Active Epistemics: Methods, Recommendations and Hypotheses to Learn

Vincent F. Hendricks
Department of Philsophy and Science Studies
Roskilde University
Denmark

Workshop on Logic, Rationality and Interaction, Beijing, China. August 5–9, 2007

1 Status

What are we concerned about at the

Workshop on Logic, Rationality and Interaction
Beijing, China. August 5–9 , 2007 ?

- **Dynamics of inquiry**

- Belief revision, dynamic epistemic and doxastic logics, stit-theory, logics of action, intention and doxastic voluntarism—in general, logics and formal models for – using an adequately covering term coined by Parikh – *social software*

- Dynamic inquiry incorporates a multiplicity of 'dynamic' parameters including for instance agent action, interaction, startegies, temporality, ...

- The result: *Multi-modal systems* ...

2 Modal Operator Theory

1. Modal operator theory was developed to study the properties of limiting convergent knowledge by combining epistemic, tense and alethic logic with rudimentary elements from formal learning theory or computational epistemology [Hendricks 01-07]

2. This talk includes

- an outline of modal operator theory and its formal framework,

- a presentation of some key results,

- some new results pertaining to learnability, methodology and the temporal behavior of hypotheses.

2.1 Worlds

1. An evidence stream ε is an ω-sequence of natural numbers, *i . e.*, $\varepsilon \in \omega^\omega$.

2. A possible world has the form (ε, n) such that $\varepsilon \in \omega^\omega$ and $n \in \omega$.

3. The set of all possible worlds $\mathcal{W} = \{(\varepsilon, n) \mid \varepsilon \in \omega^\omega, n \in \omega\}$.

4. $\varepsilon \mid n$ denotes the finite initial segment of evidence stream ε of length n.

5. Define $\omega^{<\omega}$ to be the set of all finite initial segments of elements in ω.

6. Let $(\varepsilon \mid n)$ denote the set of all infinite evidence streams that extends $\varepsilon \mid n$.

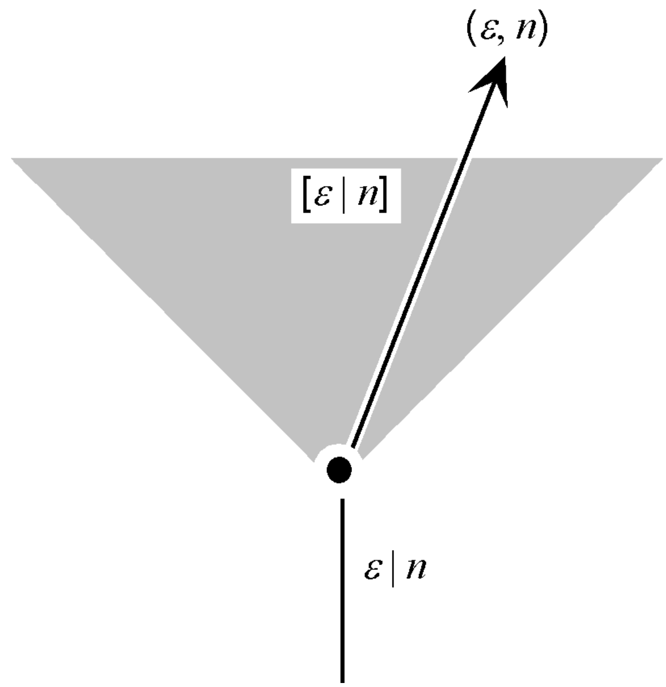
1. 

Figure 1: Refer to the finite initial segment $\varepsilon \mid n$ as the **handle** with **fan** $(\varepsilon \mid n)$. Let the **world-fan** be defined as $[\varepsilon \mid n] = \{(\tau, k) \mid k \in \omega$ and $\tau \mid k = \varepsilon \mid n\}$

2.2 Hypotheses, Methods and Recommendations

Hypotheses will be identified with sets of possible worlds. Define the set of all simple empirical hypotheses

$$\mathcal{H} = P(\omega^\omega \times \omega).$$

An empirical hypothesis h is said to be *true* in world (ε, n) iff

$$(\varepsilon, n) \in h \text{ and } \forall l \in \omega : (\varepsilon, n + l) \in h.$$

Truth requires identification and inclusion of the actual world (ε, n) in the hypothesis for all possible future states of inquiry. Attention is, **for now**, restricted to hypotheses which are stably true like laws of nature (Figure 2).

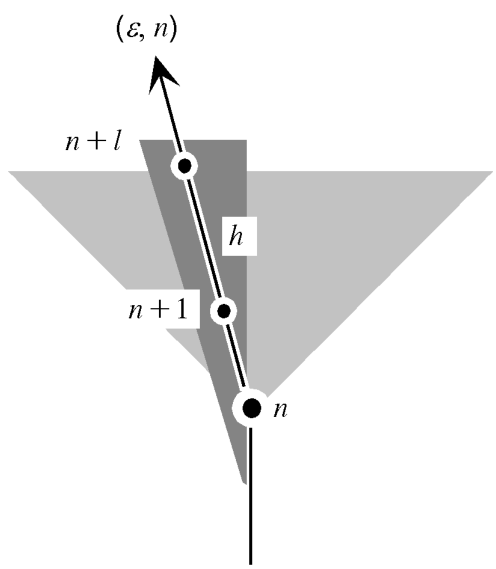


Figure 2: Truth of an empirical hypothesis in a possible world

An inquiry method, in particular a *discovery method*, conjectures hypotheses in response to the evidence received. More specifically, a discovery method δ is a function from finite initial segments of evidence to hypotheses, i.e.

$$\delta : \omega^{<\omega} \longrightarrow \mathcal{H}. \tag{1}$$

The method may be subject to various methodological recommendations, program commands or behavioral patterns. For a well-known example, say that a discovery method δ is *consistent* iff

$$\forall(\tau, n') : \Big[[\tau \mid n'] \cap \delta(\tau \mid n') \neq \emptyset\Big]. \tag{2}$$

Consistency may be strengthened to another recommendation for which the discovery method δ has *perfect memory* iff

$$\textit{if } (\mu, k) \in \delta(\varepsilon \mid n) \textit{ then}$$

$$(\mu \mid n = \varepsilon \mid n) \textit{ and } \forall l \leq k : (\mu, l) \in \delta(\varepsilon \mid n). \quad (3)$$

Furthermore a discovery method δ is said to be *consistently expectant* iff

$$\textit{if } (\mu, k) \in \delta(\varepsilon \mid n) \textit{ then } [k \geq n \textit{ and } (\mu \mid n = \varepsilon \mid n)] \quad (4)$$

and finally, say that a discovery method δ is said to be *infallible* iff iff

$$(\varepsilon, n) \in \delta(\varepsilon \mid n)$$

2.3 Operators

2.3.1 Knowledge

(ε, n) validates $K_\delta h \ iff$

1. $(\varepsilon, n) \in h$ and $\forall l \in \omega$: $(\varepsilon, n + l) \in h$

2. $\forall n' \geq n, \forall(\tau, n') \in [\varepsilon \mid n]$:

   (a) $\delta(\tau \mid n') \subseteq h$,

   (b) $(\tau, n') \in \delta(\tau \mid n')$.

2.3.2 Alethics

A hypothesis is **universally necessary** if and only it is correct in all possible worlds simpliciter:

(ε, n) validates $\Box h$ iff $\forall(\tau, m) \in \mathcal{W}$: (τ, m) validates h.

A hypothesis is said to be **empirically necessary** if and only if it is correct in all the possible worlds admitted by the background knowledge:

(ε, n) validates $\boxminus h$ iff $\forall(\tau, m) \in [\varepsilon \mid n]$: (τ, m) validates h.

A hypothesis is **temporally necessary** if it is correct for all possible state coordinates:

(ε, n) validates $\boxdot\, h$ iff $\forall k \in \omega$: (ε, k) validates h.

2.3.3 Tenses

1. (ε, n) validates Fh *iff* $\exists k > n$: (ε, k) validates h.

2. (ε, n) validates Gh *iff* $\forall k > n$: (ε, k) validates h.

3. (ε, n) validates Ph *iff* $\exists k < n$: (ε, k) validates h.

4. (ε, n) validates Hh *iff* $\forall k < n$: (ε, k) validates h.

3 Multi-Modal Logic

The above set-theoretical characterization of inquiry lends itself to a multi-modal logic. The modal language $\mathcal{L}$ used is that of the regular propositional logic augmented with the new tense, alethic and epistemic modalities.

A model $\mathbb{M} = <\mathcal{W}, \varphi, \delta>$ *consists of:*

1. *A non-empty set of* possible worlds $\mathcal{W}$,

2. *A* denotation function φ :*Proposition Letters* $\longrightarrow P(\mathcal{W})$,*i. e.*, $\varphi(a) \subseteq \mathcal{W}$.

3. *An* inquiry method $\delta : \omega^{<\omega} \longrightarrow P(\mathcal{W})$

The Boolean conditions follow the standard recursive recipe like:

1. $\varphi_{\mathbb{M},(\varepsilon,n)}(\neg A) = 1\ iff\ \varphi_{\mathbb{M},(\varepsilon,n)}(A) = 0$,

2. $\varphi_{\mathbb{M},(\varepsilon,n)}(A \wedge B) = 1\ iff\ both\ \varphi_{\mathbb{M},(\varepsilon,n)}(A) = 1\ and\ \varphi_{\mathbb{M},(\varepsilon,n)}(B) = 1$;
   $otherwise\ \varphi_{\mathbb{M},(\varepsilon,n)}(A \wedge B) = 0$.

and similarly for the remaning connectives.

Truth-conditions for the tense and alethic operators are given by:

1. $\varphi_{\mathbb{M},(\varepsilon,n)}[FA] = 1$ *iff* $\exists n' > n : (\varepsilon, n') \vDash_{\mathbb{M}} A$.

2. $\varphi_{\mathbb{M},(\varepsilon,n)}[GA] = 1$ *iff* $\forall n' > n : (\varepsilon, n') \vDash_{\mathbb{M}} A$.

3. $\varphi_{\mathbb{M},(\varepsilon,n)}[PA] = 1$ *iff* $\exists n' < n : (\varepsilon, n') \vDash_{\mathbb{M}} A$.

4. $\varphi_{\mathbb{M},(\varepsilon,n)}[HA] = 1$ *iff* $\forall n' < n : (\varepsilon, n') \vDash_{\mathbb{M}} A$.

4 Some Results

Two questions:

1. *Which epistemic axioms can be validated by an epistemic operator based on the definition of limiting convergent knowledge for discovery methods?*

2. *Does the validity of the various epistemic axioms relative to the method depend upon enforcing methodological recommendations?*

Theorem

If knowledge is defined as limiting convergence, then knowledge validates **S4** *iff the discovery method has consistent expectations.*

Thus, methodological recommendations may be classified as to whether they are

- *boosting* in the sense that a methodological recommendation is *conducive* to validating epistemic axioms and systems,

- *debilitative* in the sense that the methodological recommendation is an *impediment* to validating epistemic axioms and systems, or

- *neutral* if it is neither boosting or debilitative.

5 Partitioning the Hypotheses Space

The above results are all based on the idea that the hypotheses of interest are absolute time invariant. Modal operator theory is flexible enough

- to study dynamics with respect to the possible temporal fluctuations of the hypotheses,

- to study the learnability of such hypotheses.

5.1 Stable Hypotheses

Stable hypotheses have a partially a Kantian motivation. An apodeictic hypothesis expresses logical necessity, *i. e.* insofar the hypothesis is true it is true at all possible times (Figure 1).

Definition 1 *Stable Hypotheses*

h *is a* stable *hypothesis* iff

1. $h \subseteq \mathcal{W}$,

2. $(\varepsilon, n) \in h \Rightarrow \forall m \in \omega : (\varepsilon, m) \in h$.

Let $\varphi(p) = h$. Then realize that by definition of temporal necessity, a stable hypothesis may also be expressed model thoretically as

$$\mathbb{M} \vDash p \Rightarrow \boxdot p$$

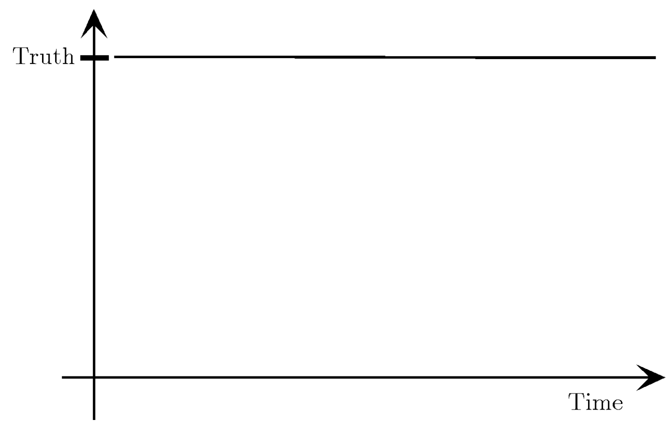


Figure 3: The constant truth-value of a stable hypothesis.

since a hypothesis is temporally necessary iff $\forall k \in \omega$: $(\varepsilon, k) \vDash h$ which is equivalent to the definition of a stable hypothesis.

5.2 Eventually Stable Hypotheses

Peirce held the view that science and its theories, *i. e.* hypotheses, may converge to the truth in the limit. This means that the hypotheses of science are free to oscillate in truth-value all they care as long as they eventually stabilize to the truth and stick with it forever after no matter how the world will turn.

Definition 2 *Eventually Stable Hypotheses*

h *is an* eventually stable *hypothesis* iff

1. $h \subseteq \mathcal{W}$,

2. $(\varepsilon, n) \in h \Rightarrow \exists k \forall m \geq k \in \omega : (\varepsilon, m) \in h$.

Given the tense-logical operators characterize the eventually stable hypotheses accordingly:

$$\mathbb{M} \models p \Rightarrow \neg \boxdot \neg Gp.$$

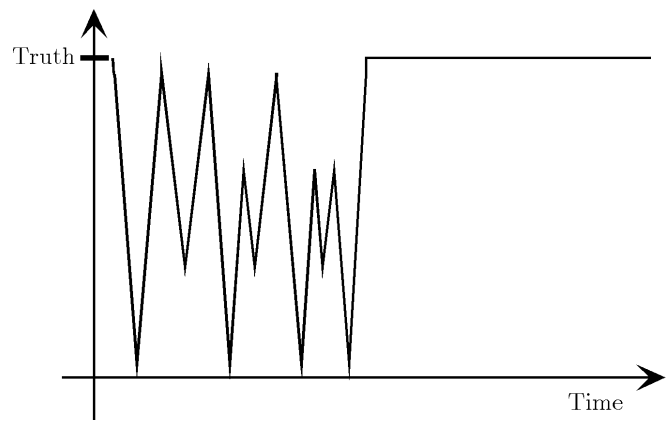


Figure 4: The eventually stabilizing truth-value.

5.3 Oscillating Hypotheses

Reading Quine radically enough suggests that not even the most basic hypotheses in the web of belief are eternally true. In this worst case define oscillating hypotheses accordingly:

Definition 3 *Oscillating Hypotheses*

h *is an* oscillating *hypothesis* iff

$(\varepsilon, k) \in h \Rightarrow \forall n \exists n_1 \exists n_2 : n_1 \geq n \wedge n_2 \geq n \wedge (\varepsilon, n_1) \notin h \wedge (\varepsilon, n_2) \in h.$

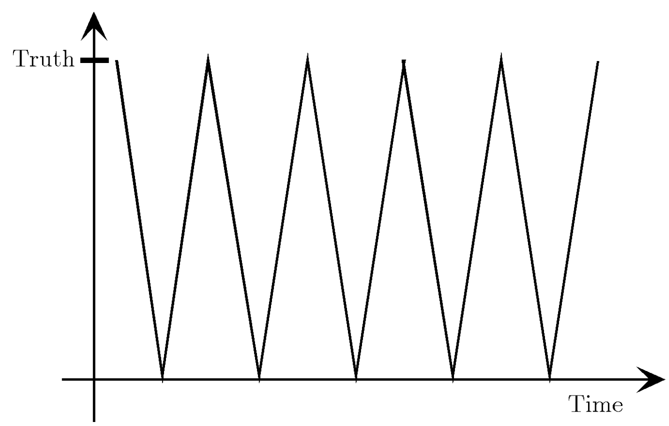


Figure 5: The forever oscillating truth-value.

The tense-logical description of oscillating hypotheses amounts to

$$\mathbb{M} \vDash p \Rightarrow \boxdot(Fp \wedge F\neg p).$$

5.4 Initially Stable Hypotheses

Kuhn was of the opinion that the truth-value of a scientific hypothesis is dependent upon the current paradigm entertained. A hypotheses true in one paradigm may come out false in another paradigm and yet turn true again in another succeeding the former.

Definition 4 *Initially Stable Hypotheses*

h *is an* initially stable *hypothesis iff*

1. $h \subseteq \mathcal{W}$,

2. $(\varepsilon, n) \in h \Rightarrow \exists k \forall m \leq k \in \omega : (\varepsilon, m) \in h$.

In other words, the initially stable hypotheses have the following tense-logical counterpart:

$$\mathbb{M} \vDash p \Rightarrow \neg \boxdot \neg Hp.$$

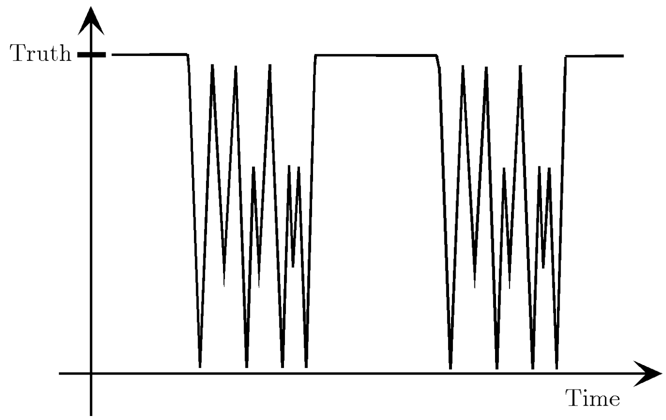


Figure 6: The stabilized truth-value up to *n* and then possibly forever stabilizing and oscillating truth-value.

5.5 Absurd Hypotheses

Then there are finally hypotheses in science that can scare the scientists more than anything. These are the ones that either look like they are eternally true, or look like they forever oscillate in truth-value. But the truth to the matter is that they will eventually drop to falsity and stick with it. Hence either the scientist may come to think that the hypothesis is true, or he will settle for the conclusion that it forever jumps between truth and falsity; *either way he fails*. Refer to these beasts as the *absurd* hypotheses:

Definition 5 *Absurd Hypotheses*

h *is an* absurd *hypothesis iff*

1. $h \subseteq \mathcal{W}$,

2. $(\varepsilon, n) \in h \Rightarrow \exists k \forall m \geq k \in \omega : (\varepsilon, m) \notin h$.

Absurd hypotheses have following tense-logical description

$$\mathbb{M} \models p \Rightarrow FG\neg p.$$

The absurd hypotheses are the only ones which do not contain an infinite data stream. However, any initial segment of such a hypothesis can be viewed as compatible with both a stable and oscillating hypotheses (Figure 7).

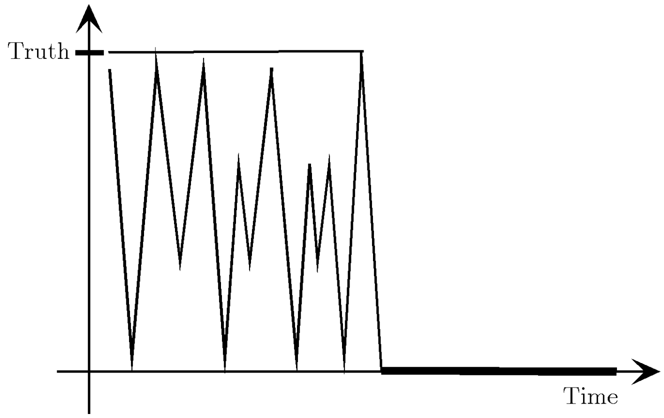


Figure 7: The eventually dropping truth-value of an absurd hypothesis.

6 Learning Hypotheses

Lemma 1 *Hypothesis Relations*

1. *stable hypotheses* $\Rightarrow$ *eventually stable hypotheses (or* $\boxdot h \Rightarrow \neg \boxdot \neg Gp$*).*

2. *stable hypotheses* $\Rightarrow$ *initially stable hypotheses (or* $\boxdot h \Rightarrow \neg \boxdot \neg Hp$*).*

3. *oscillating hypotheses* $\cap$ *initially stable hypotheses* $\neq \emptyset$.

4. *absurd hypotheses* $\cap$ *initially stable hypotheses* $\neq \emptyset$.

5. *stable hypotheses* $\cap$ *oscillating hypotheses* $= \emptyset$.

6. *eventually stable hypotheses* $\cap$ *oscillating hypotheses* $= \emptyset$.

6.1 Unidentifiability Results

Some hypotheses create grave problems for learners and their methodological recommendations:

Proposition 2 *Unidentifiability of Oscillating Hypotheses*

An oscillating hypothesis is not limiting identifiable by an infallible discovery method δ.

Corollary 3 *Unidentifiability of Initially Stable Hypotheses*

There exists initially stable hypotheses which are not limiting identifiable by an infallible discovery method δ.

Corollary 4 *Unidentifiability of Absurd hypotheses*

An absurd hypothesis is not limiting identifiable by an infallible discovery method δ.

6.2 Identifiability Results

The above propositions are based on the assumption that the discovery method is infallible. However if we choose to drop infallibility and settle for perfect memory it is possible to construct discovery methods which may identify some oscillating, initially stable and absurd hypotheses:

Proposition 5 *Consistency, Oscillating, Initially Stable and Absurd Hypotheses*

> *There exist oscillating, initially stable and absurd hypotheses which are limiting identifiable by a discovery method δ with perfect memory.*

The above sample of propositions is interesting because they balance methodological strength in learnability against the temporal complexity of the hypotheses classes.

Both infallibility and perfect memory in general are *truth-conducive methodological prescriptions for discovery methods in search of the truth of absolute time invariant empirical hypotheses*s.

This extends to some of the oscillating, initially stable and absurd hypotheses

On the other hand the infallibility prescription actually stands in the way of finding the truth when it can be found by a method furnished with the different and to some extent weaker criterion of perfect memory.

Given the definition of knowledge as limiting convergence for a discovery method one may finally ask what a discovery method based on limiting convergence can know rather than just identify.

Proposition 6 *Incompleteness over Oscillating, Initially Stable and Absurd Hypotheses*

> *Knowledge is identificationally incomplete over oscillating, initially stable and absurd hypotheses.*

7 Conclusion

With the temporal indices on the hypotheses, the epistemic game changes:

- Some hypotheses are simply beyond identificational reach even for a very powerful and truth-conducive discovery method.

- To identify and know is then relative not only to background assumptions, the truth and methodological recommendations but also to the possibly tensed nature of the hypotheses under investigation:

  An inquiry method or agent can fail to know a hypothesis not only due to either silly or too demanding methodological recommendations getting the method off-track but also due to the complex temporal nature of the hypotheses.