room for error. The agent may perhaps know that he knows, the third person may not necessarily be able to determine that the

$KK$ -thesis holds for the agent, nor that it holds for himself pertaining to the agent in question. The agent in the local environment may have more knowledge than a third person is able to ascribe to him or to the third person himself. If there is a trans-contextual third person logic of knowledge, such a logic is probably rather weak seems to be the suggestion of Levi and Lewis.

While Lewis may consider a universal third person logic rather weak there is nothing in the way of arguing for a much stronger first person logic. This is in stark contrast to the counterfactual proposal of the previous section in which the first person logic was quite weak. On the modal epistemological account all of (1)–(4) may be valid in uniform contexts for a first person knowledge operator.

### 3 ‘Elusive’ Logical Epistemology

There is a feature of Hintikka’s logical epistemology which may make it become as ‘elusive’ as any careless mainstream theory of knowledge. The principle of closure, axiom  $K$  (4), can under the certain circumstances be generalized to a stronger closure property of an agent’s knowledge considered still more unacceptable than (4) itself. *Logical omniscience*:

*Whenever an agent  $\Xi$  knows all of the formulae in a set  $\Gamma$  and  $A$  follows logically from  $\Gamma$ , then  $\Xi$  also knows  $A$ .*

In particular,  $\Xi$  knows all theorems (letting  $\Gamma = \emptyset$ ), and he knows all logical consequences of a formula which he knows (letting  $\Gamma$  consist of a single formula). Logical omniscience incorporates some generally weaker forms of omniscience like knowledge of valid formulae: Agent  $\Xi$  knows all logical truths (given rule 5) etc.<sup>24</sup>

Technical solutions to logical omniscience are either facilitated on the syntactical or semantical level. On the syntactical level, Hintikka recently suggested [28] to place suitable syntactical constraints on deductive arguments which preserve knowledge. Interesting philosophical solutions are to be found on the semantical level. The idea is here to introduce some semantical entities which account for why the agent could be accused of logical omniscience but by the end of the day is not guilty of logical omniscience. These entities are called ‘impossible, possible worlds’ by Hintikka [27]. Similar entities called ‘seemingly possible’ worlds represented by urn-models are introduced by Rantala [52]. Allowing impossible possible worlds in which

---

<sup>24</sup>See [15], [14] for a full list of logical omniscience forms.

the semantic valuation of the formulas in a certain sense is arbitrary provide the necessary means for dodging logical omniscience: The logical laws do not pass muster in the impossible possible worlds. When knowledge is evaluated with respect to all possible worlds but the logical laws do not hold in some of them, logical omniscience is simply out. In an impossible possible world a tautology  $A \rightarrow A$  may, as odd as it admittedly sounds, be false. Now the agent  $\Xi$  may all the same view that very world a possibility, so universally  $K_{\Xi}(A \rightarrow A)$  fails. In consequence, the rule of necessitation (5) is invalid in impossible possible world models. Axiom K is the victim of failure as well. In the impossible possibility both  $A$  and  $A \rightarrow A'$  may be true while simultaneously  $A'$  is false. The failure of axiom K would satisfy Nozick although he probably would consider impossible possible worlds as weird as demon worlds if not weirder. From a strictly logical point of view the epistemic logics specified by impossible worlds models are not very exciting. No real epistemic statement is valid in a universal way. The validity of the various epistemic principles may however be obtained by imposing suitable constraints on the impossibly possible models.

From a forcing perspective the introduction of impossible possible worlds is a rather curious strategy. The idea is to first inflate the local circumstances of the agent in the sense that the agent may regard some models of the (real) world possible. Then afterwards deflate the local situation because of the limited reasoning capacities of the agent. The worlds in question are really logically impossible. For example, a logical contradiction cannot be true. An agent may nevertheless not have enough resources to determine the truth-value of that contradiction and simply assume it to be true. He will consider some worlds possible, although logically they are impossible. To avoid logical omniscience let more worlds in, worlds worse than the demon worlds since the latter are at least logically possible whereas the former impossible possible worlds are not.

## 4 Epistemological Axiomatizations

There is a distinct formal feature to both Nozick's counterfactual and Lewis' contextual theories of knowledge. They are in a sense 'formal mainstream' theories as they both observe the significance of epistemic axioms drawn from Hintikka's logical epistemology and their intimate relations to the algebraic properties of the accessibility relation between possible worlds. Nozick considers the accessibility relation to be reflexive while Lewis takes it to be at least reflexive and a sort of symmetric given the rule of actuality and the

rule of resemblance respectively. Now, closure holds in uniform contexts, the  $KK$ -thesis holds, and the rule of necessitation will also immediately hold for a first person contextual epistemological logic. Using the sliding scale devised by logical epistemology to determine validity will make the first person modal epistemological logic at least have epistemic strength on the order of **S4**, perhaps even **S5** is acceptable to Lewis under certain conditions although not discussed. The third person logic of Lewis' contextualism seems to be no stronger than Nozick's first person logic validating (1) and rule (5) which by being so weak is a non-normal modal logic.

Table 1 below summarizes the results pertaining to the validity of common epistemic axioms given the first and third person perspectives on inquiry for logical, counterfactual and modal epistemology

	CE	ME	LE
N: $\frac{A}{K_{\Xi}A}$	1	1/3	1
K: $K_{\Xi}(A \rightarrow A') \rightarrow (K_{\Xi}A \rightarrow K_{\Xi}A')$		(1)/(3)	(1)
T: $K_{\Xi}A \rightarrow A$	1	1/3	1
4: $K_{\Xi}A \rightarrow K_{\Xi}K_{\Xi}A$		1	1
5: $\neg K_{\Xi}A \rightarrow K_{\Xi}\neg K_{\Xi}A$		(1)	

Table 1: CE: Counterfactual Epistemology, ME: Modal Epistemology, LE: Logical Epistemology. 1: First person perspective, 3: Third person perspective. (,): context-sensitive validity

The axioms are in turn answers to skepticism as their validity is sensitive to the forcing restrictions entertained by the various paradigms of knowledge considered above. Nozick's strategy to combat the skeptic is to impose very little relational structure on the universe of possible worlds leaving the skeptic with very little room to manoeuvre, thus limiting the skeptic's movement. The strategy of modal and logical epistemology is the opposite: To impose much more relational structure on the universe of worlds (in uniform contexts) leaving the agent with much room to manoeuvre, thus enhancing the agent's movement. To combat skepticism, force the skeptic out, either by not giving him a chance to cite distant possibilities of error as

relevant, or by making sure that whatever he cites you can reach truthfully at least from the first person perspective.

The common epistemic axioms now furnish a challenging meeting point for mainstream and formal epistemologies ... and there are many others. Some more are to be found in *Forcing Epistemology* [19] others yet uncovered. Let's join the forces and continue what Jaakko Hintikka pursued from the very beginning: To create an interactive epistemology of value to the interdisciplinary study of knowledge.

## References

- [1] Bernecker, S. and Dretske, F. (eds.) (2000). *Knowledge: Contemporary Readings in Epistemology*. Oxford: Oxford University Press.
- [2] Benthem, J.F.A.K. v. (2003). 'Fifty Years: Changes and Constants in Logic', in [20]: 35-56.
- [3] Binmore, K. and Shin, H. S. (1992). 'Algorithmic Knowledge and Game Theory', in *Knowledge, Belief, and Strategic Interaction*, Bicchieri, C. and Dalla-Chiara, M.-L. (eds). Cambridge: Cambridge University Press: 141–154.
- [4] Bull, R. and Segerberg, K. (1984). 'Basic Modal Logic', in ([12]): 1–88.
- [5] Chisholm, R. (1963). 'The Logic of Knowing', *Journal of Philosophy* **60**: 773–795.
- [6] Chisholm, R. (1977). *The Theory of Knowledge*, 2nd edn. First edition 1966. Englewood Cliffs, NJ: Prentice Hall.
- [7] DeRose, K. (1995). 'Solving the Skeptical Problem', *Philosophical Review* **104**: 1–52.
- [8] DeRose, K. (2000). 'How Can We Know that We're not Brains in Vats?', *The Southern Journal of Philosophy* **38**: 121–148.
- [9] Dretske, F. (1970). 'Epistemic Operators', *The Journal of Philosophy* **67**(1970): 1007–1022.
- [10] Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.

- [11] Fagin, R., Halpern, J. Y., Moses Y. and Vardi, M. Y. (1995). *Reasoning about Knowledge*. Cambridge: MIT Press.
- [12] Gabbay, D. and Guenther F. (1984). *Handbook of Philosophical Logic*, vol. II: *Extensions of Classical Logic*. Dordrecht: D. Reidel Publishing Company.
- [13] Girle, R. (2000). *Modal Logics and Philosophy*. Acumen Publishing.
- [14] Gochet, P. and Gribomont, P. (2003). ‘Epistemic Logic,’ forthcoming in *Handbook of the History and Philosophy of Logic*, edited by Gabbay, D.M. and Woods, J. Amsterdam: Elsevier Science.
- [15] Halpern, J. Y. (1995). ‘Should Knowledge Entail Belief?’, *Journal of Philosophical Logic* **25**: 483–494.
- [16] Hendricks, V. F. (1999). ‘Methodology in Epistemology’, *Danish Yearbook of Philosophy* **34**: 43–64.
- [17] Hendricks, V. F. (2001). *The Convergence of Scientific Knowledge—a view from the limit*. Trends in Logic: Studia Logica, vol. 9. Kluwer Academic Publishers.
- [18] Hendricks, V. F. (2003). ‘Active Agents’,  $\Phi$ NEWS (2002), vol. **2**: 5–40. A revised version of the paper is published in a special volume of the *Journal of Logic, Language, and Information*, van Benthem, J. and van Rooij, R. (eds.), vol. **12**, Autumn 2003: 469–495.
- [19] Hendricks, V. F. (2004). *Forcing Epistemology*, forthcoming. New York: Cambridge University Press.
- [20] Hendricks, V.F. and Malinowski, J. (eds.) (2003). *Trends in Logic: 50 Years of Studia Logica*. Trends in Logic: Studia Logica Library Series, Kluwer Academic Publishers. Contributions from J. van Benthem, W. Buszkowski, M.L. Dalla Chiara, M. Fitting, J.M. Font, R. Giuntini, R. Goldblatt, V. Marra, D. Mundici, R. Leporini, S.P. Odintsov, H. Ono, G. Priest and H. Wansing.
- [21] Hendricks, V. F. and Pedersen, S. A. (2000). *The Companion to Modal Operator Theory—A Program in Philosophy Online*. Department of Philosophy and Science Studies, Roskilde University, Denmark: <http://www.mot.ruc.dk>.

- [22] Hendricks, V. F. and Pedersen, S. A. (2002). *Moderne elementær logik*. Forlaget Høst & Søn.
- [23] Hendricks, V.F., Pedersen, S.A. og Jørgensen, K.F. (eds.) (2003). *Knowledge Contributors*. Synthese Library, vol. 322. Kluwer Academic Publishers. Contributions from H.v. Ditmarsch, R. Fagin, J. Halpern, J. Hintikka, W. v. d. Hoek, B. Kooi, W. Lenzen, Y. Moses, H. Rott, J. Sowa, M. Vardi and R. Wójcicki.
- [24] Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell: Cornell University Press.
- [25] Hintikka, J. (1969). ‘Semantics for Propositional Attitudes’, in *Models for Modalities*, Dordrecht: D. Reidel Publishing Company: 87–111.
- [26] Hintikka, J. (1970). ‘Knowing that One Knows’ Revisited’, *Synthese* **21**: 141–162.
- [27] Hintikka, J. (1975). ‘Impossible Possible Worlds Vindicated’, *Journal of Philosophical Logic* **4**: 475–484.
- [28] Hintikka, J. (1989). ‘Reasoning about Knowledge in Philosophy: The Paradigm of Epistemic Logic’, in *The Logic of Epistemology, and the Epistemology of Logic*, J. and M. Hintikka (eds.). Kluwer Academic Publishers: 17–35.
- [29] Hintikka, J. (1999). *Inquiry as Inquiry: A Logic of Scientific Discovery*. Kluwer Academic Publishers.
- [30] Hintikka, J. (2003). ‘A Second Generation Epistemic Logic and its General Significance’, in [23].
- [31] Hintikka, J. (2003). ‘Epistemology without Knowledge and Belief,’ *draft*: 35 pp.
- [32] Huang, Z. and Kwast, K. (1991). ‘Awareness, Negation and Logical Omniscience’, *Logics in AI, Proceedings JELIA’90*, Eijck, J. van (ed.). New York: Springer-Verlag: 282–300.
- [33] Kelly, K. (1996). *The Logic of Reliable Inquiry*. New York: Oxford University Press.
- [34] Kripke, S. (1963). ‘Semantical Analysis of Modal Logic’, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* **9**: 67–96.

- [35] Lamarre, P. and Shoham, Y. O. (1994). ‘Knowledge, Belief and Conditionalization’, *Principles of Knowledge Representation and Reasoning. Proceedings of the 4. International Conference (KR '94)*, Doyle, J., Sandewall, E. and Torraso, P. (eds.). Morgan Kaufman: 415–424.
- [36] Lemmon, E. J. (1977). *An Introduction to Modal Logic*, in collaboration with D. Scott. Oxford: Basil Blackwell Publishers.
- [37] Lenzen, W. (1978). *Recent Work in Epistemic Logic*, in *Acta Philosophica Fennica* **30**: 1–219.
- [38] Lenzen, W. (2003). ‘Knowledge, Belief, and Subjective Probability’, forthcoming in [23].
- [39] Levi, I. (1983). *The Enterprise of Knowledge*. Cambridge: Cambridge University Press.
- [40] Levi, I. (1991). *The Fixation of Belief and its Undoing*. Cambridge: Cambridge University Press.
- [41] Levi, I. (1997). ‘The Logic of Full Belief’, *The Covenant of Reason: Rationality and the Commitments of Thought*, Peters, A. K. (ed.), Cambridge: Cambridge University Press: 40–69.
- [42] Lewis, D. (1973). *Counterfactuals*. Cambridge: Harvard University Press.
- [43] Lewis, D. (1976). ‘Scorekeeping in Language Games’, *Journal of Philosophical Logic* **8**: 339–359.
- [44] Lewis, D. (1984). *On the Plurality of Worlds*. Oxford: Basil Blackwell Publishers.
- [45] Lewis, D. (1996). ‘Elusive Knowledge’, *The Australian Journal of Philosophy* **74**: 549–567. Reprinted in [1]: 366–384.
- [46] Malcolm, N. (1952). ‘Knowledge and Belief’, *Mind* **LXI**, 242. Reprinted in *Contemporary Readings in Epistemology*, Goodman, M. F. and Snyder, R. A. (eds.). Prentice Hall. New Jersey: Englewood Cliffs (1993): 272–279.
- [47] Moore, G. E. (1959). ‘Certainty’. *Philosophical Papers*. London: Allen and Unwin/Unwin Hyman. Reprinted in *Contemporary Readings in Epistemology*, Goodman, M. F. and Snyder, R. A. (eds.). Prentice Hall. New Jersey: Englewood Cliffs (1993): 257–271.



- [48] Nozick, R. (1981). *Philosophical Explanations*. Cambridge: Harvard University Press.
- [49] Pritchard, D.H. (2001). ‘Contextualism, Skepticism, and the Problem of Epistemic Descent’, *Dialectica* **55**: 327–349.
- [50] Pritchard, D.H. (2002a). ‘Contemporary Skepticism’, *Internet Encyclopedia of Philosophy*:  
<http://www.utm.edu/research/iep/s/skepcont.htm>.
- [51] Pritchard, D.H. (2002b). ‘Recent Work on Radical Skepticism’, *American Philosophical Quarterly* **39**(3): 215–257.
- [52] Rantala, V. (1975). ‘Urn Models: A New Kind of Non-Standard Model for First-Order Logic’, *Journal of Symbolic Logic* **4**: 455–474.
- [53] Shin, H.-S. and Williamson, T. (1994). ‘Representing the Knowledge of Turing Machines’, *Theory and Decision* **37**(1): 125–146.
- [54] Stalnaker, R. (1968). ‘A Theory of Conditionals,’ in *Studies in Logical Theory*, Rescher, N. (ed.). Oxford: Basil Blackwell: 98–112.
- [55] Stalnaker, R. (1983). *Inquiry*. Cambridge: MIT Press.
- [56] Stalnaker, R. (1996). ‘Knowledge, Belief and Counterfactual Reasoning in Games’, *Economics and Philosophy* **12**: 133–163.
- [57] Stalnaker, R. (1999). ‘Extensive and Strategic Form: Games and Models for Games,’ *Research in Economics* **53**: 291–293.
- [58] Voorbraak F. (1991). ‘The Theory of Objective Knowledge and Rational Belief, *Logics in AI, Proceedings of the European Workshop JELIA '90*. New York: Springer Verlag: 499–515.
- [59] Von Wright, G., H. (1951). *An Essay on Modal Logic*. Amsterdam: North-Holland Publishing Company.
- [60] Williamson, T. (2000). ‘Tennant on Knowable Truth’, *Ratio* **13**, **2**: 99–114.
- [61] Williamson, T. (2001). ‘Some Philosophical Aspects of Reasoning about Knowledge’, *Proceedings of the 8th Conference on TARK*, van Benthem, J. (ed.). San Francisco: Morgan Kaufman Publishers: 115–121.
- [62] Williamson, T. (2002). *Knowledge and Its Limits*. Oxford: Oxford University Press.