Assessing Theories, Bayes Style

Franz Huber

Philosophy, Probability, and Modeling

University of Konstanz, Germany

Alexander von Humboldt
Stiftung/Foundation

Zukunfts-
investitions-
programm
zip

Contents

1 A Theory of Theory Assessment

2 A Good Theory of Theory Assessment?

3 The Logic of Theory Assessment

Part 1

A Theory of Theory Assessment

The Problem Addressed ...

... is this:

> ... the main epistemic problem concerning science ... is the explication of how we compare and evaluate theories in the light of the available evidence ... (van Fraassen 83)

In other words: What is a good theory?

The *Assessment* Problem

- Given:

  theory H, evidence E (background B)

- Question:

  What is the „*value*“ of H in view of E?

- Answer:

  A set A of functions a s.t. $a(H, E)$ measures the value of H in view of E.

A *Very* Brief History of Confirmation

What comes closest to ass.ment: confirmation.

Two approaches in the 20th century:

– HD: $E \mid\!\sim_{\mathrm{HD}} H$ iff $H \vdash E$

– IL: $E \mid\!\sim_{\mathrm{IL}}^{r} H$ iff $Pr(H \mid E) = r$

$[Pr(H \mid E) > Pr(H)]$

Conflicting concepts of confirmation

HD *in*creases with the logical strength of H:

$$E \mid\!\sim_{\text{HD}} H,\ H' \vdash H \quad \Rightarrow \quad E \mid\!\sim_{\text{HD}} H'$$

$\Rightarrow$ logically strong (*informative*) theories

IL *de*creases with the logical strength of H:

$$E \mid\!\sim_{\text{IL}}{}^{r}\ H,\ H \vdash H' \quad \Rightarrow \quad E \mid\!\sim_{\text{IL}}{}^{s}\ H',\ s \geq r$$

$\Rightarrow$ logically weak (*plausible*) theories

Good theories are *both* true *and* informative!

$\Rightarrow$ Make the two virtues explicit in assessing a given theory H by the available data E:

- assess the *plausibility* of H in view of E: p
- assess the *informativeness* of H (rel. to E): i
- combine p and i appropriately: $a(p, i)$

Example

Your dentist tells you that

- at least one of your teeth has some desease
- the lower right eyetooth has some desease
- the lower right eyetooth has caries
- the lower right eyetooth has caries, and this is so because of your heavy consumption of chocolate icecream

How to Combine

1 f increasing in (i and) p

2 Continuity:

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, i, i', p, p' \in [0, 1]:$$

$$i > i' + \varepsilon,\ p > p' - \delta \ \Rightarrow\ f(i, p) > f(i', p')$$

3 Demarcation:

$$f(1, 0) = f(0, 1) = \beta = 0$$

The Message of Part 1

- A good theory is both *true* and *informative*.
- Truth and informativeness are *conflicting*.
- In assessing a theory by the available data one should *weigh* between these two aspects in such a way that any surplus in informativeness succeeds, if only the difference in plausibility is small enough.

Part 2

A Good Theory of Theory Assessment?

–

Confirmation, Assessment, and the Point of All This

Incremental Confirmation

Confirmation = positive statistical relevance

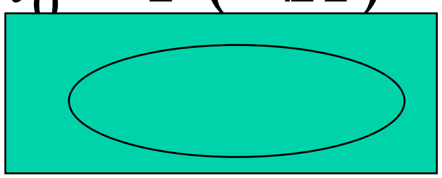
Degree of confirmation = degree to which E increases the probability of H

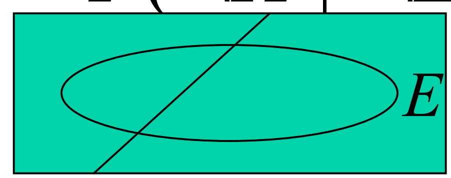
Assessing Theories, Bayes Style

Plausibility: $p = P(H \mid E)$

Informativeness: $i = i_\alpha = \alpha \cdot i_0 + (1 - \alpha) \cdot i_1$

$i_0 = P(\neg H)$ $\alpha \in [0, 1]$ $i_1 = P(\neg H \mid \neg E)$





Assessment (*confirmation*):

$a = f(i_\alpha, p)$, f continuously demarcating

Examples and Comparisons

add:	$a_0 = i_0 + p - 1$	$a_1 = i_1 + p - 1$
exp. inf.:	$e_0 = E(i_0)$	$f_1 = E(i_1)$
incr. conf.:	$d = a_0 = e_0$	$s = a_1 = e_1$
	$P(H\|E) - P(H)$	$P(H\|E) - P(H\|\neg E)$
	Earman	Joyce/Christensen

What Is the Point?

Why should one stick to theories given high values rather than to any other theories?

Traditional answer: Science seeks truth, and confirmation takes you to true theories.

21st century answer: Science seeks informative truth, and assessment takes you to the most informative true theory.

A Limit Theorem

Every continuously demarcating f reveals the true assessment structure in almost every world w when presented separating data $E_m^{\ w}$.

$$w \models H_1 \& H_2,\ H_1 \vdash H_2 \not\vdash \mathbf{T} \quad \Rightarrow \quad \exists n \forall m \geq n:$$

$$f(H_1, E_m^{\ w}) \geq f(H_2, E_m^{\ w}) > 0 = f(\mathbf{T}, E_m^{\ w})$$

…

$$w \models \neg H_1 \& \neg H_2,\ \mathbf{F} \not\dashv H_1 \vdash H_2 \ \Rightarrow\ \exists n \forall m \geq n:$$

$$f(\mathbf{F}, E_m{}^w) = 0 > f(H_1, E_m{}^w) \geq f(H_2, E_m{}^w)$$

w

informative
contingently true in w
uninformative
logically determined
informative
contingently false in w
uninformative

$\beta = 0$

The Message of Part 2

Theory assessment can answer the question why one should accept theories given high assessment values.

$\Rightarrow$ Theory assessment almost surely takes one to the most informative among all true theories, provided the data are separating.

Part 3

The Logic of Theory Assessment

–

Is There Such a Thing?

Hempel´s Conditions of Adequacy

1	$E \mid\!- H \quad \Rightarrow \quad E \mid\!\sim H$	E
2	$\{H_i : E \mid\!\sim H_i\} \mid\!- H \quad \Rightarrow \quad E \mid\!\sim H$	GC
3	E cons $\quad \Rightarrow \quad \{E\} \cup \{H_i : E \mid\!\sim H_i\}$ cons	Co
4	$E \mid\!\sim H \;\&\; H' \mid\!- H \quad \Rightarrow \quad E \mid\!\sim H'$	CC

Hempel (1945): 1-4 $\quad \Rightarrow \quad X \mid\!\sim Y$ Triv

Carnap´s Analysis (1962, §87)

Hempel was mixing up abs. and incr. conf.

1 incremental confirmation (with proviso)

2-3 absolute confirmation

4 ???

Question: What concept of confirmation did Hempel reject in his *Studies*?

The Logic of Confirmation

1945: Hempel´s Conditions

1950/62: Carnap´s Analysis

1995: Flach: $\exists$ two logics of induction

1996: Zwirn/Zwirn: $\neg\exists$ unified logic of conf.

2000: Milne: $\exists$ ``logic´´ of confirmation, but it does not deserve to be called *a logic*

Hempel Vindicated

The driving force behind 1-4 is the insight that *a good theory is both true and informative*.

1-3 confirmation aiming *truth*
(absolute confirmation *without* proviso)

4 confirmation aiming at *informativeness*
(HD-confirmation)

The Logic of Theory Assessment

Hempel also realized that these two concepts were *conflicting*, and so he gave up 4.

However, one can have the cake and eat it!

There is a logic of theory assessment (conf.) that accounts for these conflicting aspects.

According to this logic a theory is acceptable iff it is suff. informative and suff. plausible.

Carnap´s Analysis Again

Extended regular positive relevance,

$$\perp_P^{+} \cup \{<\mathbf{T},\mathbf{T}>, <\mathbf{F},\mathbf{F}>\}, \text{ P regular}$$

satisfies all principles but, not every property of positive relevance is derivable – which is as it should be, for theory assessment (conf.) is *not symmetric*.

The Message of Part 3

There is a (genuinely non-monotonic) logic of theory assessment according to which a theory is acceptable given the data iff it is sufficiently informative and suff. plausible.

Ext. reg. pos. rel. satisfies all principles, but not every property of pos. rel. is derivable.

The End

Details for Discussion: Notation

Made precise with *ranking theory* (Spohn 90).

L ... set of wffs closed under $\neg, \wedge$

Mod_L ... set of all models for the language L

$\mathrm{M}(\alpha) \subseteq \mathrm{Mod}_L$ ... set of all models of $\alpha \in L$

W ... any set of possibilities (e.g. Mod_L)

A_W ... field over W (contains $\emptyset$, closed under complem. and finite inters.)

Ranking Functions

$\kappa: W \to \mathbf{N} \cup \{\infty\}$ is a *ranking fct*. iff $\kappa^{-1}(0) \neq \emptyset$.

$\kappa^{+}: \mathrm{A}_{W} \to \mathbf{N} \cup \{\infty\}$ is a *general ranking fct*. iff

1 $\qquad \kappa^{+}(W) = 0,\ \kappa^{+}(\emptyset) = \infty$

2 $\qquad \kappa^{+}(B \cup C) = \min\{\kappa^{+}(B), \kappa^{+}(C)\}$

3 $\qquad \kappa^{+}(B \mid A) := \kappa^{+}(A \cap B) - \kappa^{+}(A)$

Assessment Models

Each rf κ induces a unique general rf κ^+:

$$\kappa^+(A) = \min\{\kappa(w) : w \in A\},\ \ A \in \mathrm{A}_W.$$

$<\mathrm{A}_W, \kappa>$ is a *rankth. assessment model for L*

iff $\quad W = \mathrm{Mod}_L$, $\mathrm{M}(\alpha) \in \mathrm{A}_W$ for $\alpha \in L$,

κ is a rf s.t. $\kappa^+(A) < \infty$ for $A \neq \emptyset$, $A \in \mathrm{A}_W$.

Consequence Relations

The *consequence rel.* $|\!\sim_\kappa \subseteq L \times L$ defined by the rankth. assessment model $<A_W, \kappa>$ for L:

$$\alpha \,|\!\sim_\kappa \beta$$

iff

$$\kappa(M(\beta) \mid M(\alpha)) \;\leq\; \kappa(M(\neg\beta) \mid M(\alpha)) \quad \&$$

$$\kappa(M(\neg\beta) \mid M(\neg\alpha)) \;\leq\; \kappa(M(\beta) \mid M(\neg\alpha)),$$

where at least one inequality is strict!

Syntax: Assessment Relations 1-5

$|\!\sim\ \subseteq L\times L$ is an *assessment relation over L* iff

$\alpha \mid\!\sim \alpha$	Refl*
$\alpha \mid\!\sim \beta,\ \alpha \dashv\vdash \gamma \quad\Rightarrow\quad \gamma \mid\!\sim \beta$	L L-Equ*
$\alpha \mid\!\sim \beta,\ \beta \dashv\vdash \gamma \quad\Rightarrow\quad \alpha \mid\!\sim \gamma$	R L-Equ*
$\alpha \mid\!\sim \beta \quad\Rightarrow\quad \alpha \mid\!/\!\!\sim \neg\beta$	Selectivity*
$\alpha \mid\!\sim \beta \quad\Rightarrow\quad \alpha \mid\!\sim \alpha\wedge\beta$	Weak Comp*

6-10

$\alpha \mid\!\sim \beta \quad \Rightarrow \quad \neg\alpha \mid\!\sim \neg\beta$ Lo&Li

$\mid\!\not- \alpha\vee\beta \quad \Rightarrow \quad \alpha\vee\beta \mid\!\sim \alpha$ or $\alpha\vee\beta \mid\!\sim \beta$

Either-Or

$\alpha\vee\beta \mid\!\not\sim \alpha,\ \mid\!\not- \alpha\vee\beta \quad \Rightarrow \quad \mathbf{T} \mid\!\sim \neg\alpha$ Neg 1

$\mathbf{T} \mid\!\sim \alpha,\ \alpha \mid\!- \beta \quad \Rightarrow \quad \mathbf{T} \mid\!\sim \beta$ Up

$\mathbf{F} \mid\!\sim \alpha,\ \alpha\vee\beta \mid\!\sim \alpha \quad \Rightarrow \quad \mathbf{F} \mid\!\sim \beta$ Down

11-ω

$\alpha \mid\!\sim \alpha\wedge\beta,\ \alpha \mid\!\sim \alpha\vee\beta \quad\Rightarrow\quad \alpha \mid\!\not\sim \neg\beta$ I

$\alpha \mid\!\not\sim \alpha\wedge\neg\beta,\ \alpha \mid\!\sim \alpha\vee\beta,\ \alpha\ \text{cont} \quad\Rightarrow\quad \alpha \mid\!\sim \beta$ II

$\alpha\vee\beta \mid\!\sim \alpha,\ \beta\vee\gamma \mid\!\sim \beta,\ \not\vdash \alpha\vee\gamma \quad\Rightarrow\quad \alpha\vee\gamma \mid\!\sim \alpha$

$\alpha\vee\beta \mid\!\sim \alpha,\ \beta\vee\gamma \mid\!\sim \beta,\ \vdash \alpha\vee\gamma \quad\Rightarrow\quad \alpha\vee\gamma \mid\!\not\sim \neg\alpha$

quasi Nr 21, suppl. Nr 21

$\alpha_i\vee\alpha_{i+1} \mid\!\sim \alpha_{i+1},\ \not\vdash \alpha_i\vee\alpha_j \quad\Rightarrow$

$\Rightarrow\quad \exists n\forall m\geq n\text{: } \alpha_m\vee\alpha_{m+1} \mid\!\sim \alpha_m$ Min

A Representation Result

The consequence relation $|\!\sim_\kappa$ defined by any rankth. assessment model $\langle A_W, \kappa\rangle$ for L is an assessment relation over L.

For every assessment rel. $|\!\sim$ over L there is a rankth. ass. model $\langle A_W, \kappa\rangle$ for L s.t. $|\!\sim\; = \;|\!\sim_\kappa$.

Proof: Extension theorem for rf (Compactness).

Further Principles

$\mid\!\not\!- \neg\alpha \quad \Rightarrow \quad \alpha \mid\!\not\!\sim \mathbf{F}$ Cons*

$\mid\!\not\!- \alpha \quad \Rightarrow \quad \alpha \mid\!\not\!\sim \mathbf{T}$ Info

$\alpha \mid\!\sim \beta \rightarrow \gamma,\ \alpha \mid\!\sim \beta \quad \Rightarrow \quad \alpha \mid\!\sim \gamma$ MPC

$\alpha \mid\!\sim \beta,\ \alpha \mid\!\sim \gamma \quad \Rightarrow \quad \alpha \mid\!\sim \beta\wedge\gamma$ or $\alpha \mid\!\sim \beta\vee\gamma$

quasi Comp

$\alpha\wedge\beta \mid\!\sim \gamma,\ \alpha\wedge\neg\beta \mid\!\sim \gamma \quad \Rightarrow \quad \alpha \mid\!\sim \gamma$ PrC, D

Non-Principles

$\alpha \mid\!- \beta \quad \Rightarrow \quad \alpha \mid\!\sim \beta$	E, Supracl*
$\beta \mid\!- \alpha \quad \Rightarrow \quad \alpha \mid\!\sim \beta$	Conversion
$\alpha \mid\!\sim \beta, \gamma \mid\!- \alpha \quad \Rightarrow \quad \gamma \mid\!\sim \beta$	Left Mon
$\alpha \mid\!\sim \beta, \beta \mid\!- \neg\gamma \quad \Rightarrow \quad \alpha \mid\!/\!\sim \gamma$	Strong Sel
$\alpha\wedge\beta \mid\!\sim \gamma, \alpha \mid\!\sim \beta \quad \Rightarrow \quad \alpha \mid\!\sim \gamma$	Cut
$\alpha \mid\!\sim \gamma, \alpha \mid\!\sim \beta \quad \Rightarrow \quad \alpha\wedge\beta \mid\!\sim \gamma$	Cautious M

Assessment *vs*. Non-mon. Cons. Rel.s

- Assessment relations are *genuinely* non-monotonic: they invalidate Left and Right Monoton(icit)y:

$$\alpha \mid\!\sim \beta,\ \beta \vdash \gamma \quad \Rightarrow \quad \alpha \mid\!\sim \gamma \qquad\qquad \text{Right Mon}$$

Why Do We Infer Non-Monotonically?

- Why are we content with a standard weaker than truth-preservation in *all* worlds (e.g. truth-preservation in all *normal* worlds)?

$\Rightarrow$ The risk of being led to a false conclusion is the price we are willing to pay for arriving at informative conclusions.