Comparing the Theories

Hannes Leitgeb

LMU Munich

February 2012

Comparing the Theories

Hanti & Kevin's theory and my own theory have *a lot* in common:

- We do not eliminate belief (nor subjective probability, of course).
- We share the formal background framework, the syntactic format, and the same amount of idealization.
- We emphasize the role of conditional belief/acceptance.
- We share a lot of "logical structure"
  (probability axioms, preferential logic).
- We rely on certain contextual parameters (thresholds, partitions).
- Both of our theories have lots of applications and allow for alternative interpretations.

Indeed, it is fair to say that our theories belong to the same *family*.

But of course there are also differences which concern the following issues:

1. Reductionism
2. Commutativity with Conditionalization
3. Rational Monotonicity
4. High Probability Constraints
5. Contextual Parameters

Reductionism

Hanti & Kevin presuppose a reductionist account of belief:

- Belief is *reducible* to probability: B_P, $*_P$

Reductionism

Hanti & Kevin presuppose a reductionist account of belief:

- Belief is *reducible* to probability: B_P, $*_P$

In contrast, my theory is not necessarily reductionist:

- It imposes *constraints* on pairs $\langle P, Bel \rangle$.

Reductionism

Hanti & Kevin presuppose a reductionist account of belief:

- Belief is *reducible* to probability: B_P, $*_P$

In contrast, my theory is not necessarily reductionist:

- It imposes *constraints* on pairs $\langle P, Bel \rangle$.

The theory becomes reductionist only if one adds a maximality or completeness axiom (just like Hilbert did in geometry):

- Given P: Belief is the *maximal* Bel', such that $\langle P, Bel' \rangle$ satisfies the constraints from before.

That is what I do in my "Reducing Belief Simpliciter to Degrees of Belief", and the rationale was to satisfy as many instances of the "$\leftarrow$" of the Lockean thesis (for a given threshold r *independent* of P).

Commutativity with Conditionalization

It is obvious to see that for a theory such as mine—*if maximality or completeness is presupposed*—only half of the commutativity diagram for conditional belief and conditionalization is satisfied:

- If $Bel^r_P(Y|X)$, then $Bel^r_{P(.|X)}(Y)$.

  But not necessarily *vice versa*.

  (Except for: if all worlds of same rank have the same probability.)

Commutativity with Conditionalization

It is obvious to see that for a theory such as mine—*if maximality or completeness is presupposed*—only half of the commutativity diagram for conditional belief and conditionalization is satisfied:

- If $Bel^r_P(Y|X)$, then $Bel^r_{P(.|X)}(Y)$.

  But not necessarily *vice versa*.

  (Except for: if all worlds of same rank have the same probability.)

In fact, Hanti & Kevin are able to prove this on one much more general grounds.

Isn't that worrisome?

Commutativity with Conditionalization

It is obvious to see that for a theory such as mine—*if maximality or completeness is presupposed*—only half of the commutativity diagram for conditional belief and conditionalization is satisfied:

- If $Bel^r_P(Y|X)$, then $Bel^r_{P(.|X)}(Y)$.

  But not necessarily *vice versa*.

  (Except for: if all worlds of same rank have the same probability.)

In fact, Hanti & Kevin are able to prove this on one much more general grounds.

Isn't that worrisome?

Not so sure—e.g., drop Maximality/Completeness, and things are just fine!

(Maximality/Completeness is not mandatory for me, since one gets the *full* Lockean Thesis with *P*-sensitive threshold anyway.)

Fix an "initial" probability measure P and update by a stream of evidence:

$$P \mapsto P_{E_1} \mapsto [P_{E_1}]_{E_2} \mapsto [[P_{E_1}]_{E_2}]_{E_3} \mapsto \cdots$$

Fix an "initial" probability measure P and update by a stream of evidence:

$$P \mapsto P_{E_1} \mapsto [P_{E_1}]_{E_2} \mapsto [[P_{E_1}]_{E_2}]_{E_3} \mapsto \cdots$$

Let the initial P determine *Bel* (that is, a particular sphere system of P-stable^r sets). And revise *Bel* iteratively, by the same stream of evidence:

$$Bel \mapsto Bel * E_1 \mapsto [Bel * E_1] * E_2 \mapsto [[Bel * E_1] * E_2] * E_3 \mapsto \cdots$$

Fix an “initial” probability measure P and update by a stream of evidence:

$$P \mapsto P_{E_1} \mapsto [P_{E_1}]_{E_2} \mapsto [[P_{E_1}]_{E_2}]_{E_3} \mapsto \cdots$$

Let the initial P determine *Bel* (that is, a particular sphere system of P-stable^r sets). And revise *Bel* iteratively, by the same stream of evidence:

$$Bel \mapsto Bel * E_1 \mapsto [Bel * E_1] * E_2 \mapsto [[Bel * E_1] * E_2] * E_3 \mapsto \cdots$$

Then for consistent $E_1 \cap E_2 \cap E_3 \cap \ldots$, and worlds excluded by evidence being excluded from all spheres,

$$\langle P, Bel \rangle, \ \langle P_{E_1}, Bel * E_1 \rangle, \ \langle [P_{E_1}]_{E_2}, [Bel * E_1] * E_2 \rangle, \ldots$$

satisfy all of my postulates (other than Maximality/Completeness)

Fix an "initial" probability measure P and update by a stream of evidence:

$$P \mapsto P_{E_1} \mapsto [P_{E_1}]_{E_2} \mapsto [[P_{E_1}]_{E_2}]_{E_3} \mapsto \cdots$$

Let the initial P determine *Bel* (that is, a particular sphere system of P-stable^r sets). And revise *Bel* iteratively, by the same stream of evidence:

$$Bel \mapsto Bel * E_1 \mapsto [Bel * E_1] * E_2 \mapsto [[Bel * E_1] * E_2] * E_3 \mapsto \cdots$$

Then for consistent $E_1 \cap E_2 \cap E_3 \cap \ldots$, and worlds excluded by evidence being excluded from all spheres,

$$\langle P, Bel\rangle,\ \langle P_{E_1}, Bel * E_1\rangle,\ \langle [P_{E_1}]_{E_2}, [Bel * E_1] * E_2\rangle, \ldots$$

satisfy all of my postulates (other than Maximality/Completeness), yet it holds:

$$Bel(Y|E_1) \text{ iff } [Bel * E_1](Y),\ \ [Bel * E_1](Y|E_2) \text{ iff } [[Bel * E_1] * E_2](Y), \ldots$$

And each belief set is determined by P and $E_1, E_2, E_3, \ldots$!

Rational Monotonicity (RM)

Hanti & Kevin do not have RM as a logical rule:

$$\frac{Bel(Z|X),\ \neg Bel(\neg Y|X)}{Bel(Z\,|\,X \wedge Y)}$$

Rational Monotonicity (RM)

Hanti & Kevin do not have RM as a logical rule:

$$\frac{Bel(Z|X),\ \neg Bel(\neg Y|X)}{Bel(Z\,|\,X \wedge Y)}$$

They like to give Gettier-like examples against RM, which I never find quite convincing :-)

Rational Monotonicity (RM)

Hanti & Kevin do not have RM as a logical rule:

$$\frac{Bel(Z|X),\ \neg Bel(\neg Y|X)}{Bel(Z\,|\,X \wedge Y)}$$

They like to give Gettier-like examples against RM, which I never find quite convincing :-)

- RM is valid in belief revision (AGM), nonmonotonic reasoning (Lehmann & Magidor), and in the logic of counterfactuals (Lewis, Stalnaker).

Rational Monotonicity (RM)

Hanti & Kevin do not have RM as a logical rule:

$$\frac{Bel(Z|X),\ \neg Bel(\neg Y|X)}{Bel(Z\,|\,X \wedge Y)}$$

They like to give Gettier-like examples against RM, which I never find quite convincing :-)

- RM is valid in belief revision (AGM), nonmonotonic reasoning (Lehmann & Magidor), and in the logic of counterfactuals (Lewis, Stalnaker).
- Semantically, RM corresponds to worlds being *totally pre-ordered*:

  Total pre-orders (preference orders) are not just presupposed in belief revision, nonmonotonic reasoning, and for counterfactuals, but also in decision theory, social choice, Popper functions,. . . .

Rational Monotonicity (RM)

Hanti & Kevin do not have RM as a logical rule:

$$\frac{Bel(Z|X),\ \neg Bel(\neg Y|X)}{Bel(Z\,|\,X \wedge Y)}$$

They like to give Gettier-like examples against RM, which I never find quite convincing :-)

- RM is valid in belief revision (AGM), nonmonotonic reasoning (Lehmann & Magidor), and in the logic of counterfactuals (Lewis, Stalnaker).
- Semantically, RM corresponds to worlds being *totally pre-ordered*:

  Total pre-orders (preference orders) are not just presupposed in belief revision, nonmonotonic reasoning, and for counterfactuals, but also in decision theory, social choice, Popper functions,. . . .
- In particular, for counterfactuals, a rule for *negated counterfactuals* is needed: What substitute do Hanti & Kevin offer?

Remark (given a logically finite language):

As things stand, Hanti & Kevin cannot get a strong completeness result for the logic they prefer, that is, system P in nonmonotonic reasoning.

KLM (1990) showed that for that purpose one actually needs to strictly partially order *states* that are *labelled* by worlds, not worlds themselves:

One needs to allow the same state description (e.g., $p \wedge q$) to occur at different places in the ordering!

Remark (given a logically finite language):

As things stand, Hanti & Kevin cannot get a strong completeness result for the logic they prefer, that is, system P in nonmonotonic reasoning.

KLM (1990) showed that for that purpose one actually needs to strictly partially order *states* that are *labelled* by worlds, not worlds themselves:

One needs to allow the same state description (e.g., $p \wedge q$) to occur at different places in the ordering!

$\hookrightarrow$ Of course, Hanti & Kevin could apply their theory to a probability measure on *states* (rather than worlds or hypotheses).

But what is the interpretation of the *probability of a state*?

Remark (given a logically finite language):

As things stand, Hanti & Kevin cannot get a strong completeness result for the logic they prefer, that is, system P in nonmonotonic reasoning.

KLM (1990) showed that for that purpose one actually needs to strictly partially order *states* that are *labelled* by worlds, not worlds themselves:

One needs to allow the same state description (e.g., $p \wedge q$) to occur at different places in the ordering!

$\hookrightarrow$ Of course, Hanti & Kevin could apply their theory to a probability measure on *states* (rather than worlds or hypotheses).

But what is the interpretation of the *probability of a state*?

Remark: If I applied my theory to a *set* of probability measures, as some would prefer, then I would also fall back upon P.

High Probability Constraints

Consider an example:

$$P(\{w_0\}) = \frac{1}{10},\ P(\{w_1\}) = \ldots = P(\{w_{18}\}) = \frac{1}{20}$$

High Probability Constraints

Consider an example:

$$P(\{w_0\}) = \frac{1}{10},\ P(\{w_1\}) = \ldots = P(\{w_{18}\}) = \frac{1}{20}$$

For, say, fixed $t \geq 1$, and Hanti & Kevin's

$$u < v \text{ iff } P(\{u\}) > t \cdot P(\{v\})$$

being in place, their theory predicts

$$Bel(\{w_0\}), \text{ while } P(\{w_0\}) = \frac{1}{10}$$

High Probability Constraints

Consider an example:

$$P(\{w_0\}) = \tfrac{1}{10},\ P(\{w_1\}) = \ldots = P(\{w_{18}\}) = \tfrac{1}{20}$$

For, say, fixed $t \geq 1$, and Hanti & Kevin's

$$u < v \text{ iff } P(\{u\}) > t \cdot P(\{v\})$$

being in place, their theory predicts

$$Bel(\{w_0\}), \text{while } P(\{w_0\}) = \frac{1}{10}$$

So the "$\rightarrow$" of the Lockean thesis is invalidated.

High Probability Constraints

Consider an example:

$$P(\{w_0\}) = \tfrac{1}{10},\ P(\{w_1\}) = \ldots = P(\{w_{18}\}) = \tfrac{1}{20}$$

For, say, fixed $t \geq 1$, and Hanti & Kevin's

$$u < v \text{ iff } P(\{u\}) > t \cdot P(\{v\})$$

being in place, their theory predicts

$$Bel(\{w_0\}), \text{ while } P(\{w_0\}) = \frac{1}{10}$$

So the "$\rightarrow$" of the Lockean thesis is invalidated.

In my theory, P-stability^r yields a total pre-order $<$ so that

$$P(\{u\}) > \sum_{v:\, u<v} \frac{r}{1-r} \cdot P(\{v\})$$

and "$\rightarrow$" of the Lockean thesis holds (for threshold $r \geq \frac{1}{2}$).

Friendly suggestion to Hanti & Kevin:

- If you insist on presupposing merely a strict partial order $<$ on worlds, then you could still adapt my sum condition to such orders:

$$P(\{u\}) > \sum_{v:\, u<v} \frac{r}{1-r} \cdot P(\{v\})$$

where now $<$ is *not* demanded to result from a total pre-order.

Then, and only then, you are guaranteed the "$\rightarrow$" of the Lockean thesis.

(Equivalently: Look for P-stable^r subsets of *proper subsets* of W!)

Friendly suggestion to Hanti & Kevin:

- If you insist on presupposing merely a strict partial order $<$ on worlds, then you could still adapt my sum condition to such orders:

$$P(\{u\}) > \sum_{v:\, u<v} \frac{r}{1-r} \cdot P(\{v\})$$

where now $<$ is *not* demanded to result from a total pre-order.

Then, and only then, you are guaranteed the "$\rightarrow$" of the Lockean thesis.

(Equivalently: Look for P-stable^r subsets of *proper subsets* of W!)

And in my view, it is in fact not good enough to merely *allow* for models in which believed propositions have a high enough probability. It should be a "quasi-logical constraint" that this is so: for me,

$$Bel(X) \text{ and } P(\neg X) \geq P(X)$$

is *analytically false*.

Remark:

If they followed my suggestion—how would they determine the strict partial order from P (in line with their reductive account)?

Say, the "$\rightarrow$" of the Lockean thesis was taken care of.

Then this would still leave Hanti & Kevin with the following issue:

$$
\begin{array}{ccccc}
 & w_2\colon \frac{11}{100} & w_4\colon \frac{11}{100} & w_6\colon \frac{11}{100} & w_8\colon \frac{11}{100} \\
 & | & | & | & | \\
w_0\colon \frac{8}{100} & w_1\colon \frac{12}{100} & w_3\colon \frac{12}{100} & w_5\colon \frac{12}{100} & w_7\colon \frac{12}{100}
\end{array}
$$

('$|$' means $<$; thresholds from/to w_0 are so that no $<$-connections emerge).

This yields:

- $Bel(\{w_0, w_1, w_3, w_5, w_7\})$, and $P(\{w_0, w_1, w_3, w_5, w_7\}) = \frac{56}{100}$ ✓
- In fact, the "$\rightarrow$" of the Lockean thesis holds (for threshold $\frac{1}{2}$). ✓

Say, the "$\rightarrow$" of the Lockean thesis was taken care of.

Then this would still leave Hanti & Kevin with the following issue:

$$
\begin{array}{ccccc}
 & w_2\colon \frac{11}{100} & w_4\colon \frac{11}{100} & w_6\colon \frac{11}{100} & w_8\colon \frac{11}{100} \\
 & | & | & | & | \\
w_0\colon \frac{8}{100} & w_1\colon \frac{12}{100} & w_3\colon \frac{12}{100} & w_5\colon \frac{12}{100} & w_7\colon \frac{12}{100}
\end{array}
$$

('$|$' means $<$; thresholds from/to w_0 are so that no $<$-connections emerge).

This yields:

- $Bel(\{w_0, w_1, w_3, w_5, w_7\})$, and $P(\{w_0, w_1, w_3, w_5, w_7\}) = \frac{56}{100}$ ✓
- In fact, the "$\rightarrow$" of the Lockean thesis holds (for threshold $\frac{1}{2}$). ✓
- But: $\neg Bel(\{w_1, \ldots, w_8\})$, even though $P(\{w_1, \ldots, w_8\}) = \frac{92}{100}$?

That is: they still don't have the "$\leftarrow$" of the Lockean thesis for any $r \geq \frac{1}{2}$.

- And they *could not* have the "$\leftarrow$" of the Lockean thesis, unless their theory collapses into my theory!

A fortiori, they could not have the full Lockean thesis (not even with a threshold depending on P).

- And they *could not* have the "$\leftarrow$" of the Lockean thesis, unless their theory collapses into my theory!

A fortiori, they could not have the full Lockean thesis (not even with a threshold depending on P).

But:

> *In many cases a quantitative concept corresponds to a classificatory concept. Thus temperature corresponds to the property Warm... (Carnap 1950)*

- And they *could not* have the "$\leftarrow$" of the Lockean thesis, unless their theory collapses into my theory!

A fortiori, they could not have the full Lockean thesis (not even with a threshold depending on P).

But:

> *In many cases a quantitative concept corresponds to a classificatory concept. Thus temperature corresponds to the property Warm... (Carnap 1950)*

It seems that for *temperature* and *warm* we *do* get a "Lockean thesis".

So why not for *degree of belief* and *belief*?

Contextual Parameters

We all rely on contextual parameters, such as partitions and thresholds.

But I need *one* threshold r, while they need *many*:

$$u < v \text{ iff } P(\{u\}) > / \geq t_{u,v} \cdot P(\{v\})$$

- If $t_{u,v}$ is constantly 1 (for all $u, v \in W$), then any of Hanti & Kevin's $<$ must result from a total pre-order.

Contextual Parameters

We all rely on contextual parameters, such as partitions and thresholds.

But I need *one* threshold r, while they need *many*:

$$u < v \text{ iff } P(\{u\}) > / \geq t_{u,v} \cdot P(\{v\})$$

- If $t_{u,v}$ is constantly 1 (for all $u, v \in W$),
  then any of Hanti & Kevin's $<$ must result from a total pre-order.
- If $t_{u,v} = t_v$ only depends on v (for all $v \in W$), then certain strict total orders that do not result from a total pre-order must be excluded.

Contextual Parameters

We all rely on contextual parameters, such as partitions and thresholds.

But I need *one* threshold r, while they need *many*:

$$u < v \text{ iff } P(\{u\}) > / \geq t_{u,v} \cdot P(\{v\})$$

- If $t_{u,v}$ is constantly 1 (for all $u, v \in W$),
  then any of Hanti & Kevin's $<$ must result from a total pre-order.
- If $t_{u,v} = t_v$ only depends on v (for all $v \in W$), then certain strict total orders that do not result from a total pre-order must be excluded.
- So in order to deviate properly from total pre-orders,
  they *need* thresholds $t_{u,v}$ that vary both with u and v.

  In a nutshell: probabilities are totally ordered, which is why it is hard to get rid of totality for worlds.

Contextual Parameters

We all rely on contextual parameters, such as partitions and thresholds.

But I need *one* threshold r, while they need *many*:

$$u < v \text{ iff } P(\{u\}) > / \geq t_{u,v} \cdot P(\{v\})$$

- If $t_{u,v}$ is constantly 1 (for all $u, v \in W$),
  then any of Hanti & Kevin's $<$ must result from a total pre-order.
- If $t_{u,v} = t_v$ only depends on v (for all $v \in W$), then certain strict total orders that do not result from a total pre-order must be excluded.
- So in order to deviate properly from total pre-orders,
  they *need* thresholds $t_{u,v}$ that vary both with u and v.

  In a nutshell: probabilities are totally ordered, which is why it is hard to get rid of totality for worlds.

For, say, 8 worlds (hypotheses): where do these $8 \cdot 7 = 56$ numbers come from?

Another reason why I stick to the more simple(-minded) solution.