Introduction to Mathematical Philosophy

Hannes Leitgeb, Stephan Hartmann

June 19, 2015

Contents

Week 4: Overview

Overview of Lecture 4: If-then

4.1 Introduction: Conditionals are 'if-then' sentences in natural language. We distinguish between two kinds of conditionals: indicative conditionals and subjunctive conditionals. By considering some examples, we find that whether a conditional seems acceptable to us may also depend on whether the conditional is stated in the indicative or in the subjunctive mood. For the rest of the lecture, we focus on indicative conditionals.

4.2 Analyzing 'if A then B' as 'not A or B' (Part 1): In mathematics and in elementary logic, the indicative conditional 'if A then B' is understood as 'not A or B' (or equivalently as: 'not (A and not B)', 'not A or (A and B)'). We determine what kind of proposition is expressed by 'if A then B', once it is analyzed in this manner.

4.3 Analyzing 'if A then B' as 'not A or B' (Part 2): We argue that understanding 'if A then B' as 'not A or B' is fine as far as the purposes of mathematicians are concerned, and we mention in passing that some ancient Greek philosophers understood conditionals in a similar way. But we also point out that, given this analysis, some strange-looking indicative conditionals actually turn out to be true.

4.4 Conditional Belief: Just as sentences can be conditional, also belief states can be conditional. A person can have a belief in a proposition on the supposition of another proposition: we explain, at first in in informal terms, what 'on the supposition of' means; we repeat some fact about subjective probability from the last lecture; and we mention that also the converse of a theorem from Lecture 3 is true.

4.5 Conditional Probability I: Then we turn to a proposal of what a rational person's degree of belief in a proposition Y on the supposition of another proposition X should be like in probabilistic terms – where P is the person's degree of belief function, the answer is: the conditional probability of Y given X. In this part, we motivate and justify this proposal by gradually approaching the official definition of conditional probability.

4.6 Conditional Probability II: We are finally able to define the notion of conditional probability in precise terms. We draw some basic conclusions about conditional probabilities and the corresponding process of conditionalizing a probability measure on a proposition X (that is being supposed); and we run through a couple of examples.

4.7 Degrees of Acceptability for Conditionals: Taking together what we have learned about indicative conditionals and conditional probability, we put forward two theses on how to measure the rational degree of acceptability (to measure how strongly one should accept a conditional). Assuming that both of these theses are true, we derive an equation of the form: $P(X \to Y) = P(Y|X)$. That is: The (unconditional) subjective probability of the proposition $X \to Y$ equals the conditional probability of Y given X.

4.8 Lewis' Triviality Theorem: By a theorem of David Lewis, given some plausible background assumptions, it follows from this equation that the degree of acceptability of a conditional always equals the degree of belief in its then-part, which is absurd. We prove Lewis' theorem in a fairly detailed manner.

4.9 Reponses to the Triviality Theorem: A diagnosis of what has led us to the absurd conclusion of Lewis' theorem reveals different options of what assumptions to give up: One such option corresponds to the Suppositional Theory of conditionals according to which indicative conditionals do not express propositions, they are not true or false, their degrees of acceptability are given by corresponding conditional probabilities, but their degrees of acceptability are not determined in any way by whatever unconditional degrees of belief in conditional propositions.

4.10 Conclusions: We found that in mathematics an indicative conditional of the form 'if A then B' is analyzed as 'not A or B'. But it is questionable whether this is also a successful analysis of indicative conditionals as they are used outside of mathematics. Over and above sentences, also beliefs can be conditional; we made this precise in terms of conditional probabilities, and we stated and proved a famous theorem by David Lewis which is sometimes interpreted as showing that indicative conditionals do not express propositions at all and do not have truth values (at least not in the usual sense of the word), but their degrees of acceptability can still be determined by means of conditional probabilities.

Chapter 4

Week 4: If-then

4.1 Introduction (09:13)

Welcome to the fourth lecture of our Introduction to Mathematical Philosophy! In the third lecture, we investigated the rationality of beliefs, and of degrees of belief, with the help of mathematical methods. In particular, we found that if one distributes one's degree of beliefs over propositions rationally, then this distribution must satisfy the mathematical laws of subjective probability.

Today we will build on this when we turn to another classical topic of philosophy: if-then sentences – conditionals. We will see that we do not just assert conditionals in mathematics, science, philosophy, and in our daily communicative practice, but indeed our beliefs themselves can be conditional. And how acceptable a conditional in natural language seems to us might depend on what our conditional beliefs are like. From this we will be able to draw some surprising conclusions about conditionals with the help of probability theory again. If that isn't fun, then this is no conditional.

(Slide 1)

A conditional is an 'If… then…' sentence: a sentence that is of the form

- If A, then B

or which has the same meaning as a sentence of that form, such as, e.g.,

- If A, B
- B if A
- In case that A it holds that B

  ⋮

In the following, if we want to talk about an arbitrary conditional, we will simply say

(Slide 2)

In short:

- $A \to B$

So putting the '$\to$' symbol between a descriptive sentence A and another descriptive sentence B is just a short way of saying: if A then B.

(Slide 3)

For instance:

- If Conny comes to the party, then I will talk to her there.
- If Oswald did not kill Kennedy, someone else did.
- If Shackleton had known how to ski, then he would have reached the South Pole in 1909.
- If Oswald had not killed Kennedy, someone else would have.

> The Oswald conditionals were introduced as examples into the literature by the US-American philosopher Ernest Adams. We will return to his work later in this lecture.
> The Shackleton example is due to the British philosopher Jonathan Bennett. Ernest Shackleton was an Irish explorer. In 1901 he advertised one of his expedition in the Times as follows: "Men wanted for hazardous journey. Small wages. Bitter cold. Long months of complete darkness. Constant danger. Safe return doubtful. Honour and recognition in case of success."

As grammarians would say, the first two of these conditionals are formulated in the indicative mood:

If Conny comes to the party, then I will talk to her there.

If Oswald did not kill Kennedy, someone else did.

They are of a form that makes it clear that their if-parts are meant to speak about what is actually case: If it actually turns out that Conny comes to the party, then I will talk to her there. If it actually turns out that Oswald did not kill Kennedy, then clearly someone else did. Such conditionals are typically constructions of the kind does-will or did-did or is-is or the like: 'If ... does so and so, then ... will so and so', 'If ... did so and so, then ... did so and so', 'If ... is so and so, then ... is so and so', and so on.

In contrast, the other two conditionals from before are stated in the subjunctive mood:

If Shackleton had known how to ski, then he would have reached the South Pole in 1909.

If Oswald had not killed Kennedy, someone else would have.

They are of a form that makes it clear that their if-parts are not meant to speak about what is actually case: in normal circumstances, when I assert

If Shackleton had known how to ski, then he would have reached the South Pole in 1909.

I convey, first of all, that Shackleton did not actually know how to ski; but additionally I say: if he had known how to ski, and hence something would have been the case that is not actually the case, then under these merely possible conditions he would have reached the South Pole in 1909.

Accordingly, when I assert

If Oswald had not killed Kennedy, someone else would have.

I convey that, according to the available evidence, Oswald did kill Kennedy; but if that were not so, if he had not killed Kennedy, then someone else would have killed Kennedy, or at least that is what I claim when I assert this if-then sentence – whether I am right about this or not. So the if-parts of subjunctive conditionals are usually assumed or presupposed to be false: that is why one often speaks of such subjunctive conditionals as 'counterfactuals': for their if-parts run counter to the facts. Such conditionals typically involve the terms 'had' or 'were' in the if-part and the term 'would' in the then-part: as in, for example, 'had known how to ski' and 'would have reached the South Pole in 1909'.

Later we will be interested especially in the conditions under which a conditional is acceptable to us: when we would be happy to assert it, to reason upon it, and the like. What is more, we will wonder later how we can measure how acceptable a conditional is to us: how we can determine the degree of acceptability of a conditional. And for questions like that, the grammatical mood of a conditional can actually play a significant role.

(Slide 4)

Reconsider the *indicative*

- If Oswald did not kill Kennedy, someone else did.

  ↪ Acceptable

in comparison with the *subjunctive*

- If Oswald had not killed Kennedy, someone else would have.

  ↪ Not acceptable

The first sentence, the indicative conditional, is acceptable from my point of view: I firmly believe that Kennedy was actually killed; so if Oswald did not do it, then someone else did.

However, the second conditional, the subjunctive one, is not acceptable, at least to the best of my knowledge: if the world had evolved differently than it actually did, if Oswald had been stopped before shooting Kennedy, or if he had changed his mind in time, or the like, and if therefore Oswald had not killed Kennedy, I am not sure at all what would have happened: for all I know, Kennedy might have lived a happy life and would have died only much later from a heart attack. Claiming that

If Oswald had not killed Kennedy, someone else would have.

is thus not acceptable to me.

We find that, in this case, the indicative conditional is acceptable while the subjunctive one is not, even though the if-parts of both of them talk about Oswald not killing Kennedy, and the then-parts of both of them talk about someone else killing Kennedy. The only difference between the two conditionals is their grammatical mood. This shows that whether or not a conditional in natural language is acceptable to us does not just depend on the propositions that are expressed by its if-part and its then-part, respectively, but the acceptability of a conditional may also depend on the mood in which the conditional is formulated.

Subjunctive conditionals constitute a really interesting topic, and they are particularly important in metaphysics, as they concern mere possibilities: possible worlds that differ from the actual world. Formal methods are again crucial if one wants to understand them properly: in order to determine under what conditions subjunctive conditionals are true and how we ought to reason with them, methods from logic and from the so-called possible worlds semantics of subjunctive conditionals – something very much as our reconstruction of propositions and belief in terms of sets of possible worlds in the last lecture – have become standard tools in today's metaphysics, therefore.

But I will nevertheless put subjunctive conditionals to one side now: we cannot deal with everything here that is philosophically fascinating – unfortunately. Instead, for the rest of this lecture, I will focus solely on indicative conditionals: those conditionals whose if-parts concern, in a pretty much straight-forward sense, the actual world, rather than some non-actual possibilities. This might sound as if indicative conditionals should be much simpler to handle philosophically than subjunctive conditionals, but we will see that it is already difficult enough to understand them properly.

If you want to read more about conditionals, take a look at Dorothy Edgington's excellent corresponding entry in the Stanford Encyclopedia:
http://plato.stanford.edu/entries/conditionals/
(But I would recommend doing so only after this lecture, as the entry is not easy).

Quiz 31:
Compare (this is an example from Jonathan Bennett):
(1) If Shakespeare did not write Hamlet, then some aristocrat did.
(2) If Shakespeare had not written Hamlet, then some aristocrat would have.
Which one is subjunctive, which one is indicative?
Which one sounds acceptable to you?
Solution

4.2 Analyzing 'if A then B' as 'not A or B' (Part I) (09:59)

It is not so easy to understand indicative conditionals as far as they are used in natural language discourse. In certain more regimented scientific languages, however, indicative conditionals are often much less problematic to analyze. Take the language of modern mathematics: mathematicians learn to understand the indicative if-then – subjunctive conditionals do not play much of a role in mathematics anyway – as follows:

When a mathematician says

(Slide 5/1)

$A \to B$:

in a theorem or in a proof, then they can always be taken to be saying

(Slide 5/2)

- not (A and not B)
- (not A) or B
- (not A) or (A and B)

or written slightly more formally:

(Slide 6)

- $\neg(A \wedge \neg B)$
- $\neg A \vee B$
- $\neg A \vee (A \wedge B)$

You might think: why do I state three kinds of sentences here? When they assert 'if A then B', can mathematicians be taken to be saying

not (A and not B)

or rather

not A or B

or

not A or (A and B)?

The point is: it does not matter really. If a mathematician says '$A \to B$' then what this means can be reformulated in each one of these three ways equivalently – for the proposition that is expressed by

$\neg(A \wedge \neg B)$

is the same as the proposition expressed by

$\neg A \vee B$

which in turn is the same as the proposition expressed by

$\neg A \vee (A \wedge B)$.

What is that proposition? Let us apply what we have learned in the last lecture: First we fix a non-empty set W of possible worlds again; propositions are then considered to be sets of possible worlds, that is, subsets of W; the usual logical operations on propositions coincide with certain set-theoretic operations on these propositions or sets: negation corresponds to taking set-theoretic complements, conjunction – the 'and' – corresponds to taking intersections of sets, and disjunction – the 'or' – corresponds to unions of sets. Descriptive sentences express propositions in this sense, that is, they express certain sets of worlds. We also said that a proposition is true if and only if it includes the actual world as a member, but presently this is less important; what we need to do in the present case is to consider two descriptive sentences – A on the one hand, B on the other – which we then put together by means of logical symbols, and finally want we determine what propositions are expressed by the resulting logically complex sentences.

So let us assume that

(Slide 7)

A expresses the set X of worlds.

B expresses the set Y of worlds.

The proposition expressed by

- $\neg(A \wedge \neg B)$

is nothing else than

- $W \setminus (X \cap (W \setminus Y))$

(See Figure 4.1.)

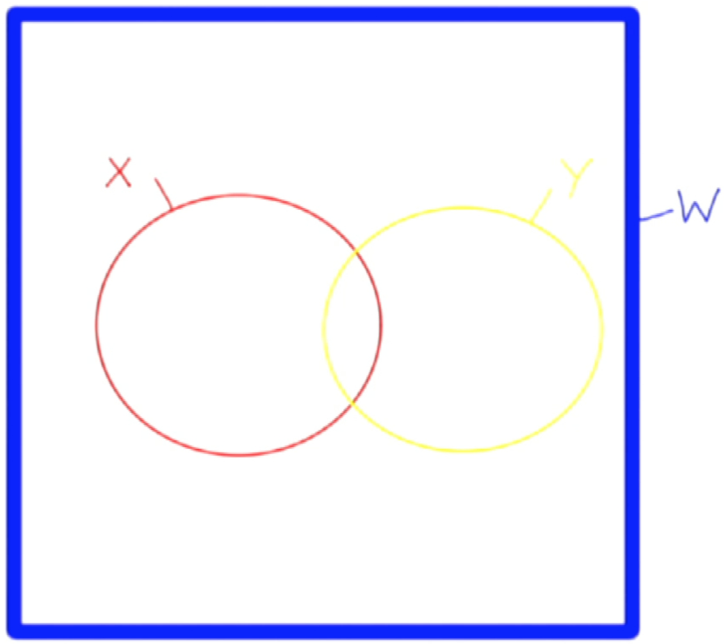

Figure 4.1: X, Y as subsets of W

that is:

The negation of B expresses the complement of Y with respect to W, that is, the set W without Y; A expresses X; $A \wedge \neg B$ expresses therefore the intersection of X with taking W without Y; and of that intersection we need to take the complement again with respect to the set W of all possible worlds, in correspondence with the negation symbol that is applied to the whole conjunction $A \wedge \neg B$. In the picture (see Figure 4.2) you see which set of worlds results from that, and that is the proposition that is expressed by $\neg(A \wedge \neg B)$.

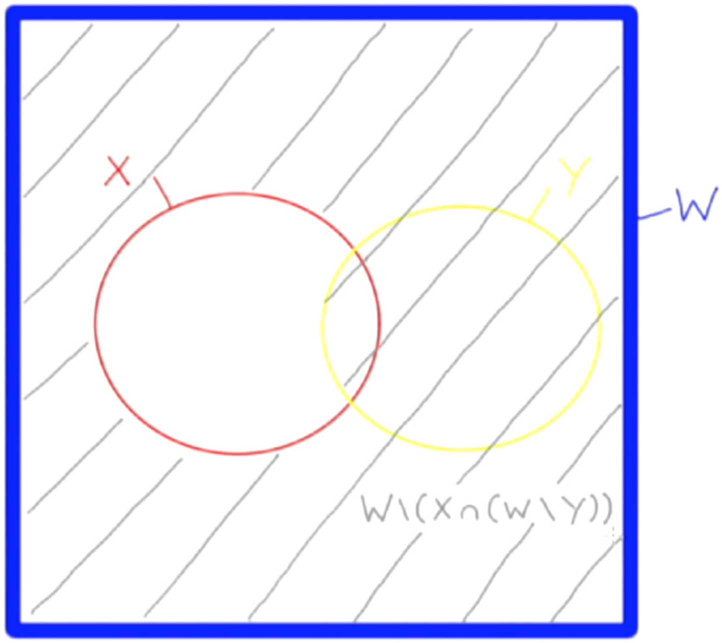

Figure 4.2: $W \setminus (X \cap (W \setminus Y))$

And now, if you do the same for the second sentence, $\neg A \vee B$, then you will find that it expresses the set

$(W \setminus X) \cup Y$

of possible worlds, which is – precisely the same set of possible worlds again (see Figure 4.2)!

And if you determine the proposition that is expressed by the third sentence, $\neg A \vee (A \wedge B)$, then that proposition is the set

$(W \setminus X) \cup (X \cap Y)$

which is once again the same proposition (see Figure 4.2).

So the three descriptive sentences

(Slide 10)

$A \to B$:

- $\neg(A \wedge \neg B)$
- $\neg A \vee B$
- $\neg A \vee (A \wedge B)$

all express one and the same proposition; in fact they do so whatever the propositions are like that are expressed by A and by B, respectively: there is no way that these three sentences could not express the same proposition. Since the proposition expressed by each one of the sentences is identical to the proposition expressed by each of the other, it also holds that the proposition expressed by one of the sentences is a subset of the proposition expressed by any other of them, and the other way around; in other words: according to what we had said in the last lecture, the three sentences

$\neg(A \wedge \neg B)$

$(\neg A) \vee B$

$(\neg A) \vee (A \wedge B)$

are pairwise logically equivalent, which means that as far as matters of truth or falsity are concerned, it does not really matter whether one asserts one of them or the other, since they are true in precisely the same possible worlds anyway, and they are false in precisely the same possible worlds.

Remark on $\neg A$ or B: In the logical literature, the conditional 'if A then B' that is defined by $\neg A$ or B (or in a logically equivalent manner) is called the 'material conditional'. That is also the kind of conditional that one learns about first in introductory logic courses. And in the area of conditional logic – one of the areas of philosophical logic – one introduces and studies types of conditionals that get much closer to the 'if-then' of natural language than the material conditional does.

Remark on logical equivalence: Logical equivalence is simply logical implication in both directions. Sentence A is logically equivalent to sentence B if and only if A logically implies B, and B logically implies A. That is: Sentence A is logically equivalent to sentence B if and only if the following holds: if A is true then it must be the case that B is true as well, and if B is true then it must be the case that also A is true.
As I had mentioned early on in this course, the precise details on logical implication are the subject of logic courses and/or logic textbooks. In a logic course, it is shown how one can define the relationsship of logical implication, and hence also the relation of logical equivalence, in exact formal terms.

O.k., now back to my claim from above: my claim was that when a mathematician states ‘if A, then B’ then what they express is the very proposition, the very set of possible worlds, that can be expressed in terms of

$\neg(A \wedge \neg B)$

or alternatively by means of

$\neg A \vee B$

or alternatively by means of

$\neg A \vee (A \wedge B)$

amongst various other logically equivalent formulations.

But why should anyone mean one of these three sentences when they say ‘if A then B’? Basically, the thought is this: Generally, when we assert an indicative conditional ‘if A then B’, we really only care about what is the case when A, the if-part, is true: after all, the if-part of an indicative conditional is meant to apply to the actual world: if that is so, if A is true, then we want to say something interesting and informative about what is the case then, namely, that B holds; we can be right about this or wrong, but in any case that is what we want to say. In other words: in the case where the if-part A of an indicative conditional ‘if A then B’ is true, we think that the conditional applies, and the conditional tells us more about this case then, namely that also B is the case. In the other possible case, when the if-part A of an indicative conditional ‘if A then B’ is not true, the conditional does not really apply, and accordingly we do not care then at all about the conditional anymore.

Now the idea of interpreting

if A then B

as

$(\neg A) \vee (A \wedge B)$

– the third of our sentences from before – is precisely this: by asserting 'if A then B' one really says that, first case, either the conditional does not apply, so that's the uninteresting case – the case when A is false, that is, $\neg A$ is true – or, second case, the interesting one – the conditional does apply in which case one claims that also B is true. So 'if A then B' amounts to:

(Slide 11)

So 'if A then B' amounts to:

- Either case 1: The conditional does not apply ($\neg A$).
- Or case 2: The conditional does apply (A) and B is true (B).

Or more briefly:

$\neg A \vee (A \wedge B)$

According to this view of indicative conditionals, when we say 'if A then B' we mean

Either A is not true anyway; or it is true and B holds.

The other two sentences from above are simply alternative formulations that ultimately capture the same idea:

(Slide 12)

$A \to B$:

- $\neg(A \wedge \neg B)$
- $\neg A \vee B$

If one says 'if A then B' one wants to exclude one logically possible case: the case that we care about, that is, when A is true, and where additionally B is false: for we want to say that in the very case when A is true also B is true. So we want to rule out that $A \wedge \neg B$: saying 'if A then B' amounts to

$\neg(A \wedge \neg B)$

The other sentence, $\neg A \vee B$, is really just a shorter way of saying $\neg A \vee (A \wedge B)$ again: the one is true if and only if the other one is. Indeed we have already seen before that all of these three sentences express the same propositions, have the same meaning, are true in precisely the same worlds, anyway.

Quiz 32:
Please show that $\neg A \vee B$ is logically equivalent to $\neg A \vee (A \wedge B)$. What you have to do is: First assume that $\neg A \vee B$ is true, and then prove on that basis that also $\neg A \vee (A \wedge B)$ is true. Secondly, the other way around: assume that $\neg A \vee (A \wedge B)$ is true, and then prove on that basis that also $\neg A \vee B$ is true.
Solution

4.3 Analyzing ‘if A then B’ as ‘not A or B’ (Part II) (12:34)

Other than such general considerations, there are also reasons for treating indicative conditionals in terms of $\neg A \vee B$ which are more specific to the practice of mathematics:

When a mathematician asserts $A \to B$, then this is normally in a context where she has also proven $A \to B$.

(Slide 13/1)

Assume that a mathematician has proven $A \to B$, that is,

- $\neg A \vee B$

Now, say, she is also able to prove $\neg A$ later: what will happen? Nothing in particular. Since for her the sentence $A \to B$ is nothing else really than

$(\neg A) \vee B$

using one of the equivalent versions from above,

she will happily embrace her newly proven sentence $\neg A$, which is in fact more informative then the disjunction

$(\neg A) \vee B$

that she had already proven.

Beforehand she had shown that $\neg A$ is the case or something else is the case, while now she is even able to prove something more specific, that is, that $\neg A$ is the case. That’s nice, and nothing bad follows from this. She can simply forget about $A \to B$, that is, $(\neg A) \vee B$, if she wants to, since by now she knows more than that, and that is exactly what we would expect to be the case, since we are dealing with the uninteresting case of an indicative conditional here, the case in which the if-part of ‘if A then B’ is false, as follows from the mathematician having proven $\neg A$.

(Slide 13/2)

Case 1:

- She also proves $\neg A$.

  Nothing bad follows.

Now consider the second possible case: the mathematician has already proven $A \to B$ and later she also manages to prove A.

(Slide 14)

Assume that a mathematician has proven $A \to B$, that is,

- $\neg A \lor B$

Case 2:

- She also proves A.

  $$\left.\begin{array}{l} A \\ A \to B \\ \hline B \end{array}\right\} \text{logically valid}$$

That's then the interesting case that we do care about: by proving A she knows that A is true, in which case the conditional $A \to B$ that she had proven before applies: and now the conditional should tell her that B is the case. From taking the proofs of $A \to B$ and of A together and applying the logically valid argument

$$\begin{array}{l} A \\ A \to B \\ \hline B \end{array}$$

she should be able to prove B. And that is exactly what follows if she is analyzes '$A \to B$' as '$(\neg A) \lor B$'. For the argument

$$\begin{array}{l} A \\ \neg A \lor B \\ \hline B \end{array}$$

is indeed logically valid: given that A is true and also $\neg A \lor B$ is true, it follows that $\neg A \lor B$ must be true because B is true, since $\neg A$ has already been ruled out by A being true; so given the truth of both A and $\neg A \lor B$, the truth of B follows logically.

Once again, everything is just as it should be: we are dealing with the interesting case of an indicative conditional here, the case in which A is true, and by understanding conditionals as suggested before, the mathematician can do precisely what she was meant to do.

There is one final possible case: the mathematician has already proven '$A \to B$', but she is neither able to prove A nor is she able to prove $\neg A$.

(Slide 16)

Assume that a mathematician has proven $A \to B$, that is,

- $\neg A \lor B$

Case 3:

- She neither proves A nor $\neg A$.

  Nothing bad follows.

That is perfectly possible: after all, provability is not the same a truth. One of A or not A is true, but it might well be that our mathematical principles, our axioms, do not tell us which of them is true: they leave that question open. This is just like in the case of rational belief where we found that it is rationally possible to suspend judgment on a subject matter: neither to believe A nor to believe 'not A'. And similarly our axioms of mathematics might, as it were, neither believe A nor believe 'not A'.

Anyway: say, she has already proven $A \to B$, but she is neither able to prove A nor is she able to prove $\neg A$. What is going to happen? Not much really. Since she is not able to prove A, she will not be able to derive B by applying the conditional $A \to B$; and because she is not able to prove $\neg A$, she will not be able to replace $A \to B$, that is, $\neg A \lor B$, by the more specific information $\neg A$ either. But once again nothing bad follows from this: $A \to B$, that is, $\neg A \lor B$, will belong to her stock of proven sentences, but she will not be able to infer much from this, or in any case, nothing that would affect the practice of mathematicians in any problematic manner. Indeed, the logical rules that are valid for $\neg A \lor B$ correspond very nicely to the logical rules that mathematicians want to apply to $A \to B$ in any of their proofs.

Remark on proving A and proving $\neg A$: What about the remaining case, that is, when the mathematician proves both A and $\neg A$? That would mean that mathematics is inconsistent, which we believe is not the case. And if it were the case, then we would have greater problems than worrying about $A \to B$.

We find that at least as far as proofs and provability in mathematics are concerned, treating $A \to B$ as

$\neg A \lor B$

or as any logically equivalent reformulation thereof is perfectly respectable: it does the job, and it does not lead to any trouble for mathematicians. That is why it is the standard convention in mathematics these days to analyze the indicative ‘if-then’ in this way.

And there is even very traditional philosophical support for this: already some philosophers of the so-called Dialectical School, one of the famous ancient Greek schools of philosophy in the fourth and the third century before Christ, understood conditionals in precisely that way; most famously, Philo the logician did so, when he claimed that a conditional is false precisely in the case when its if-part is true and its then-part is false, and it is true in all other cases; that is:

$A \rightarrow B$

really says that one possible case is excluded, that is, the case in which A is true and B is false. That is, $A \rightarrow B$ really says

$\neg(A \wedge \neg B)$

just as today’s mathematicians would have it.

Remark on the Dialectical School: If you want to know more about the Dialectical School and its understanding of conditionals, take a look at
http://plato.stanford.edu/entries/dialectical-school/
and
http://plato.stanford.edu/entries/logic-ancient/#DioCroPhiLog.

So if some of the ancient Greek philosophers were happy with this, and if it is common practice in modern mathematics, why not think that the indicative conditional

$A \rightarrow B$

is logically equivalent to

$\neg(A \wedge \neg B)$

across the board, for all indicative conditionals in natural language, rather than just for indicative conditionals in certain parts of language, such as the language of mathematics?

Indeed, some philosophers have argued that the indicative if-then can always be analyzed logically in the way that was explained before, ‘if A then B’ expresses precisely the proposition that we described before, and – taking up Lecture 2 again – a Tarskian truth condition for indicative conditionals in a language L can be stated as follows:

(Slide 18)

- (case: x = ‘if’ + y + ‘then’ + z)

  if there is a sentence y of L and a sentence z of L, such that x is the result of putting together ‘if’, with y, with ‘then’, and with z, then

  x is true if and only if y is not true or z is true;

  (Equivalently:

  if x is the result of putting together ‘if’, with a sentence y of L, with ‘then’, and with a sentence z of L, then

  x is true if and only if y is not true or z is true.)

  (Equivalently, and most precisely:

  for all sentences y of L, for all sentences z of L, if x is the result of putting together ‘if’, with y, with ‘then’, and with z, then

  x is true if and only if y is not true or z is true.)

That is: The indicative conditional $A \rightarrow B$ is true if and only if A is not true or B is true, or in other words, if and only if $\neg A \vee B$ is true.

So is there anything that does speak against this view, at least outside of mathematics? Yes.

Firstly, the view entails that each of the following indicative conditionals turns out to be true:

(Slide 19)

- If the moon is made of green cheese, then 2+2=4.
- If the moon is made of green cheese, then it is not the case that 2+2=4.
- If the moon is not made of green cheese, then 2+2=4.

The first two of these conditionals turn out to be true, because their if-part is false: or, put differently, if we are asserting them, then we are really claiming that

The moon is not made of green cheese or 2+2=4.

The moon is not made of green cheese or it is not the case that 2+2=4.

(Slide 20)

- If $\underbrace{\text{the moon is made of green cheese}}_{A}$, then $\underbrace{\text{2+2=4}}_{B}$.
- If $\underbrace{\text{the moon is made of green cheese}}_{A}$, then $\underbrace{\text{not 2+2=4}}_{B}$.

correspond to, respectively:

- $\underbrace{\text{The moon is not made of green cheese}}_{\neg A}$ or $\underbrace{\text{2+2=4}}_{B}$.
- $\underbrace{\text{The moon is not made of green cheese}}_{\neg A}$ or $\underbrace{\text{not 2+2=4}}_{B}$.

Since the moon is in fact not made of green cheese, each of these or-sentences, each of these disjunctions, is true, and hence the corresponding indicative conditionals are true, if we analyze indicative conditionals in that way.

Remark on 'The moon is not made of green cheese or $2 + 2 = 4$': Actually, that sentence is even true for two reasons, as it were: The moon is indeed not made of green cheese; and $2 + 2$ is indeed equal to 4.

And the third of the conditionals is true, because its then-part is; or, in other words: we are really asserting

The moon is (not not) made of green cheese or $2 + 2 = 4$

which is true, simply because it is true that $2 + 2 = 4$.

(Slide 21)

- If $\underbrace{\text{the moon is not made of green cheese}}_{A}$, then $\underbrace{\text{2+2=4}}_{B}$.

corresponds to:

- $\underbrace{\text{The moon is not not made of green cheese}}_{\neg A}$ or $\underbrace{\text{2+2=4}}_{B}$.

So if the 'if A then B' is understood as 'not A or B' then each of these three indicative conditionals comes out true, which sounds odd – to say the least.

As I said, this distribution of truth values is odd. But it is not quite as odd as one might think. For this only means that, e.g., the material conditional 'If the moon is made of green cheese, then it is not the case that $2 + 2 = 4$' comes out as true (because 'The moon is not made of green cheese or it is not the case that $2 + 2 = 4$' is true), while it does *not* mean that anyone would be forced to assert in everyday communication that conditional 'If the moon is made of green cheese, then it is not the case that $2 + 2 = 4$' with this material conditional understanding. Compare the following situation: Assume that I know A to be true and I also know B to be false. And then I assert, while talking to someone: 'A or B'. I have asserted a true sentence (since A is true). But what I say is still misleading to my communication partner: for if she hears that I say 'A or B', she assumes that I do not know which of the two – A or B – is true, for if I knew which one were true, then – being a good communication partner – I would have given her the more specific information from the start (such as: 'A'). By asserting 'A or B' I therefore violate a norm of good communication, even though what I say is true. This shows that in order to be helpful in one's communication, it is not good enough to assert truths: one should assert truths that are also not misleading in any way to one's communication partners. For the same reason, it does not automatically follow from a material understanding of 'If the moon is made of green cheese, then it is not the case that $2 + 2 = 4$' that one would ever be obligated to assert that conditional in a reasonable discourse. If anything, knowing that the moon is not made of green cheese, and given the material understanding of the conditional, I should assert 'The moon is not made of green cheese' instead of the less specific 'If the moon is made of green cheese, then it is not the case that $2 + 2 = 4$' (that is: the moon is not made of green cheese or it is not the case that $2 + 2 = 4$). Such norms on good communication, and what pragmatic inferences (called 'implicatures') we may draw from them, was studied in detail by the British philosopher H. Paul Grice: check out
http://plato.stanford.edu/entries/implicature/
and
http://plato.stanford.edu/entries/grice/
for more on this.

Furthermore, the argument in favour of the analysis, the argument that was specific to mathematics – why this analysis is fine within mathematics – was based on the fact that in mathematics the only thing that counts ultimately is what one can prove; and in that respect we found the analysis to be unproblematic. But in everyday contexts we usually assert everyday indicative conditionals that we cannot prove: what we assert are not mathematical conditionals, normally, anyway; and we are usually not certain about them either, nor about their if-parts, nor about their then-parts. We might accept an indicative conditional to a more or less high degree, and we might assign certain degrees of belief to their if-parts and their then-parts, but usually these degrees will be strictly between 1, 100%, and 0, 0%.

So the fact that understanding

$A \to B$

in terms of

not A or B

is fine as far as matters of provability are concerned does not show that it is fine also as far as matters over and above provability are concerned. Maybe for everyday reasoning in terms of everyday indicative conditionals the analysis can be harmful in some way, who knows?

We will see at the end of this lecture that there is actually an additional argument that might be interpreted as showing that indicative conditionals cannot be analyzed in terms of 'not A or B' or the like – in fact, the argument might even be taken to show that indicative conditionals cannot be analyzed in terms of anything that expresses a proposition, in terms of anything that is true or false. Sounds surprising? Well, it sure is – it came as a big surprise also to the logicians and philosophers of language who were thinking about conditionals in the 1970s when the argument was discovered.

But before we turn to that argument, we first need to continue our story on belief, and especially on degrees of belief, from the last lecture by adding to it a conditional component: for also our beliefs can be conditional, and we need to understand what is meant by this, and what purposes such conditional beliefs might serve.

Quiz 33:
(1): Picture the set W of all possible worlds in terms of a square again; draw two distinct but intersecting circles in the square – one representing the proposition X, the other one representing the proposition Y. Assume that sentence A expresses proposition X and sentence B expresses proposition Y. Now consider every logically possible position that the actual world might have in the diagram (there are four regions to consider): determine for each position the corresponding truth values of A (or X), B (or Y), and $\neg A \vee B$ (or $\neg X \vee Y$), where 'true' means 'true at the actual world', and 'false' means 'false at the actual world'.
(2): When e.g. a philosopher puts forward an argument of the form
(P1) Not B.
(P2) If A then B.
Therefore: (C) Not A.
we regard the argument as logically valid. (Remember the additional problem set for Lecture 1.) Prove that indeed this argument form turns out to be valid, if we understand 'If A then B' as 'Not A or B' ($\neg A \vee B$).
Solution

4.4 Conditional Belief (08:09)

Many of the sentences by which we communicate are conditional. But also our mental attitudes towards propositions can be conditional. I desire to stay at home tomorrow if the weather is lousy then. I believe that the weather will be lousy tomorrow given that the weather forecast says so. This does not mean that I desire to stay at home tomorrow unconditionally – I only have that desire on the assumption that the weather will be lousy. And it doesn't mean either that I believe the weather to be lousy tomorrow unconditionally – that is: I do not believe this to be so in the sense discussed in the last lecture – but I only believe it on the supposition that the weather forecast will predict the weather to be lousy. In fact, if one thinks about it, many of our desires and beliefs may be expected to be conditional in that sense: we desire or believe something given that something else is the case. Since we do not know as yet whether that something else will be the case or not, we do not have the corresponding unconditional desires or beliefs: instead maybe we are merely prepared to have such unconditional desires or beliefs in case we find out that this something else is the case. So, this is a little like in the case of the hypothetical mathematician that we were dealing with before: once she has proven $A \to B$, she is prepared to react in some way once more information comes in; in particular, if she is also able to prove A, then she is prepared to conclude B. But normally she could not prove B just by having proven $A \to B$. As far as our mental states are concerned, it seems plausible to assume that in many case we are in analogous conditional states of "preparedness", and conditional desires and conditional beliefs are simply special cases of such conditional attitudes. And what I have just said about all-or-nothing conditional desires and all-or-nothing conditional beliefs should also hold when we turn to conditional degrees of desire and conditional degrees of belief, much as in the last lecture when we were able to complement our systematic account of rational all-or-nothing belief by a systematic account of rational degrees of belief.

In the following, we will work out an account of conditional belief just for conditional degrees of belief: conditional subjective probability. And just as in the last lecture, we will only be interested again in inferentially perfectly rational persons and their conditional degrees of belief.

Remark on 'inferentially perfectly rational': From this point we will mostly suppress this qualification, since by now it should have become clear that we are concentrating on inferentially perfectly rational persons and that the degree of belief function of such persons are subjective probability measures, that is, obey the laws of probability, as argued in Lecture 3.

Let us start by returning to our example of a subjective probability measure from the last lecture.

Back then, we were presupposing a set W of possible worlds which consisted of 8 worlds: $\{w_1, \ldots, w_8\}$. And we considered a mapping B which looked like this:

(Slide 22)

$B(w_1) = 1/15$

$B(w_2) = 1/3$

$B(w_3) = 1/15$

$B(w_4) = 1/15$

$B(w_5) = 1/3$

$B(w_6) = 1/15$

$B(w_7) = 1/15$

$B(w_8) = 0$

The sum of all these values of B,

$1/15 + 1/3 + 1/15 + 1/15 + 1/3 + 1/15 + 1/15 + 0$

was 1; and each single value is a real number not less than 0 and not greater than 1.

Then we assumed the degree of belief function of an inferentially perfectly rational person to be determined by that mapping B so that the person's degree of belief in whatever proposition X was the sum of the values of B on worlds in X:

(Slide 23)

For all propositions X (over W):

$$P(X) = \sum_{w \text{ in } X} B(w).$$

Indeed we formulated as a theorem that every rational degree of belief function P on sets of worlds – where the total set W of possible worlds is finite – must arise from some such function B on worlds in this manner, and in our example P was assumed to arise from the function B that is given above.

Here is a fact that I did not stress in the last lecture, just to keep things as simple as possible: it is not just so that for every rational degree of belief function P on propositions there is such a unique corresponding function B on worlds that determines P in the manner explained before, but this also works the other way around: given a function B on worlds in a finite total set W of possible worlds, such that summing up the B-values of all worlds yields the numerical value 1 and each single B-value is greater than or equal to 0 and less than or equal to 1, one can always determine a function P on propositions by means of

$P(X)$ is defined as the sum of all $B(w)$, for which w is a member of X

and the P function on propositions that is defined in this way will then automatically satisfy all of the properties that we had postulated to hold for the degree of belief functions of inferentially perfectly rational persons, that is, the axioms of probability:

(Slide 24)

For all propositions X (over W):

$$P(X) = \sum_{w \text{ in } X} B(w).$$

- $P(W) = 1$;
- for all propositions X (over W):

  $P(X)$ is a real number, such that $0 \leq P(X) \leq 1$;
- for all propositions X, Y (over W):

  if $X \cap Y = \{\}$, then $P(X \cup Y) = P(X) + P(Y)$.

So it doesn't just hold that for every P with these properties there is a unique corresponding B, but also: for every such B, one can define a corresponding P and that P is again uniquely determined.

Remark on this converse theorem: This is like in the case of all-or-nothing belief, where we first stated and proved in 3-8 a theorem that led us from Bel to B_W, and then in the quiz at the end of 3-8 we also saw how to get conversely from B_W to Bel.

Alright: Let P now be the degree of belief function that corresponds to the B function from our example – so P looks like this (see Figure 4.3):

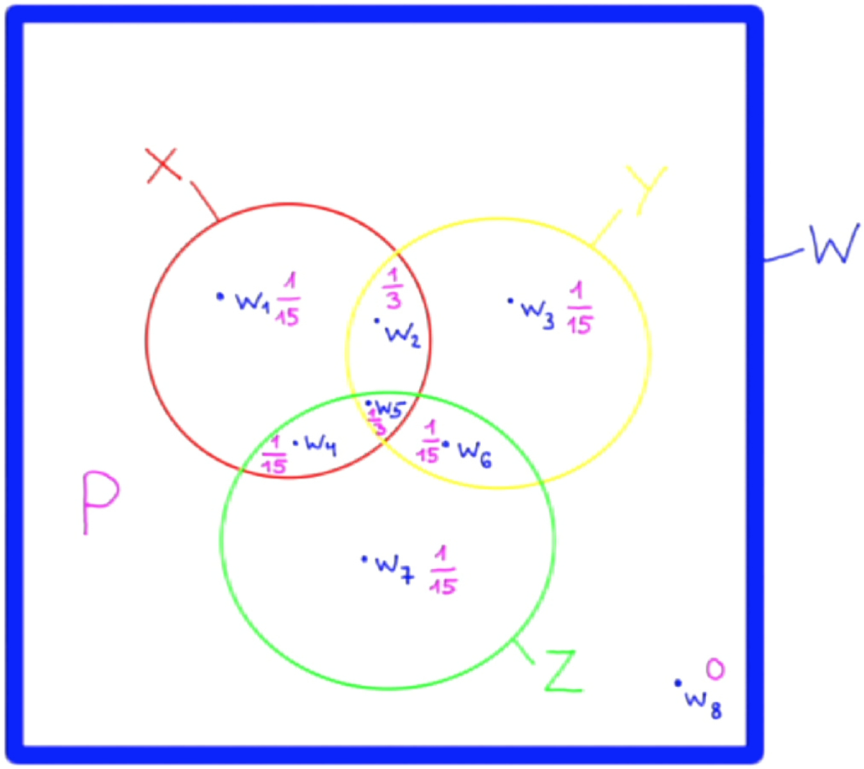

Figure 4.3: P

For instance, as explained already in the last lecture, $P(X)$, where X is the proposition $\{w_1, w_2, w_4, w_5\}$, is nothing but the sum $B(w_1) + B(w_2) + B(w_4) + B(w_5)$ which is nothing but 0.8: X is believed to degree 0.8 or with 80% certainty.

(Slide 25)

E.g., $P(\underbrace{\{w_1, w_2, w_4, w_5\}}_{X}) = B(w_1) + B(w_2) + B(w_4) + B(w_5) =$

$$= 2/3 \ + \ 2/15 = 0.8$$

In our example, that set X was the proposition expressed by the descriptive sentence 'Socrates is a philosopher': so a person with that degree of belief function regards it as quite likely that Socrates is a philosopher, but she is not certain, that is, 100% certain, of this. (By the way, we are always assuming here that it is clear which Socrates is meant – namely, the famous philosopher Socrates – our inferentially perfectly rational person is just not sure whether that person was a philosopher or not. She does not quite know enough about Socrates and the history of philosophy, that is.)

So far this was all about unconditional degrees of belief: unconditional subjective probability.

Now we want to determine what conditional degrees of belief such a person might have; that is: what degrees of belief in propositions she might have on the supposition of other proposition.

Quiz 34:
Prove what is essentially the converse of the theorem that had we had proven in Lecture 3-12. That is:
Let W be a finite non-empty set of possible worlds. Let B be an assignment of non-negative real numbers to the worlds in W – that is, each world w in W is assigned a non-negative real number $B(w)$ – such that summing up the B-values of all worlds in W yields the numerical value 1. Finally, assume that for all propositions X (that is, all subsets X of W), $P(X)$ is defined as the sum of all $B(w)$ for which w is a member of X:
$P(X) = \sum_{w \text{ in } X} B(w)$.
Prove:
(i) $P(W) = 1$.
(ii) For all propositions X: $P(X)$ is a real number, such that $0 \leq P(X) \leq 1$.
(iii) For all propositions X, Y: if $X \cap Y = \{\}$, then $P(X \cup Y) = P(X) + P(Y)$.
Solution

4.5 Conditional Probability I (13:24)

But what does 'on the supposition of a proposition' mean probabilistically? Here is the idea: Given P as in our example, the person in question believes X to a certain degree, 0.8. Now let us assume that the person initiates a certain kind of reasoning process: Say, the person supposes that X is the case. So this is hypothetical reasoning: it is not that she has actually learned X to be case, she only says to herself: hm, suppose X. So this is like a thought experiment in which she purports, for the time being, that X is the case. Or it is like when one makes a contingency plan: different things can happen; let us run through each of them and check what will be the case then. First case: suppose so and so happens; then... second case: suppose so and so happens; then...; and so on. Supposing X is like considering one of these possible ways the actual world might turn out to be. And for the same reason for which it can be very useful to come up with a good contingency plan, it can be very useful more generally to suppose a proposition and to entertain what follows from this. In our example, X is that proposition that is supposed or assumed by our inferentially perfectly rational person: and under that supposition the person then begins to reason. Of course, the supposition will have some kind of effect on how she will reason: in particular, we should expect the supposition of X to have an influence on how the person distributes her degrees of belief over propositions in the light of what the supposed proposition X is like. Supposing X, purporting X to be the case, will certainly "color" her degrees of belief in various propositions in some way, or otherwise supposing X would be utterly superfluous. But in which way?

The person pretends X to be the case, hypothetically: so, first of all, under that supposition all possible worlds in which X is false ought to be ruled out or ignored in the reasoning processes that are going to be based on that supposition. The non-X worlds should not play a role anymore within the boundaries of the supposition of X, within that thought experiment. It would be strange to say: "suppose X is the case; now, consider the following possibility in which not-X holds..." That should not be a serious possibility anymore in a context in which X has been supposed already. But what does it mean to rule out a world probabilistically: by what we found in the last lecture, it means that the B-value of any world w that is not a member of X, of any world w at which X is false, must be set to 0; and the corresponding probability of the proposition that has w as its only member is set to 0 accordingly whenever w is not a member of X. Assigning 0 to any such w or to the set including w and only w corresponds to being certain that w is not the actual world: and under the supposition that X, that is, under the supposition that the actual world is in X, one can be certain that any world that is not in X cannot be the actual world – in the context of the supposition.

For X being the set $\{w_1, w_2, w_4, w_5\}$ again, applying this first step of supposing the proposition X to the degree of belief function P thus leads to the following result:

(Slide 26)

$X = \{w_1, w_2, w_4, w_5\}$

w_1: $1/15 \to 1/15$
w_2: $1/3 \to 1/3$
w_3: $1/15 \to 0$
w_4: $1/15 \to 1/15$
w_5: $1/3 \to 1/3$
w_6: $1/15 \to 0$
w_7: $1/15 \to 0$
w_8: $0 \to 0$

So at this stage we have, e.g.: the B-value of w_1 stays the same, since w_1 is a member of X. But the B-value of w_3 is set, by hand as it were, from previously $1/15$ to 0, because w_3 is not and X-world: it is not a member of X, X is false at w_3, which is why w_3 needs to be ruled out at least hypothetically given the supposition of X. For the same reason, also the values for w_6 and w_7 shrink to 0. The value for w_8 is also 0, so it is also ruled out, but that had been so even before X was supposed. From the viewpoint of the person whose degree of belief function P we are considering right now, w_8 had not been a serious candidate for the actual world even prior to her supposition of X.

But we are not done as yet. Or rather: our inferentially perfectly rational person is not done as yet. She supposes the proposition X to be the case; and then she intends to reason rationally under that assumption. But rational reasoning presupposes that her degree of belief function conforms to the laws of probability again, as we had also seen in the last lecture: however, after changing the B-values of worlds in the manner discussued, they no longer sum up to 1 anymore:

$1/15+1/3+0+1/15+1/3+0+0+0 = 2/3+2/15$, that is, 0.8: the original degree of belief in the proposition X of our example. That this sum of new values of worlds is $P(X)$ should not come as a surprise by now: the person in question is supposing X, all values for worlds outside of X have been set to 0, which is why, obviously, the sum of all the new values of worlds is now identical to the sum of the initial values of worlds in X, which was 0.8. But we know already that this new assignment of numbers to worlds does not determine a probability measure anymore: by the theorem from the last lecture, every probability measure is determined by a unique assignment of numbers to worlds where these numbers are not less than 0, not greater than 1, and where the numbers sum up to 1: to 100%. But now we are facing a function whose sum of values is 0.8: merely 80%, if one wants to state it like that. The sum of values for worlds in the total set W of worlds is 0.8, but in order to be rational, the degree of belief of a person in W should always be 1, as we know from the last lecture. So how can we change that 0.8 into 1, accordingly?

Here is a proposal: first change the values of worlds outside of X to 0, as we have done already; but then, at a second stage, multiply all of the resulting values with one and the same number, so that after multiplying them with that number, the values of worlds in W do sum up to 1 again. Can we do so? And if so, what is that magic number?

It is in fact very easy to see that there is such a number, and only one such number, as long as the probability of the proposition that is supposed by our inferentially perfectly rational person is greater than 0: in our case, this is not a problem, since the probability of X is 0.8 which is indeed greater than 0. In that case, what we need to do is simply to multiply all values of possible worlds with $1/0.8$, $1/P(X)$, or in our example, with the number 1.25. Or equivalently: divide all values of worlds by $P(X)$, by 0.8.

If we do so, the following happens:

(Slide 27)

$X = \{w_1, w_2, w_4, w_5\}$

w_1: 1/15	$\to$	1/15	$\to$	(1/15)/0.8	= 0.08333...
w_2: 1/3	$\to$	1/3	$\to$	(1/3)/0.8	= 0.41666...
w_3: 1/15	$\to$	0	$\to$	0/0.8	= 0
w_4: 1/15	$\to$	1/15	$\to$	(1/15)/0.8	= 0.08333...
w_5: 1/3	$\to$	1/3	$\to$	(1/3)/0.8	= 0.41666...
w_6: 1/15	$\to$	0	$\to$	0/0.8	= 0
w_7: 1/15	$\to$	0	$\to$	0/0.8	= 0
w_8: 0	$\to$	0	$\to$	0/0.8	= 0

The sum of the resulting values is now $2 \cdot 0.08333... + 2 \cdot 0.41666...$ which is indeed 1, as intended. And the additional benefit of that second stage of manipulating numbers is that the proportions between the B-values of worlds in X – of the worlds within the proposition that had been assumed – are not affected in any way, since multiplying all of their values with the same constant does not change any proportions between them. For instance: originally, $B(w_1)$ had been 1/15 and $B(w_2)$ had been 1/3: so the value of w_2 had been 5 times the value of $B(w_1)$. Then the values of members of non-X had been set to 0; clearly, this did not affect the values of worlds in X in any way. And ultimately all values got multiplied with one and the same number, in our case, 1/0.8: in this way the value of w_1 was changed to 0.08333..., whereas the value of w_2 was changed to 0.41666... But still the value of w_2 is 5 times the value of w_1, as it had been the case before X was supposed. The proportion between the values of worlds in X remain invariant by the whole two-stage operation.

But note that all of that was only possible because $P(X)$ was greater than 0: if $P(X)$ had been 0, then we could not have divided the values of worlds by $P(X)$, by 0, since dividing numbers by 0 is not well-defined mathematically: it is not determined mathematically what result dividing by 0 is meant to produce. But this is not much of a restriction in our present context anyway: in our case, an inferentially perfectly rational person supposes X, and the restriction of demanding $P(X)$ to be greater than 0 means that X had not been ruled out completely by that person initially. If she is already 100% certain that X is false, there is normally not much of a point in supposing X anyway. So when we are dealing with suppositions of a proposition X, let us simply assume that we only consider cases in which $P(X) > 0$.

> Remark on division by zero: A very nice discussion of the problem of division by zero can be found in section 8.5 of the following classic:
> P. Suppes, *Introduction to Logic*, New York: Van Nostrand Reinhold Company, 1957.
> If you can get hold of this book, I very much recommend it for further study after this course.

We find that changing the B-values of worlds outside of X to 0 and then dividing all values of worlds in W by $P(X)$ has the following attractive consequences: (i) all worlds in not-X are ruled out as serious candidates for being the actual world in the suppositional context, in line with the intended act of supposing X; (ii) as far as worlds in X are concerned, things stay the same, at least as far as the proportions of values of worlds in X are concerned; and ultimately (iii) the sum of the resulting values of worlds in the full set W is 1 again, which means that the ultimate assignment of values to worlds in W determines a probability measure again.

This rational reconstruction of the act of supposing X leads, therefore, from an initial degree of belief function, P, to a new one, P_X, which is a probability measure again, and in which the new degrees of belief in propositions do reflect the fact that X has been supposed.

(Slide 28)

$X = \{w_1, w_2, w_4, w_5\}$

$$
\begin{array}{llllll}
w_1\colon 1/15 & \rightarrow & 1/15 & \rightarrow & (1/15)/0.8 & = 0.08333\ldots \\
w_2\colon 1/3 & \rightarrow & 1/3 & \rightarrow & (1/3)/0.8 & = 0.41666\ldots \\
w_3\colon 1/15 & \rightarrow & 0 & \rightarrow & 0/0.8 & = 0 \\
w_4\colon 1/15 & \rightarrow & 1/15 & \rightarrow & (1/15)/0.8 & = 0.08333\ldots \\
w_5\colon 1/3 & \rightarrow & 1/3 & \rightarrow & (1/3)/0.8 & = 0.41666\ldots \\
w_6\colon 1/15 & \rightarrow & 0 & \rightarrow & 0/0.8 & = 0 \\
w_7\colon 1/15 & \rightarrow & 0 & \rightarrow & 0/0.8 & = 0 \\
w_8\colon 0 & \rightarrow & 0 & \rightarrow & 0/0.8 & = 0
\end{array}
$$

Initial degree of belief function: P

Final degree of belief function: P_X

In a sense, P_X remains as close as possible to the original subjective probability measure P, except of course that the supposition of X is supposed to have the relevant effect on degrees of belief, or otherwise supposing X would be superfluous again.

Quiz 35:
Say, we are given a degree of belief function P that satisfies the laws of probability; so P is a subjective probability measure. And assume X to be a proposition, such that $P(X) > 0$. Then we determine P_X as explained in the lecture (which is a probability measure again). Now suppose we would apply the same procedure to P_X again: that is, we determine $(P_X)_X$. What would happen?
Solution

4.6 Conditional Probability II (14:41)

We submit that this is the answer to our question "what does 'on the supposition of a proposition X' mean probabilistically"? It means: determining a new degree of belief function P_X from one's given degree of belief function P as follows:

(Slide 29)

Let W be a non-empty and finite set of possible worlds.

Let P be a probability measure (over W), and assume P to be determined uniquely by B as explained in the previous lecture.

Then we can define: for all propositions X with $P(X) > 0$, for all worlds w in W,

- $B_X(w) = 0$ if w is not a member of X;
- $B_X(w) = \frac{B(w)}{P(X)}$ if w is a member of X.

B_X is simply the new assignment of values to worlds determined from B once X has been supposed, as described before.

And from that function B_X on worlds we can then determine the new probability measure P_X on propositions as usual:

(Slide 30)

For all propositions Y:

$$P_X(Y) = \sum_{w \text{ in } Y} B_X(w).$$

And, as mentioned before, since the values of B_X sum up to 1 again, and they are greater than or equal to 0 and less than or equal to 1, P_X is indeed a probability measure again.

The definition of P_X can also be reformulated, if we want to:

For all propositions Y:

(Slide 31)

$$P_X(Y) = \sum_{w \text{ in } X \cap Y} \frac{B(w)}{P(X)} + \sum_{w \text{ not in } X} 0$$

The second sum is the sum of values over those worlds that are not members of X: their values are 0 once X has been supposed – they are ruled out in the context of that supposition – and their resulting sum is 0, too. The first sum is the sum of values over those worlds that are, first of all, members of X – they are not ruled out by supposing X – and which additionally are also members of Y, or otherwise their values should not be used to determine the new degree of belief in Y. Hence, this first sum concerns the values of worlds that lie in the intersection of X with Y. As it is the case generally for worlds in X, also the values of worlds in $X \cap Y$ remain intact except that they are all divided by $P(X)$, that is, by the original probability of the proposition X that is being supposed – 0.8 in our previous example.

In fact we can further simplify this:

First of all, by placing the constant factor $1/P(X)$ outside of the first sum, and simply dropping the second sum:

(Slide 32)

$$P_X(Y) = \sum_{w \text{ in } X \cap Y} \frac{B(w)}{P(X)} + \sum_{w \text{ not in } X} 0$$

$$P_X(Y) = \frac{1}{P(X)} \cdot \sum_{w \text{ in } X \cap Y} B(w)$$

And since the sum of the initial B-values of worlds in $X \cap Y$ is nothing but the initial probability $P(X \cap Y)$, as we know already from the last lecture, we can also rewrite this as:

(Slide 33)

$$P_X(Y) = \frac{1}{P(X)} \cdot P(X \cap Y)$$

that is

(Slide 34)

$$P_X(Y) = \frac{P(X \cap Y)}{P(X)}$$

Since $P(X)$ has been assumed to be greater than 0, this is perfectly well-defined. Call $P_X(Y)$ the conditional probability of Y given X as determined by P. Or, more typically, in the probabilistic literature, this conditional probability of Y given X relative to P is denoted like this:

(Slide 35)

Definition:

For all probability measures P (over W), for all propositions X with $P(X) > 0$, for all propositions Y:

$$P(Y|X) = \frac{P(X \cap Y)}{P(X)}$$

In words: The conditional probability of Y given X as determined by P is $P(X \cap Y)/P(X)$.

So to the right of the vertical line, we find the proposition denoted that is supposed, X; to the left of the vertical line, we find the proposition denoted whose probability we want to determine on the supposition of X; the proposition Y; and P is the probability measure for which we are doing all of that. Thus, with every probability measure P, with every unconditional degree of belief function P, there comes also a conditional degree of belief function, which takes two propositions as input, X and Y, and which gives us as output the conditional probability of Y given X, of Y on the supposition of X, as being determined by the unconditional probability measure P. That conditional probability measure is determined from the unconditional probability measure by means of the ratio formula that you see on the slide:

The conditional probability of Y on the supposition of X is the ratio between the unconditional probability of $X \cap Y$ and the unconditional probability of X.

Remark on conditional probability: That is also how conditional probability is defined formally in probability theory, that is, in mathematics. The axioms of probability from Lecture 3 and the formal definition of conditional probability as discussed in this lecture are explained in detail in every textbook/lecture notes about probability theory. It is just that probability measures in probability theory are studied in more general contexts than ours: in particular, W is not always assumed to be finite; and not every subset of W needs to count as a proposition (or 'event', as probability theorists say).
One additional remark: In this lecture we still refer to what we called a B-function in order to determine a probability measure P. But in the future this won't be necessary anymore: I hope you will have understood by then how this works; and one can always drop the B-function and do everything in terms of P right from the start, since we already know from Lecture 3 that always $P(\{w\}) = B(w)$. That is: the B-function can be read off from the probability measure P anyway.
We will also liberalize things in the future in another direction: by now you should have seen how propositions and sentences correspond to each other. Therefore, in the future, it should not cause much of a problem if we assign probabilities to sentences in certain contexts, rather than to the propositions that they express.

And if we keep X fixed, if we fix the very proposition X that is supposed, the new unconditional probability measure P_X that is defined by

(Slide 36)

(If $P(X) > 0$:)

$$P_X(Y) = P(Y|X) = \frac{P(X \cap Y)}{P(X)}$$

has all of the properties that we expect the result of supposing X to have for a given initial unconditional probability measure P. For example: what is

$P_X(X)$?

What is the degree of belief in X on the supposition of X, the conditional probability of X given X?

By definition, this is

(Slide 37)

(If $P(X) > 0$:)

$$P_X(X) = P(X|X) = \frac{P(X \cap X)}{P(X)}$$

and since the set $X \cap X$ is of course X again, this is just

(Slide 38)

(If $P(X) > 0$:)

$$P_X(X) = P(X|X) = \frac{P(X \cap X)}{P(X)}$$

$$= \frac{P(X)}{P(X)} = 1$$

And that is exactly as it should be: in a context in which X has been supposed to be true, the degree of belief in X should be maximal, 1, X should become a certainty; this does not mean that the person whose degree of belief function we are considering is certain of X outside of the context of supposing X – in our example from above, where X was the set $\{w_1, w_2, w_4, w_5\}$, the person regarded it as likely to degree 0.8 that Socrates is a philosopher but not more than that. But given X, once X is assumed to be true, the effect of that assumption on the degree of belief in X is that X becomes a certainty. When the person is done with her thought experiment, with her suppositional reasoning on the basis of X, the person may well return to her original degree of belief in X, such as 0.8, but within the thought experiment that she is entertaining she cannot rationally doubt what she herself had presupposed to be true for the sake of the argument. Hence, in our example from above, the degree of belief in X outside of the suppositional context, 0.8, is changed to 1 inside the suppositional context.

Moreover it holds, e.g.: $P(X|X \cap Y) = 1$, since $\frac{P(X \cap Y)}{P(X \cap Y)}$ is equal to 1 again: which makes good sense; once $X \cap Y$ is assumed, of course the degree of belief in the logically weaker proposition X must become 1 as well:

(Slide 39)

E.g., by

$$P(Y|X) = \frac{P(X \cap Y)}{P(X)}$$

it holds:

$$P(X|X \cap Y) = \frac{P(X \cap Y \cap X)}{P(X \cap Y)} = \frac{P(X \cap Y)}{P(X \cap Y)} = 1$$

Or we have: $P(X|\neg X \cap Y) = 0$, since $\frac{P(\neg X \cap Y \cap X)}{P(\neg X \cap Y)} = \frac{P(\{\})}{P(\neg X \cap Y)} = 0$, according to the laws of probability from the last lecture. And again this sounds right: given that $\neg X \cap Y$, the proposition X can be ruled out: after all, its negation has been supposed.

(Slide 40)

E.g., by

$$P(Y|X) = \frac{P(X \cap Y)}{P(X)}$$

we also have:

$$P(X|\neg X \cap Y) = \frac{P(\neg X \cap Y \cap X)}{P(\neg X \cap Y)} = \frac{P(\{\})}{P(\neg X \cap Y)} = 0$$

(We will use these trivial facts concerning conditional probabilities later in the proof of a theorem in the last part of this lecture.)

As another example, take P as in our example from before, with X being the assumed proposition $\{w_1, w_2, w_4, w_5\}$ again (see Figure 4.4).

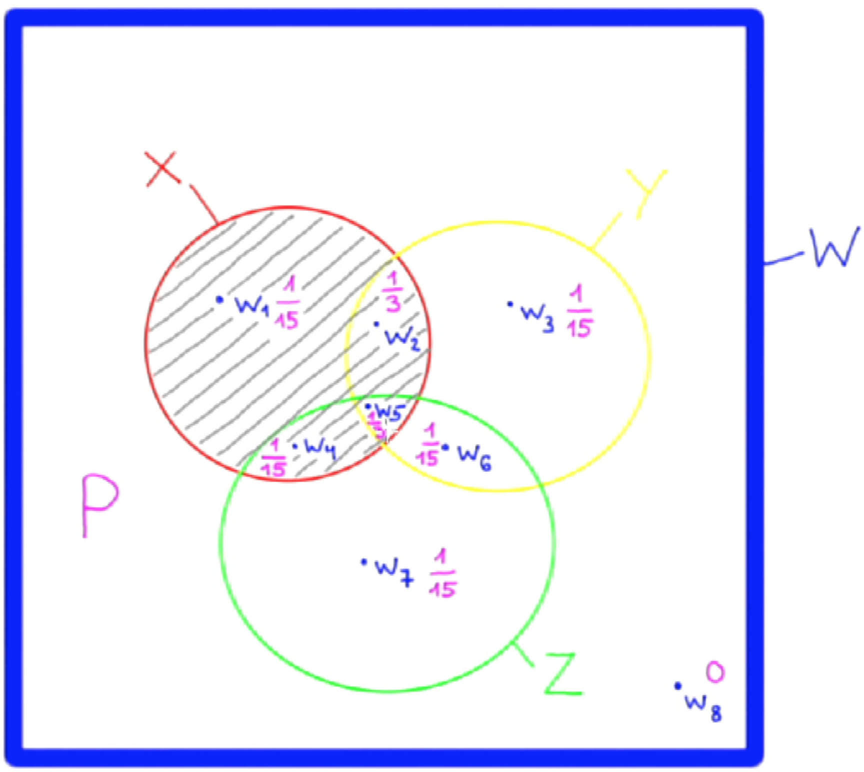

Figure 4.4: Example: X is assumed

Then P_X looks like this (see Figure 4.5):

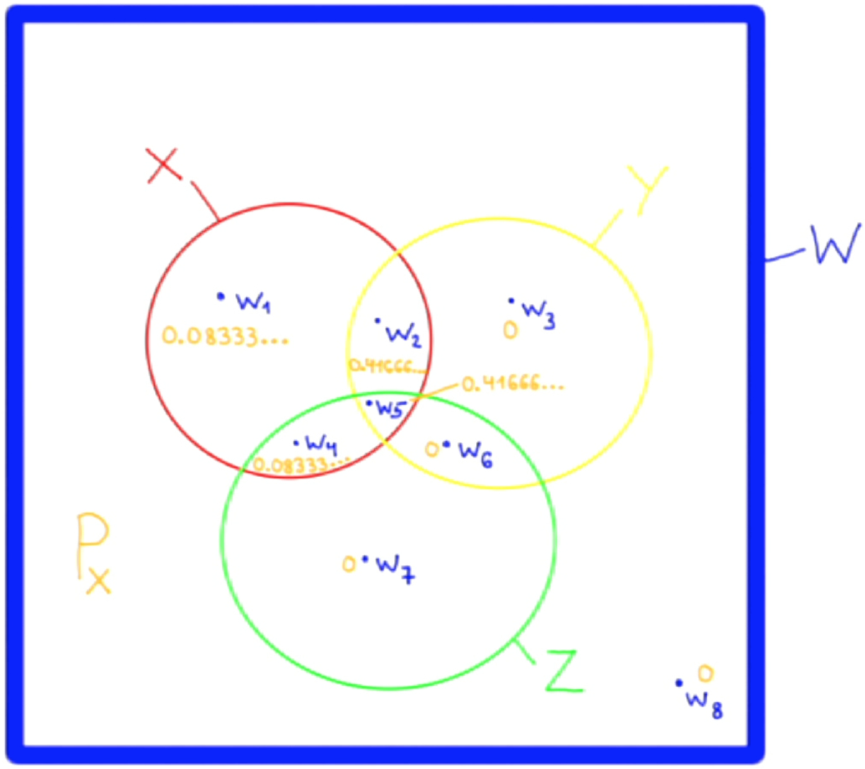

Figure 4.5: P_X

For instance, the initial degree of belief in the proposition $Y \cap Z$, the set $\{w_5, w_6\}$ was: $1/3 + 1/15$, that is, $6/15$: 0.4:

(Slide 41)

Initial degree of belief in $Y \cap Z = \{w_5, w_6\}$:

$$P(Y \cap Z) = P(\{w_5, w_6\}) = 1/3 + 1/15 = 0.4$$

(See Figure 4.6.)

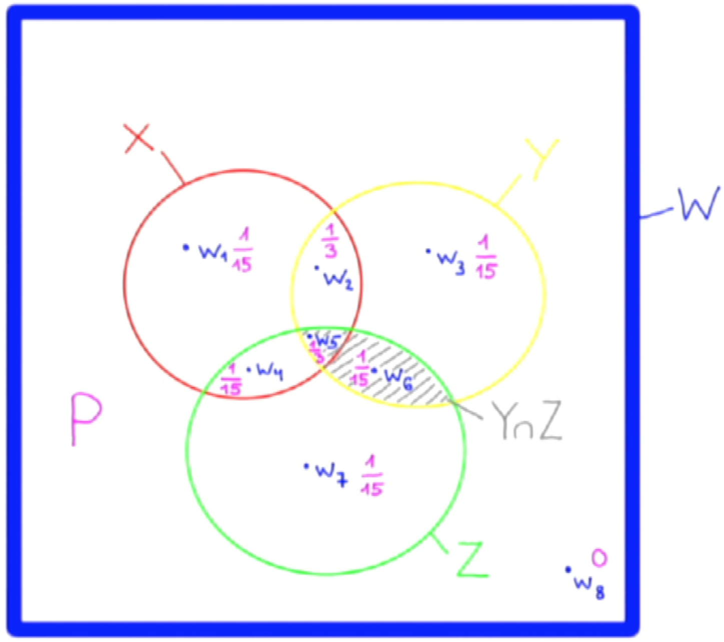

Figure 4.6: $P(Y \cap Z)$

When the person supposes X, when she supposes that Socrates is a philosopher, the degree of belief in $Y \cap Z$ changes only very slightly from 0.4 to 0.416666...:

(Slide 42)

Degree of belief in $Y \cap Z = \{w_5, w_6\}$
on the supposition of X:

$$P_X(Y \cap Z) = P(Y \cap Z|X) = P(X \cap Y \cap Z)/P(X)$$

$$= P(w_5)/0.8 = \frac{1/3}{0.8} = 0.416666\ldots$$

(See Figure 4.7.)

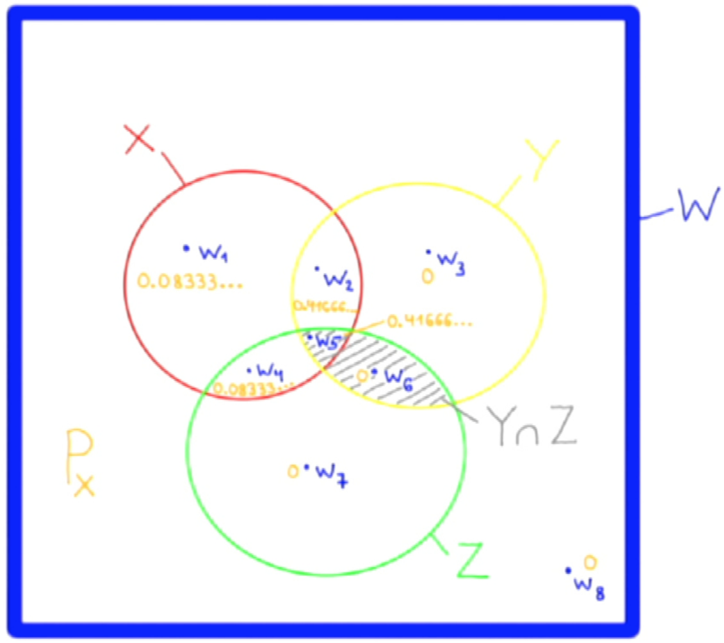

Figure 4.7: $P_X(Y \cap Z)$

In this case, for a person with a degree of belief function P as in our example, assuming that Socrates is a philosopher makes it slightly more likely that $Y \cap Z$ is the case, but really there is almost no change.

Now take a different kind of supposition: say, Z, the set $\{w_4, w_5, w_6, w_7\}$ is being supposed. We have already determined that the initial probability of $Y \cap Z$ was 0.4. But once Z is supposed that probability increases quite drastically:

(Slide 43)

Degree of belief in $Y \cap Z = \{w_5, w_6\}$ on the supposition of

$$Z = \{w_4, w_5, w_6, w_7\}:$$

$$P_Z(Y \cap Z) = P(Y \cap Z|Z) = P(Z \cap Y \cap Z)/P(Z)$$

$$= P(Y \cap Z)/0.53333\ldots = \frac{0.4}{0.53333\ldots} = 0.75$$

Supposing X does not have a lot of impact on the degree of belief in $Y \cap Z$, but supposing Z does have: one might say that Z confirms $Y \cap Z$ in a way that X does not. Stephan will deal in detail with the topic of confirmation in the next lecture, but for our present purposes it is sufficient to see how supposing a proposition can change a person's distribution of degrees of beliefs over propositions. Given Z it is quite likely that the actual world is in $Y \cap Z$ – the conditional probability of $Y \cap Z$ given Z is 0.75 and hence quite high – whilst given X it is much less likely that the actual world is in $Y \cap Z$ – the conditional probability of $Y \cap Z$ given X is only 0.416666...

This process of moving from a probability measure P to the probability measure P_X is also called 'conditionalizing P on X' in the relevant literature.

(Slide 44)

Conditionalizing P on X leads to: P_X

And other than the act of supposing X it can also be used to make precise in probabilistic terms the effect that learning X, receiving a new piece of evidence X, has on the probability measure P: so it tells us what happens to one's degree of belief function when one perceives that X is the case, or when one is told by a fully trustworthy source that X is the case, or the like. For many applications of subjective probability theory in philosophy the learning interpretation is much more salient than the interpretation in terms of suppositions. But for our present purposes the suppositional interpretation of conditionalization is much more important, simply because supposing A is closer to the 'If A then...' in a conditional than learning X is. It is this interpretation of conditionalization in terms of supposition that we will apply next when we return to our initial topic: indicative conditionals.

Quiz 36:
Let P be given, as in the lecture, by the following B assignment:
$B(w_1) = 1/15$
$B(w_2) = 1/3$
$B(w_3) = 1/15$
$B(w_4) = 1/15$
$B(w_5) = 1/3$
$B(w_6) = 1/15$
$B(w_7) = 1/15$
$B(w_8) = 0$
Determine:
(i) $P(\{w_1, w_3\}|\{w_1, w_3, w_4, w_6, w_7\})$,
(ii) $P(\{w_1, w_3\}|\{w_3, w_4, w_7\})$,
(iii) $P(\{w_1\}|\{w_1, w_2\})$.
Solution

4.7 Degrees of Acceptability for Conditionals (13:09)

We are now ready to put the two parts of this lecture so far together – conditional sentences in the indicative mood on the one hand, and conditional degrees of belief on the other.

Other than asking under what conditions a sentence is true, as we did in the second lecture, we may also be interested in the question how strongly we accept a sentence, or how acceptable we find a sentence. Say, a descriptive sentence A expresses the proposition that X. How acceptable is A to me, or to what degree do I accept it to be true: well, simply determine my degree of belief in the proposition X, in the proposition expressed by A, and you get the answer. And that degree of belief can be determined, at least in principle, by studying my betting behavior, as we discussed in the third lecture. Or how strongly do I accept the conjunction sentence, the 'and' sentence, $A \wedge B$? No problem: if A expresses the proposition X again, and B expresses the proposition Y, then $A \wedge B$ expresses the conjunction proposition $X \cap Y$, as we know from the last lecture, and my subjective probability for that proposition should be the answer again: the degree with which $A \wedge B$ is acceptable in my eyes is nothing but my degree of belief in the proposition that is expressed by $A \wedge B$, my probability for $X \cap Y$.

Now let us ask the same questions concerning indicative conditionals: Other than asking under what conditions an indicative conditional $A \to B$ is true, for which the analysis in terms of $\neg A \vee B$ gave us at least some kind of interesting answer – though maybe ultimately not a completely satisfactory one – we may also be interested in the question how strongly we accept an indicative sentence, or how acceptable an indicative conditional is to us.

But now it seems that there are two kinds of answers to this question, not just one. Let me explain.

First of all, assume the indicative conditional $A \to B$ to express a proposition; if A expresses the proposition X again, and B expresses Y again, take the conditional $A \to B$ to express the proposition $X \to Y$, where in a context that is concerned with propositions, '$\to$' is not a logical symbol anymore that connects two sentences, but instead it is an operation or a function that maps the two propositions X and Y, in that order, to their corresponding conditional proposition $X \to Y$, and where that conditional proposition is some subset of a given set W of possible worlds again. If indicative conditionals express propositions, and at least at first glance we do not have any reason to doubt them doing so, things should be just like in the case of, say, conjunction before: for instance, we regarded $X \wedge Y$ as simply yet another proposition, a subset of W, namely the result of applying the conjunction operation to the two proposition X and Y, where that conjunction operation consisted in mapping the two propositions X and Y to their set-theoretic intersection. In the case of conjunction, the ordering of X and Y did not actually matter, while, presumably, in the case of $X \to Y$ the order is important – 'if A, then B' usually does not express the same proposition as the reversed 'If B, then A' – but other than that the situation should be perfectly analogous. But if so, then we should be able to give the analogous kind of answer as before also to our new question: How strongly do I accept the indicative conditional $A \to B$? Well, determine my degree of belief in the proposition $X \to Y$, and you get the answer. And also that degree of belief could be determined, at least in principle, in terms of my betting behaviour again.

Let me put this on record for later:

(Slide 45)

Thesis 1:

(i) There is a conditional operation $\to$ that can take any two propositions as input, which maps them to a proposition as output, and which has the following property:

For every indicative conditional $A \to B$, where A expresses the proposition X, and B expresses the proposition Y, it holds that:

$$A \to B \text{ expresses the proposition } X \to Y.$$

(Slide 46)

Thesis 1: [CONTINUED]

(ii) For every indicative conditional $A \to B$, where A expresses the proposition X, and B expresses the proposition Y, and for every probability measure P on propositions, it holds that:

the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(X \to Y)$

where $X \to Y$ is the proposition expressed by $A \to B$ as explained in (i).

Remark on 'rel. to P' ('relative to P'): When we speak here of "the degree of acceptability for the indicative conditional $A \to B$ (rel. to P)" we mean simply the degree by which the indicative conditional $A \to B$ is acceptable to a perfectly rational person *in a context in which the person's degree of belief function is P* – that is, we mean the person's degree of acceptability for $A \to B$ *relative to P*.

For instance, if the indicative conditional $A \to B$ is analyzed in terms of $\neg A \vee B$, as it is the case within mathematics, then the operator $\to$ above on propositions X and Y is simply

$X \to Y = \neg X \cup Y$

and according to (ii) of Thesis 1 the degree of acceptability for the indicative conditional $A \to B$ would be identical to $P(\neg X \cup Y)$. But that is just one possibility consistent with Thesis 1; Thesis 1 is not that specific really: it just says that indicative conditionals express propositions of some sort and their degrees of acceptability are determined in the same way as e.g. the degree of acceptability of A is determined, that is, by means of the probability of the proposition that gets expressed. What is the degree of acceptability of an indicative conditional? It is the degree of belief in that indicative conditional being true.

But that is not yet the end of the story. There is a second highly plausible answer to the question 'How acceptable is the indicative conditional $A \to B$ to me?' It is an answer that had been suggested already by Frank Plumpton Ramsey – yes, the British philosopher that we had the pleasure to meet already in the last lecture – and he did so in a famous footnote to one of his papers.

In that footnote he says:

(Slide 47)

Frank P. Ramsey:

> 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.

The idea should sound quite familiar from the last part of this lecture on conditionalization. What Ramsey suggests is this: if you want to determine to what degree the indicative conditional 'if A then B' is acceptable to you, then, firstly, suppose A to be the case – add it hypothetically to your stock of knowledge; and then determine your degree of belief in B under the supposition of A – argue about B on that basis of supposing A. In other words: the acceptability of the indicative conditional $A \to B$ is tied to a property of a particular act of suppositional reasoning. And this sounds very plausible indeed: it would be surprising if the 'if A then...' in natural language were not closely related in some way to suppositions of the sort: 'Suppose that A:...' And Ramsey tries to clarify, at least to some extent, in what way they are related.

For instance: why is 'If Oswald did not kill Kennedy, then someone else did' very much acceptable to me? Because when I suppose that Oswald did not kill Kennedy, then under that supposition my degree of belief in someone else having killed Kennedy becomes very high indeed; and this is so because I believe very firmly indeed that someone at least did kill Kennedy. In probabilistic terms: my subjective conditional probability in someone else than Oswald having killed Kennedy under the supposition that Oswald did not do it is very high. While Ramsey does not refer to conditional probabilities in his footnote explicitly, it is quite clear from the last section of this lecture that conditionalization – determining degrees of belief on the supposition of propositions – is a very natural way of making Ramsey's intuitive thought formally precise. And from the rest of Ramsey's probabilistic work it is pretty clear that something like that must have been on his mind as well.

Remark on Ramsey's footnote: What is described by this footnote is now called the 'Ramsey test for conditionals'; the footnote is contained in a paper by F. P. Ramsey titled "General propositions and causality" which was written in 1929. The paper can be found in:
Ramsey, F. P., *The Foundations of Mathematics and other Logical Essays*, ed. by R. B. Braithwaite, New York: Routledge and Kegan Paul, 1931.

So let us call the following Thesis 2 concerning the acceptability of indicative conditionals:

(Slide 48)

Thesis 2:

> For every indicative conditional $A \to B$, where A expresses the proposition X, and B expresses the proposition Y, and for every probability measure P on propositions, it holds that:
>
> *the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(Y|X)$ (or $P_X(Y)$).*

the conditional probability of Y given X as determined by P.

So what is the degree of acceptability of an indicative conditional according to Thesis 2? It is a conditional degree of belief.

Remark on '$P(Y|X)$': By Thesis 2, the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(Y|X)$. As explained before, if $P(X) = 0$, then '$P(Y|X)$' is undefined. Hence, by Thesis 2, if $P(X) = 0$, then also the degree of acceptability for $A \to B$ (rel. to P) is undefined.
Additional remark on '$P(Y|X)$': this is often read, and rightly so, as 'the probability of Y given X'; but 'given' does sound very much like 'if' (as in 'the probability of Y if X'), doesn't it? No wonder philosophers started relating the acceptability of conditionals to conditional probabilities.

Therefore, we end up with two answers to our question ‘How acceptable is $A \to B$ to me?’: One is given by Thesis 1, and it is:

$P(X \to Y)$,

my unconditional probability in the conditional proposition $X \to Y$ being true.

The other one is given by Thesis 2, and it is

$P(Y|X)$,

my conditional probability in Y being true on the supposition that X is true.

Both theses sound very plausible indeed.

And there does not seem to be any reason whatsoever why they couldn’t both be true: if both are true, then this means simply that

$P(X \to Y) = P(Y|X)$.

For if Thesis 1 is true, then the degree of acceptability of $A \to B$ is identical to $P(X \to Y)$, and if Thesis 2 is true, the degree of acceptability of $A \to B$ is identical to $P(Y|X)$, so if both theses are true, then the degree of acceptability of $A \to B$, a certain number, must be identical at the same time to $P(X \to Y)$ and $P(Y|X)$, which means that $P(X \to Y)$ must be identical to $P(Y|X)$.

If so, then determining the degree of belief in the proposition $X \to Y$ on the one hand, and determining the conditional degree of belief in Y given X on the other, are merely two different ways of determining one and the same outcome. It is like calculating $1 + 3$ on the one hand, and $2 + 2$ on the other: the outcome is always 4, and it could not be otherwise, even though the manner in which the outcome 4 is determined is different in each case.

I should add that I am not telling a fictitious story here. Some excellent logicians and philosophers of language indeed suggested back in the 1970s that

$P(X \to Y)$ should be equal to $P(Y|X)$

for pretty much the reasons that I have explained.

But they have been proven wrong, as least if none of the other assumptions that we have made along the way are given up – such as, propositions are sets of possible worlds, and the like. This follows from a theorem to which we will now turn – a theorem that came as a big surprise back then, and which one could not just as have “seen” in some more immediate way without making proper philosophical use of mathematics.

Remark on $P(X \to Y) = P(Y|X)$: This equality – sometimes called "the equation" – was defended, e.g., by the US-American philosopher and logician Robert Stalnaker in the 1970s. The discovery that, given some plausible background assumptions, the equality cannot hold generally, came as a surprise – it was the "bombshell", as Dorothy Edgington calls it in her paper "On Conditionals", *Mind* 104 (1995), 235-329.

Quiz 37:
Assume Thesis 2 from the lecture to be true; that is, for all descriptive sentences A, B, where A expresses the proposition X, and B expresses the proposition Y, it holds for every subjective probability measure P:
the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(Y|X)$.
Given that this is so, what is the degree of acceptability that a (perfectly rational) person ought to assign to an indicative conditional of the form $A \to A$ (where A expresses some proposition X, such that $P(X) > 0$)?
Solution

4.8 Lewis' Triviality Theorem (12:09)

The theorem in question was shown by the US-American philosopher David Lewis, who died only quite recently in 2001, and who was one of the foremost philosophers of the 20th century.

Let me now state one version of Lewis' famous so-called triviality theorem on conditionals:

Remark on Lewis' theorem: It is contained in the following article:
D. Lewis, "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review* 85 (1976), 297-315.
The article was reprinted later in D. Lewis, *Philosophical Papers*, Volume 2, Oxford: Oxford University Press, 1986.

(Slide 52/1)

Theorem:

Let W be a given non-empty set of possible worlds. By propositions we mean subsets of W again, as usual.

In line with the first part of thesis 1 from above, we assume:

(Slide 52/2)

(Ass. 1) There is a conditional operation $\rightarrow$ that can take any two propositions over W as input and which maps them to a uniquely determined proposition over W as output.

In line with our findings from the last lecture, we also assume:

(Slide 52/3)

(Ass. 2) Every rational degree of belief function P (on W) is a probability measure.

And in line with our present interest in conditionals and with the second part of Thesis 1 as well as with Thesis 2 taken together, we assume:

(Slide 52/4)

(Ass. 3) Every rational degree of belief function P (on W) satisfies:

For all propositions X, Y: $P(X \rightarrow Y) = P(Y|X)$,

where $X \rightarrow Y$ is as described by Ass. 1.

Finally, in line with the last part of this lecture on conditionalization and its important role for degrees of belief in general, we assume:

(Slide 53)

(Ass. 4) The set of all rational degree of belief functions P (on W) is closed under conditionalization: that is,

for every rational degree of belief function P (on W), for every proposition X for which $P(X) > 0$, the conditionalization P_X of P on X is a rational degree of belief function (on W) as well.

And we presuppose that P_X was defined like this:

(If $P(X) > 0$:)

For all propositions Y: $P_X(Y) = P(Y|X) = P(X \cap Y)/P(X)$.

From these Assumptions 1-4, and presupposing our previous definition of conditional probabilities in terms of the ratio formula, it follows:

(Slide 54)

Conclusion 1:

For all propositions X, Y, Z, for all rational degree of belief functions P (on W) for which it holds that $P(X) > 0$, $P(X \cap Y) > 0$:

$$P(Y \to Z|X) = P(Z|X \cap Y).$$

Conclusion 2:

For all propositions X, Y, Z, for all rational degree of belief functions P (on W) for which it holds that $P(Z) > 0$, $P(\neg Z) > 0$, $P(Z \cap Y) > 0$, $P(\neg Z \cap Y) > 0$:

$$P(Y \to Z) = P(Z).$$

Conclusion 1 will merely be an intermediate conclusion that we will need for the proof of Conclusion 2. But Conclusion 2 is the actual killer: it means that if all of our Assumptions 1-4 from above are satisfied, that is, essentially, if our Thesis 1 and 2 from before concerning the degrees of acceptability for indicative conditionals are satisfied simultaneously, then an inferentially perfectly rational person would have to believe every conditional proposition

$Y \to Z$

with the same strength with which she believes the proposition

Z

as long as

$P(Z) > 0, P(\neg Z) > 0, P(Z \cap Y) > 0, P(\neg Z \cap Y) > 0$, that is, for all interesting non-trivial cases of degrees of beliefs: where $P(Z)$ is neither the extreme value 0 nor the extreme value 1, and the like, as it is the case for most propositions that concern everyday matters.

But that Conclusion 2 is absurd: Why would anyone who is inferentially perfectly rational be forced to assign the same degree of belief to

If Conny comes to the party, then I will talk to her there

as to

I will talk to Conny at the party

?

Or why believe in

If Oswald did not kill Kennedy, someone else did

to the same degree to which one believes in

Someone else [other than Oswald] killed Kennedy

?

This is just plain nonsense. So something has gone awfully wrong in the Assumptions 1-4 from above, or as far as our Theses 1 and 2 are concerned that had justified these assumptions. And that is what Lewis' theorem shows.

In fact, if we turned the theorem into an argument again, in which Assumptions 1-4 would figure as the premises and in which Conclusion 2 would be the overall conclusion of the argument, then it would be fair to call that argument a paradox again. This is not what Lewis's theorem is called normally in the relevant literature, but it might well have been called like that. The reason why Lewis' theorem is sometimes called a triviality theorem is because only very restricted probability measures – only very trivial ones – satisfy Conclusion 2: for all propositions Y and Z, $P(Y \to Z) = P(Z)$.

Here is the proof of the theorem:

(Slide 55)

Proof: Concerning Conclusion 1.

$$
\begin{aligned}
P(Y \to Z|X) &= P_X(Y \to Z) && \text{(def. of } P_X\text{)} \\
&= P_X(Z|Y) && \text{(Ass. 4 and 3)} \\
&= \frac{P_X(Y \wedge Z)}{P_X(Y)} && \text{(def. of cond. prob.)} \\
&= \frac{P(Y \wedge Z|X)}{P(Y|X)} && \text{(def. of } P_X\text{)} \\
&= \frac{\frac{P(X \wedge Y \wedge Z)}{P(X)}}{\frac{P(X \wedge Y)}{P(X)}} && \text{(def. of.cond. prob.)} \\
&= \frac{P(X \wedge Y \wedge Z)}{P(X \wedge Y)} && \text{(calculate)} \\
&= P(Z|X \wedge Y) && \text{(def. of cond. prob.)}
\end{aligned}
$$

Now we turn to the proof of Conclusion 2. Here we presuppose Conclusion 1, and we will also apply the so-called Addition Theorem for probabilities that we had proved already in the last lecture on the basis of the laws of probability: whatever the propositions X and Y, it holds that $P(Y) = P(X \cap Y) + P(\neg X \cap Y)$.

(Slide 56)

Concerning Conclusion 2.

$$
\begin{aligned}
P(Y \to Z) &= P(Z \wedge (Y \to Z)) + P(\neg Z \wedge (Y \to Z)) && \text{(Add. Th., by Ass. 2)} \\
&= P((Y \to Z)|Z)P(Z) + \\
&\quad\; P((Y \to Z)|\neg Z)P(\neg Z) && \text{(def. of cond. prob., calc.)} \\
&= P(Z|Z \wedge Y)P(Z) + \\
&\quad\; P(Z|\neg Z \cap Y)P(\neg Z) && \text{(Conclusion 1)} \\
&= 1 \cdot P(Z) + 0 \cdot P(\neg Z) && \text{(def. of cond. prob., Ass. 2)} \\
&= P(Z) && \text{(calculate)}
\end{aligned}
$$

(See Figure 4.8, 4.9, 4.10, and 4.11 for the relevant calculations concerning the second line of the calculation on Slide 56; and see Figure 4.12 for the transition from line 3 to line 4.)

Concerning Conclusion 2.

$\frac{P(Z\cap(Y\to Z))}{P(Z)}P(Z)$

$$
\begin{aligned}
P(Y\to Z) &= P(Z\wedge(Y\to Z)) + P(\neg Z\wedge(Y\to Z)) && \text{(Add. Th., by Ass. 2)}\\
&= P((Y\to Z)|Z)P(Z) + \\
&\quad P((Y\to Z)|\neg Z)P(\neg Z) && \text{(def. of cond. prob., calc.)}\\
&= P(Z|Z\wedge Y)P(Z) + \\
&\quad P(Z|\neg Z\cap Y)P(\neg Z) && \text{(Conclusion 1)}\\
&= 1\cdot P(Z) + 0\cdot P(\neg Z) && \text{(def. of cond. prob., Ass. 2)}\\
&= P(Z) && \text{(calculate)}
\end{aligned}
$$

Figure 4.8: Calculation 1

Concerning Conclusion 2.

$P(Z\cap(Y\to Z))$

$$
\begin{aligned}
P(Y\to Z) &= P(Z\wedge(Y\to Z)) + P(\neg Z\wedge(Y\to Z)) && \text{(Add. Th., by Ass. 2)}\\
&= P((Y\to Z)|Z)P(Z) + \\
&\quad P((Y\to Z)|\neg Z)P(\neg Z) && \text{(def. of cond. prob., calc.)}\\
&= P(Z|Z\wedge Y)P(Z) + \\
&\quad P(Z|\neg Z\cap Y)P(\neg Z) && \text{(Conclusion 1)}\\
&= 1\cdot P(Z) + 0\cdot P(\neg Z) && \text{(def. of cond. prob., Ass. 2)}\\
&= P(Z) && \text{(calculate)}
\end{aligned}
$$

Figure 4.9: Calculation 2

Concerning Conclusion 2.

$$\begin{aligned}
P(Y \to Z) &= P(Z \wedge (Y \to Z)) + P(\neg Z \wedge (Y \to Z)) && \text{(Add. Th., by Ass. 2)} \\
&= P((Y \to Z)|Z)P(Z) + \\
&\quad P((Y \to Z)|\neg Z)P(\neg Z) && \text{(def. of cond. prob., calc.)} \\
&= P(Z|Z \wedge Y)P(Z) + \\
&\quad P(Z|\neg Z \cap Y)P(\neg Z) && \text{(Conclusion 1)} \\
&= 1 \cdot P(Z) + 0 \cdot P(\neg Z) && \text{(def. of cond. prob., Ass. 2)} \\
&= P(Z) && \text{(calculate)}
\end{aligned}$$

$\frac{P(\neg Z \cap (Y \to Z))}{P(\neg Z)} P(\neg Z)$

Figure 4.10: Calculation 3

Concerning Conclusion 2.

$$\begin{aligned}
P(Y \to Z) &= P(Z \wedge (Y \to Z)) + P(\neg Z \wedge (Y \to Z)) && \text{(Add. Th., by Ass. 2)} \\
&= P((Y \to Z)|Z)P(Z) + \\
&\quad P((Y \to Z)|\neg Z)P(\neg Z) && \text{(def. of cond. prob., calc.)} \\
&= P(Z|Z \wedge Y)P(Z) + \\
&\quad P(Z|\neg Z \cap Y)P(\neg Z) && \text{(Conclusion 1)} \\
&= 1 \cdot P(Z) + 0 \cdot P(\neg Z) && \text{(def. of cond. prob., Ass. 2)} \\
&= P(Z) && \text{(calculate)}
\end{aligned}$$

$P(\neg Z \cap (Y \to Z))$

Figure 4.11: Calculation 4

Concerning Conclusion 2.

$$
\begin{aligned}
P(Y \to Z) &= P(Z \wedge (Y \to Z)) + P(\neg Z \wedge (Y \to Z)) && \text{(Add. Th., by Ass. 2)} \\
&= P((Y \to Z)|Z)P(Z) + \\
&\quad\; P((Y \to Z)|\neg Z)P(\neg Z) && \text{(def. of cond. prob., calc.)} \\
&= P(Z|Z \wedge Y)P(Z) + \\
&\quad\; P(Z|\neg Z \cap Y)P(\neg Z) && \text{(Conclusion 1)} \\
&= 1 \cdot P(Z) + 0 \cdot P(\neg Z) && \text{(def. of cond. prob., Ass. 2)} \\
&= P(Z) && \text{(calculate)}
\end{aligned}
$$

Figure 4.12: Calculation 5

In case you wonder why we did not have to apply Ass. 1 from above: Ass. 1 was a presupposition of Ass. 3; and Ass. 1 also implied that the laws of probability as covered by Ass. 2 do in fact also apply to, e.g., $Y \to Z$. So Ass. 1 is implicit in some of our applications of Ass. 3 and 2.

This concludes the proof of the theorem.

Quiz 38:
Please run through the proof of Lewis' theorem again; please do it slowly, please do it several times if necessary, and ideally you would take a pen and a piece of paper and write and justify each single line for yourself (guided by the lecture).

4.9 Responses to the Triviality Theorem (11:20)

Time for diagnosis. What has gone wrong? Clearly we cannot continue to take each of our Assumptions 1-4 for granted. Ass. 2. – rational degree of belief functions are probability measures – is supported very well, as we discussed already in the last lecture; it cannot be the problem: the problem is really to do with how we have been treating conditionals. Similarly, Ass. 4 cannot be the issue: if P is a rational degree of belief function, why should the conditionalization of P on some proposition X, for which $P(X) > 0$, not count as a rational degree of belief function? As we have seen before, P_X is the result of supposing the proposition X in a context in which P is the initial degree of belief function: is it irrational to make suppositions? That can't be it.

So the troublemaker must be either Ass. 1, or Ass. 3 for which Ass. 1 was a presupposition, or both.

If we give up Ass. 3,

(Slide 57)

(Ass. 3) Every rational degree of belief function P (on W) satisfies:

For all propositions X, Y: $P(X \to Y) = P(Y|X)$,

where $X \to Y$ is as described by Ass. 1.

then it cannot be the case that both of our original Theses 1 and 2 are the case:

(Slide 58)

Thesis 1:

(i) There is a conditional operation $\to$ that can take any two propositions as input, which maps them to a proposition as output, and which has the following property:

For every indicative conditional $A \to B$, where A expresses the proposition X, and B expresses the proposition Y, it holds that:

$$A \to B \text{ expresses the proposition } X \to Y.$$

This became effectively Ass. 1 later in the theorem, although we were only focusing on the propositions $X \to Y$ then.

(Slide 59)

Thesis 1: [CONTINUED]

(ii) For every indicative conditional $A \to B$, where A expresses the proposition X, and B expresses the proposition Y, and for every probability measure P on propositions, it holds that:

the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(X \to Y)$

where $X \to Y$ is the proposition expressed by $A \to B$ as explained in (i).

And Thesis 2 was:

(Slide 60)

Thesis 2:

> For every indicative conditional $A \to B$, where A expresses the proposition X, and B expresses the proposition Y, and for every probability measure P on propositions, it holds that:
>
> *the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(Y|X)$ (or $P_X(Y)$).*

Thesis 1 and 2 combined gave us Ass. 3, so if we drop Ass. 3, we need to drop either of Thesis 1 or 2 or maybe both of them.

Giving up Thesis 1 would really mean to give up that indicative conditionals in natural language express propositions, and hence, as only sentences that express propositions are true or false, it would also mean to give up that indicative conditionals are true or false. That would be a very drastic consequence: if this is the right way out of the problem that became transparent to us through Lewis' theorem, then indicative conditionals are not actually descriptive sentences: they do not express propositions, and they do not have truth values, at least not in the standard sense of being true or false. It would also follow from this that indicative conditionals $A \to B$ could not be analyzed as $\neg A \vee B$, since $\neg A \vee B$ does express a proposition and is true or false; but the consequences are more general then that: indicative conditionals $A \to B$ could not be analyzed as anything that expresses a proposition, and which thus has a standard truth value, anything whatsoever. You might ask yourself: if an indicative conditional did not express a proposition, wouldn't that mean that these conditionals did not have any meaning? Which could not be right, as 'If Conny comes to the party, then I will talk to her there' is perfectly meaningful. So giving up Thesis 1 would have to be a non-starter. But that would be much too quick: take 'Yippee' as an example. This is a perfectly meaningful linguistic expression: by uttering it I express my state of mind that I am very happy; you understand me. But 'Yippee' is not a descriptive sentence; it might be appropriate to utter 'Yippee' in one context and inappropriate to utter it in another, but it is not true or false, and it does not express a proposition, a set of possible worlds at which it is true. Nevertheless it is possible to communicate in terms of uttering 'Yippee', and we sometimes do. Or take an imperative or a question: these are not true or false either, and they do not express propositions, but they are still meaningful linguistic objects which we use on a daily basis. Now much the same might be true of indicative conditionals: maybe they are not true or false, maybe they do not express propositions, but we can still use them sensibly in communication. In particular, we might use them to express our conditional beliefs in some way, much as 'Yippee' expresses my happy state of mind. The degree of acceptability of an indicative conditional might have to be measured in terms of its corresponding conditional degree of belief, and not in terms of an unconditional degree of belief in a conditional proposition, simply because there is not even a conditional proposition that the indicative conditional would express. If so, Thesis 1 needs to be given up.

Remark on indicative conditionals expressing states of mind: You might remember that at the beginning of Lecture 2 I gave you a link to the Stanford Encyclopedia entry
http://plato.stanford.edu/entries/moral-cognitivism/
in which the thesis is discussed that, perhaps, normative sentences are not true or false but that they express mental states, such as emotions or the like. Now we are considering the analogous view that indicative conditionals are not true or false but that they express mental states, such as conditional degrees of belief.

In contrast, abandoning Thesis 2 would amount to saying that determining the degrees of acceptability of indicative conditionals does not correspond to determining some properties of acts of suppositional reasoning in the way that Ramsey had suggested. 'If A then...' and 'Suppose A. It holds then that...' would turn out to be much more different than one might have expected.

Or one might try giving up both Theses 1 and 2; or maybe one would at least have to rephrase both Theses 1 and 2 in some way. For example: maybe there are two kinds of degrees of acceptability for indicative conditionals: the degree of belief of $A \to B$ being true, as suggested by Thesis 1, and the suppositional degree of acceptability of $A \to B$, as suggested by Thesis 2, and the mistake was to assume that these two degrees of acceptance would always coincide. Instead, perhaps one would have to reformulate Thesis 1 in terms of a notion of degree-of-acceptance$_1$ and Thesis 2 in terms of a notion of degree-of-acceptance$_2$, where the two kinds of degrees of acceptance would have different mental and communicative functions and where they could not always coincide in value – as shown by Lewis' theorem. Once the two kinds of degrees of acceptability are disentangled like that, there is no problem anymore. But of course one would need to have good arguments that there are indeed two sensible types of degree of acceptability for indicative conditionals which do reflect two ways of handling such conditionals in our mental lives and in our communication.

Or one drops some of the background assumptions that had been used above: $\to$ was assumed to be one and the same operation on propositions whatever the degree of belief function P was like that would then be applied to, say, the proposition $X \to Y$. But maybe what proposition an indicative conditional A $\to B$ expresses, what propositional operation $\to$ the logical symbol $\to$ expresses, is not independent of the degree of belief function that one considers? If so, then what proposition an indicative conditional $A \to B$ expresses is strongly context-dependent – it also depends on what the relevant person's belief state, what her degree of belief function, is like.

Remark on ‘what proposition an indicative conditional $A \to B$ expresses might depend on P’: This proposal was put forward, as a way out of the problem that had been highlighted by Lewis’ theorem, by the Dutch philosopher Bas van Fraassen in an article from 1976: Van Fraassen, B., “Probabilities of Conditionals”, in: Harper, W. and Hooker, C. (eds.), *Foundations of Probability theory, Statistical Inference, and Statistical Theories of Science*, Volume I, Dordrecht: Reidel, 1976, pp. 261-308.

Or maybe propositions are not simply sets of worlds but more complicated objects: say, sets of pairs, sets of pairs of worlds and something else, such as linguistic contexts of assertion or the like; or propositions are maybe not just true or false at worlds, but they can have other values at worlds as well: they can be undefined; or they have some kinds of numerical values between 0 and 1 at worlds; and so on.

We cannot go into further details here. Let me just emphasize that once again I am not telling a fictitious story here: All of these options are being discussed in the philosophy of language; they have been discussed for years, and they are discussed intensively again as we speak. For instance, the option of dropping Thesis 1 while retaining Thesis 2 – maintaining that indicative conditional do not express propositions and are not true or false, and measuring their degrees of accaptability in terms of their corresponding conditional degrees of belief – has been worked out in detail within the so-called Suppositional Theory of Conditionals which has been developed by the US-American philosopher Ernest W. Adams and the British philosopher Dorothy Edgington.

(Ernest Adams has already passed away, unfortunately; but Dorothy Edgington is very active.)

Ernest Adams came up with the idea of measuring the degree of acceptability of $A \to B$ in terms of the corresponding conditional degree of belief, and he did so already in the 1960s and 1970s, and he is also the philosopher who put forward the famous Oswald-Kennedy example that shows that the acceptability of conditionals may depend on their grammatical mood. The Suppositional Theory of Conditionals, which was very much further developed and defended by Dorothy Edgington in later years, is a wonderfully rich philosophical theory which rests on the formal theory of subjective probability and which regards Lewis’ theorem as providing additional support for this theory of conditionals. It is not without critics, of course, but it is certainly a serious contender in the friendly battle for the best philosophical arguments concerning indicative conditionals.

Remark on the Suppositional Theory of conditionals: Please take a look at
http://plato.stanford.edu/entries/conditionals/#SupThe
for more on this theory.
According to the Suppositional Theory, an indicative conditional does not express a proposition and is not true or false (at least according to the most straight forward version of the theory). That is also why we spoke of the "degree of acceptability" of an indicative conditional in this lecture (as many philosophers do in this area); if we had said "degree of belief" of an indicative conditional, then this could have been understood as the result of applying a degree of belief function P to a proposition of the form $X \to Y$, which would have contradicted the Suppositional Theory from the start. It is better to stay neutral initially by speaking of "degrees of acceptability" and only to determine later what philosophical conclusions one can derive from what philosophical assumptions.

Quiz 39:
Assume Thesis 2 from the lecture to be true again; rational degrees of acceptability for indicative conditionals are measured in terms of the corresponding conditional probabilities. Now consider a fair six-sided die; say, your degree of belief function P reflects the fairness of the six-sided die: what is then your degree of acceptability for the indicative conditional 'If in the next throw the die rolls an even number, then it will roll a 2'?
Solution

4.10 Conclusions (07:44)

Today we have dealt with 'if-then': conditionals and conditional degrees of belief. We distinguished between conditionals in the indicative and the subjunctive mood, and we found that in mathematics an indicative conditional of the form 'if A then B' can be understood as the descriptive sentence '$\neg A \vee B$' or something that is logically equivalent to it. While this is a perfectly respectable way of understanding if-then in the context of mathematical theorems and proofs, we also saw that it is questionable whether this is a plausible understanding of the indicative if-then in natural language. Then we turned to conditional degrees of belief, we introduced formally the concept of conditional probability, and we argued that the conditionalization of a given degree of belief function on a proposition could be considered as capturing the effect that the supposition of that proposition has on the given degree of belief function. Finally, we asked ourselves how to measure the degree of acceptability of an indicative conditional: we discovered two natural ways of answering that question, but with Lewis we also proved it to be impossible for both of them to be true. So we ended up with the open question which of the two answers is the right one, if any of them, which is where we left the discussion.

As always, let me wrap up with a couple of remarks and with responses to some worries that you might have.

First of all: what we considered today was a chapter of the philosophy of language; more particularly: of the philosophy of conditional sentences in language. But it is also possible to interpret Lewis' theorem as an insight into some of our proposition attitudes: one might interpret it as saying that beliefs in conditionals do not generally coincide with conditional beliefs. From that point of view, the theorem teaches us something that might be relevant to the philosophy of mind and even to empirical science, especially cognitive psychology: initially one might have thought that believing a conditional to be true is the same mental state as believing a proposition to be true on the supposition of another, but by Lewis' theorem that does not seem to be so. One can easily imagine that at some points in history, when philosophers of mind or cognitive psychologists referred to beliefs in a conditional, they actually should have referred to conditional beliefs, and vice versa. They did not realize that they were dealing with two phenomena here, not just with one. If so, Lewis' theorem may well help us to avoid potential misunderstandings and errors also in areas outside of the philosophy of language. And by now it might not be surprising for you to hear that Lewis' paper on his theorem is heavily cited e.g. by computer scientists who are interested in the representations of conditional knowledge in data bases.

Secondly: Within the philosophy of language, there is no doubt that Adams' proposal of relating the acceptability of indicative conditionals to conditional probabilities, and Lewis' theorem on probabilities of conditionals vs conditional probabilities led to enormous progress on the whole subject matter. Without probability theory, without mathematics, that kind of progress would not have been feasible. It would have been impossible to detect and justify Lewis' result without using mathematical terms and without applying probabilistic reasoning. And without making Ramsey's footnote precise in terms of conditional probability, it would have remained unclear what kind of conclusions we would be entitled to draw from it: "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" is a nice suggestion, but without sharpening that thesis probabilistically we would not have been able to see that it actually excludes indicative conditionals from expressing propositions and from being true or false, at least given the usual background assumptions.

Thirdly: In the last lecture we considered both all-or-nothing belief and numerical degrees of belief, and we discussed their properties when they are rational. Today we dealt a lot with conditional degrees of belief. But how about conditional all-or-nothing belief? Isn't there anything like that at all? Don't worry: there is. There is a very well-developed theory of conditional belief – believing Y on the supposition of X – that can be extracted from what is called the theory of belief revision in formal epistemology and also from what is called nonmonotonic reasoning in theoretical computer science. These theories can also be used to give alternative all-or-nothing formalizations of Ramsey's footnote, and there are similar kinds of triviality results that one can derive from assuming that all-or-nothing beliefs in conditionals always coincide with conditional all-or-nothing beliefs.

Remark on belief revision theory and nonmonotonic reasoning: In these areas one does not study numerical conditional degrees of belief (such as $P(Y|X)$) but rather all-or-nothing types of conditional belief that one can express, e.g., in terms of 'Revising my present beliefs in light of evidence X leads to belief in Y' or 'If X is the case, then plausibly (or normally) also Y is the case'. If you want to know more about belief revision theory, please take a look at
http://plato.stanford.edu/entries/logic-conditionals/,
http://plato.stanford.edu/entries/logic-belief-revision/,
and
http://plato.stanford.edu/entries/formal-belief/#BelRevThe.
If you like to read more about nonmonotonic reasoning, please check out
http://plato.stanford.edu/entries/logic-nonmonotonic/
and
http://plato.stanford.edu/entries/formal-belief/#NonRea.

Finally, you might worry now about the indicative conditionals as used in this course: as everyone else I have been speaking in terms of conditional sentences a lot even within the context of these lectures. What did I mean by them? Fortunately, there is not much of a worry as far as these lectures are concerned: when I asserted such a conditional, you understood me quite well on a pre-theoretic level, simply because you are a competent speaker of English, and for pretty much everything that we discussed so far this was absolutely sufficient in order to follow. What is more: for all of the indicative conditionals $A \to B$ that I employed in the various bits of philosophical theory here and there, you could think of them adequately as $\neg A \vee B$ again, just as mathematicians would have it: as far as I can see, nothing bad at all will follow from this. Similarly, if you wanted to reconstruct in precise terms the usual theories in, say, physics, then the indicative if-then as used there can also always be understood as $\neg A \vee B$ without causing any problems. This is shown by the vast literature on logical reconstructions of physical theories by philosophers of science who do analyze the if-then in physics in terms of $\neg A \vee B$ again. It seems that understanding $A \to B$ in terms of $\neg A \vee B$ is not as black as it is often painted. Or maybe theoretical discourse in physics, just as theoretical discourse in philosophy if guided by mathematical reasoning, are both sufficiently regimented in order for the analysis of $A \to B$ by means of $\neg A \vee B$ to be applicable.

In any case: if this is a conditional, then that must be the end of this lecture.

Here are some references to the relevant background literature:
On conditionals in general:
Bennett, J., *A Philosophical Guide to Conditionals*, Oxford: Clarendon Press, 2003.
Edgington, D., "On Conditionals", *Mind* 104 (1995), 235-329.
On conditional probability from the philosophical point of view:
Hájek, A., "What conditional probability could not be", *Synthese* 137/3 (2003), 273-323. (The paper also discusses conceptions of conditional probability according to which conditional probability is not defined in terms of the usual ratio formula. This is particularly interesting if one wants to have a well-defined notion of conditional probability of the form $P(Y|X)$ even when $P(X)$ equals 0.)
On the Suppositional Theory of conditionals:
See chapter 4 of Bennett (2003), Edgington (1995), but also:
Adams, E. W., *The Logic of Conditionals*, Dordrecht: Reidel, 1975.
Adams, E. W., *A Primer of Probability Logic*, Stanford: CLSI Publications, 1998. (Adams' book from 1998 I can especially recommend to you, as it is very accessible and it is written very nicely.)
On Lewis' theorem:
See chapter 5 of Bennett (2003), Edgington (1995), but also:
Hájek, A., Hall, N., "The Hypothesis of the Conditional Construal of Conditional Probability", in: Eells, E. and Skyrms, B., *Probability and Conditionals*, Cambridge: Cambridge University Press, 1994, pp. 75-111.
Hájek, A., "Triviality on the Cheap?", in Eells and Skyrms (1994), pp. 113-140.

Appendix A

Quiz Solutions Week 4: If-then

Quiz 31:

Compare (this is an example from Jonathan Bennett):

(1) If Shakespeare did not write Hamlet, then some aristocrat did.

(2) If Shakespeare had not written Hamlet, then some aristocrat would have.

Which one is subjunctive, which one is indicative?

Which one sounds acceptable to you?

SOLUTION Quiz 31:

(1) is in the indicative mood, (2) is in the subjunctive mood. Acceptability is a subjective state, so you will have to check this for yourself, but at least to me (1) is acceptable (I regard it as very likely that Shakespeare wrote Hamlet; but if he did not, since someone must have written Hamlet around that time, probably some aristocrat did); and to me (2) is not acceptable (I think that Shakespeare wrote Hamlet; but if the world had developed differently, so that Shakespeare had not written Hamlet, then I am very sure that no one else would have written Hamlet, that is, no one would have written the play Hamlet as we know it).

Back to quiz

Quiz 32:

Please show that $\neg A \vee B$ is logically equivalent to $\neg A \vee (A \wedge B)$. What you have to do is: First assume that $\neg A \vee B$ is true, and then prove on that basis that also $\neg A \vee (A \wedge B)$ is true. Secondly, the other way around: assume that $\neg A \vee (A \wedge B)$ is true, and then prove on that basis that also $\neg A \vee B$ is true.

SOLUTION Quiz 32:

Solution: Assume $\neg A \vee B$ is true. Since we understand 'or' inclusively, so that 'or' leaves open the possibility that both $\neg A$ and B are the case, precisely one of the following three possibilities must be the case:

(i) $\neg A$ is true, B is false;

(ii) $\neg A$ is false, B is true;

(iii) $\neg A$ is true, B is true.

In case (i), $\neg A \vee (A \wedge B)$ is true because $\neg A$ is true.

In case (ii), $\neg A \vee (A \wedge B)$ is true since A is true (because $\neg A$ is false) and also B is true, which is why $A \wedge B$ is true.

In case (iii), $\neg A \vee (A \wedge B)$ is true as $\neg A$ is true.

So in every possible case in which $\neg A or B$ is true, also $\neg A \vee (A \wedge B)$ is true.

Now the other direction: Assume $\neg A \vee (A \wedge B)$ is true. Here there are really only two possible cases, as it cannot be the case that both $\neg A$ is true and $A \wedge B$ is true (for $\neg A$ and A cannot both be true). So two possible cases remain:

(i) $\neg A$ is true, $A \wedge B$ is false.

(ii) $\neg A$ is false, $A \wedge B$ is true.

In case (i), $\neg A \vee B$ is true because $\neg A$ is true.

In case (ii), $\neg A \vee B$ is true since B is true (as even $A \wedge B$ is true).

Thus, in every possible case in which $\neg A \vee (A \wedge B)$ is true, also $\neg A \vee B$ is true.

Taking both directions together we have: $\neg A \vee B$ is logically equivalent to $\neg A \vee (A \wedge B)$.

Back to quiz

Quiz 33:

(1): Picture the set W of all possible worlds in terms of a square again; draw two distinct but intersecting circles in the square – one representing the proposition X, the other one representing the proposition Y. Assume that sentence A expresses proposition X and sentence B expresses proposition Y. Now consider every logically possible position that the actual world might have in the diagram (there are four regions to consider): determine for each position the corresponding truth values of A (or X), B (or Y), and $\neg A \vee B$ (or $\neg X \vee Y$), where 'true' means 'true at the actual world', and 'false' means 'false at the actual world'.

(2): When e.g. a philosopher puts forward an argument of the form

(P1) Not B.

(P2) If A then B.

Therefore: (C) Not A.

we regard the argument as logically valid. (Remember the additional problem set for Lecture 1.) Prove that indeed this argument form turns out to be valid, if we understand 'If A then B' as 'Not A or B' ($\neg A \vee B$).

SOLUTION Quiz 33:

Solution (1):

Case (i) A is true, B is true: $\neg A \vee B$ is true. Case (ii) A is true, B is false: $\neg A \vee B$ is false. Case (iii) A is false, B is true: $\neg A \vee B$ is true. Case (iv) A is false, B is false: $\neg A \vee B$ is true.

See Figure A.1.

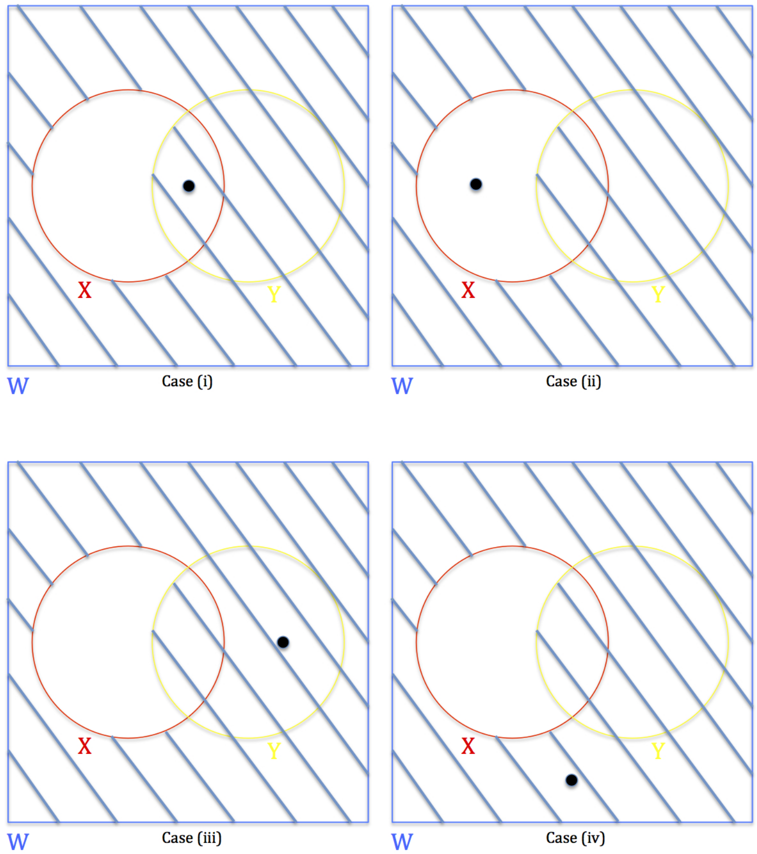

Figure A.1: Cases (i)-(iv)

What we have determined in this way is the so-called truth table of the material conditional. A material conditional $A \to B$ is false in precisely one case, that is, in case A is true and B is false; in all other cases, it is true.

Solution (2): So we are considering an argument of the form

(P1) Not B.

(P2) Not A or B.

Therefore: (C) Not A.

Assume P1 and P2 to be true: hence, by P1, B is false, and by P2, $\neg A$ is true or B is true. But the truth of B had just been ruled out by P1, from which it follows that $\neg A$ must be true, which is what the conclusion C says.

The argument form is therefore logically valid.

Back to quiz

Quiz 34:

Prove what is essentially the converse of the theorem that had we had proven in Lecture 3-12. That is:

Let W be a finite non-empty set of possible worlds. Let B be an assignment of non-negative real numbers to the worlds in W – that is, each world w in W is assigned a non-negative real number $B(w)$ – such that summing up the B-values of all worlds in W yields the numerical value 1. Finally, assume that for all propositions X (that is, all subsets X of W), $P(X)$ is defined as the sum of all $B(w)$ for which w is a member of X:

$P(X) = \sum_{w \text{ in } X} B(w)$.

Prove:

(i) $P(W) = 1$.

(ii) For all propositions X: $P(X)$ is a real number, such that $0 \leq P(X) \leq 1$.

(iii) For all propositions X, Y: if $X \cap Y = \{\}$, then $P(X \cup Y) = P(X) + P(Y)$.

SOLUTION Quiz 34:

(i) By the definition of 'P', $P(W) = \sum_{w \text{ in } W} B(w)$. By assumption, the sum of B-values of worlds in W is 1. Therefore, $P(W) = 1$.

(ii) This follows immediately from the definition of 'P', from the assumption that each number $B(w)$ is a non-negative real number, and from the assumption that the sum of B-values of all worlds in W is 1. (In the case where X is the empty set, one understands the corresponding "empty" sum to yield the numerical value 0.)

(iii) Suppose $X \cap Y = \{\}$: By the definition of 'P', $P(X \cup Y) = \sum_{w \text{ in } X \cup Y} B(w)$. Since $X \cap Y = \{\}$ – the set X does not overlap with the set Y – we can split this sum into two sums: $\sum_{w \text{ in } X \cup Y} B(w) = \sum_{w \text{ in } X} B(w) + \sum_{w \text{ in } Y} B(w)$. And by the definition of 'P' again, we can rewrite these two sums as follows: $\sum_{w \text{ in } X} B(w) + \sum_{w \text{ in } Y} B(w) = P(X) + P(Y)$. Hence, we have: $P(X \cup Y) = P(X) + P(Y)$.

Back to quiz

Quiz 35:

Say, we are given a degree of belief function P that satisfies the laws of probability; so P is a subjective probability measure. And assume X to be a proposition, such that $P(X) > 0$. Then we determine P_X as explained in the lecture (which is a probability measure again). Now suppose we would apply the same procedure to P_X again: that is, we determine $(P_X)_X$. What would happen?

SOLUTION Quiz 35:

Nothing would happen. When P_X was determined, all values of worlds outside of X were set to 0, and the values of worlds in X all got multiplied by $\frac{1}{P(X)}$. The so-determined degree of belief function P_X is a probability measure that satisfies the additional property: $P_X(X) = 1$ (since all the worlds outside of X have value 0, the sum of values of all worlds is 1, and in fact in the present situation the sum of the values of worlds in X must already be 1; we will return to this point in 4-6). Setting the worlds outside of X to 0 once again does not change anything. And multiplying the values of worlds in X by $P_X(X)$ does not change anything either, since $P_X(X) = 1$. So $(P_X)_X$ is simply identical to P_X: nothing happens by supposing X once again.

Back to quiz

Quiz 36:

Let P be given, as in the lecture, by the following B assignment:

$B(w_1) = 1/15$

$B(w_2) = 1/3$

$B(w_3) = 1/15$

$B(w_4) = 1/15$

$B(w_5) = 1/3$

$B(w_6) = 1/15$

$B(w_7) = 1/15$

$B(w_8) = 0$

Determine:

(i) $P(\{w_1, w_3\}|\{w_1, w_3, w_4, w_6, w_7\})$,

(ii) $P(\{w_1, w_3\}|\{w_3, w_4, w_7\})$,

(iii) $P(\{w_1\}|\{w_1, w_2\})$.

SOLUTION Quiz 36:

By the definition of conditional probability, we have:

(i) $P(\{w_1, w_3\}|\{w_1, w_3, w_4, w_6, w_7\}) = \frac{P(\{w_1,w_3,w_4,w_6,w_7\}\cap\{w_1,w_3\})}{P(\{w_1,w_3,w_4,w_6,w_7\})} = \frac{P(\{w_1,w_3\})}{P(\{w_1,w_3,w_4,w_6,w_7\})} = \frac{2/15}{5/15} = \frac{2}{5}$;

(ii) $P(\{w_1, w_3\}|\{w_3, w_4, w_7\}) = \frac{P(\{w_3,w_4,w_7\}\cap\{w_1,w_3\})}{P(\{w_3,w_4,w_7\})} = \frac{P(\{w_3\})}{P(\{w_3,w_4,w_7\})} = \frac{1/15}{3/15} = \frac{1}{3}$;

(iii) $P(\{w_1\}|\{w_1, w_2\}) = \frac{P(\{w_1,w_2\}\cap\{w_1\})}{P(\{w_1,w_2\})} = \frac{P(\{w_1\})}{P(\{w_1,w_2\})} = \frac{1/15}{1/15+1/3} = \frac{1/15}{6/15} = \frac{1}{6}$.

Back to quiz

Quiz 37:

Assume Thesis 2 from the lecture to be true; that is, for all descriptive sentences A, B, where A expresses the proposition X, and B expresses the proposition Y, it holds for every subjective probability measure P:

the degree of acceptability for the indicative conditional $A \to B$ (rel. to P) is identical to $P(Y|X)$.

Given that this is so, what is the degree of acceptability that a (perfectly rational) person ought to assign to an indicative conditional of the form $A \to A$ (where A expresses some proposition X, such that $P(X) > 0$)?

SOLUTION Quiz 37:

The degree of belief function of a perfectly rational person is a subjective probability measure P. By Thesis 2, the degree of acceptability for $A \to A$ relative to a subjective probability measure P is identical to $P(X|X)$. And $P(X|X) = 1$ (for $P(X) > 0$). So the person ought to accept $A \to A$ with maximal degree 1, which sounds plausible indeed.

Back to quiz

Quiz 39:

Assume Thesis 2 from the lecture to be true again; rational degrees of acceptability for indicative conditionals are measured in terms of the corresponding conditional probabilities. Now consider a fair six-sided die; say, your degree of belief function P reflects the fairness of the six-sided die: what is then your degree of acceptability for the indicative conditional 'If in the next throw the die rolls an even number, then it will roll a 2'?

SOLUTION Quiz 39:

Since the die is six-sided, we may assume the relevant set W of possible worlds to consist of precisely six worlds that correspond to: die rolls 1; die rolls 2;...; die rolls 6. Since the six-sided die is also fair, your degree of belief function P ought to have the following property: $P(\{\text{die rolls 1}\}) = P(\{\text{die rolls 2}\}) = \ldots = P(\{\text{die rolls 6}\}) = 1/6$. By Thesis 2, your degree of acceptability for 'If in the next throw the die rolls an even number, then it will roll a 2' must be your conditional probability of the die rolling a 2 in the next throw given that the die rolls an even number in the next throw. That conditional probability is $P(\{\text{die rolls 2}\}|\{\text{die rolls 2, die rolls 4, die rolls 6}\}) = \frac{P((\{\text{die rolls 2, die rolls 4, die rolls 6}\}\cap\{\text{die rolls 2}\})}{P(\{\text{die rolls 2, die rolls 4, die rolls 6}\})} = \frac{P(\{\text{die rolls 2}\})}{P(\{\text{die rolls 2, die rolls 4, die rolls 6}\})} = \frac{1/6}{3/6} = \frac{1}{3}$.

Please check for yourself: Does this sound plausible to you? Would you have regarded 'If in the next throw the die rolls an even number, then it will roll a 2' as more likely, as more acceptable than 1/3? Or as less likely/acceptable?

Back to quiz

List of Figures