

Philosophy as Rational Construction and The Role of Mathematics in Philosophy

Hannes Leitgeb

LMU Munich

June 2019

- Science aims at the description, explanation, and prediction of *nature*.

- Science aims at the description, explanation, and prediction of *nature*.
- Is philosophy, or should it be, nothing but a “high-level” science?
(*Naturalism*; Quine; Williamson; . . .)

- Science aims at the description, explanation, and prediction of *nature*.
- Is philosophy, or should it be, nothing but a “high-level” science?

(*Naturalism*; Quine; Williamson; . . .)

↪ That is what I want to argue *against!*

Philosophy should share certain *methods* with science, but not its *aims*.

Metaphilosophical proposal: Let us understand philosophy as devoted to

- *rational intellectual constructions* that do not aim to describe nature but *that aim to realize norms of rationality as such* (and no other norms).

Call this proposal: *Rational Constructionism*.

Ultimately, the hope is:

this proposal is not too revisionary, it clarifies our self-image, and it may improve philosophy (just a bit) in the long run.

(It is *not* my goal to capture everything that was called 'philosophy' once; nor to *police* philosophers.)

As I am going to argue, mathematical methods are bound to play a major role if rational constructionism is a good proposal.

Mathematics is something that philosophy and science can share.

Plan:

- 1 Rational Intellectual Constructions. . .
- 2 . . . Aiming to Realize Norms of Rationality as such
- 3 Mathematics for Rational Construction: A Case Study
- 4 Mathematics for Rational Construction: A Survey
- 5 Conclusions and the Future

Rational Intellectual Constructions...

First part of the proposal: Let us understand philosophy as aiming at

- *rational intellectual constructions* (of a certain kind).

For a start, it is useful to focus on special rational constructions:

- rational *reconstructions*.

I take the term 'rational reconstruction' from Rudolf Carnap (e.g. in his *Aufbau*, 1928); it was used later also by Reichenbach, Habermas,...

(From 1945 Carnap also spoke of 'explication'. Explication is restricted to concepts, while rational reconstruction is broader.)

I will explain first 'rational reconstruction' and then turn to 'construction proper'.

Rational reconstruction:

- '*Re-*' means that one starts with a given X upon which one reflects: X is an intellectual human/cultural product with rationality features.
- '*-construction*' means that one is taking apart X : studying, amending, and reassembling it—a kind of engineering.
- '*rational*' means that, when doing so, one is taking a particular normative stance:
 - (i) one determines X 's rational/irrational features;
 - (ii) one evaluates X 's rational/irrational features;
 - (iii) if sufficiently rational, one leaves X as it is; if overly irrational, one rejects X ; else, one corrects X 's irrational features and enhances X 's rational features by replacing X by a similar but more rational X' .

Philosophy is thus viewed as a certain kind of “rational therapy”—Wittgenstein!
The “medicine” for the “therapy” crucially involves language.

Philosophy is thus viewed as a certain kind of “rational therapy”—Wittgenstein! The “medicine” for the “therapy” crucially involves language.

But, *contrary to Wittgenstein*:

- the object of the “therapy” need not be linguistic: one may also rationally reconstruct *concepts, frameworks, beliefs, reasoning, decisions, methods, norms, ideals, plans of institutions, how they all relate to each other, . . .*—any type of intellectual entity!
- the “therapy” need not be exhausted by a “diagnosis” (clarification through examples, analysis) or “amputation” (rejection) but may involve actual “treatment”:

precisification (where necessary), *systematization*, and various kinds of *rationality-improvement* mediated by: the definition of concepts, the defense of philosophical theses, the construction of philosophical theories and models, the development of new conceptual frameworks, applications of logical-mathematical methods, applications to concrete circumstances.

- Diachronically, *analysis* is typically the first stage of reconstruction:

Stage 1: Analyze X
Stage 2: Destruct X / Replace by X' } Rational reconstruction

- The outputs of rational analysis and rational reconstruction *coincide*, when clarification indicates that not much “rationality repairing/criticism” needs to be done:

Stage 1: Analyze X
Stage 2: [] } Rational analysis

- But often rational reconstruction *should* go beyond plain analysis:

E.g., an inferential pattern, or the reasons for an action, or the presuppositions of a question, might be perfectly clear (and exact and systematic), but *wrong*, in which case they ought to be *rejected* / *rectified*.

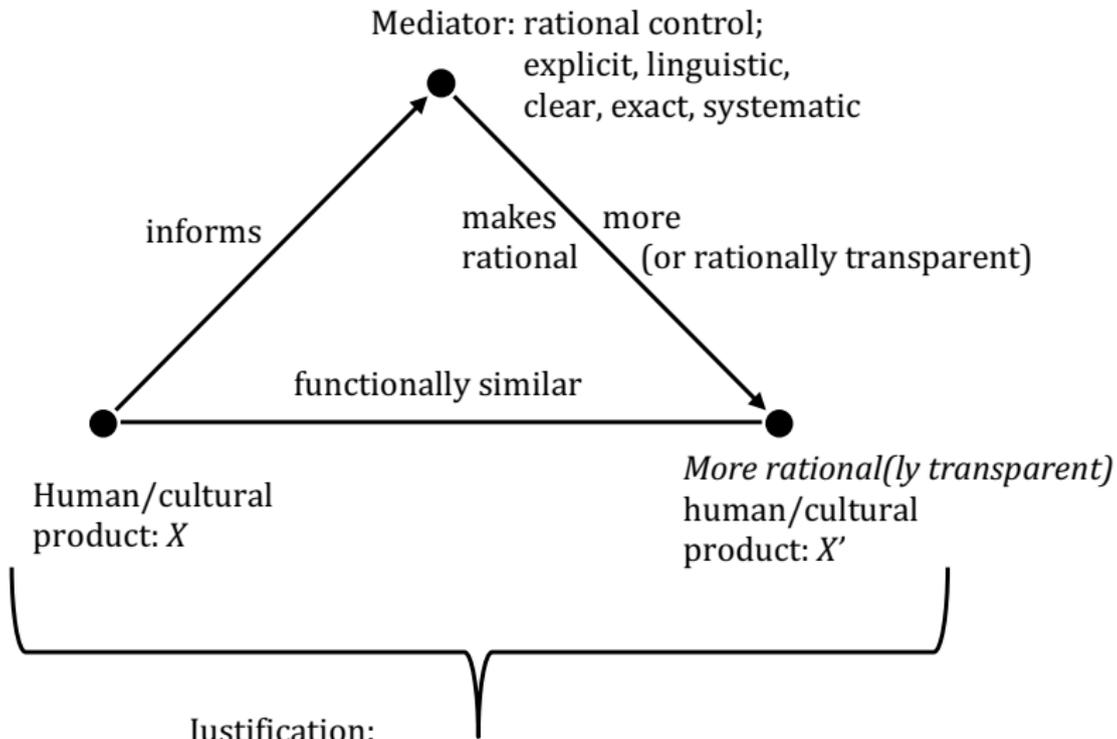
- Rational construction: either rational analysis of X , or rational reconstruction of X by X' , or a *completely new* X' gets constructed:

Stage 1: []
Stage 2: Construct X' } Rational construction (proper)

Clarification, precisification (where necessary), and systematization *facilitate* rational criticism and improvement of X , by improving our understanding of X , the informativity of X , and the order/structure of X .

But the ultimate goal is rationality-improvement itself:

- X' should be more rational than (a partially irrational) X but at the same time functionally similar enough to X for the relevant purposes.
- The 'rational' in 'rational construction' is to be understood very broadly (theoretical and practical; not necessarily just instrumental).
- The study of the norms and presuppositions of rationality informs this kind of normative project and is therefore crucial for it. The norms and presuppositions are themselves subject to rational (re-)construction again.



- Similarity
- Theoretical rationality: norms concerning truth
- Practical rationality: norms concerning fruitfulness, power, simplicity, morality, aesthetics,...

Example: What do mean by mean by the predicate 'true (sentence)'?

Example: What do we mean by the predicate 'true (sentence)'?

Famously, Tarski (1933) suggested an answer, and he did so by *rationaly reconstructing* truth (this is one of Carnap's 1950 examples of explication):

- Tarski starts by looking at examples and the history of the subject; he detects a pattern: all instances of the truth scheme

'A' is true if and only if A

seem assertable and acceptable, and they capture in some sense the vague "truth as correspondence" idea; he points to paradoxes (Liar!).

Example: What do we mean by the predicate 'true (sentence)'?

Famously, Tarski (1933) suggested an answer, and he did so by *rationaly reconstructing* truth (this is one of Carnap's 1950 examples of explication):

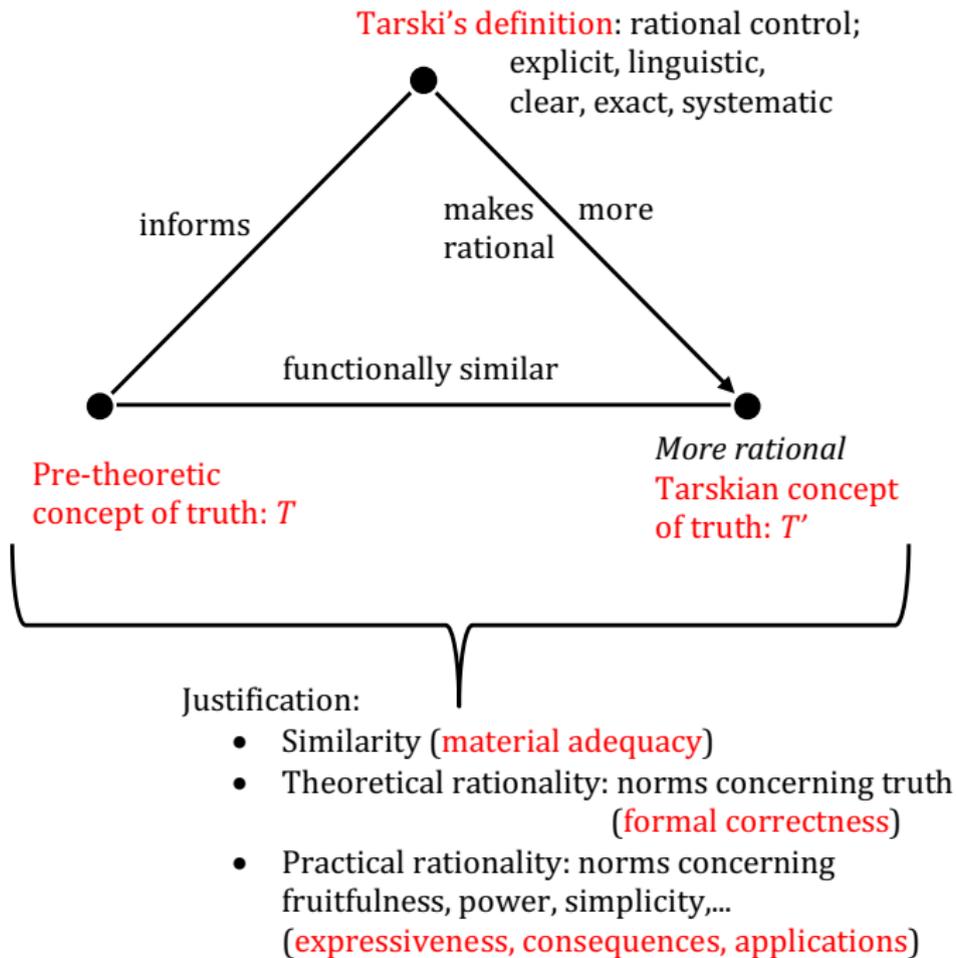
- Tarski starts by looking at examples and the history of the subject; he detects a pattern: all instances of the truth scheme

'A' is true if and only if A

seem assertable and acceptable, and they capture in some sense the vague "truth as correspondence" idea; he points to paradoxes (Liar!).

And then he develops a way of doing better than that.

- He shows how we can define truth for a great variety of formalized fragments of natural or scientific language in a precise framework of syntax and higher-order logic, such that the definition is materially adequate but no paradoxical claims follow. The standard laws of truth can be derived, the theory is provably consistent, and it turned out to be enormously fruitful in logic, philosophy of language, and linguistics.



... Aiming to Realize Norms of Rationality as such

Second part of the proposal: Let us understand philosophy as consisting of

- rational intellectual constructions *that aim to realize norms of rationality as such* (and no other norms).

Aiming to realize norms of rationality as such:

- aiming to construct an X' that satisfies A , where rationality requires that it *ought to be the case that A* ;
- such that the corresponding norms are *general*: satisfying these norms does not commit oneself epistemically to particular ways that nature is like, nor ontically to any particular universe of discourse.

(E.g.: X' *ought to be clear, precise, systematic, consistent, . . .*).

According to that proposal, none of the following should count as philosophical:

- Constructions of scientific concepts/theories in physics, biology, . . . for scientific purposes:

For their aim is not merely to realize norms of rationality as such but rather norms such as *you ought to describe/explain/predict observations*: realizing such norms commits oneself epistemically to *particular ways that nature is like* and ontically to *particular universes of discourse*.

(This also applies to empirical investigations of rationality in cognitive psychology. \leftrightarrow Psychologism debate \checkmark)

- Constructions of bridges, glasses, artificial hearts, . . . :

For their aim is not rational *intellectual* constructions.

- Constructions of formal concepts, theories, and algorithms in mathematics and theoretical computer science:

For, again, their aim is not merely to realize norms of rationality as such.

E.g., Menger's explication of *dimension* in topology, which influenced Carnap's views on explication, also aimed to realize *geometric* norms.

But, as intended, the proposal to understand philosophy as aiming at

- rational intellectual constructions that realize norms of rationality as such (and no other norms)

does match features that have been ascribed to philosophy in the past:

- For philosophy has been claimed to be meta, normative, apriori, general, increasing our self-understanding. ✓
- Compare: Socratic maieutics; the emphasis on clear and distinct ideas in Descartes, Leibniz, . . . ; Kant on philosophy as “Die Gesetzgebung der menschlichen Vernunft”; philosophy as *Geisteswissenschaft*; Neo-Kantian roots of rational reconstruction; logical construction in early analytic philosophy. ✓

What can be rationally constructed in that way (that is, rationally analyzed, rationally reconstructed, or properly rationally constructed)?

Everything that is to be found in philosophy!

- Logical constants, form, and validity (logic). Meaning, truth, and communication (philosophy of language). Belief, justification, theoretical rationality, and knowledge (epistemology). Scientific concepts, theories, methods, and confirmation (philosophy of science). Moral action, moral attitudes, and moral maxims (ethics). Preference, decision-making, and desire-belief rationality (decision theory). Historical philosophical concepts and positions (history of philosophy). ... ✓

- The only role for “intuitions” in rational construction is as starting points. (cf. Cappelen 2012, Deutsch 2015)

In the constructional system, we shall... reconstruct these manifestations in a rationalizing or schematizing fashion; intuitive understanding is replaced by discursive reasoning. (Carnap 1928, §54)

About progress:

- Both science and philosophy make progress.

However:

- Scientific progress leads to (some degree of) *convergence* because the natural phenomena “out there” stabilize theory-choice.
- Philosophical rational constructions have improved a lot over time (their clarity, exactness, systematicity, their rationality features, the methods by which they are carried out, new possibilities, . . .). Philosophical activities, which aim at such constructions, find better targets and hit their targets better. But not by describing a stabilizing natural phenomenon “out there”!

That is why philosophical progress does *not necessarily lead to convergence* (cf. Chalmers 2015).

Philosophical construction is driven by choices that are rational but “free” in the sense of not being committed to particular ways nature might be or to any particular universe of discourse. That explains why rational construction is *pluralistic*: not uniquely determined and open-ended.

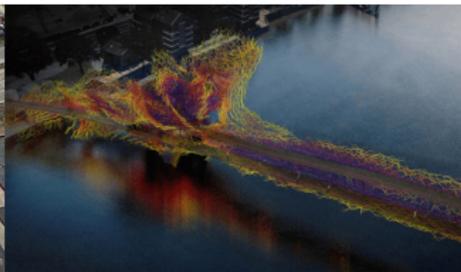
Philosophy as rational construction is compatible with progress. . .



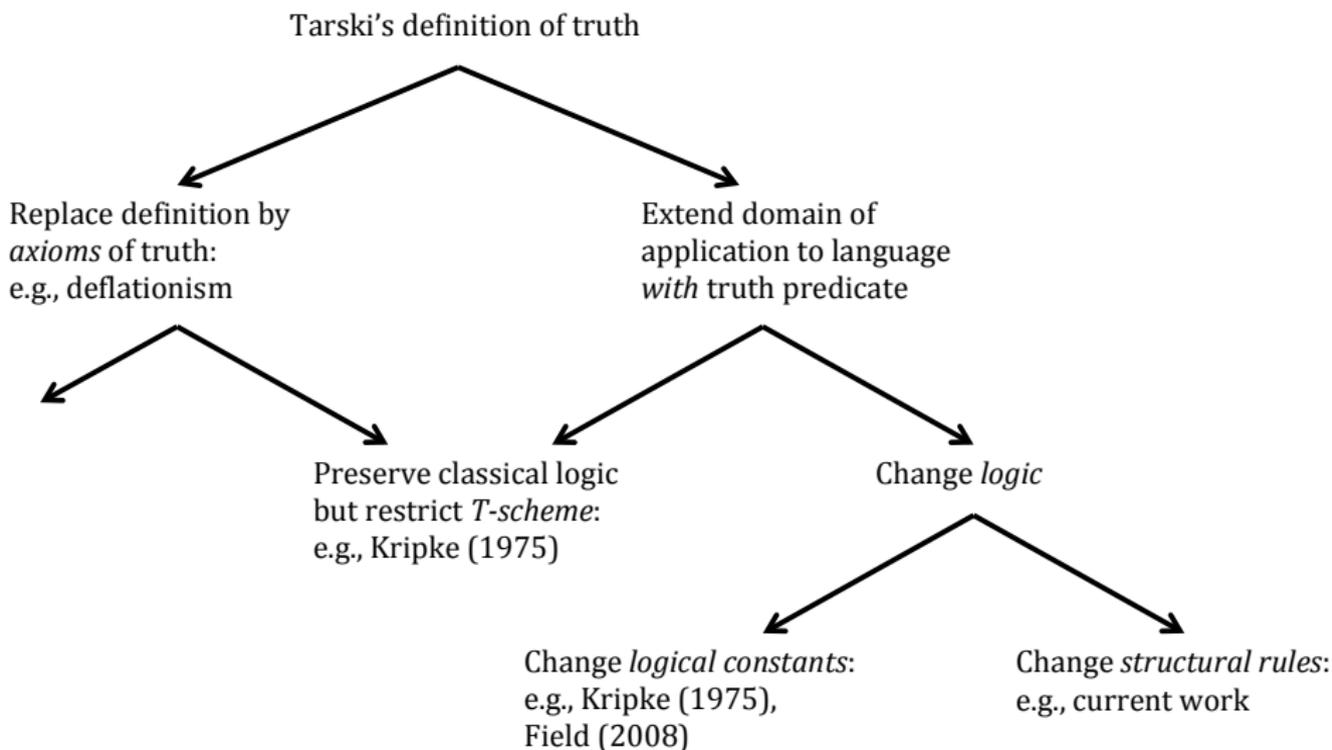
Philosophy as rational construction is compatible with progress. . .



without convergence. . .



Non-uniqueness and open-endedness in the case of the concept of *truth*:



(No end in sight: on Monday I will mention a new way of improving Kripke.)

Logical-mathematical methods (higher-order logic, set theory) did play a crucial role in the case of Tarski's rational reconstruction/explication of *truth*.

As I am now going to argue, this is a *typical* case:

- Logical-mathematical methods are bound to play a major role in the philosophical rational construction of concepts, reasoning, decisions, . . .

- Carnap on explication (rational reconstruction):

By the procedure of explication we mean the transformation of an inexact, prescientific concept, the explicandum, into a new exact concept, the explicatum. [...] A concept must fulfill the following requirements in order to be an adequate explicatum for a given explicandum; (1) similarity to the explicandum, (2) exactness, (3) fruitfulness, (4) simplicity. (R. Carnap, Logical Foundations of Probability, 1950)

Mathematics for Rational Construction: A Case Study

- Carnap on explication (rational reconstruction):

By the procedure of explication we mean the transformation of an inexact, prescientific concept, the explicandum, into a new exact concept, the explicatum. [...] A concept must fulfill the following requirements in order to be an adequate explicatum for a given explicandum; (1) similarity to the explicandum, (2) exactness, (3) fruitfulness, (4) simplicity. (R. Carnap, Logical Foundations of Probability, 1950)

I will concentrate on the following example now:

- the explication of the rational acceptability of conditionals.
explicandum

(Whenever I discuss methodological features, I will put them in red.)

QUESTION: When should a conditional 'if A , then B ' be acceptable to a rational agent?

This question is not "cooked up" in any way, nor is it a question that makes sense only with some mathematical framework already in place.

QUESTION: When should a conditional 'if A , then B ' be acceptable to a rational agent?

This question is not “cooked up” in any way, nor is it a question that makes sense only with some mathematical framework already in place.

Here is one reason why we should care about this:

- Pragmatic meaning of a sentence:
the constraint imposed on a receiver by the assertion of the sentence.
- Pragmatic meaning of $A \Rightarrow B$:
receiver, please be such that $A \Rightarrow B$ is acceptable to you!

So in order to understand the pragmatic meaning of conditionals, we first need to know what it means for a conditional to be acceptable.

F.P. Ramsey hinted at an answer to our question in his “General propositions and causality” (1929):

If two people are arguing ‘If p will q?’ and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q. . .

No mathematization as yet.

F.P. Ramsey hinted at an answer to our question in his “General propositions and causality” (1929):

If two people are arguing ‘If p will q?’ and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q. . .

No mathematization as yet.

The Ramsey test is plausible, because it ties the acceptance of conditionals to suppositional reasoning. However, if we want to draw any non-trivial philosophical conclusions from this, then we need to be more informative about ‘knowledge’, ‘adding hypothetically to’, ‘arguing on that basis’ (“*exactness*”!).

F.P. Ramsey hinted at an answer to our question in his “General propositions and causality” (1929):

If two people are arguing ‘If p will q?’ and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q. . .

No mathematization as yet.

The Ramsey test is plausible, because it ties the acceptance of conditionals to suppositional reasoning. However, if we want to draw any non-trivial philosophical conclusions from this, then we need to be more informative about ‘knowledge’, ‘adding hypothetically to’, ‘arguing on that basis’ (“exactness”!).

And for that purpose we first need to look at examples.

the explicandum. . . should be made as clear as possible by informal explanations and examples (Carnap 1950)

For instance, Ernest Adams (1970) did so in terms of his Oswald pair:

- *Indicative*: We accept 'If Oswald didn't kill Kennedy, someone else did'.
- *Subjunctive*: We do not accept 'If Oswald hadn't killed Kennedy, someone else would have'.

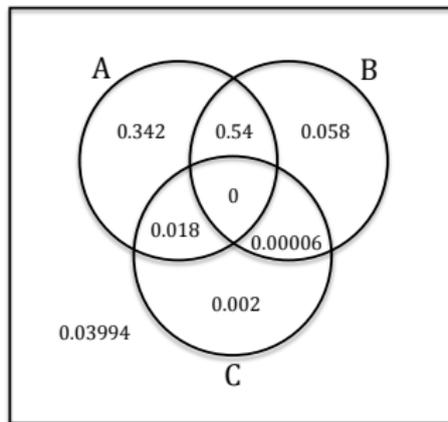
For instance, Ernest Adams (1970) did so in terms of his Oswald pair:

- *Indicative*: We accept 'If Oswald didn't kill Kennedy, someone else did'.
- *Subjunctive*: We do not accept 'If Oswald hadn't killed Kennedy, someone else would have'.

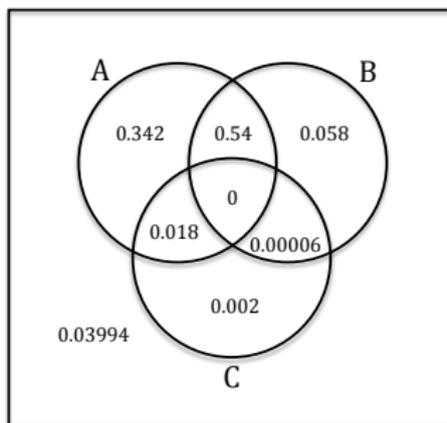
Hence, in whatever way one makes the Ramsey test more precise, it needs to be sensitive to the grammatical mood of conditionals.

One way to make it more precise: by means of *probability theory*.

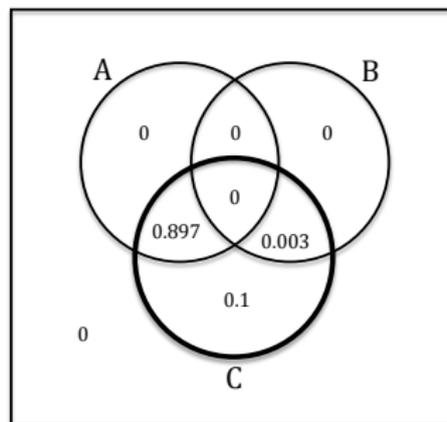
E.g., a probability measure P :



E.g., a probability measure P :



P conditionalized on C :



For indicative conditionals, Adams (1975) made acceptance precise as follows:

- For every subjective probability measure P , for all sentences A, B :

The degree of acceptability $Acc_P(A \rightarrow B)$ (rel. to P) equals $P_A(B)$,

where: $P_A(B) = P(B|A)$.

What's happening here? Belief (Ramsey: "knowledge") is analyzed on a quantitative scale in terms of subjective probabilities, which may be justified independently. The indicative 'if-then' is translated into *matter-of-fact* supposition, which gets formalized by conditionalization. Carnap (1950): "introduce the explicatum into a well-connected system of scientific concepts".

Other choices would have been possible: e.g., in the theory of belief revision, belief is treated on an ordinal scale; cf. Carnap's §4 on scales of concepts.

Now additionally assume indicative conditionals to be true or false at worlds:

Then presumably, the degree of acceptability of an indicative conditional ought to be equal to the degree of belief of that conditional to be true.

That is:

- Stalnaker's Thesis:

For every subjective probability measure P , for all sentences A, B :

$$P(B|A) = P(A \rightarrow B)$$

Now additionally assume indicative conditionals to be true or false at worlds:

Then presumably, the degree of acceptability of an indicative conditional ought to be equal to the degree of belief of that conditional to be true.

That is:

- Stalnaker's Thesis:

For every subjective probability measure P , for all sentences A, B :

$$P(B|A) = P(A \rightarrow B)$$

But by David Lewis' (1976) Triviality Theorem *that's impossible!*

This is not fictional. Stalnaker (1970) actually proposed the above as a thesis, and Lewis' theorem was a big surprise.

In Dorothy Edgington's (1995) words: it was "the bombshell".

First derive that $P(B \rightarrow C|A) = P(C|B \wedge A)$ by means of Stalnaker's thesis.

Then show: $P(B \rightarrow C) =$

$$= P(C \wedge (B \rightarrow C)) + P(\neg C \wedge (B \rightarrow C)) \quad \text{(Addition Theorem)}$$

$$= P((B \rightarrow C)|C)P(C) + P((B \rightarrow C)|\neg C)P(\neg C) \quad \text{(Ratio formula, Ax.)}$$

$$= P(C|B \wedge C)P(C) + P(C|B \wedge \neg C)P(\neg C) \quad \text{(From above)}$$

$$= 1 \cdot P(C) + 0 \cdot P(\neg C) \quad \text{(Axioms)}$$

$$= P(C) \quad \text{Absurd!}$$

First derive that $P(B \rightarrow C|A) = P(C|B \wedge A)$ by means of Stalnaker's thesis.

Then show: $P(B \rightarrow C) =$

$$= P(C \wedge (B \rightarrow C)) + P(\neg C \wedge (B \rightarrow C)) \quad \text{(Addition Theorem)}$$

$$= P((B \rightarrow C)|C)P(C) + P((B \rightarrow C)|\neg C)P(\neg C) \quad \text{(Ratio formula, Ax.)}$$

$$= P(C|B \wedge C)P(C) + P(C|B \wedge \neg C)P(\neg C) \quad \text{(From above)}$$

$$= 1 \cdot P(C) + 0 \cdot P(\neg C) \quad \text{(Axioms)}$$

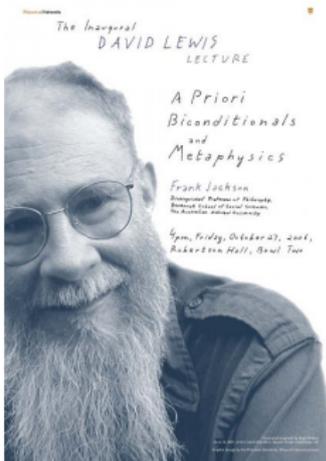
$$= P(C) \quad \text{Absurd!}$$

One possible conclusion: Indicative conditionals are not true or false, they do not express propositions, and their pragmatic meaning does not derive from their semantic meaning. (The Suppositional Theory of Conditionals: Adams, Edgington, . . .) Or one modifies some implicit assumptions: maybe indicatives are true in a world *and a context*, . . .

How could one have found anything like that without mathematization?

And mainstream philosophy *does* take this up (cf. Bennett 2003).

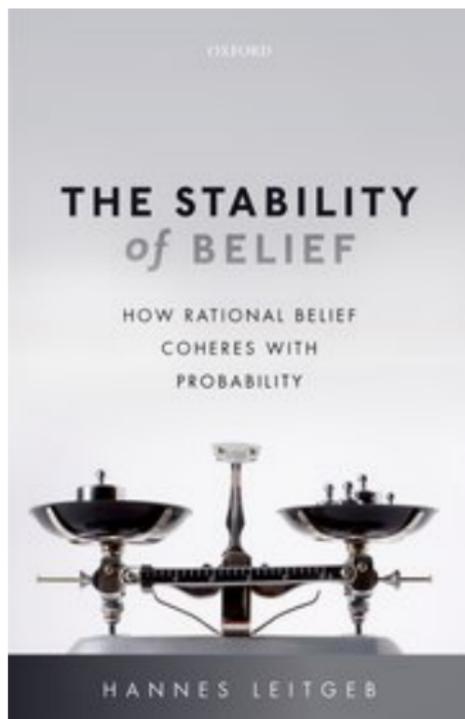
“The explicatum is to be a *fruitful* concept. . .” (Carnap 1950).



What one can learn from this case study:

- *Exactness* and *fruitfulness* pull towards mathematization.
- Mathematization also adds to *interaction with science*. (Lots of citations of Lewis' Triviality paper in computer science journals! A lot of psychological work on conditionals is based on Adams and Edgington.)
- *Similarity* is a rough-and-ready form of *structural* similarity (belief, assumption, indicative/subjunctive, . . .). If the structure is complex enough, mathematization pays off and will normally even be *necessary* for progress (as exemplified by the Triviality theorem).
- If different explications of the same concept lead to analogous results, then the philosophical conclusions thereof are *robust*.
(E.g.: Triviality results both for *probabilistic* conditional belief and for *logic-governed binary* conditional belief. Convergence on the same logic.)

One can also study rational belief and the acceptability of conditionals *simultaneously* in logical and probabilistic terms; e.g.:



Mathematics for Rational Construction: A Survey

Traditionally, mathematical methods have helped scientists to achieve their aims of describing, explaining, and predicting nature.

But mathematical methods (including logical methods) can also help us to achieve successful rational constructions in philosophy.

- By improving the clarity and exactness of a rational construction:

Philosophical Concept =*df* ...[[Logical-Mathematical Concepts](#)]...

E.g., Tarski's definition of truth by [logical](#) methods,

the explication of the rational acceptability of conditionals, scientific confirmation,... by [probabilistic](#) methods (Bayesianism),

⋮

- By systematizing and justifying a rational construction:

Philosophical Axiom 1.

Philosophical Axiom 2.

⋮

(Mathematical Theorems.)

Philosophical Conclusion.

E.g.: Justifying

degrees of belief ought to conform to the axioms of probability

by Dutch book arguments, decision-theoretic representation theorems, minimizing inaccuracy, ... (see e.g. Joyce 1998).

- By ruling out a rational construction through inconsistency:

Philosophical Axiom 1.

Philosophical Axiom 2.

⋮

(Mathematical Theorems.)

Contradiction.

E.g.: Lewis (1976) on ‘the degree of acceptability of an indicative conditional is given by conditional probability’ and ‘indicative conditionals have truth conditions’.

- By proving the consistency of a rational construction through building a **model** in the model-theoretic sense:

E.g.: Kripke (1975), Field (2008) on truth for a semantically closed language with the unrestricted truth scheme.

- By informing us about a rational construction through a (toy-)model (as in the physicist's sense), that is, a set of idealized assumptions about some system:

E.g.: Zollman (2010) on the epistemic benefits of special kinds of diversity.

(Models in that sense allow for the application of computer simulations.)

There are many examples of each type:

- Logical form, logical consequence, logical validity, logical concepts, logical rules, non-classical logic ([proof theory/model theory](#)).
- Ontological argument for the existence of God ([logic](#)).
- Model-theoretic arguments about realism/antirealism ([model theory](#)).
- Truth, meaning (extensions, intensions), concepts, propositions, modalities, pragmatics ([possible worlds semantics](#), [graph theory](#), [modal logic](#)).
- Rational belief, aggregating beliefs, coherence of sets of propositions ([scales of belief](#), [doxastic logic](#), [ranking functions](#), [nonmonotonic reasoning](#), [subjective probability theory](#)).
- Acceptability of conditionals ([probability theory](#), [probability logic](#), [conditionalization vs imaging](#)).
- Inductive and abductive reasoning, (degrees of) confirmation, (degrees of) explanatory power, learning propositions and conditionals ([subj. probability theory](#), [belief revision theory](#), [statistical learning theory](#)).
- Knowledge ([epistemic logic](#)).

- Definitions, scientific theories, measurement ([logic](#), [model theory](#)).
- Metaethics ([deontic logic](#), [conditional norms](#), [defeasible norms](#)).
- Rational decision, rational preference, consequentialism, pragmatics, social philosophy ([decision theory](#), [social choice](#), [game theory](#)).
- Causality, causal explanation ([structured equations](#), [Bayesian networks](#)).
- Infinity, Zeno's Paradoxes ([calculus](#), [set theory](#)).
- Abstraction ([logical construction](#), [abstraction principles](#), [mereotopology](#)).
- Philosophy of time ([temporal logic](#)).
- Metaphysical necessity ([modal logic](#)), counterfactuals ([conditional logic](#), [total pre-orders](#), [choice functions](#)).
- Objective chance ([probability theory](#), [reflection principles](#)).
- Similarity, resemblance nominalism ([graph theory](#)).
- Mereology ([Boolean algebras](#)).
- Mental states ([computability theory](#), [Ramsification](#)).
- Properties, natural properties, dispositions ([higher-order logic](#), [convexity](#), [conditional logic](#)). And so forth!

Since mathematics

- is not committed epistemically to particular ways that nature is like,
- and can be used as a general theory of structure that is not tied to any particular universe of discourse,

its application in philosophy does not undermine the aim of realizing norms of rationality as such.

Of course, logical-mathematical methods are also important in the *philosophy of mathematics*:

- the **incompleteness theorems** and ‘mechanistic thesis $\rightarrow K \neq T$ ’, **modal logic** and potential infinity, **provability logic** and provability, **second-order logic** and structuralism, **abstraction principles** and neo-logicism, . . .

And *foundational work in mathematics* itself often overlaps with logical and philosophical work:

- foundational **set theory**, **homotopy-type-theory**, . . .

Mathematical philosophy (which is *not* just philosophy of mathematics):

- the application of logical and mathematical methods in philosophy.

↪ the core idea of the *Munich Center for Mathematical Philosophy (MCMP)*:

(<http://www.mcmp.philosophie.uni-muenchen.de>

About 50 people working there. 500+ video lectures.

Coursera course on “Introduction to Mathematical Philosophy”.)



Conclusions and the Future

- Rational constructionism might itself amount to a fruitful rational reconstruction of philosophy.
- Philosophy differs in its aims from science, but both make progress and both can be helped by mathematical methods.
- In the future: why not use our philosophical know-how to rationally (re-)construct outputs, programs, and methods of AI and machine learning—making them more transparent, understandable, systematic, criticizable, and coherent?

If philosophy finds a new role in the rational reconstruction of AI and machine learning, then only with the help of mathematical methods!

The only way to rectify our reasonings is to make them as tangible as those of the Mathematicians, so that we can find our error at a glance, and when there are disputes among persons, we can simply say: Let us calculate [calculemus]. (Leibniz 1685)