What is Mathematical Philosophy?

Hannes Leitgeb and Stephan Hartmann

Munich Center for Mathematical Philosophy (MCMP)

Summer School on Mathematical Philosophy for Female Students
Munich, 27 July 2014

Motivation

- Mathematical Philosophy is the study of **philosophical problems** with the help of **mathematical methods**.

Motivation

- Mathematical Philosophy is the study of **philosophical problems** with the help of **mathematical methods**.
- Mathematical is used in the sciences. Physicists have, of course, always used it. Meanwhile mathematical and statistical methods are also of tremendous use in many other sciences such as the biomedical and social sciences.
- Q: But why should one also use mathematics in philosophy?
- A: It works! Mathematical methods help us to solve problems that cannot be solved otherwise.

Motivation

- Deeper reason: Mathematics is not actually about number-crunching! Modern mathematics is the study of formal structure by means of proof: whenever an object or a class of objects exhibits enough structure, mathematics allows us to describe these objects in precise terms, to extract information from these descriptions, and to transform this information into conclusions.
- But structure can be found everywhere in philosophy: the concept of truth has a formal structure, every system of rational beliefs must have a particular kind of structure, the classification of acts into morally right and morally wrong ones has a certain structure, and so on.
- Using mathematics in philosophy has many practical advantages: it allows us to formulate a problem in precise terms, it works as an inference machine, it allows us to tackle complex problems that are too hard to tackle otherwise,...

Motivation

- Which methods? There is no need to stick to one (mathematical) method. Which method is useful will depend on the problem at hand. The mathematical philosopher is a consumer of methods developed in other fields.
- Which problems? Mathematical Philosophy addresses problems from many different philosophical fields (such as epistemology, metaphysics, ethics, political philosophy, etc.).
- Many problems are directly related to problems in the sciences (e.g. in the decision sciences and in cognitive psychology). Mathematical Philosophy is, at least to some extent, an interdisciplinary endeavor.
- Historical remark: Many of the great philosophers were indeed also mathematicians (think of Leibniz!)

Two Ways of Doing Mathematical Philosophy

1. The Mathematician's Way
2. The Physicist's Way

Two Ways of Doing Mathematical Philosophy

1. The Mathematician's Way
2. The Physicist's Way

The proof of the pudding is in its eating. . .

In the remainder of this talk, we will illustrate these two ways by looking at one concrete research field. You will see many more examples during this week!

I. The Mathematician's Way

Mathematical Philosophy by Explication

One way of doing mathematical philosophy ("the mathematician's way"):

- By Carnapian explication:

> *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)*

Mathematical Philosophy by Explication

One way of doing mathematical philosophy ("the mathematician's way"):

- By Carnapian explication:

> *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 deal with just one example:

- the explication of the $\underbrace{\textit{acceptability of conditionals}}_{\text{explicandum}}$ and its consequences.

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

The Acceptance of Conditionals

QUESTION: What does it mean for a a conditional to 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.

The Acceptance of Conditionals

QUESTION: What does it mean for a a conditional to 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 x:
  the constraint imposed on a receiver by asserting x.
- 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.

The Ramsey Test for Conditionals

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 for Conditionals

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 precise about ‘knowledge’, ‘adding hypothetically to’, ‘arguing on that basis’ (*“exactness”*!).

The Ramsey Test for Conditionals

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 precise 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)*

Indicative vs. Subjunctive Conditionals

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’.

Indicative vs. Subjunctive Conditionals

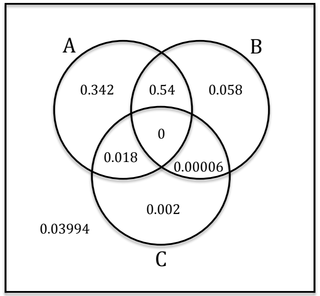
For instance, Ernest Adams (1970) did so in terms of his Oswald pair:

- *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'.

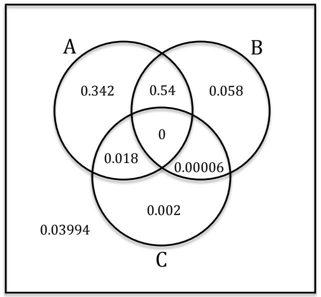
Hence, in whatever way one makes the Ramsey test more precise, it needs to be sensitive to the grammatical mood of conditionals.

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

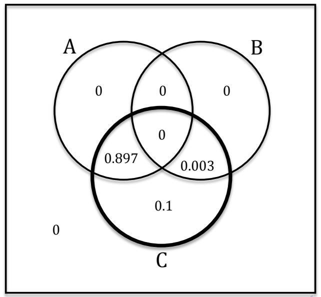
E.g., a probability measure P:



E.g., a probability measure P:



P conditionalized on C:



The Ramsey Test for Indicative Conditionals

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

The Ramsey Test for Indicative Conditionals

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

- For every subjective probability measure $\mathfrak{Cr}$, for all sentences A, B:

  The degree of acceptability $Acc_{\mathfrak{Cr}}(A \to B)$ *(in* $\mathfrak{Cr}$*) equals* $\mathfrak{Cr}_A(B)$*,*

  where: $\mathfrak{Cr}_A(B) = \mathfrak{Cr}(B|A)$.

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 $\mathfrak{Cr}$, for all sentences A, B:

$$\mathfrak{Cr}(B|A) = \mathfrak{Cr}(A \to B)$$

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 $\mathfrak{Cr}$, for all sentences A, B:

$$\mathfrak{Cr}(B|A) = \mathfrak{Cr}(A \to B)$$

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 $\mathfrak{Cr}(B \to C|A) = \mathfrak{Cr}(C|B \wedge A)$ by means of Stalnaker's thesis.

Then show: $\mathfrak{Cr}(B \to C) =$

$$= \mathfrak{Cr}(C \wedge (B \to C)) + \mathfrak{Cr}(\neg C \wedge (B \to C)) \quad \text{(Addition Theorem)}$$

$$= \mathfrak{Cr}((B \to C)|C)\,\mathfrak{Cr}(C) + \mathfrak{Cr}((B \to C)|\neg C)\,\mathfrak{Cr}(\neg C) \quad \text{(Ratio formula, Ax.)}$$

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

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

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

First derive that $\mathfrak{Cr}(B \to C|A) = \mathfrak{Cr}(C|B \wedge A)$ by means of Stalnaker's thesis.

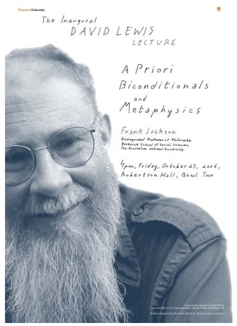
Then show: $\mathfrak{Cr}(B \to C) =$

$= \mathfrak{Cr}(C \wedge (B \to C)) + \mathfrak{Cr}(\neg C \wedge (B \to C))$	(Addition Theorem)
$= \mathfrak{Cr}((B \to C)|C)\,\mathfrak{Cr}(C) + \mathfrak{Cr}((B \to C)|\neg C)\,\mathfrak{Cr}(\neg C)$	(Ratio formula, Ax.)
$= \mathfrak{Cr}(C|B \wedge C)\,\mathfrak{Cr}(C) + \mathfrak{Cr}(C|B \wedge \neg C)\,\mathfrak{Cr}(\neg C)$	(From above)
$= 1 \cdot \mathfrak{Cr}(C) + 0 \cdot \mathfrak{Cr}(\neg C)$	(Axioms)
$= \mathfrak{Cr}(C)$	Absurd!

One possible conclusion: Indicative conditionals are not true or false, they do not express propositions, and their pragmatic meaning is not derivative 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).



Some General Remarks on Explication

- Explication is not *conceptual analysis*. (One does not encounter Adams' and Edgington's probabilistic theory of the acceptability of conditionals just by "zooming" into natural language concepts.)
- Is an explicatum a *"model"* or does it "live" in one? Probably neither. (Explicata are of concepts; they improve explicanda.)
- *Exactness* and *fruitfulness* pull towards mathematization. Mathematization also adds to *continuity 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).

- Explication is *open-ended*; Carnap mentions A. Naess on series of consecutive precisifications of a given concept.

  (Conditionals are still a hot topic!)
- If different explications of the same concept lead to analogous results, then the philosophical conclusions thereof are *robust*.

  (E.g.: Triviality results for *probabilistic* conditional belief and for *ordinal* conditional belief. Convergence on the same logical system.)
- There are various other successful philosophical explications.

  (E.g., explications of: logical consequence, truth, rational belief, confirmation, rational decision,. . . $\hookrightarrow$ summer school!)
- For subjunctive conditionals there must be a different explication of acceptance (remember Adam's examples); and there is.

II. The Physicist's Way

Modeling and Simulation

Another way of doing mathematical philosophy ("the physicist's way").

- By modeling and simulation

  Make a number of (often highly idealized) assumptions and explore their consequences with mathematical and computational methods.

Modeling and Simulation

Another way of doing mathematical philosophy ("the physicist's way").

- By modeling and simulation

  Make a number of (often highly idealized) assumptions and explore their consequences with mathematical and computational methods.

I will deal with just one example.

- how we learn an indicative conditional.

Motivation

- We often learn indicative conditionals of the form "if A, B".
- Here we ask: How should we change our beliefs in the light of this evidence?
- I will discuss several examples below that show that learning a conditional sometimes makes us change our beliefs, and sometimes not. But what is the rational procedure for this?
- There are several proposals discussed in the literature, but in a recent survey, Igor Douven (2012) concludes that a proper general account of probabilistic belief updating by learning (probabilistic) conditional information is still to be formulated.
- Our goal is to provide such an account.

Conditionalization and the Material Conditional

- How should we change our beliefs in the light of this evidence?
- If we want to use Bayesian conditionalization, then we have to formally represent the conditional.
- Perhaps naturally, we use the material conditional and identify $\mathrm{A} \to \mathrm{B}$ with $\neg\mathrm{A} \vee \mathrm{B}$. Popper and Miller (1983) have shown that then

$$P^*(\mathrm{A}) := P(\mathrm{A}|\mathrm{A} \to \mathrm{B}) < P(\mathrm{A})$$

  if $P(\mathrm{A}) < 1$ and $P(\mathrm{B}|\mathrm{A}) < 1$.
- This leads to counterintuitive consequences as, e.g., the sundowners example below demonstrates.
- However, if we do not use the material conditional, then we cannot express the conditional in Boolean terms, and hence we cannot apply conditionalization.
- Question: What can be done?

Stalnaker's Thesis

Stalnaker proposed to identify the probability of a conditional with the conditional probability:

Stalnaker's Thesis

$P(\mathrm{A} \to \mathrm{B}) = P(\mathrm{B}|\mathrm{A})$

- This thesis, which Stalnaker found trivial, has been criticized, most famously perhaps by Lewis who came up with various triviality results.
- Note, however, that Stalnaker's thesis (even if it were true) cannot be applied directly to learning a conditional via Bayes' Theorem. It simply does not tell us how to do this.

A Proposal for a General Recipe

- One way to proceed is to use the conditional probability assignment as a constraint on the new probability distribution P'. Apart from satisfying the constraint, P' has to be as close as possible to the old distribution P, i.e. we want to change our beliefs conservatively.
- Technically, this is done by minimizing the Kullback-Leibler divergence between the posterior and the prior distribution.
- While this might sound like a reasonable (and practicable) proposal, van Fraassen and others have confronted it with counterexamples, most famously the Judy Benjamin example.
- Here I want to show that the proposed procedure works if one additionally makes sure that the causal structure of the problem at hand is properly taken into account.

The Kullback-Leibler Divergence

- Let $\mathrm{S}_1, \ldots, \mathrm{S}_\mathrm{n}$ be the possible values of a random variable S over which probability distributions P and P' are defined.
- The Kullback-Leibler divergence between P' and P is then given by

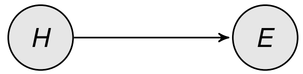
$$D_{KL}(P'||P) := \sum_{i=1}^{n} P'(\mathrm{S_i}) \log \frac{P'(\mathrm{S_i})}{P(\mathrm{S_i})}.$$

- Note that the KL divergence is not symmetrical. So it is not a distance.
- Gabriele Kern-Isberner first applied the KL divergence to the problem we discuss here.

Gabriele Kern-Isberner

Conditionalization

- We introduce the binary propositional variables H and E:
  H: "The hypothesis holds", and ¬H: "The hypothesis does not hold".
  E: "The evidence obtains", and ¬E: "The evidence does not obtain".
- The probabilistic relation between H and E can be represented in a Bayesian Network:



- We set $P(\mathrm{H}) = h$ and $P(\mathrm{E}|\mathrm{H}) = p, P(\mathrm{E}|\neg\mathrm{H}) = q$.

Conditionalization

- Calculate the prior distribution over H and E. ($\overline{x} := 1 - x$)

$$\begin{aligned} P(\mathrm{H},\mathrm{E}) &= hp \quad , \quad P(\mathrm{H},\neg\mathrm{E}) = h\overline{p} \\ P(\neg\mathrm{H},\mathrm{E}) &= \overline{h}q \quad , \quad P(\neg\mathrm{H},\neg\mathrm{E}) = \overline{h}\,\overline{q}. \end{aligned}$$

- Next, we learn that E obtains, i.e. $P'(\mathrm{E}) = 1$.
- We assume that the network stays the same as before. Hence

$$\begin{aligned} P'(\mathrm{H},\mathrm{E}) &= h'p' \quad , \quad P'(\mathrm{H},\neg\mathrm{E}) = h'\overline{p'} \\ P'(\neg\mathrm{H},\mathrm{E}) &= \overline{h'}q' \quad , \quad P'(\neg\mathrm{H},\neg\mathrm{E}) = \overline{h'}\,\overline{q'}. \end{aligned}$$

- From $P'(\mathrm{E}) = h'p' + \overline{h'}q' = 1$, we conclude that $p' = q' = 1$.
- Minimize the KL divergence: $P'(\mathrm{H}) = P(\mathrm{H}|\mathrm{E})$
- Note that Jeffrey conditionalization obtains if one learns E with $P'(\mathrm{E}) =: e' < 1$.

Learning a Conditional

- Let us now apply this methodology to learning a conditional. We start with a prior probability distribution P over the variables A and B. Next, we learn that $\mathrm{A} \to \mathrm{B}$, i.e. we impose the constraint $P'(\mathrm{B}|\mathrm{A}) = 1$ on the new distribution P'. Minimizing the KL divergence between P' and P leads to $P'(\mathrm{A}) \leq P(\mathrm{A})$. In fact, one obtains exactly the same result as for conditioning on the material conditional.
- Note also that $P(\mathrm{A}|\mathrm{B}, \mathrm{A} \to \mathrm{B}) = P(\mathrm{A}|\mathrm{B})$, if A → B is the material conditional. One obtains the same result if one minimizes KL after having learned B and A → B. This observation is related to the *Old Evidence Problem*, which is considered to be a challenge for Bayesianism.

Challenge 1: The Ski Trip Example

Harry sees his friend Sue buying a skiing outfit. This surprises him a bit, because he did not know of any plans of hers to go on a skiing trip. He knows that she recently had an important exam and thinks it unlikely that she passed. Then he meets Tom, his best friend and also a friend of Sue, who is just on his way to Sue to hear whether she passed the exam, and who tells him,

If Sue passed the exam, then her father will take her on a skiing vacation.

Challenge 1: The Ski Trip Example

Harry sees his friend Sue buying a skiing outfit. This surprises him a bit, because he did not know of any plans of hers to go on a skiing trip. He knows that she recently had an important exam and thinks it unlikely that she passed. Then he meets Tom, his best friend and also a friend of Sue, who is just on his way to Sue to hear whether she passed the exam, and who tells him,

If Sue passed the exam, then her father will take her on a skiing vacation.

Recalling his earlier observation, Harry now comes to find it more likely that Sue passed the exam.

Ref.: Douven and Dietz (2011)

Challenge 2: Judy Benjamin Problem

A soldier is dropped with her platoon in a territory that is divided in two parts, the Red Territory (R) and the Blue Territory (¬R) where each territory is also divided in two parts, Second Company (S) and Headquarters Company (¬S), forming four sections of almost equal size. The platoon is dropped somewhere in the middle so she finds it equally likely to be in one section as in any of the others, i.e. $P(\mathrm{R},\mathrm{S}) = \mathrm{P}(\mathrm{R},\neg\mathrm{S}) = \mathrm{P}(\neg\mathrm{R},\mathrm{S}) = \mathrm{P}(\neg\mathrm{R},\neg\mathrm{S}) = 1/4$. Then they receive a radio message:

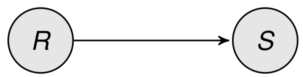
I can not be sure where you are. If you are in Red Territory the odds are 3:1 that you are in the Secondary Company.

How should Judy Benjamin update her belief function based on this communication?

Ref.: van Fraassen (1981)

Judy Benjamin Problem

- We introduce two binary propositional variables. The variable R has the values R: "Judy lands in Red Territory", and ¬R: "Judy lands in Blue Territory". The variable S has the values S: "Judy lands in Second Company", and ¬S: "Judy lands in Headquarters".
- The probabilistic relation between the variables:



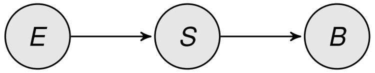
- Learning: $P'(\mathrm{S}|\mathrm{R}) = k \neq 1/2$
- Assume that the network does not change. Then minimizing the Kullback-Leibler divergence yields $P'(\mathrm{R}) < P(\mathrm{R})$, which is not intuitive.

The Challenges Met: The Ski Trip Example

We define three variables:

- E: Sue has passed the exam.
- S: Sue is invited to a ski vacation.
- B: Sue buys a ski outfit.

The causal structure is given as follows:



Additionally, we set $P(\mathrm{E}) = e$ and

$$P(\mathrm{S}|\mathrm{E}) = p_1 \quad , \quad P(\mathrm{S}|\neg\mathrm{E}) = q_1$$
$$P(\mathrm{B}|\mathrm{S}) = p_2 \quad , \quad P(\mathrm{B}|\neg\mathrm{S}) = q_2 .$$

Note that the story suggests that $p_1 > q_1$ and $p_2 > q_2$.

The Ski Trip Example

- Learning: $P'(\mathrm{B}) = 1$ and $P'(\mathrm{S}|\mathrm{E}) = 1$.
- Again, the causal structure does not change.

  Theorem 1: *Consider the Bayesian Network above with the prior probability distribution. Let*

  $$k_0 := \frac{p_1\, p_2}{q_1\, p_2 + \overline{q_1}\, q_2} .$$

  We furthermore assume that (i) the posterior probability distribution P' is defined over the same Bayesian Network, (ii) the learned information is modeled as constraints on P', and (iii) P' minimizes the Kullback-Leibler divergence to P. Then $P'(\mathrm{E}) > P(\mathrm{E})$, iff $k_0 > 1$.

- The same result obtains for the material conditional.

The Ski Trip Example: Assessing k_0

1. Harry thought that it is unlikely that Sue passed the exam, hence e is small.
2. Harry is surprised that Sue bought a skiing outfit, hence

   $$P(\mathrm{B}) = e\,(p_1\, p_2 + \overline{p_1}\, q_2) + \overline{e}\,(q_1\, p_2 + \overline{q_1}\, q_2)$$

   is small.
3. As e is small, we conclude that $q_1\, p_2 + \overline{q_1}\, q_2 := \varepsilon$ is small.
4. p_2 is fairly large (≈ 1), because Harry did not know of Sue's plans to go skiing, perhaps he even did not know that she is a skier. And so it is very likely that she has to buy a skiing outfit to go on the skiing trip.
5. At the same time, q_2 will be very small as there is no reason for Harry to expect Sue to buy such an outfit in this case.
6. p_1 may not be very large, but the previous considerations suggest that $p_1 \gg \varepsilon$.

The Ski Trip Example: Assessing k_0

We conclude that

$$k_0 : = \frac{p_1\, p_2}{q_1\, p_2 + \overline{q_1}\, q_2} = \frac{p_1}{\varepsilon} \cdot p_2$$

will typically be greater than 1. Hence, $P'(\mathrm{E}) > P(\mathrm{E})$.

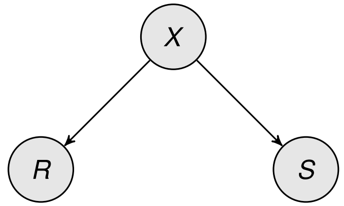
No Causal Structure

- What if no causal structure is imposed?
- We computed this case, i.e. we considered only the three variables B, E and S and modeled the learning of $P'(\mathrm{B}) = 1$ and $P'(\mathrm{S}|\mathrm{E}) = 1$ in the usual way.
- Minimizing the KL divergence then leads to $P'(\mathrm{E}) < P(\mathrm{E})$, i.e. to the wrong result.

The Judy Benjamin Example

We define:

- R: The platoon is dropped in the Red Territory.
- S: The platoon is dropped in the Secondary Company.
- X: Wind comes from a certain direction (or any other cause that comes to mind).



The Judy Benjamin Example

- Learning: $P'(\mathrm{S}|\mathrm{R}) = k \neq 1/2$
- Then the following theorem holds:

Theorem 2: *Consider the Bayesian Network above with a suitable prior probability distribution P. We furthermore assume that (i) P' is defined over the same Bayesian Network, (ii) the learned information is modeled as a constraint on P', and (iii) P' minimizes the Kullback-Leibler divergence to P. Then $P'(\mathrm{R}) = P(\mathrm{R})$.*

No Causal Structure

- Perhaps it is enough to only take the existence of a third variable X into account, without imposing a causal structure. Let us compute this case!
- We find that imposing $P'(\mathrm{S}|\mathrm{R}) = k \neq 1/2$ as a constraint on the posterior distribution and minimizing the KL divergence leads to $P'(\mathrm{R}) < P(\mathrm{R})$, i.e. to the wrong result.

Some General Remarks on Modeling

- Models involve idealizations. Some models are more idealized ("toy models"), others less.
- Many models are embedded in a theory ("models of a theory"), others are not ("phenomenological models"). In our case, Bayesianism is the relevant theory.
- Models are typically connected to data (broadly construed). Sometimes one considers "stylized facts", sometimes human judgments,. . .
- Often computers help to solve the equations of a model ("computer simulations").

References

References

1. Adams, E.W., *The Logic of Conditionals*, Dordrecht: Reidel, 1975.
2. Bennett, J., *A Philosophical Guide to Conditionals*, Oxford: Clarendon Press, 2003.
3. Bovens, L. and S. Hartmann, *Bayesian Epistemology.* Oxford: Oxford University Press, 2003.
4. Carnap, R., *Logical Foundation of Probability*, Chicago: The University of Chicago Press, 1950.
5. Diaconis, P. and S. Zabell, "Updating Subjective Probability", *Journal of the American Statistical Association* 77 (1982): 822–830.
6. Douven, I., "Learning Conditional Information", *Mind & Language* 27 (3) (2012): 239–263.
7. Douven, I., "The Evidential Support Theory of Conditionals", *Synthese* 164 (2008): 19–44.

References

8. Douven, I, "Indicative Conditionals", in: L. Horsten and R. Pettigrew (eds.): *A Companion to Philosophical Logic*, London: Continuum Press, 2011.
9. Douven, I. and R. Dietz, "A Puzzle about Stalnaker's Hypothesis", *Topoi* 30 (2011): 31–37
10. Douven, I. and J. W. Romeijn, "A New Resolution of the Judy Benjamin Problem", *Mind* 120 (479) (2012): 637–670.
11. Edgington, D., "On Conditionals", *Mind* 104 (1995), 235–329.
12. Fraassen, B. van, "A Problem for Relative Information Minimizers in Probability Kinematics, *British Journal for the Philosophy of Science* 32 (1981): 375–379.
13. Kern-Isberner, G. *Conditionals in Nonmonotonic Reasoning and Belief Revision.* Berlin: Springer, 2001.

References

14. Lewis, D.K., “Probabilities of Conditionals and Conditional Probabilities”, *The Philosophical Review* 85 (1976), 297–315. Reprinted in Lewis, D.K., *Philosophical Papers*, Vol. II. Oxford: Oxford University Press, 1986, 133–152.
15. Popper, K. and D. Miller, “A Proof of the Impossibility of Inductive Probability”, *Nature* 302 (1983): 687–688.
16. Ramsey, F.P, “General Propositions and Causality”, 1929. In his *The Foundations of Mathematics and other Logical Essays*, edited by R. B. Braithwaite, London: Kegan Paul, Trench & Trubner, 1931, 237–255.
17. Stalnaker, R.C., “Probability and Conditionals.” *Philosophy of Science* 37 (1970), 64–80.
18. Williamson, J., *In Defence of Objective Bayesianism.* Oxford: Oxford University Press, 2010.