Belief, Methodology and Knowledge

Vincent Fella Hendricks
Department of Philosophy
University of Copenhagen
Njalsgade 80
DK 2300 Copenhagen S, Denmark
vincent@coco.ihi.ku.dk

Stig Andur Pedersen
Departments of Philosophy and Mathematics
University of Roskilde
P. O. Box 260
DK 4000 Roskilde, Denmark
sap@ruc.dk

May 4, 1999

In Honor of Peter Gärdenfors' 50th Birthday

Peter Gärdenfors's accomplishments are very impressive and we were very honored when given the opportunity to contribute to this *Festschrift* acknowledging his 50th birthday. Of the many accomplishments one outstanding one is his formulation of the AGM-axioms of *belief revision*. Belief revision has now become a prosperous enterprise for epistemologists, methodologists and logicians. In 1995 we started, along with Kevin T. Kelly and Oliver Schulte from Carnegie Mellon University, to view belief revision from a formal learning theoretical perspective. This resulted in some articles one of which was presented by Kevin T. Kelly at the *10. International Conference on Logic, Methodology and Philosophy of Science* in Florence 1995 which Peter Gärdenfors also attended [Kelly et al. 96]. In 1997 Vincent F. Hendricks received his PhD with a dissertation [Hendricks 97] on the learning theoretical consequences of belief revision for ideal and computable agents with Peter Gärdenfors on the thesis-committee.

Gärdenfors' theory and formal learning theory are two different perspectives on scientific inquiry. The AGM-axioms seek to minimize the damage done to the current beliefs in light of new information while formal learning theory seeks to uncover the truth whatever it might be. In turn they both address epistemology and methodology and what better occasion than this to further look at their intersection.

1 Modal Learning: Belief and Knowledge

Formal learning theory has for years remained a rather marginal research programme. Methodologists and philosophers of science alike have had a hard time seeing how for instance Gold's application of formal learning theory to the language acquisition paradigm [Gold 67] had any philosophical relevance. Additionally Osherson et al.'s impressive list of further language acquisition [Osherson, et al. 86] results has generally not been looked upon as being of importance to classical problems in the philosophy of science and epistemology. Under the rubric of the *logic of reliable inquiry* [Kelly 96], Kelly has now formulated a framework in which learning theory plays a significant role in uncovering the structural problems of induction. Furthermore the logic of reliable inquiry can be used to study the reliable, and sometimes unreliable, nature of the hypothetically advocated methodological principles under which "rational" scientific inquiry should proceed. Formal learning theory has however not yet been tuned to cover classical epistemological issues like the definition of knowledge.

Modal learning theory [Hendricks & Pedersen 98a, 98b, 98c, 98d, 99a, 99b] applies the logic of reliable inquiry to study the definition of knowledge. Modal simply refers to the fact that unary modal operators are used to represent epistemic attitudes like knowledge and belief. Given the above premisses we ask the following question:

> *How do the methodological principles aid the believing method in getting to the correct answer?*

Review once more the standard tripartite definition of knowledge; it suggests that some scientist, or scientist applying a method, δ knows a hypothesis h if the following conditions obtain:

Definition 1 *The Standard Definition of Knowledge*

Method δ knows $h \Leftrightarrow$

1. *δ believes h,*
2. *h is true,*
3. *δ is justified in believing h.*

The three components of definition 1 have received unequal attention. The fact that a method is required to believe the hypothesis to which it wants to commit itself has never received any serious challenge. In order to model belief assume that the scientist interacts with the world through data streams: A data stream ε is an $\omega-$sequence of natural numbers, *i. e.,* $\varepsilon \in \omega^\omega$. Hence, a data stream $\varepsilon = (a_0, a_1, a_3 ..., a_n, ...)$ consists of code numbers of evidence; at each stage i in inquiry a_i is the code number of all evidence acquired at this stage. Continue to define a possible world. A possible world is a pair consisting of a data stream ε and a time index n, (ε, n), such that $\varepsilon \in \omega^\omega$ and $n \in \omega$. The set of all possible worlds $\mathcal{W} = \{(\varepsilon, n) \mid \varepsilon \in \omega^\omega, n \in \omega\}$. Let $(\varepsilon \mid n)$ denote the finite initial segment of length n of data stream ε. Furthermore $\omega^{<\omega}$ denotes the set of all finite initial segments of elements in ω. Let $[\varepsilon \mid n]$ denote the set of all infinite data streams that extends $(\varepsilon \mid n)$. Refer to the finite initial segment $(\varepsilon \mid n)$ as the *handle* with *fan* $[\varepsilon \mid n]$. The world-fan is defined as $\widetilde{[\varepsilon \mid n]} = [\varepsilon \mid n] \times \omega$. The background knowledge $[\varepsilon \mid n]_\mathcal{K}$ of accessible possible worlds is defined as the set of all worlds such that $[\varepsilon \mid n]_\mathcal{K} = \left\{ (\widetilde{[\varepsilon \mid n]} \mid \mathcal{K}) \mid \mathcal{K} \subseteq \mathcal{W} \right\}$.

The method that the scientist applies conjectures hypotheses in response to the evidence seen so far. In accordance with standard practice identify hypotheses with sets of possible worlds. The set of all empirical hypotheses $\mathcal{H} = P(\omega^\omega) \times P(\omega)$ such that an empirical hypothesis h is a member of $\mathcal{H}$. Finally a scientific discovery method is a function from finite evidence sequences to hypotheses: $\delta : \omega^{<\omega} \longrightarrow \mathcal{H}$.

Now, define belief in the following way:

Definition 2 *Belief*

> $B_\delta h$ *is* true *at* (ε, n) $(i.e., (\varepsilon, n) \models B_\delta h))$ *iff*
> $\forall n' \geq n, \forall (\tau, n') \in [\varepsilon \mid n]_\mathcal{K} : \delta(\tau \mid n') \subseteq h.$

What belief amounts to is the existence of a time such for all worlds in accordance with the method's background assumptions, δ performs a consistent conjecture entailing hypothesis h even though h is not required to be true.

Given definition 1 knowledge requires belief that h, but satisfaction of condition 1 alone does not suffice for knowledge of h, since the belief in h could be incorrect. This explains the presence of condition 2. Since Plato's *Theatetus* it has been a standard assumption that knowledge implies a condition of correctness. The standard analysis thus suggests, as a necessary condition, that knowledge of h implies that h is correct.

Definition 3 *Correctness*

> *Hypothesis* h *is* correct *in world* (ε, n) $(i.e., (\varepsilon, n) \models h) \Leftrightarrow$
> $[\varepsilon \mid n]_\mathcal{K} \cap h \neq \emptyset.$

Realize that h is considered to be correct in a world (ε, n) if and only if:

- *the hypothesis (or proposition, i. e. set of possible worlds) corresponding to h, is verified by evidence up to n,*
- *possibly the hypothesis will be verified by future evidence in the sense that there exists possible worlds in the proposition corresponding to h which are consistent with existing evidence.*

The conception of correctness presented here is rather weak because it only guarantees possible correctness in the future. Conversely, incorrectness is proportionally strong since a hypothesis is incorrect if it is actually falsified by existing evidence, *i. e.* $[\varepsilon \mid n]_{\mathcal{K}} \subseteq \overline{h}$. [1]

The justification condition has probably received the greatest attention of the three components from Plato to the contemporary theories of knowledge. In definition 1, knowledge requires that the satisfaction of the belief condition 1 is "adequately" connected to the satisfaction of the truth condition 2. Conditions 1 and 2 are jointly insufficient to secure knowledge: Some correct beliefs may be the result of arbitrary insights, accidental inferences, evidence collected under obscured perceptual circumstances etc. Such beliefs should on the standard analysis not count as knowledge since clause 1 and 2 are inadequately connected to each other. This is due to the questionable means or *methods* by which the (*de facto*) correct beliefs have been derived. According to condition 3, if some argument or other justificational structure can be provided that describes why the two conditions are adequately connected, then the scientist may be said to have adequate indication that a known proposition is correct. So condition 3 essentially concerns methodology, *i. e.* the study of the methods by which science arrives at its posited correct answers and hence the methodological principles imposed on the methods in order to get it right. Methodological principles like simplicity, consistency with the evidence, AGM-revision, unification etc. *are important because they prescribe how the method should behave 'rationally' in terms of tracking the truth or getting to the correct answer.*

Assume that whether belief may become knowledge is dependent upon justification. Justification is then exhausted in following truth-tracking methodological prescriptions. Now adopt the following methodological principle:

Definition 4 *Weak Epistemic Soundness (WES)*

$$\delta \textit{ is } \text{weak epistemically sound} \Leftrightarrow \forall (\tau, n') \;\; [[\tau \mid n']_{\mathcal{K}} \cap \delta(\tau \mid n') \neq \emptyset] \,.$$

In other words it is required that δ's conjecture "touches" the fan, or can be made true within the fan of possible worlds, *i. e.* background knowledge. Hence the criterion of weak epistemic soundness ensures that the method is barred from making any wild conjectures outside the background knowledge.

[Lenzen 78] admits given Gettier-examples [Gettier 63] and the troublesome justification condition 3 of definition 1 that:

> "The search for the correct analysis of knowledge, while certainly of extreme importance and interest to *epistemology*, seems not significantly to affect the object of epistemic logic, *i. e.* the question of validity of certain epistemic-logical principles". [Lenzen 78], p. 34.

[1] We settle for the lowest common denominator with respect to the notion of truth or correctness. Correctness here is essentially a mixture of a closed world assumption, Kuhnian paradigmatics, verficationism and empirical adequacy [Van Fraassen 80]. From now on we will often speak of correctness rather than truth.

The contributions of epistemic logic are extremely useful but are however all suffering from the same shortcomings as admitted by Lenzen. How does the logical contributions relate to the standard epistemological tripartite definition of knowledge? How are belief and correctness adequately connected given justification to equal knowledge? Most contributions of epistemic logic define knowledge and belief separately ([Hintikka 62], [Halpern 95]) largely neglecting the delicate connection between the two, besides the fairly obvious entailment property (that knowledge implies belief). So the philosophical shortcomings are partially caused by the formal shortcomings. It seems that an important task of epistemic logic is to provide a logical but also philosophical account of how belief may become knowledge.

A standard way to measure epistemic strength is to determine the modal system corresponding to the introduced epistemic operator. Rational belief is typically taken to satisfy at least *KD4:*

$$KD4 = \left\{ \begin{array}{l} B_\delta(h \Rightarrow l) \Rightarrow (B_\delta h \Rightarrow B_\delta l) \\ B_\delta h \Rightarrow \neg B_\delta \neg h \\ B_\delta h \Rightarrow B_\delta B_\delta h \end{array} \right. \left\| \begin{array}{l} (K) \\ (D) \\ (4). \end{array} \right.$$

In [Hendricks & Pedersen 98c] it was demonstrated that the notion of belief presented in definition 2 satisfies *KD4*. Knowledge on the other hand have had many axiomatizations which all are at least as strong as *S4*. The *S4*-system is obtained by keeping (K) and (4) but replacing (D) with the stronger, very natural axiom given definition 1, (T) $K_\delta A \Rightarrow A$. Again, we need to explain methodologically how belief can "jump" from *KD4* to *S4* and hence become knowledge. The idea is to make the belief forming method obey weak epistemic soundness because then one may aside from *KD4* satisfy $B_\delta A \Rightarrow A$ and insofar get *S4* for free.

Proposition 5 *Weak epistemic soundness, Belief and Truth*

> *If the discovery method satisfies weak epistemic soundness, then* $B_\delta A \Rightarrow A$ *is satisfiable.*

Proof. To see that $B_\delta A$ satisfies $B_\delta A \Rightarrow A$ re-run the definition of belief remembering the criterion of weak epistemic soundness: Then the following conditions obtain:

$\forall n' \geq n, \forall(\tau, n') \in [\varepsilon \mid n]_\mathcal{K}$:

1. $[\tau \mid n']_\mathcal{K} \subseteq [\varepsilon \mid n]_\mathcal{K}$. (by def. of background knowledge)
2. $[\tau \mid n']_\mathcal{K} \cap \delta(\tau \mid n') \neq \emptyset$. (by def. of weak epistemic soundness)
3. $\delta(\tau \mid n') \subseteq h$. (by def. of strong belief in $\mathcal{K}$)

But condition 2 and 3 imply $[\tau \mid n']_\mathcal{K} \cap h \neq \emptyset$. Then in turn $[\tau \mid n']_\mathcal{K} \subseteq [\varepsilon \mid n]_\mathcal{K}$ and $[\tau \mid n']_\mathcal{K} \cap h \neq \emptyset$ implies $[\varepsilon \mid n]_\mathcal{K} \cap h \neq \emptyset$ which is the definition 3 of correctness of h in world (ε, n).

■

2 Believing in Correct Methodology

A few important lessons are taught by proposition 5. It is an interesting result that balances epistemic strength against methodological principles. The proposition reveals something about the nature of belief: If "rational" belief has to respect the available evidence, the methodological principles and the background knowledge then belief eventually reach a point where it collapses into knowledge. And this collapse is due to the methodological behavior of the method, *i .e.* fleshing out the justification condition.

Another point related to the standard definition of knowledge is hitherto supported. The introduction revealed that the justification condition has to do with methodology, *i. e.* how science arrives at its posited truths. And arriving at the correct answer is accomplished by forcing the methods to obey certain "rational" behavioral patterns like consistency with the evidence, unification, conservativeness, and the AGM-axioms, just to mention a few of the most prominent recommendations. Now, a central theme in formal and modal learning theory, to which methodologists have remained fairly indifferent, is to investigate whether "rational" methodological inductive maxims stand in the way of finding correctness or not. Say that a methodological principle is *restrictive* if the correct answer could have been found in a logically reliable way by a method violating it. Then there is nothing particularly rational about the maxim. Inquiry may proceed forever rationally according to the dictates and yet chase its own tail heading nowhere near correctness. For instance it has been shown in [Kelly 94] and [Kelly 96] that Popper's methodology of conjectures and refutations is by all means a sub-optimal strategy in terms of getting to the correct answer and not even a very reliable one. There exists another strategy which is demonstrably complete over a certain set of problems for which Popper's recommendation fails. Furthermore, in [Kelly et al. 96] we showed that the AGM-axioms of belief revision are restrictive if the method starts out with the wrong kind of beliefs. On the positive side, starting out with the correct kind of beliefs the AGM-architecture is complete in the ideal limiting case. But in [Hendricks 97] it was demonstrated how the AGM-axioms are restrictive if the method is to be effective.

Observe first that the belief forming method alone neither should nor could satisfy axiom (T). Nor was it required to reliably be tracking correctness. *Correctness is, however, logically implied by weak epistemic soundness, background knowledge and belief as demonstrated in proposition* 5. Hence, epistemic soundness is a rational prescription because it does not present an impediment to finding correctness - on the contrary.

So what distinguishes knowledge from belief is whether belief is justified and insofar this obtains knowledge and belief become indistinguishable in accordance with knowledge as justified true belief. But the recipe for this result is

$$\boxed{\textit{Belief} + \textit{Methodology} = \textit{Knowledge}}$$

rather than *Truth + Belief + Justification = Knowledge.* In conclusion, Lenzen says that there is a vagueness in notions like "being justified in believing"

[Lenzen 78], p. 28 but epistemic logic does not have to worry significantly about it because *it is not its proper field of research to say what knowledge is* but rather how it behaves once it is defined. We disagree! Epistemic logic is as much committed to tell a reasonable story about what knowledge is as good epistemology is. The vagueness of justification is no more an impediment to epistemic logic than it is to epistemology: Let methodology ensure that you get it right.

References

[Gettier 63] Gettier, E. (1963). "Is Justified True Belief Knowledge?", *Analysis* 23.6.

[Gold 67] Gold, E. M. "Language Identification in the Limit", *Information and Control* **10**: 447-474.

[Goldman 79] Goldman, A. (1986). "What is Justified True Belief", *Justification and Knowledge*, ed. Pappas, G. S. D. Reidel Publishing Company: 1-23.

[Halpern 95] Halpern, J. Y. (1995). "Should Knowledge Entail Belief?", *Journal of Philosophical Logic* **25**: 483-494.

[Hendricks 97] Hendricks, V. F. (1997). *Epistemology, Methodology and Reliability*, Doctoral Thesis. Department of Philosophy, University of Copenhagen.

[Hendricks & Pedersen 98a] Hendricks, V. F. & Pedersen, S. A. (1998). "The Epistemology of Limiting Convergence", draft.

[Hendricks & Pedersen 98b] Hendricks, V. F. & Pedersen, S. A. (1998). "Assessment and Discovery in the Limit Scientific Inquiry", to be presented at *LMPS99*.

[Hendricks & Pedersen 98c] Hendricks, V. F. & Pedersen, S. A. (1998). "Belief + Methodology = Knowledge", submitted for publication in *Erkenntnis*.

[Hendricks & Pedersen 98d] Hendricks, V. F. & Pedersen, S. A. (1998). "Discovery, Knowledge and Reliable Limiting Convergence - the $\mathbf{K_aLC}$-paradigm. Presented at *Model-based Reasoning in Scientific Discovery - MBR'98*, Pavia, Italy, December 17-19.

[Hendricks & Pedersen 99a] Hendricks, V. F. & Pedersen, S. A. (1999). "Scientific Respect in Limiting Scientific Inquiry", submitted for publication in *Synthese*.

[Hendricks & Pedersen 99b] Hendricks, V. F. & Pedersen, S. A. (1999). "Modal Learning in Branching Time", to be presented at *LMPS99*.

[Hintikka 62] Hintikka, J. (1962). *Knowledge and Belief.* Cornell: Cornell University Press.

[Kelly 94] Kelly, K. (1994). "Reliable Methods", *XI International Congress on Logic, Methodology and Philosophy of Science X*. Dordrect: Kluwer: 353 - 381.

[Kelly 96] Kelly, K. (1996). *The Logic of Reliable Inquiry.* New York: Oxford University Press.

[Kelly et al. 96] Kelly, K., O. Schulte and V. F. Hendricks (1996). "Reliable Belief Revision", *Logic, Methodology, and Philosophy of Science X*, Dordrecht: Kluwer.

[Lenzen 78] Lenzen, W. (1978). *Recent Work in Epistemic Logic*, *Acta Philosophica Fennica*, vol. **30**.

[Osherson, et al. 86] Osherson, D., M. Stob & S. Weinstein (1986). *Systems that Learn.* Cambridge: M.I.T. Press.

[Van Fraassen 80] Van Fraassen, B. (1980). *The Scientific Image.* Oxford: Clarendon Press.