

Introduction to Mathematical Philosophy

Hannes Leitgeb, Stephan Hartmann

June 19, 2015

Contents

Overview	v
3 Rational Belief	1
3.1 Introduction (05:51)	1
3.2 Propositions and Possible Worlds (08:26)	4
3.3 Propositions as Sets of Possible Worlds I (11:56)	7
3.4 Propositions as Sets of Possible Worlds II (10:14)	15
3.5 Logical Implication and the Subset Relation (06:06)	28
3.6 Belief in Propositions (09:51)	34
3.7 Postulates on Rational Belief (05:34)	38
3.8 Proving a Theorem on Rational Belief (08:35)	40
3.9 Philosophical Implications (10:10)	45
3.10 Rational Degrees of Belief (07:50)	51
3.11 Postulates on Rational Degrees of Belief (12:54)	59
3.12 A Theorem on Rational Degrees of Belief (07:06)	64
3.13 Rational Degrees of Belief and Bets (15:16)	68
3.14 The Lottery Paradox (10:33)	74
3.15 Conclusions (07:51)	81
A Quiz Solutions Week 3: Rational Belief	85

Week 3: Overview

Overview of Lecture 3: Rational Belief

3.1 Introduction: Belief is a major topic in epistemology and the philosophy of mind. It is a certain kind of attitude towards propositions; or expressed differently: the belief that so-and-so has the proposition that so-and-so as its content. In order to understand belief properly, we first need to understand propositions a bit better. In philosophy, propositions have three roles to play: they are the meanings of descriptive sentences; they are the bearers of truth values; and they are the contents of many of our mental states.

3.2 Propositions and Possible Worlds: Propositions are true or false at possible worlds; one of these possible worlds is the actual world, and a proposition is true at it if and only if the proposition is true. We argue for two principles on possible worlds and propositions: worlds are identical if and only if the same propositions are true at them; and propositions are identical if and only if they are true at the same worlds.

3.3 Propositions as Sets of Possible Worlds I: If one defines propositions as sets of possible worlds, so that every proposition is identical to the set of worlds at which it is true, then these two principles on possible worlds and propositions can be derived from that definition (using elementary set theory). Furthermore, the definition enables us to understand negation ('not') for propositions in terms of set-theoretic complement, conjunction ('and') of propositions in terms of set-theoretic intersection, and disjunction ('or') of propositions in terms of set-theoretic union.

3.4 Propositions as Sets of Possible Worlds II: We look at various examples of logical operations on propositions which now coincide with certain set-theoretic operations. And we can picture the results of these logical operations in the same way in which one pictures the results of set-theoretic operations (that is, in terms of geometrical regions).

3.5 Logical Implication and the Subset Relation: We find that the subset relation is the set-theoretic counterpart of the relationship of logical implication between propositions.

3.6 Belief in Propositions: If one believes that so-and-so, one normally acts as if so-and-so is the case, one is willing to assert that so-and-so, that so-and-so belongs to one's picture

of the world, and one draws logical inferences based on this belief that so-and-so. In the following we concentrate on the inference/reasoning-related aspects of beliefs; we do so for perfectly rational persons and their beliefs, and we restrict ourselves, for the sake of simplicity, to cases of beliefs in propositions where the underlying set of possible worlds is a finite set.

3.7 Postulates on Rational Belief: We argue for four postulates on rational belief which say, essentially, that the beliefs of a perfectly rational person are closed under the rules of logic. We state a theorem that characterizes what the belief system of a perfectly rational person looks like given our four postulates: there is always some least or smallest believed proposition (which we call ' B_W ') in the sense of the subset relation, and the person in question believes a proposition if and only if B_W is a subset of that proposition (B_W logically implies that proposition).

3.8 Proving a Theorem on Rational Belief: Then we prove that theorem in full detail.

3.9 Philosophical Implications: The theorem implies that rationality has an information-compressing effect: in order to determine all of a perfectly rational person's beliefs, one does not need to determine for each proposition whether it is believed or not, instead it suffices to determine for each possible world whether it is a member of B_W or not. By comparing a proposition with B_W it also becomes very easy to determine whether the proposition is believed by the person in question, or whether it is disbelieved (its negation is believed), or whether the person suspends judgement on the proposition (does not have an opinion about it nor about its negation). We find that rational belief is subject to the same kinds of mathematical structures that one can find also in one of the traditional areas of mathematics, that is, algebra.

3.10 Rational Degrees of Belief: Other than the three-fold distinction concerning all-or-nothing belief – believing, disbelieving, or suspending judgement on a proposition – one can also assign a numerical degree of belief to a proposition that measures one's strength of belief in the proposition. We draw an analogy between such degrees of belief and the areas of regions in the set-theoretic diagrams that we used to picture propositions.

3.11 Postulates on Rational Degrees of Belief: Based on this analogy, we argue for three postulates on rational degrees of belief, which say essentially that the distribution of a perfectly rational person's degrees of belief over propositions satisfies the mathematical laws of probability. From the three postulates we can derive also some further plausible rationality constraints on degrees of belief.

3.12 A Theorem on Rational Degrees of Belief: We state a theorem that characterizes what the degree of belief function of a perfectly rational person looks like given our three postulates: there is always an assignment (which we call ' B ') of non-negative real numbers to worlds where the sum of all these numbers is 1 and where the degree of belief in a proposition is equal to the sum of values assigned to the worlds that are members of

the proposition (the worlds at which the proposition is true). It follows once again that rationality has an information-compressing effect: in order to determine all of a perfectly rational person's degrees of beliefs, one does not need to determine for each proposition its degree of belief, instead it suffices to determine for each world its B -value.

3.13 Rational Degrees of Belief and Bets: Our three postulates on degrees of belief can be justified by an argument from a rational person's betting behavior: if a person does not distribute her degrees of belief over propositions in line with the laws of probability, then it is always possible to construct a sequence of bets, such that the person will consider each of the bets as fair given her degrees of belief, but where accepting all of the bets leads to sure loss overall.

3.14 The Lottery Paradox: It is not clear how rational all-or-nothing belief relates to rational degrees of belief. The Lottery Paradox shows that combining seemingly plausible assumptions on rational belief and degrees of belief logically implies an absurdity. It is a matter of present research in formal epistemology which of these seemingly plausible assumptions ought to be given up.

3.15 Conclusions: Based on a set-theoretic conception of propositions, a theory of rational belief can be developed in terms of concepts that correspond to mathematical concepts known from algebra. And rational degrees of belief can be argued to obey the mathematical laws of probability. It is an open question what a joint theory of rational belief and rational degrees of belief should look like.

Chapter 3

Week 3: Rational Belief

3.1 Introduction (05:51)

Welcome to the third lecture of our Introduction to Mathematical Philosophy! In the second lecture, we defined truth, first recursively and then explicitly, for a little formal toy language, we proved some basic laws of truth by means of our definition of truth and the principle of complete induction over sentences, and while doing all of that, we managed to steer clear of the Liar paradox. That lecture was a tough one, wasn't it? At the beginning of the last lecture we also found that it was important to distinguish carefully 'being true' from 'believing to be true' – today we will turn our attention to the second of these concepts, belief. We will see that when belief is rational, it exhibits well-known mathematical structures again: concepts from logic, set theory, algebra, and probability theory will become the means by which we can philosophically study, and systematically apply, the concept of rational belief. And once again some of the properties of rational belief we will be able to justify with the help of mathematical proof.

Belief is a particular kind of mental state; it is one of the so-called propositional attitudes: mental attitudes towards propositions.

(Slide 1)

- I believe that *Bayern Munich wins the Champions League next year*.
 - I hope that *Bayern Munich wins the Champions League next year*.
 - I desire that *Bayern Munich wins the Champions League next year*.
 - Others fear that *Bayern Munich wins the Champions League next year*.
- ⋮

These are all instances of different attitudes – belief, hope, desire, fear – but all of these instances of attitudes are directed towards propositions as their contents. In my examples, in fact, directed to one and the same proposition: one and the same thing is believed, hoped, desired, by myself, and feared, by others, and that thing is the proposition that Bayern Munich wins the Champions League next year.

We had already mentioned propositions early in the last lecture: On the one hand, propositions are what gets expressed by descriptive sentences – e.g., the sentence ‘Bayern Munich wins the Champions League next year’ expresses the proposition that Bayern Munich wins the Champions League next year; accordingly, the same proposition is expressed also by the German sentence ‘Bayern München gewinnt die Champions League nächstes Jahr’.

(Slide 2)

First role of propositions:

- They are the meanings of descriptive sentences:

‘Bayern Munich wins the Champions League next year’ expresses the same proposition as ‘Bayern München gewinnt die Champions League nächstes Jahr’.

Both express *that Bayern Munich wins the Champions League next year*.

So different sentences, even from different languages, may express the same proposition, as long as these sentences are synonymous with each other: one role that propositions play in philosophy is that of the meanings of descriptive sentences, and instead of saying that different sentences have the same meaning, we can also say that they express one and the same proposition.

Secondly, as we had also seen in the last lecture, we may ascribe truth or falsity either to descriptive sentences or to the propositions that are expressed by them: we may either say that the sentence ‘Bayern Munich wins the Champions League next year’ is true, or alternatively that the proposition that Bayern Munich wins the Champions League next year is true, or shorter: that Bayern Munich wins the Champions League next year is true. So the second role of propositions is that of bearers of truth values; they belong to the kinds of things of which we can say sensibly and informatively that they are true or false.

(Slide 3)

Second role of propositions:

- They are bearers of truth values:

That Bayern Munich wins the Champions League next year is true or false.

We might even think that propositions are the primary bearers of truth values: the reason why both of the two sentences ‘Bayern Munich wins the Champions League next year’ and ‘Bayern München gewinnt die Champions League nächstes Jahr’ are true is that the two of them express the same proposition, and that proposition is true – so the truth of descriptive sentences might be said to derive from the truth of the propositions that they express, and not the other way around. On the other hand, it is much easier to say what a sentence is – a string of symbols put together in line with certain grammatical rules – while it is much harder to say what a proposition is. Propositions are simply pretty strange abstract objects; which is why it may still be preferable to first define truth for descriptive sentences and not for propositions, which is exactly what we did in the last lecture. But never mind, that issue is not so important for today.

For now it will be important that propositions are precisely the kinds of things that are believed, which is also why propositions will play a more salient role than sentences in this lecture.

(Slide 4)

Third role of propositions:

- They are the contents of belief:

I believe *that Bayern Munich wins the Champions League next year*.

For the same reason we also need to get a somewhat clearer picture of propositions: we need to find a sensible way of thinking about them and of how different propositions relate to each other logically, before we can turn to the question of what it means to believe a proposition, and in particular, what it means to believe a proposition rationally. Gaining at least a basic understanding of the contents of propositional attitudes is a presupposition for understanding the propositional attitudes themselves. Therefore, let us focus on propositions first before we turn to belief in a proposition.

Remark on propositions: There are various different theories of propositions in philosophy, and there are different types of propositions that serve different theoretical purposes. We will focus on just one theory and just one type here: propositions as sets of possible worlds. Propositions in this sense are quite coarse-grained objects: e.g., every two distinct logically true sentences, such as ‘Snow is white or snow is not white’ and ‘It is not the case that (snow is white and snow is not white)’ are true in the same possible worlds – namely, in all possible worlds – and hence they end up expressing the same proposition, that is, the same set of possible worlds. So whatever differences in meaning there might be between these two sentences is abstracted away by turning to their corresponding sets of possible worlds. And propositions as sets of possible worlds are also unstructured objects: a proposition just by itself does not come with any kind of ordering of worlds or of other objects, since it is merely a set of worlds. But for some purposes it might make better sense to assume that propositions are some kind of structured objects. The main reason for conceiving of propositions as sets of possible worlds is a combination of simplicity and, nevertheless, theoretical power; and for the purposes of our lectures, propositions in this sense will be perfectly adequate (while for other purposes this would not be so). If you want to know more about propositions in general, and about more fine-grained and structured types of propositions in particular, take a look at

<http://plato.stanford.edu/entries/propositions/>

and

<http://plato.stanford.edu/entries/propositions-structured/.>

By the way, there are also philosophers who deny the existence of propositions altogether: Willard van Orman Quine is the most famous example; see

<http://plato.stanford.edu/entries/quine/>

if you like to read more about his work.

3.2 Propositions and Possible Worlds (08:26)

Consider the proposition that Socrates is a philosopher. We believe that the proposition is true. But if things had been different, it might not have been true: for instance, what if Socrates’ parents had been killed very early in his life, and because of that he would not have received any formal education? Consequently, he would not have been able to read, say, Anaxagoras as an adolescent, and maybe he would not have become a philosopher then. Here is a way of expressing that thought:

The proposition that Socrates is a philosopher is true, that is, it is true in the actual world: in the world that we actually inhabit. ‘True in the actual world’ is simply a long-winded way of saying ‘true’. However, in some other possible world, one that is different from ours, the proposition that Socrates is a philosopher is false: in that world, e.g., Socrates’ parents died early, he could not benefit from proper education, and ultimately he did not

become a philosopher. Accordingly, the proposition that Socrates is not a philosopher is false at the actual world, while it is true in that other hypothetical world. So propositions are true in some possible worlds, and false in others.

(Slide 5)

The proposition *that Socrates is a philosopher* is true at the actual world, but it is not true at every possible world.

In the first lecture, we had already mentioned that the term ‘infinitely powerful’ as ascribed to God might mean something like: there are infinitely many possible ways the world might be – infinitely many possible worlds – and for each of them God would in principle have the power to make it real or actual. What we add to this now is that, accordingly, propositions are true or false at such possible worlds. The actual world is a possible world that is distinguished by the feature that a proposition is true at it if and only if the proposition is true. But propositions are also true or false at the merely possible worlds, that is, at the possible worlds that are not actual, that differ from the actual world.

If you want to know more about possible worlds, as they are studied in metaphysics, please check out

<http://plato.stanford.edu/entries/possible-objects/>.

We might even take one further step: if one possible world w is distinct from another possible world, w' , then they must also differ in terms of the propositions that are true at them: there must be a proposition that is true at the one world but not true at the other. For instance: the actual world and the merely possible world that I described before are distinct, and indeed their difference shows up in the fact that the proposition that Socrates is a philosopher is true in the one but false in the other world. In fact, that is precisely the respect – or, in any case, one respect – in which I explained to you what the two worlds were all about and why and how they would differ. Two distinct worlds which would be completely indistinguishable in terms of the truths and falsehoods that hold at them would seem very much like a distinction without a difference.

And accordingly: if one proposition X is distinct from another proposition, Y , then they must also differ in terms of the possible worlds at which they are true: there must be a world at which the one proposition is true while the other one is not. For example, the proposition that Socrates is a philosopher differs from its negation, the proposition that Socrates isn’t a philosopher. And indeed there is a world in which the one proposition is true while the other one is not: for example, the proposition that Socrates is a philosopher is true at the actual world, while the proposition that Socrates is not a philosopher is not true at the actual world. Once again, two distinct propositions which would nevertheless be indistinguishable in terms of the worlds at which they are true would seem very much like a distinction without a difference.

Therefore, the following two principles on possible worlds and propositions look plausible:

(Slide 6)

- (IdWorlds) For all possible worlds w , for all possible worlds w' :
 $w = w'$ if and only if for all propositions X : X is true at w if and only if X is true at w' .
- (IdProp) For all propositions X , for all propositions Y :
 $X = Y$ if and only if for all possible worlds w : X is true at w if and only if Y is true at w .

The point of these principles is their right-to-left directions, not their left-to-right directions: if the left-hand side of either of these principles is true, then the corresponding right-hand side holds trivially; we would not even have to postulate that it holds by means of some special principle: e.g., if w is identical to w' , then of course, by logic, w and w' have the same properties, including that if w has the property that a proposition X is true at it, then w' has the same property that X is true at it, and the other way around. Similarly in the case when a proposition X is identical to a proposition Y : if X has the property that it is true at a world w , then, by logic, also Y has the property of being true at w , and vice versa.

Remark on the logic of the ‘=’ symbol: What is required here is the simple logical principle of the indiscernibility of identicals: if x is identical to y , then x has precisely the same properties as y ; that principle is part of standard (so-called first-order) predicate logic, which one learns about in introductory logic courses. The converse principle is Leibniz’ principle of the identity of indiscernibles: if you like to read more about it, take a look at <http://plato.stanford.edu/entries/identity-indiscernible/>.

The principle of identity of indiscernibles is part of so-called second-order logic, about which you can find more here:

<http://plato.stanford.edu/entries/logic-higher-order/>

(Watch out: that entry presupposes really some knowledge of first-order logic; but please don’t worry about it: that material is completely optional, and our course does not rely on it at all.)

So the actual meat of the two principles is given by their right-to-left directions: if for all propositions X it holds that X is true at w if and only if X is true at w' , that is, if the same propositions are true at w and w' , then by the first principle, the possible world w is identical to the possible world w' .

And analogously, if for all possible worlds w it holds that X is true at w if and only if Y

is true at w , that is, if X and Y are true at the same possible worlds, then by the second principle, the proposition X is identical to the proposition Y .

We have identified two plausible principles of identity for possible worlds and propositions. One should not interpret them as some kind of definitions of identity for possible worlds or propositions: we already know what the identity symbol ‘ $=$ ’ means, what ‘is identical to’ means; the two principles are simply meant to be true and informative statements which are of some interest to us; that’s all. Do not read too much into them. In fact, there is a lot that these principles do not tell us: most importantly, they do not tell us what a possible world is, nor what a proposition is. We are not told what a possible world is “made of”, as it were – of concrete matter in space and time, or of some abstract entities, or of some other “stuff” – nor do the principles hand us any information about how propositions are built up: from concrete objects and properties or relations, maybe, or from something else. But then again the principles are not completely empty either: they express how possible worlds and propositions relate to each other, how they come as a structured package; they involve what possible worlds and propositions, as it were, “do”: that propositions are true at possible worlds; and they give us a means of distinguishing between possible worlds as well as a means of distinguishing between propositions: possible worlds can be distinguished in terms of the propositions that are true at them, and propositions can be distinguished in terms of the worlds at which they are true.

Quiz 19:

Take our principles ($\text{Id}_{\text{Worlds}}$) and (Id_{Prop}) from the lecture as given.

- (1): Assume that a proposition X is true at world w but it is not true at world w' . Can we determine from this whether w is identical to/distinct from w' ?
- (2): Assume that propositions X and Y are true at precisely the same possible worlds. Can we determine from this whether X is identical to/distinct from Y ?

[Solution](#)

3.3 Propositions as Sets of Possible Worlds I (11:56)

Now let me take one final step: I am going to give you an account of propositions based on the notion of possible world, such that the two principles from above will follow from this account if combined with set theory. Here is the idea: simply identify a proposition with the set of possible worlds at which it is true. In other words:

(Slide 7)

Definition:

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

- (i) X is a proposition (over W) if and only if X is a subset of W .
- (ii) If X is a proposition (over W) and w is a world in W , then

X is true at w if and only if w is a member of X .

According to this set-theoretic account of propositions in terms of possible worlds, propositions are sets of possible worlds, and truth at a world coincides with including the world as a member (see Figure 3.1).

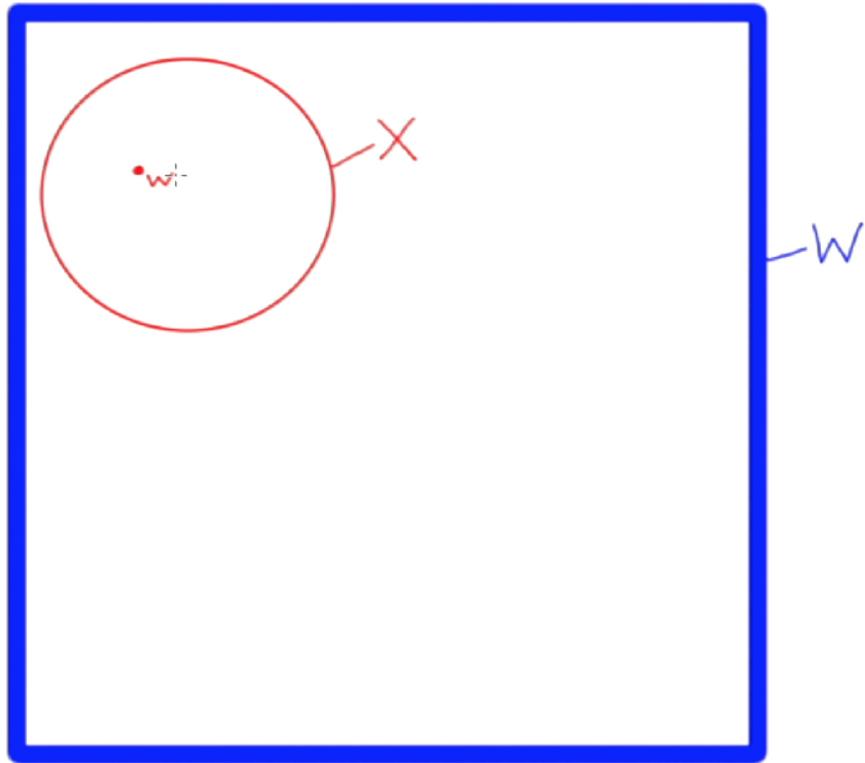


Figure 3.1: X is a subset of W and w is a member of X

Let us see what this implies.

By (i), every proposition X (over W) is a set, namely a subset of W , which is why it holds trivially that X is the set of all w for which it is the case that w is a member of X ; that is, X is the set of all w such that w is a member of X :

(Slide 8)

$$X = \{w: w \text{ is a member of } X\}$$

After all, that holds trivially for every set whatsoever: every set is precisely the set of its members.

And by (ii) from before we can replace ‘ w is a member of X ’ by ‘ X is true in w ’, from which we get, as promised: X is the set of w s such that X is true at w :

(Slide 9)

$$X = \{w: X \text{ is true at } w\}$$

According to (i) and (ii) from above, every proposition X (over W) is the set of worlds in W at which it, X , is true.

Remark on ‘over W ’: When I say ‘proposition (over W)’ in the last definition, then the ‘over W ’ is just a little reminder that from this point we always use the term ‘proposition’ in a context in which already some set W of possible worlds has been determined, so that propositions are subsets of that set W . But it’s not very important really.

Now let us assume a non-empty set W of possible worlds to be given, and let us take (i) and (ii) for granted: then we can always replace ‘proposition’ by ‘subset of W ’, and ‘is true at w ’ by ‘includes w as a member’. In this way, our original principles from before translate into:

(Slide 10)

- (Id'_{Worlds}) For all members w of W , for all members w' of W :

$w = w'$ if and only if for all subsets X of W : w is a member of X if and only if w' is a member of X .

- (Id'_{Prop}) For all subsets X of W , for all subsets Y of W :

$X = Y$ if and only if for all members w of W : w is a member of X if and only if w is a member of Y .

And we find that both (Id'_{Worlds}) and (Id'_{Prop}) follow quite trivially now from the principles of set theory. We only need to check out their right-to-left directions again: for (Id'_{Worlds}) , assume that w and w' are members of W , such that they are also members of precisely the

same subsets of W ; then w must be identical to w' : because, if w were distinct from w' , then w would e.g. be a member of the set $\{w\}$, the set that contains only w as a member, as explained in the first lecture, while w' would not be a member of $\{w\}$, and hence w and w' would not be members of precisely the same subsets of W after all, in contradiction with the assumption.

Similarly, for (Id'_{Prop}) : assume that X and Y are subsets of W , such that precisely the same worlds in W are members of them; since X and Y are both subsets of W , this means that they have precisely the same members overall, whether within W or outside of W ; but then by the extensionality principle for sets, which we had mentioned in the first lecture, X must be identical to Y . For if sets X and Y have precisely the same members, then they are identical.

So by identifying a proposition with a set, more particularly, with the set of possible worlds at which it is true, we are able to derive the two principles (Id_{Worlds}) and (Id_{Prop}) , which we had found to be plausible beforehand. This gives us at least a defeasible reason for regarding our account of propositions as sets of possible worlds to be plausible.

But there is more that this account has to offer: If propositions are subsets of some given set W of possible worlds, what is then the negation of a proposition? For instance, in our example from before, the sentence ‘Socrates is a philosopher’ was meant to express a proposition, which we now regard as a particular set of possible worlds, but the same holds for the negation of that sentence, that is, for the sentence ‘Socrates is not a philosopher’: that sentence should also express a proposition, a certain set of possible worlds. What is that set?

The answer is easy:

Consider a proposition X , some subset of W . Let us denote the negation of that proposition by: $\neg X$. $\neg X$ is another proposition and hence must be some subset of W again. Clearly, X and $\neg X$ cannot both be true at one and the same possible world: so there cannot be a world w , such that w is both a member of X and a member of $\neg X$. Moreover, for every possible world in W it must be the case that either X or $\neg X$ is true at w : hence, every possible world in W is either a member of X or a member of $\neg X$. Taking these two statements together gives us the answer to our question: $\neg X$ must be the set of possible worlds in W that are not members of X .

(Slide 12)

$\neg X$ is the set of possible worlds in W that are not members of X .

Set-theoretically: $\neg X$ is the complement of X (with respect to W)

Equivalently: $\neg X = \{w \text{ in } W: w \text{ is not a member of } X\}$

Equivalently: $\neg X = W \setminus X$ (read: ‘ W without X ’)

(See Figure 3.2.)

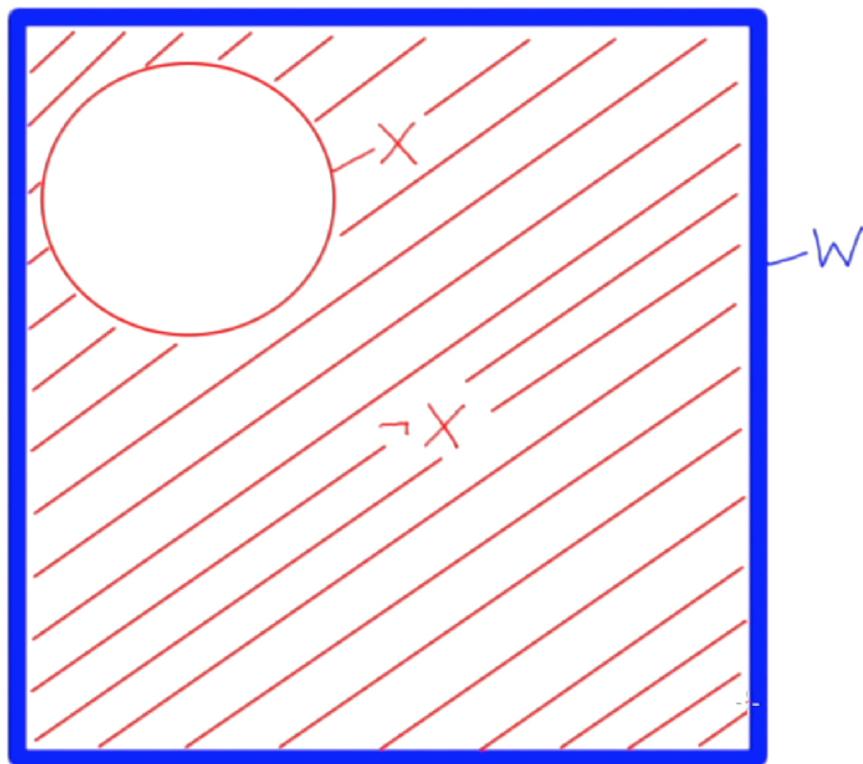


Figure 3.2: $\neg X$

Accordingly, the proposition that is expressed by ‘it is not the case that Socrates is a philosopher’ is the set-theoretic complement of the proposition expressed by ‘Socrates is a philosopher’ with respect to our given set W of possible worlds.

In a similar way, we can determine other logical operations on propositions: for instance, let X be the proposition again expressed by the sentence ‘Socrates is a philosopher’; and let Y be the proposition expressed by the sentence ‘Plato is a teacher of Aristotle’; syntactically, as explained in the last lecture, we can put the two sentences together by inserting an ‘and’ between them – in this way we get the so-called conjunction of these two sentences, their ‘and’-connection:

Socrates is a philosopher and Plato is a teacher of Aristotle

Now what is the conjunctive proposition, say, $X \wedge Y$, determined somehow from the propositions X and Y , which is expressed by that conjunctive sentence? It must be a proposition that is true precisely at those worlds at which both X and Y are true, for that is what the logical symbol ‘and’ expresses.

(Slide 14)

$X \wedge Y$ is the set of possible worlds in W that are members of both X and Y .

Set-theoretically: $X \wedge Y$ is the intersection of X and Y

Equivalently: $X \wedge Y = \{w \text{ in } W: w \text{ is a member of } X \text{ and } w \text{ is a member of } Y\}$

Equivalently: $X \wedge Y = X \cap Y$ (read: ‘ X intersected with Y ’)

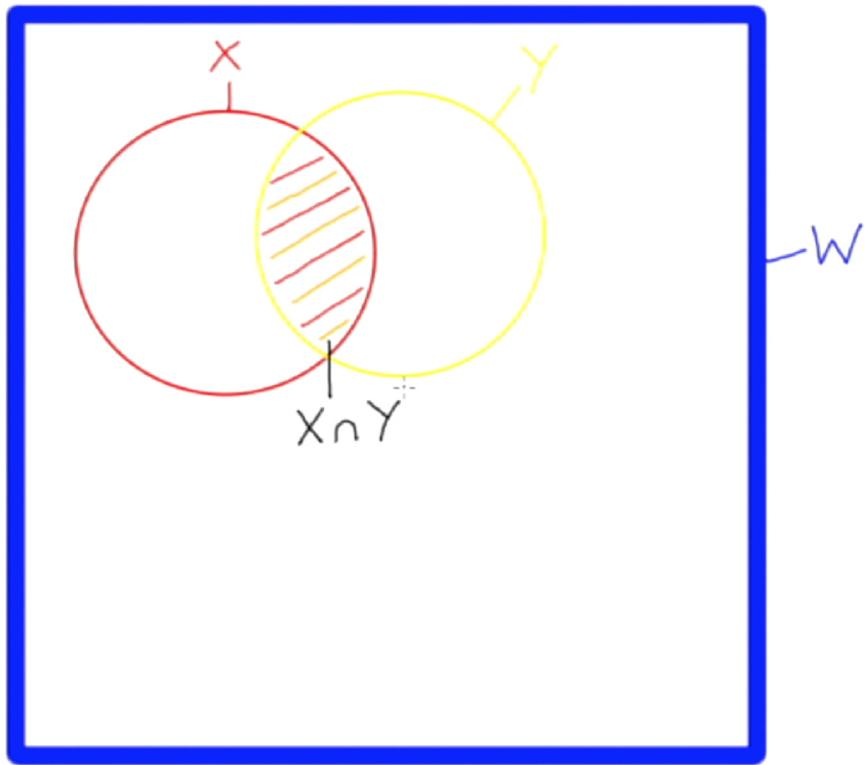
In set-theoretic terms: the conjunctive proposition $X \wedge Y$ is the set of all worlds that are members of X and Y at the same time: it is the set-theoretic intersection of X and Y .

$X \wedge Y$ is the intersection of X and Y :

Equivalently: $X \wedge Y = \{w \text{ in } W: w \text{ is a member of } X \text{ and } w \text{ is a member of } Y\}$

Equivalently: $X \wedge Y = X \cap Y$ (read: ‘ X intersected with Y ’)

(See Figure 3.3.)

Figure 3.3: $X \cap Y$

Accordingly, we can also formulate the so-called disjunction of two sentences, their ‘or’-connection:

Socrates is a philosopher or Plato is a teacher of Aristotle

Now what is the disjunctive proposition, say, $X \vee Y$, expressed by that disjunctive sentence? It must be a proposition that is true precisely at those worlds at which X is true or Y is true, for that is what the logical symbol ‘or’ expresses.

(Slide 16)

$X \vee Y$ is the set of possible worlds in W that are members of X or Y

Set-theoretically: $X \vee Y$ is the union of X and Y .

Equivalently: $X \vee Y = \{w \text{ in } W: w \text{ is a member of } X \text{ or } w \text{ is a member of } Y\}$

Equivalently: $X \vee Y = X \cup Y$ (read: ‘ X united with Y ’)

(See Figure 3.4.)

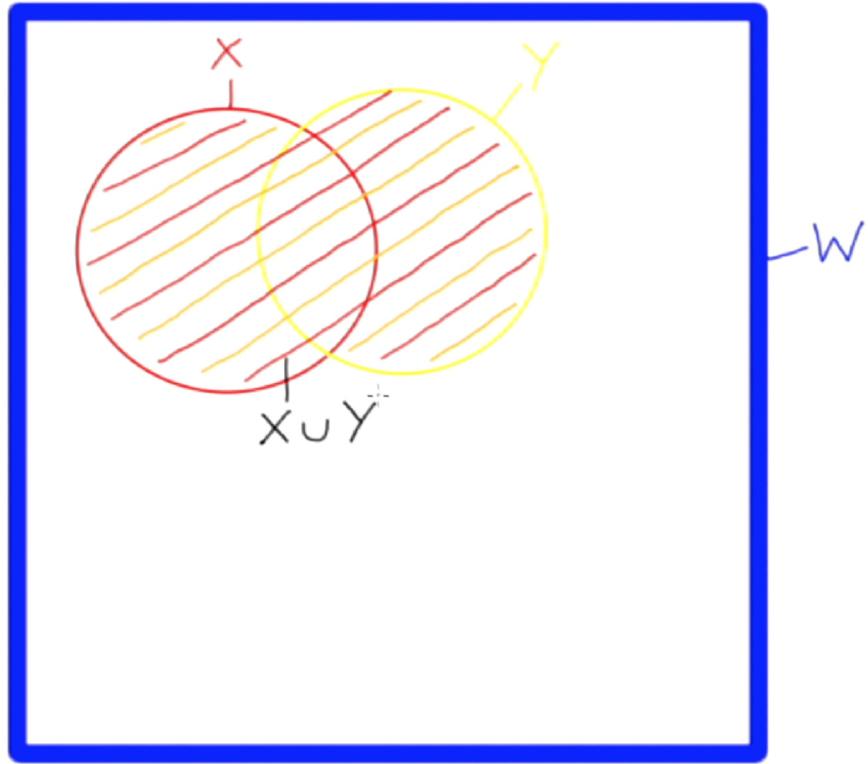


Figure 3.4: $X \vee Y$

As you can see, we interpret ‘ X or Y ’ as: X or Y or both of them. That is: we follow the usual practice of mathematicians, logicians, and philosophers in not reading ‘or’ as an

exclusive or: in order for $X \vee Y$ to be true at a world, it is perfectly fine if both X and Y are true there. ‘or’ simply means: at least one of X and Y is true, either just X , or just Y , or both of them.

Remark on logical vs set-theoretic symbols: ‘ \neg ’, ‘ \wedge ’, ‘ \vee ’ are the usual logical symbols for negation (‘not’), conjunction (‘and’), disjunction (‘or’), respectively. ‘ \setminus ’, ‘ \cap ’, ‘ \cup ’ are the usual set-theoretic symbols denoting relative complement (so ‘ $W \setminus X$ ’ denotes the complement of X relative to W , or ‘ W without X ’), intersection, union, respectively. By our set-theoretic account of propositions, it does not really matter in the following whether we apply the logical symbols or the corresponding set-theoretic symbols to names of propositions. (For instance, we will be able to denote one and the same proposition by ‘ $X \vee Y$ ’ and by ‘ $X \cup Y$ ’.

Remark on the diagrams: Such diagrams in which sets show up as funny blobs are sometimes called ‘Euler-Venn diagrams’, since both the Swiss mathematician Leonhard Euler and the British logician John Venn used variants of such diagrams in their work. More about the history of such logical/set-theoretic diagrams and the subtle differences between their different variants can be found at

<http://plato.stanford.edu/entries/diagrams/>.

Quiz 19:

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 ; finally, determine (and color) graphically (i) the proposition $W \setminus (X \cup Y)$, and also (ii) the proposition $(W \setminus X) \cap (W \setminus Y)$. What does the graphical representation in (i) look like in comparison with that in (ii)?

[Solution](#)

3.4 Propositions as Sets of Possible Worlds II (10:14)

This gives us a very neat picture of how the logical symbols with which we dealt in the last lecture relate to standard set-theoretic operations:

(Slide 17)

- *negation* (of descriptive sentences) corresponds to *complement* (of propositions/sets)
- *conjunction* (of descriptive sentences) corresponds to *intersection* (of propositions/sets)
- *disjunction* (of descriptive sentences) corresponds to *union* (of propositions/sets)

Here is another way of stating the same point – let us abbreviate ‘the proposition expressed by sentence A ’ by means of

$Prop(A)$.

Then we get, first of all, according to the view developed before:

(Slide 19)

- $Prop(\neg A) = \neg Prop(A) = W \setminus Prop(A)$

Here ‘ $\neg A$ ’ is the negation of the sentence A : $\neg A$ is a negation sentence. $Prop(\neg A)$ is the proposition that is expressed by that sentence. What is that proposition? It is the negation of the proposition expressed by A . What is the negation of the proposition expressed by A ? It is W without the proposition expressed by A , that is, the complement of that proposition with respect to our given set W of possible worlds. We use the same symbol ‘ \neg ’ here to express the negation of sentences and the negation of propositions: we should use different symbols really, since negating a sentence is not the same thing as negating a proposition, but since it will always become clear from the context what exactly we mean by ‘ \neg ’, it is handy just to use one symbol for both of them. E.g.: on the left, in ‘ $Prop(\neg A)$ ’, we negate a sentence, whereas on the right of the first identity symbol, in ‘ $\neg Prop(A)$ ’, a proposition is negated.

Similarly, we use ‘ \wedge ’ and ‘ \vee ’ both for conjunctions and disjunctions of sentences and of propositions; the context will make it clear again what is meant. That being in place, our account from above amounts to:

(Slide 20)

- $Prop(A \wedge B) = Prop(A) \wedge Prop(B) = Prop(A) \cap Prop(B)$
- $Prop(A \vee B) = Prop(A) \vee Prop(B) = Prop(A) \cup Prop(B)$

We could even turn these statements into a definition of ‘ $Prop$ ’ – of ‘the proposition expressed by’ – if we added a specification of the propositions that are expressed by descriptive sentences without any logical symbols: e.g., we would have to define also the proposition expressed by ‘Socrates is a philosopher’, the proposition expressed by ‘Plato is a teacher of Aristotle’, and so on. If we did, then we could actually state a recursive definition of ‘ $Prop$ ’ as applied to any of the sentences in our little toy language L_{simple} from the last lecture; and the definition would be recursive in exactly the sense that had also been explained in that lecture. But that is not our goal now. It should have become clear enough how $Prop$ maps sentences that result from applying logical symbols to other sentences to propositions that result from applying logical operations to other propositions; and these logical operations on propositions coincide with certain set-theoretic operations; that is all that I wanted to make clear at this point.

Let us take a look at a concrete example:

Assume W to consist of precisely eight possible worlds:

(Slide 21)

$$W = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8\}$$

So this is nothing like the infinitely many possible worlds that, maybe, God would be able to make actual; but never mind.

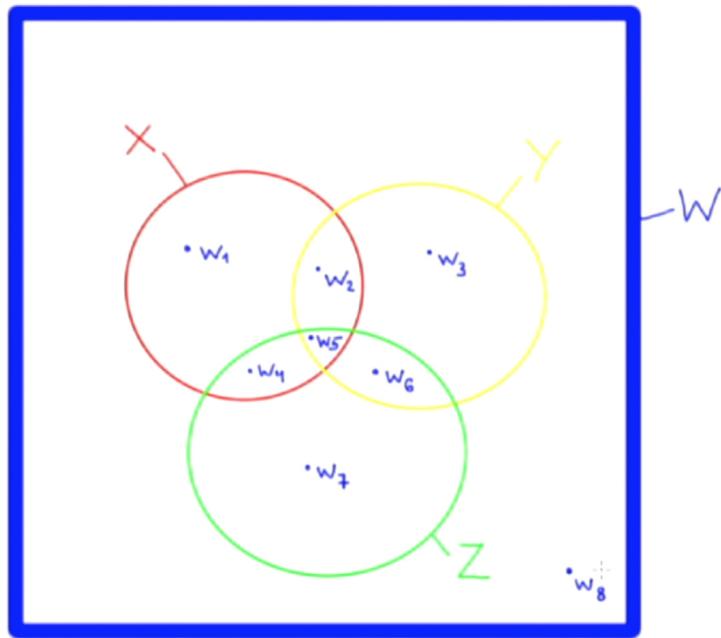


Figure 3.5: Distribution of w_1, \dots, w_8

In the diagram (see Figure 3.5), each of these eight possible worlds corresponds to one of the eight minimal regions. The three circles denoted by ‘ X ’, ‘ Y ’, ‘ Z ’ highlight three propositions, that is, three sets of possible worlds, on which I want you to focus now:

(Slide 22)

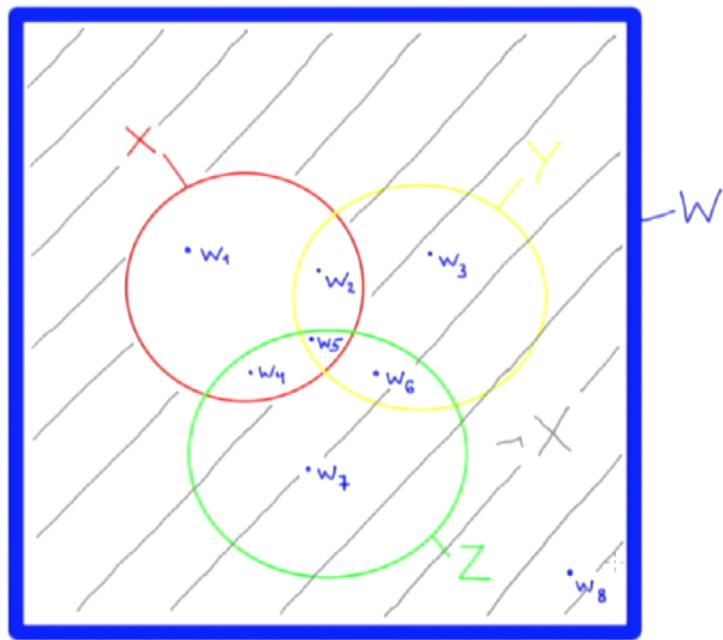
$$\begin{aligned} X &= \{w_1, w_2, w_4, w_5\} \\ Y &= \{w_2, w_3, w_5, w_6\} \\ Z &= \{w_4, w_5, w_6, w_7\} \end{aligned}$$

w_8 is not a member of any of the three sets.

E.g., X might be the proposition expressed by the sentence ‘Socrates is a philosopher’ again, Y the proposition expressed by ‘Plato is a teacher of Aristotle’, and Z might be some other proposition expressed by some other sentence, although at this point we do not even care that much anymore what sentence would express Z , if any sentence at all. Propositions are sets of possible worlds, and Z is such a set; that’s good enough. The actual world, in which things are as they are actually, might e.g. be w_2 , whereas the hypothetical world in which Socrates did not become a philosopher might be w_3 , assuming that Plato still managed to become Aristotle’s teacher in that world, even though Socrates is not a philosopher in this world and hence, presumably, did not teach Plato. We see that X is true in w_2 , but X is not true in w_3 .

Now we can apply our set-theoretic account of the logical operations on propositions:

e.g., the negation of X is $\{w_3, w_6, w_7, w_8\}$; that proposition is true in w_3 , while it is not true in the actual world, that is, in w_2 (see Figure 3.6).

Figure 3.6: $\neg X$

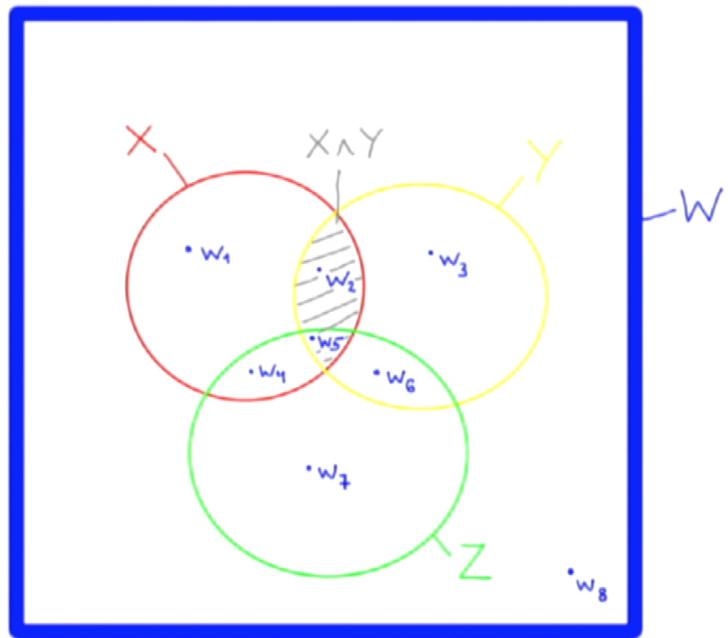
What is the negation of the negation of X ? The complement of the complement of X ? It is simply X again. While the sentence ‘Socrates is not not a philosopher’ is different from the sentence ‘Socrates is a philosopher’ – the one is, e.g., longer than the other – the two sentences do express the very same proposition X .

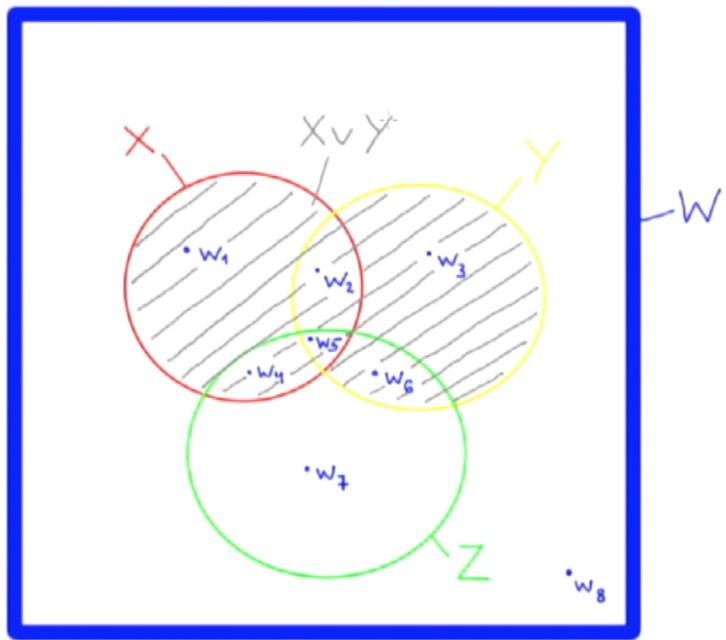
(Slide 24)

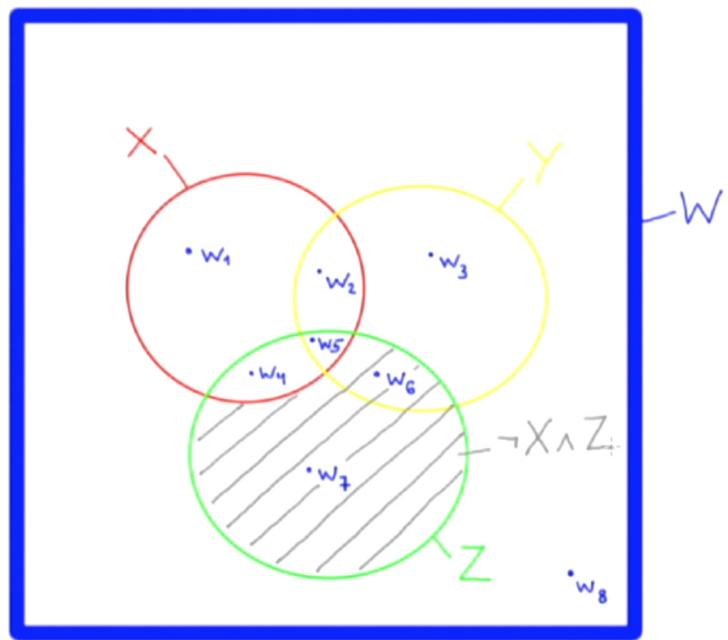
$$W = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8\}$$

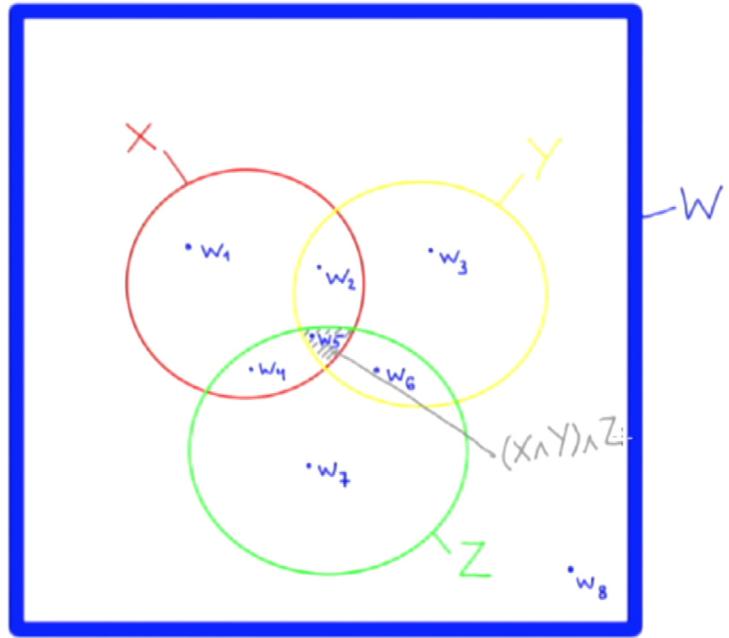
$$\begin{aligned} X &= \{w_1, w_2, w_4, w_5\} \\ Y &= \{w_2, w_3, w_5, w_6\} \\ Z &= \{w_4, w_5, w_6, w_7\} \\ \neg X &= \{w_3, w_6, w_7, w_8\} \\ X \wedge Y &= \{w_2, w_5\} \\ X \vee Y &= \{w_1, w_2, w_3, w_4, w_5, w_6\} \\ \neg X \wedge Z &= \{w_6, w_7\} \\ (X \wedge Y) \wedge Z &= \{w_5\} \\ \neg X \wedge \neg Y \wedge \neg Z &= \{w_8\} \end{aligned}$$

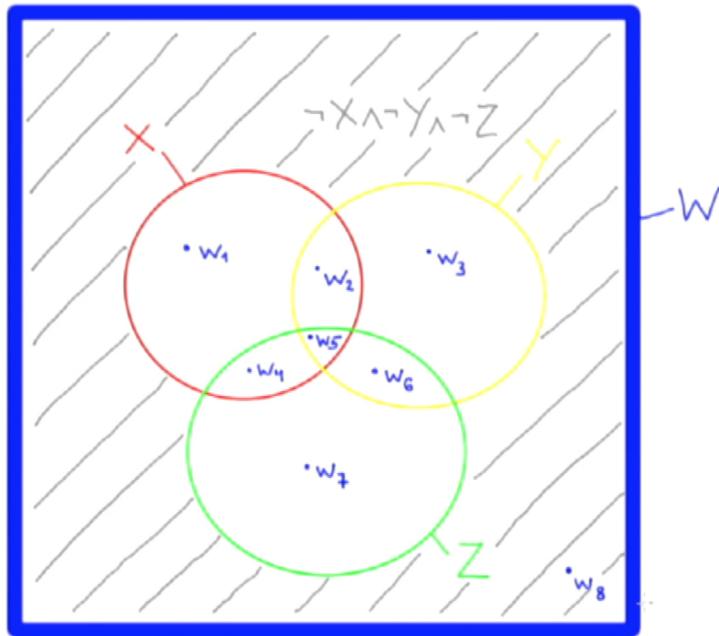
The conjunction of the two propositions X and Y is $\{w_2, w_5\}$ (see Figure 3.7), the disjunction of X and Y is $\{w_1, w_2, w_3, w_4, w_5, w_6\}$ (see Figure 3.8), the conjunction of $\neg X$ with Z is $\{w_6, w_7\}$ (see Figure 3.9), the conjunction of the conjunction of X and Y with Z is $\{w_5\}$ (see Figure 3.10): so that is the conjunction of all of X and Y and Z . $\neg X \wedge \neg Y \wedge \neg Z$ is the proposition $\{w_8\}$ (see Figure 3.11); and so on and so forth.

Figure 3.7: $X \wedge Y$

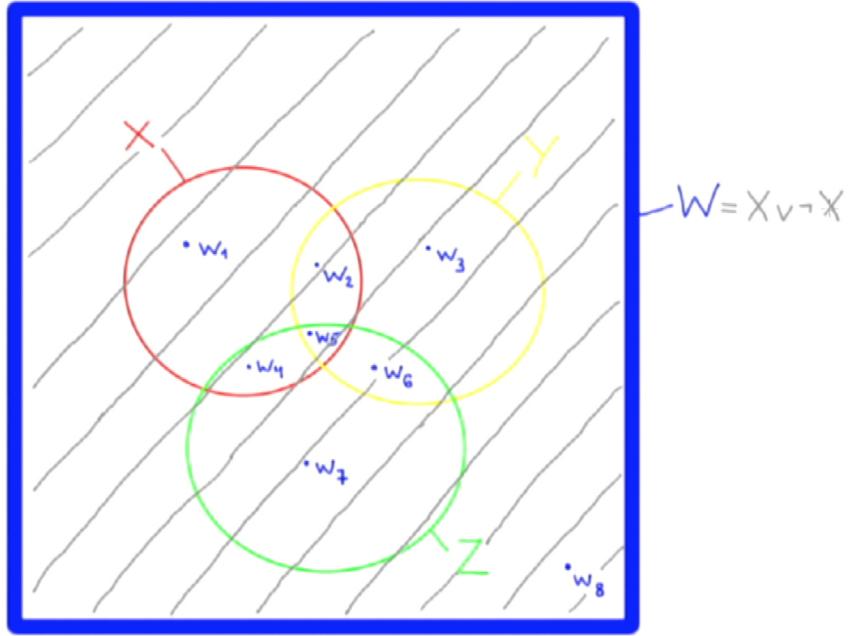
Figure 3.8: $X \vee Y$

Figure 3.9: $\neg X \wedge Z$

Figure 3.10: $(X \wedge Y) \wedge Z$

Figure 3.11: $\neg X \wedge \neg Y \wedge \neg Z$

How about the descriptive sentence ‘Socrates is a philosopher or Socrates is not a philosopher’: this is a logical truth, a logical law – that sentence will be true whatever the world is like; it does not matter really whether Socrates was a philosopher or whether this was not so, the sentence ‘Socrates is a philosopher or Socrates is not a philosopher’ must be true. This can be seen also by checking out the proposition that is expressed by that sentence: it is the proposition $X \cup \neg X$, the set of worlds in W which are members of X taken together with the set of worlds in W that are not members of X – of course that set is nothing else than our given set W of all possible worlds (see Figure 3.12).

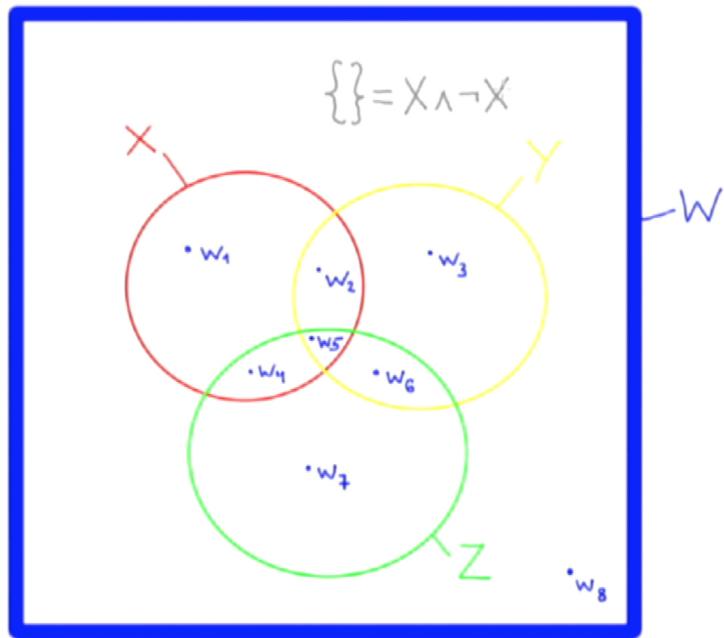
Figure 3.12: $W = X \vee \neg X$

(Slide 25)

$$X \vee \neg X = X \cup \neg X = W$$

And that's exactly as expected: 'Socrates is a philosopher or Socrates is not a philosopher' cannot be false; there is no possible world at which it is false; accordingly, the proposition expressed by it cannot be false either; there is no world at which it is false; the proposition is simply the total set W of all possible worlds.

For analogous reasons, the sentence 'Socrates is a philosopher and Socrates is not a philosopher' is a contradictory sentence which cannot be true at any world; and if you turn to the proposition that is expressed by it, you find that this proposition is nothing but $X \cap \neg X$, which is the empty set: the one and only subset of W that does not include any worlds as members (see Figure 3.13).

Figure 3.13: $X \cap \neg X = \{\}$

(Slide 26)

$$X \cap \neg X = X \cap \neg X = \{\}$$

We had encountered the empty set already in our first lecture, and now it reappears in the form of a contradictory proposition: a proposition that is not true at any possible world whatsoever; just as what one would expect of the proposition that is expressed by the contradictory sentence ‘Socrates is a philosopher and Socrates is not a philosopher’.

Quiz 20:

Just as in the lecture, let $W = \{w_1, \dots, w_8\}$, $X = \{w_1, w_2, w_4, w_5\}$, $Y = \{w_2, w_3, w_5, w_6\}$, $Z = \{w_4, w_5, w_6, w_7\}$.

- (1): Determine $(X \cap Y) \cap \neg Z$.
- (2): Determine $\neg(X \cap Z) \cup \neg\neg Y$

[Solution](#)

3.5 Logical Implication and the Subset Relation (06:06)

There is still more that we can read off the diagram given this account of propositions as sets of possible worlds: for instance, you will remember that we were dealing with logically valid arguments already in our first lecture:

(Slide 27)

(P1) ...

(P2) ...

:

(Pn) ...

(C) ...

The idea was that an argument like that is logically valid if and only if the following is the case: whenever all of the premises are true, then also the conclusion is true; if all of the premises are true, then necessarily also the conclusion must be true. It was not important whether the premises were actually true: if they were true, then this conclusion would have to be true, too. Although we did not define this notion of logical validity precisely – that would be the job of a proper introduction into logic – I hope that the rough idea of this notion became clear enough.

Now, instead of saying that the argument ‘P1,P2,...,Pn. Therefore: C’ is logically valid, one can also say that the premises P1,P2,..,Pn taken together logically imply the conclusion C; that is just an alternative way of expressing the same thought. P1,P2,...,Pn taken together stand in the relation of logical implication or logical consequence to C.

Consider now the case where there is just one premise: where ‘P. Therefore: C’ is logically valid, or where the premise P logically implies the conclusion C.

Here is an example:

(Slide 30)

$$A \text{ logically implies } A \vee B$$

Whatever the descriptive sentence A might be, it logically implies the descriptive sentence $A \vee B$: for given that A is true, $A \vee B$ is true as well – remember our truth definition from the last lecture: $A \vee B$ is true if and only if A is true or B is true. If at least one of A and

B is true, then $A \vee B$ is true as well, and by assumption, by the premise being true, A is true: so $A \vee B$ is true. A logically implies $A \vee B$.

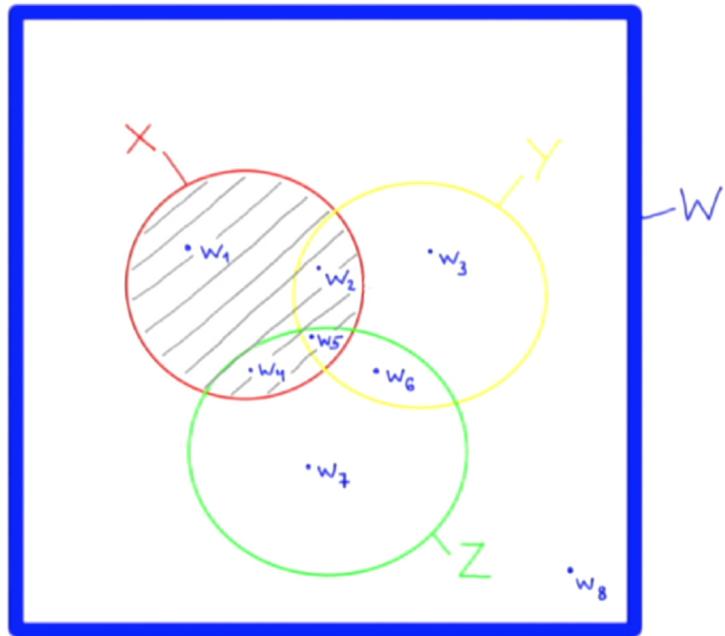
There is a very nice way of seeing this in our diagram, once we take the step from descriptive sentences to the propositions that they express: e.g., let us assume that the sentence A expresses the proposition X , whereas the sentence B expresses the proposition Y . The sentence A might be ‘Socrates is a philosopher’ again, B ‘Plato is a teacher of Aristotle’, and therefore $A \vee B$ would be the sentence ‘Socrates is a philosopher or Plato is a teacher of Aristotle’.

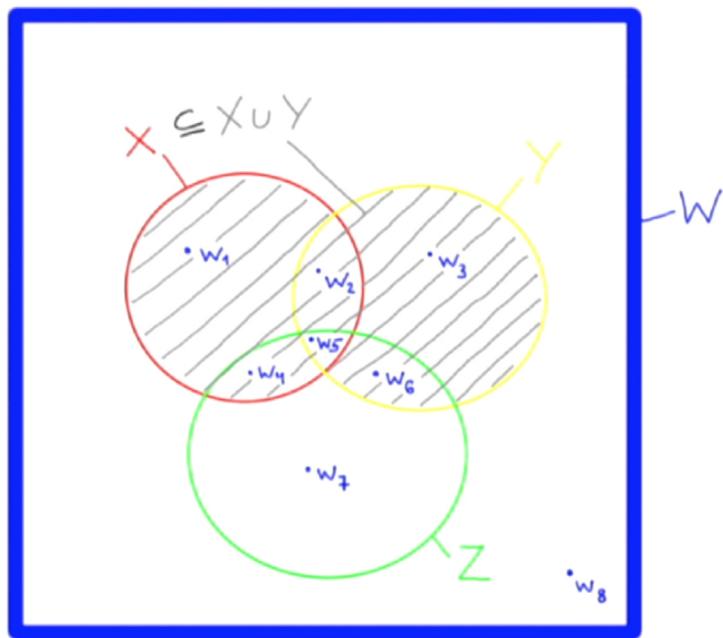
(Slide 31)

A logically implies $A \vee B$

X is a subset of $X \cup Y$

If you now consider the proposition that is expressed by A , the set X of possible worlds (see Figure 3.14), and then also the proposition that is expressed by $A \vee B$, that is, the proposition $X \cup Y$, you find that X is a subset of $X \cup Y$ (see Figure 3.15): every world that is a member of X is also a member of $X \cup Y$. In other words: every world at which X is true is also a world at which $X \cup Y$ is true. Whenever X is true, also $X \cup Y$ is true. But that is nothing but a restatement of the notion of logical implication or consequence, but this time on the level of propositions instead of sentences: back then we said, whenever A is true, $A \vee B$ is true; now we say, in every world in which X is true, also $X \cup Y$ is true.

Figure 3.14: X

Figure 3.15: $X \cup Y$

Thus we also have:

(Slide 32)

- *logical implication* (of descriptive sentences) corresponds to the *subset relation* (between propositions/sets)

That is:

(Slide 33)

A logically implies B if and only if $Prop(A)$ is a subset of $Prop(B)$.

Remark on logical implication and propositions: For the record: I am cheating here a little bit in order to keep things simple. The right-to-left direction of this principle actually holds only if W can be regarded as the set of all logically possible worlds as far as the sentences A and B are concerned. It will not be important for anything that follows in the lecture, so if you like you can simply skip this. If so, please just click through the following remark...

Let us have a look at an example: Consider the two sentences C ('The sun was shining at least once in Munich on August 12th 2013') and D ('The MCMP had about 50 members on August 12th 2013'). Let W be a set of three possible worlds, $W = \{w_1, w_2, w_3\}$, such that C is true in w_1, w_2 but not in w_3 , whereas D is true in w_1 , but not in w_2 or w_3 . So w_1 is a world in which the sun was shining at least once in Munich on August 12th 2013 and the MCMP had about 50 members on August 12th 2013. w_2 is a world in which the sun was shining at least once in Munich on August 12th 2013 and the MCMP did not have about 50 members on August 12th 2013. And w_3 is a world in which the sun was not shining at least once in Munich on August 12th 2013 and the MCMP did not have about 50 members on August 12th 2013. The proposition that is expressed by C is the set $\{w_1, w_2\}$, while the proposition expressed by D is the set $\{w_1\}$; accordingly, the proposition expressed by $C \wedge D$ is the set $\{w_1\}$ again. We find that e.g. $Prop(D)$ is a subset of $Prop(C \wedge D)$ (since $\{w_1\}$ is a subset of $\{w_1, w_2\}$). But we would not actually want to say that 'The MCMP had about 50 members on August 12th 2013' logically implies 'The sun was shining at least once in Munich on August 12th 2013 and the MCMP had about 50 members on August 12th 2013'. The problem is that as far as C and D are concerned, we are lacking one logically possible world in W : a world w_4 in which C is false but D is true. Only with this additional world being in place, all logically possible combinations of truth values for C and D can be generated. If we add such a world w_4 to our set of logically possible worlds, then $Prop(D)$ ends up being the set $\{w_1, w_4\}$, while $Prop(C \wedge D)$ ends up being the set $\{w_1\}$ again. Thus, $Prop(D)$ is no longer a subset of $Prop(C \wedge D)$ (as intended). What still does hold is that $Prop(C \wedge D)$ is a subset of $Prop(D)$, and that is exactly as it should be: for $C \wedge D$ does logically imply D (whenever the former is true, the latter must be true as well).

And if we want, we can read the right-hand side also as: the proposition expressed by A logically implies the proposition expressed by B . That is:

(Slide 34)

For all propositions X, Y :

X logically implies Y if and only if X is a subset of Y .

Here we presuppose our conception of propositions as sets of possible worlds again, of course.

For example, we also find:

The sentence $A \wedge B$ logically implies the sentence A . Which corresponds to: $X \cap Y$ is a subset of X (see Figure 3.16).

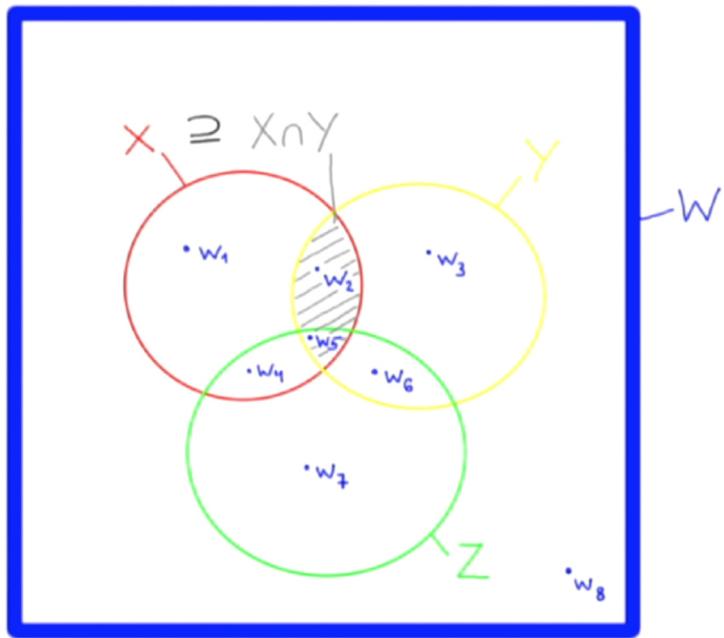


Figure 3.16: $X \cap Y \subseteq X$

Or in other words: the proposition $X \cap Y$ logically implies the proposition X .

And so on.

Remark on $X \cap Y$ logically implying X : Of course, $X \cap Y$ also logically implies Y .

By defining propositions to be sets of possible worlds, we have not just been able to derive the two principles ($\text{Id}_{\text{Worlds}}$) and (Id_{Prop}) that we had considered plausible beforehand, we can also make good sense of the negation, conjunction, and disjunction of propositions in terms of set theoretic operations, where these operations are nothing but the propositional counterparts of the negation, conjunction, and disjunction of descriptive sentences; and we are able to determine the relation of logical consequence between propositions to be the subset relation, which in turn is the propositional counterpart of the relation of logical consequence between descriptive sentences.

This being in place, we are now ready to study one of the central attitudes that one can take toward propositions, that is, towards sets of possible worlds: belief.

Quiz 21:

Show that if X is a subset of Y and X is a subset of Z , then X is also a subset of $Y \cap Z$.

Solution

Remark on possible worlds: In case you worry about them – please do not! Quite independently of how metaphysicians might think about possible worlds, in our lectures reference to possible worlds will just be a handy way of referring to possibilities which, given the evidence that a person has, the person cannot rule out. For instance, say, in probability theory one wants to describe the possible outcomes of throwing a six-sided die; since the die is yet to be thrown, one does not know about the actual outcome as yet. In such a case, a probability theorist would distinguish between six possible outcomes, say, w_1, w_2, \dots, w_6 : at w_1 the die rolls 1, ..., at w_6 the die rolls 6. One of the six possibilities corresponds to what actually happens, but it is only after the throw that we will know which of them is actual.

That's just like what we do presently in our lecture. The only difference is that we say ‘possible world’ instead of ‘possible outcome’ (the set W of possible worlds would be $\{w_1, \dots, w_6\}$), and we say ‘actual world’ instead of ‘actual outcome’. That's all really.

3.6 Belief in Propositions (09:51)

I believe that Socrates is a philosopher, I believe that the weather will be nice tomorrow, I believe that $2 + 2 = 4$, and there is much more that I believe. Obviously, we are not concerned with religious belief here, with faith: I do like the nice weather, but I do not feel any kind of religious attachment to it. Rather believing a proposition is the same as holding that proposition to be true. A belief does not have to be active in one’s conscious mind either in order to count as a belief: for example, I bet you believe that water is wet, and you believed so even before I brought this up a moment ago, and before you started thinking about it: the belief state had been there already, it merely had not occurred to you consciously before I alerted you to it.

What kind of attitude is holding something to be true? What kind of mental function do beliefs serve?

First of all, if I believe that A is the case, I normally act as if A were the case; for instance, since I believe that the weather will be nice tomorrow, I might make plans for a hike tomorrow with a couple of friends. So one role that beliefs play in our mental lives is that decisions for action are based upon them. Or really our beliefs lead to such decisions if

certain other circumstances are also the case; e.g., if I am also subject to certain desires, such as the desire to go hiking tomorrow. If I desire to be hiking tomorrow, and I believe that nothing is going to interfere with executing that desire, then it seems like a rational decision to actually go hiking. My belief makes me disposed to go hiking given that the desire and maybe some other conditions are present.

One especially important case of acting based on beliefs is the act of asserting a sentence: for instance, you ask me ‘Who is this Socrates that you have been talking about?’: In normal circumstances at least, I would assert ‘Socrates is a philosopher’ in order to answer your question, and I would thereby express my belief that Socrates is a philosopher. This does not work in each and every case: maybe I want to hide the information from you, and that is why I lie to you by asserting ‘Socrates is a showmaster’; or the like; but normally assertions of descriptive sentences are being made in line with one’s beliefs.

Other than the practical role that beliefs have in decision-making, they also have a more theoretical role to play: our beliefs constitute our picture of the world, our theory about the world, as it were; and much as it is the case with scientists, even in our everyday life we want that picture or theory to be accurate: belief aims at the truth, just as science aims at the truth. In abstract terms: my aim is that if I believe the proposition X , then the actual world is a member of that set X ; my aim is for X to be true. In more concrete terms: If I learn a new piece of evidence, then if I am rational, I turn that piece of evidence into a belief of mine, simply because I take the piece of evidence to be true. Moreover, if I am rational, I might reason on the basis of beliefs to new beliefs by means of rules of inference that I take to be truth-preserving, that is, logically valid, because these rules are truth-preserving, and I am interested in the conclusion: if the premise beliefs are true, then the conclusion belief must be true as well. For example: I believe that the weather will be nice tomorrow, because I have seen what the wind and the clouds are like above the mountain close by, and taking this together with my previous experience with the local weather conditions makes me believe that the weather will be nice tomorrow. Furthermore, you tell me that according to the general weather forecast that you saw on tv, either the weather will not be nice already tomorrow, or, it will still be nice tomorrow but not nice the day after tomorrow; I trust you and start to believe what you told me. And then, by putting the two beliefs together, I infer that the weather will not be nice the day after tomorrow.

For I do believe X ,

(Slide 37)

X : that the weather will be nice tomorrow

- Believed: X

and I also start to believe $\neg X \vee (X \wedge \neg Y)$, because of what you told me, where

(Slide 38)

X : that the weather will be nice tomorrow

Y : that the weather will be nice on the day after tomorrow

- Believed: X
- Believed: $\neg X \vee (X \wedge \neg Y)$

Since I believe X , I can rule out the left half of the disjunctive proposition $\neg X \vee (X \wedge \neg Y)$, which is why the right half remains, which I may thus conclude:

(Slide 39)

X : that the weather will be nice tomorrow

Y : that the weather will be nice on the day after tomorrow

- Believed: X
- Believed: $\neg X \vee (X \wedge \neg Y)$
- Believed: $X \wedge \neg Y$

And from the belief in that conjunctive proposition I can infer its right-hand part

$\neg Y$: the weather will not be nice the day after tomorrow.

(Slide 40)

X : that the weather will be nice tomorrow

Y : that the weather will be nice on the day after tomorrow

- Believed: X
- Believed: $\neg X \vee (X \wedge \neg Y)$
- Believed: $X \wedge \neg Y$
- Believed: $\neg Y$

At least that is what is going on when I am extracting information from my two beliefs in a rational manner.

So beliefs play a major role for decision-making on the one hand and for theorizing about the world, for learning and reasoning, on the other.

Let us focus now just on the second aspect of belief: the less practical, and more theoretical role of belief. Occasionally, it might happen that we create, abandon, or change our beliefs irrationally: we simply mess up. For example, I might ignore evidence although I take it

to be true; or I start to believe something without any good reason whatsoever, I start believing it without aiming at the truth. In such cases I fail: I don't do what I rationally ought to do, or I do what I am not rationally permitted to do. All of these kinds of failures are scientifically interesting; for example, when cognitive psychologists study, in terms of experiments, how we actually reason, they are particularly interested in such kinds of failures. In contrast, philosophers are not so much interested in the failures nor, more generally, in how we actually reason – which is an empirical question – but much more in how we rationally ought to reason or how we are rationally permitted to reason – which are normative questions. Whether or not humans live up to these normative standards: what would a perfectly rational person be like as far as her inferences from premise beliefs to conclusion beliefs are concerned? What can we say about inferentially perfectly rational believers?

In the following I am going to introduce a couple of postulates which all seem highly plausible and which tell us something about such inferentially perfectly rational believers: essentially, the postulates express that such a person's beliefs are closed under the rules of logic; if a proposition follows logically from some believed propositions, then that logically implied proposition is also believed by the rational believer. In order to state these postulates on rational belief precisely, we presuppose again a set W of possible worlds from which we can determine the set of all propositions, the set of all subsets of W , as explained before. The intended interpretation of W in the present context is that of the set of possible worlds that are, in principle, entertainable at all by our given perfectly rational person. It's the set of ways the world might be that this perfectly rational believer can make sense of. By her cognitive architecture, she simply could not have any thoughts that would not correspond to any proposition over W , that is, any subset of W .

We will also make one simplifying assumption: let us presuppose again that W is a finite set of possible worlds, not an infinite one. Thus, we are dealing with persons again who, unlike God perhaps, can only distinguish between finitely many ways the world might be. You might think: how could such a being be perfectly rational in the first place? But we are only concerned here with persons who are inferentially perfectly rational: it is not that these persons are supposed to be perfectly rational in each and every respect, they are just perfectly rational with respect to the inferences that they are able to draw; and within their cognitive boundaries, they are perfectly rational in terms of their inferences. Such persons are certainly not uninteresting or silly in any sense: the set of possible worlds for them might be enormously large, just not infinitely large. If we humans are finite beings in the relevant sense, too, then we should not feel particularly alienated by these persons' finite powers of discrimination: and such a person can still be perfectly rational within the limits of these finite powers.

Remark on belief: If you want to know more about the mental state of belief as studied in the area of philosophy called the philosophy of mind, please take a look here:

<http://plato.stanford.edu/entries/belief/>

And here you can find more about the topic of justified (or rational) belief as studied in epistemology:

<http://plato.stanford.edu/entries/epistemology/>

3.7 Postulates on Rational Belief (05:34)

These are thus our postulates on rational belief:

(Slide 41)

- Rational Belief 1: If a person is inferentially perfectly rational (with W as her set of entertainable possible worlds), then she believes W .

Remember that the set W of all possible worlds is the proposition that is expressed by any logical law of the form $A \vee \neg A$. Of course, a perfectly rational person believes that proposition to be the case.

(Slide 42)

- Rational Belief 2: If a person is inferentially perfectly rational, then she does not believe $\{\}$.

As determined before, the empty set $\{\}$ of possible worlds is the proposition that is expressed by a logical contradiction of the form $A \wedge \neg A$. Clearly, a perfectly rational person does not believe that proposition to be the case.

(Slide 43)

- Rational Belief 3: If a person is inferentially perfectly rational, if she believes X , and if X is a subset of Y , then she also believes Y .

Assume our perfectly rational person believes the proposition X . If X is a subset of another proposition Y , then this means that X logically implies Y , as we had also discussed before. But if the person believes X to be true, and the inference from X to Y is truth-preserving, by the standards of logic, then a perfectly rational person will believe Y , too. Or in slightly different terms: if X is believed, then this really amounts to the belief that the actual world is a member of X . But if X is a subset of Y , then the actual world must also be a member of Y : hence a perfectly rational person will also believe Y .

(Slide 44)

- Rational Belief 4: If a person is inferentially perfectly rational, if she believes X , and if she believes Y , then she also believes $X \cap Y$.

This is similar to the case before, the main difference is that we are considering a case of two premise beliefs now, and not just one. Assume our perfectly rational person believes the proposition X , and she also believes the proposition Y . If X is true and Y is true, then necessarily also $X \cap Y$ is true. That is: the two propositions X and Y taken together logically imply the conjunctive proposition $X \cap Y$. But then a perfectly rational believer should be such that if she already believes both X and Y to be true, she also believes $X \wedge Y$, the conjunctive proposition, the intersection of X with Y , to be true. Or again in other words: if the person believes that the actual world is in X , and she also believes that the actual world is in Y , then she must believe that the actual world is in $X \cap Y$: for that's then the only “place” in which the actual world could be.

None of this will sound particularly surprising. But already from such reasonably weak assumptions on rational belief, we can derive an interesting property of rational belief. It is captured by the following little theorem:

(Slide 45)

Theorem:

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

If the Rational Belief postulates 1–4 are the case (for the given set W of possible worlds), then for every inferentially perfectly rational person p (for whom W is the set of entertainable possible worlds), there is a non-empty proposition B_W , such that the following holds:

For all propositions X (over W),

person p believes X if and only if B_W is a subset of X .

This means the following: whatever W , whatever inferentially perfectly rational person we consider, if the Rational Belief postulates from above are satisfied, then there must be a distinguished set of possible worlds, call it B_W , a special subset of the given set W of all possible worlds, such that every proposition X that is believed by that person is logically implied by B_W , and vice versa, if B_W logically implies a proposition X , then X is believed. It is as if B_W completely summarizes that perfectly rational believer's belief system: the distinguished proposition B_W is itself believed by the person, since of course B_W is a subset of itself – every set is a subset of itself – and everything else that is believed by the person

has B_W as a proper subset. B_W generates such a person's belief system in a bottom-up fashion: all, and only, the propositions that are logically implied by B_W are believed by that person. B_W is characterized by being the least, the smallest believed proposition overall. For the same reason, B_W is also determined uniquely given the assumptions in the theorem.

Quiz 22:

Assume our postulates Rational Belief 1-4.

Show that the following is the case: If a person is inferentially perfectly rational, then if she believes a proposition X , she does not also believe the negation of X .

[Solution](#)

Our postulates Rational Belief 1-4 (and what follows from them) correspond to principles of so-called doxastic logic and epistemic logic: the logic of rational belief and the logic of knowledge. These logics are special cases of what is called modal logic in philosophy: the logic of modalities. The belief operator ' B_p ' ('it is believed by person p that') is a so-called doxastic modality, and the knowledge operator ' K_p ' ('it is known by person p that') is a so-called epistemic modality. For instance, in the previous quiz we have shown that if a person is inferentially perfectly rational, then if she believes a proposition X , she does not also believe the negation of X ; in the language of doxastic logic this would be expressed in terms of: $B_p X \rightarrow \neg B_p \neg X$. The set B_W that is mentioned in our little theorem is called the set of doxastically accessible worlds (from the viewpoint of a given person) in the possible worlds semantics of doxastic logic. See

<http://plato.stanford.edu/entries/logic-epistemic/>

for more on doxastic and epistemic logic, and see

<http://plato.stanford.edu/entries/logic-modal/>

for more on modal logic in general.

3.8 Proving a Theorem on Rational Belief (08:35)

Here is the proof:

Let a finite, non-empty set W of possible worlds be given, let p be an inferentially perfectly rational person, and assume our Rational Belief postulates 1-4 from before to hold.

The proposition B_W that our theorem claims to exist can be constructed as follows:

(Slide 46)

Proof:

Let B_W be the conjunction of all propositions believed by p .

We will show that this so-constructed proposition has precisely the properties that are asserted by the theorem.

First of all, let us make sure that what we want to construct here exists: I said that B_W is meant to be the intersection of all believed propositions. But are there any such believed propositions at all of which we can then take their intersection?

Yes! By postulate Rational Belief 1, p believes at least one proposition: the set W of all possible worlds. So taking B_W to be the intersection of all believed propositions is well-defined as the intersection of something, since there is at least one such believed proposition.

Secondly, you might wonder: in what order should I intersect p 's believed propositions in order to determine the intended set B_W ? Fortunately, it does not matter really:

One can show easily, and in general, that intersection satisfies the law of commutativity:

(Slide 48/1)

Law of commutativity:

- $X \cap Y = Y \cap X$

and it also satisfies the law of associativity:

(Slide 48/2)

Law of associativity:

- $X \cap (Y \cap Z) = (X \cap Y) \cap Z$

The situation is like when you are taking the product of numbers: e.g.:

(Slide 49)

Compare:

$$2 \cdot 3 = 3 \cdot 2$$

$$2 \cdot (3 \cdot 5) = (2 \cdot 3) \cdot 5$$

From these two general laws for the intersection of sets it follows that the ordering in which one takes multiple intersections does not matter: whatever the order, the outcome will always be the same.

That is also the reason why mathematicians would write immediately:

(Slide 50)

$$X_1 \cap X_2 \cap \dots \cap X_n$$

without worrying about parentheses at all, and why they would not mind at all writing instead

(Slide 51/1)

$$X_2 \cap X_1 \cap \dots \cap X_n$$

or maybe:

(Slide 51/2)

$$X_n \cap X_{n-1} \cap \dots \cap X_1$$

or the like.

So we see that it does not matter in which manner the propositions believed by our inferentially perfectly rational person p are intersected: the result will always be one and the same set, and that set we have called ' B_W '.

Alright. Now let us prove that B_W has the intended features.

Most importantly,

(Slide 52)

We need to show:

For all propositions X (over W),

person p believes X if and only if B_W is a subset of X .

Assume the left-hand side: person p believes X .

We need to show: B_W is a subset of that X .

Since X is one of the propositions believed by p , B_W resulted from taking the intersection of propositions that included X . In other words:

(Slide 53)

If X_1, \dots, X_n are all the propositions believed by p , then

$$B_W = X_1 \cap X_2 \cap \dots \cap X_n$$

where one of X_1, \dots, X_n is X .

(We know that there are only finitely many propositions that can be believed by p , since a finite set such as W has only finitely many subsets.)

E.g., if X_1 is X , then

(Slide 54/1)

$$B_W = X \cap X_2 \cap \dots \cap X_n$$

and analogously if X were one of the other sets. In any case, the set denoted on the right must be a subset of X : if $n = 1$, then B_W would be identical to X and hence a subset of X , since every set is a subset of itself. And for $n > 1$, we are dealing with the intersection of X with something else here, which also determines a subset of X . Therefore, because B_W was defined by an expression like the one on the right, B_W must be a subset of X . But that is what we needed to show.

(Slide 54/2)

In any case: B_W is a subset of X . ✓

(Slide 55)

Now for the other direction of:

For all propositions X (over W),

person p believes X if and only if B_W is a subset of X .

Assume the right-hand side: B_W is a subset of X .

We need to show: person p believes X .

We do so in two steps. First we prove that B_W is itself a proposition believed by p . Then we conclude from this that also X is believed by p .

Remember that B_W was defined like this:

(Slide 56/1)

We have:

$$B_W = X_1 \cap X_2 \cap X_3 \cap \dots \cap X_n$$

where X_1, \dots, X_n are all the propositions believed by p .

By what I said before about intersections in general, we can reformulate this a bit, as follows:

(Slide 56/2)

$$B_W = (\dots ((X_1 \cap X_2) \cap X_3) \cap \dots \cap X_n)$$

Nothing has happened really, except that I have reintroduced parentheses in some way; we know that it does not actually matter in what way, as the outcome will always be the same.

But now we can apply our postulate Rational Belief 4, and we actually do so several times:

(Slide 57)

$$B_W = (\dots ((X_1 \cap X_2) \cap X_3) \cap \dots \cap X_n)$$

By postulate Rational Belief 4:

- Since X_1 is believed by p , X_2 is believed by p , also $(X_1 \cap X_2)$ is believed by p .
 - Since $X_1 \cap X_2$ is believed by p , and X_3 is believed by p , also $((X_1 \cap X_2) \cap X_3)$ is believed by p .
 - Since $((X_1 \cap X_2) \cap X_3)$ is believed by p , and X_4 is believed by p , also $((((X_1 \cap X_2) \cap X_3) \cap X_4)$ is believed by p .
- ⋮
- $\underbrace{(\dots ((X_1 \cap X_2) \cap X_3) \cap \dots \cap X_n)}_{B_W}$ is believed by p .

In other words:

B_W is believed by p .

Now for the second step: We have just shown that p believes B_W . By assumption, B_W is a subset of X . But now we can apply our postulate Rational Belief 3:

(Slide 58)

By postulate Rational Belief 3:

Since B_W is believed by p , and B_W is a subset of X , also X is believed by p . ✓

Which is what we needed to show. We have proven the two sides of the equivalence claim to be actually equivalent: the one implies the other and vice versa.

So we are done, except for one final point: in the theorem we also claim B_W to be non-empty. But that follows from what we have already shown before together with postulate Rational Belief 2: we have already shown that p believes B_W . But Rational Belief 2 says that the empty proposition, the empty set of possible worlds, is not believed by p ; therefore, B_W cannot be the empty set.

This concludes the proof of our little theorem.

Quiz 23:

Our theorem says (I suppress some details, such as concerning p being inferentially perfectly rational and W being the set of all worlds): if Rational Belief postulates 1-4 hold, then there is a non-empty proposition B_W , such that for all propositions X , person p believes X if and only if B_W is a subset of X .

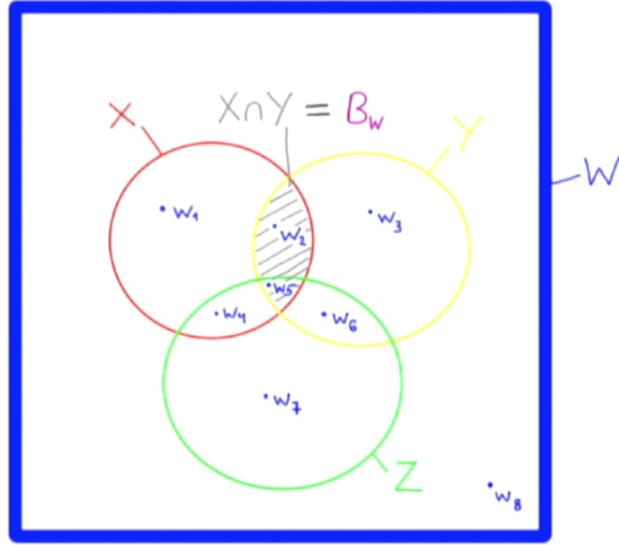
But actually one can also prove the converse of this: if there is a non-empty proposition B_W , such that for all propositions X , person p believes X if and only if B_W is a subset of X , then Rational Belief postulates 1-4 hold. Please prove this converse statement.

[Solution](#)

3.9 Philosophical Implications (10:10)

Now what does this tell us about the belief system of an inferentially perfectly rational being? Say, we are right in taking our Rational Belief postulates 1-4 for granted. Then, by the theorem, an inferentially perfectly rational belief system always looks like this: there is a non-empty set of worlds B_W , such that the propositions that the person believes are all the sets of worlds that have B_W as a subset.

For instance, let W again be our set of eight possible worlds from before; then e.g. B_W could be the set $X \cap Y$, that is, $\{w_2, w_5\}$ (see Figure 3.17).

Figure 3.17: $B_W = X \cap Y$

(Slide 61)

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

$$Y = \{w_2, w_3, w_5, w_6\}$$

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

E.g., assume belief to be generated by $B_W = X \cap Y = \{w_2, w_5\}$.

Believed: $X, Y, X \cup Y, \dots$

Not believed: $Z, \{w_5\}, \{w_8\}, \dots$

If for our given inferentially perfectly rational person's belief system, that set B_W is the one claimed to exist by the theorem, then we know how to determine that person's beliefs from B_W : e.g., X would be believed by her, Y would be believed, and $X \cup Y$ would be believed, as $\{w_2, w_5\}$ is a subset of each of them.

On the other hand, neither Z nor $\{w_5\}$ nor $\{w_8\}$ would be believed, as $\{w_2, w_5\}$ is not a subset of them.

By the theorem, whatever the belief system of an inferentially perfectly rational person on

this set of worlds is like, it is always generated by some special subset of the full set of worlds in such manner.

If I wanted to tell you what that person p 's belief system is like, it would suffice to inform you about her B_W set: in the present case, e.g., I would tell you ' B_W is $\{w_2, w_5\}$ '. You could determine everything else logically or set-theoretically. We can see that rationality has an effect of information compression, or reduction of complexity: instead of needing to determine for each subset of W , for each set of possible worlds, whether it is believed by person p or not, it is enough to determine one set: B_W . In our example: instead of determining for each of the 256, that is 2^8 , subsets of W whether they are believed or not, it is sufficient to determine for each world whether it is a member of B_W or not. That's a reduction from 2^8 bits – 1 meaning a set is believed, 0 meaning it is not believed – to 8 bits – 1 meaning the world is a member of B_W , 0 meaning that it is not.

What else can we conclude from our theorem together with our postulates on rational belief? As far as the role of belief in decision-making is concerned, an inferentially perfectly rational person will always act as if the actual world is a member of a certain distinguished set of possible worlds, the set B_W . She acts as if the actual world is identical to one of the members of B_W . As far as the role of belief in the person's theorizing about the world is concerned, her picture or theory of the world really consists in a set of candidates for what the actual world might be like: the candidates being the members of B_W .

We can also derive a very useful and plausible way of classifying propositions from the little theorem above in combination with the Rational Belief postulates 1-4: We know that an inferentially perfectly rational person's belief system, her set of believed propositions, is generated by just one set, B_W . But for every proposition X , every subset X of W , there are just three ways in which X can relate to B_W logically or set-theoretically:

(Slide 62)

Assume belief to be generated by B_W .

If X is a proposition (a subset of W), then there are three possible cases:

- B_W is a subset of X : X is believed.
- $B_W \cap X = \{\}$: $\neg X$ is believed (X is disbelieved).
- B_W is not a subset of X , $B_W \cap X \neq \{\}$:

Neither X is believed, nor $\neg X$ is believed.

Ad Case 2: B_W and X have empty intersection. In that case, it follows that B_W must be a subset of the negation of X , of the set W without X : so in that case, $\neg X$ is believed. Or, as we might also say: X is disbelieved, that is, the negation of X is believed.

Ad Case 3: B_W is not a subset of X , but B_W and X have a non-empty intersection. In that case X and B_W have non-empty overlap, but so do $\neg X$ and B_W . B_W is not a subset of X , but B_W is not a subset of the negation of X either. In that case, neither is X believed, nor the negation of X is believed: the inferentially perfectly rational person in question suspends judgement on X . As I had mentioned at the beginning of the last lecture, this is perfectly possible in the case of belief, while it would not be possible in the case of truth, since for truth it holds that either X is true or the negation of X is true.

Cases 1-3 cover all logically possible cases: for every X , one of them obtains, and which case does obtain can be read off immediately from what B_W and X are like.

In our example from before (see slide 63 and Figure 3.17 again), for instance, X is believed, $\{w_4, w_7\}$ is disbelieved, since its negation $\{w_1, w_2, w_3, w_5, w_6, w_8\}$ is believed, and Z is a case of suspended judgement: neither Z nor its negation $\neg Z$, that is, $\{w_1, w_2, w_3, w_8\}$ is believed.

(Slide 63)

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

$$Y = \{w_2, w_3, w_5, w_6\}$$

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

E.g., assume belief to be generated by $B_W = X \cap Y = \{w_2, w_5\}$.

Believed: e.g., X

Disbelieved: e.g., $\{w_4, w_7\}$

Neither believed nor disbelieved: e.g., $Z, \neg Z$

One final remark on all of that: as you might have realized by now, I try to keep funny symbols and technical terminology to an absolute minimum in these lectures, whether philosophical or mathematical symbols or terminology (well, some are required, of course). But right now I cannot resist the temptation to throw a couple of funny technical expressions at you, and in this case, mathematical expressions: the point is that all of the formal structures that we have encountered today are well-known from a part of one of the standard areas of modern mathematics: a part of algebra.

(Slide 64)

Boolean algebra of sets

The set of propositions in our sense of the word – the set of subsets of a given non-empty set W of possible worlds – is called a Boolean algebra of sets, or a Boolean field of sets. In such

an algebra, propositions, sets, are ordered according to the subset relation – the relation of logical consequence, in our terminology – every proposition or set has its complement or negation, for every two propositions there is their intersection or conjunction, and for every two propositions there is their union or disjunction.

Remark on Boolean algebras: If you want to know more about the mathematics of Boolean algebra, please take a look at
<http://plato.stanford.edu/entries/boolalg-math/>,
 (but please note that this Stanford Encyclopedia entry is mathematically very sophisticated and we won't need this at all in what follows), and if you want to read more about the history of applications of algebraic concepts and methods in logic, then please consider
<http://plato.stanford.edu/entries/algebra-logic-tradition/>.

(Slide 65)

Boolean algebra of sets

Filter

Furthermore, the Rational Postulates 1-4 characterize what is called a filter in algebra: a set of propositions, subsets of W , such that the largest proposition W is in the filter; the empty set is not in the filter; for every set in the filter, if the set is a subset of another one, then also that latter set is in the filter; and for every two sets in the filter, their intersection is in the filter, too. Our little theorem from before expresses then: every filter of subsets of a finite set W is generated by a non-empty subset of W : B_W , in our terminology. Not that this is a deep mathematical result in any sense; but that's not the point right now.

Remark on filters: See e.g.
<http://mathworld.wolfram.com/Filter.html>
 for the definition. But note that some textbooks define 'filter' in the way that a filter may contain the empty set as a member.

(Slide 66)

Boolean algebra of sets

Filter

Ultrafilter

Even truth for propositions corresponds to a well-known algebraic structure: the set of all propositions that are true at the actual world is what algebraists call an ultrafilter:

the set of all true propositions obeys the same laws as a filter, but additionally for every proposition, for every subset of W , either the set or its negation, its complement, is a member of the ultrafilter.

Remark on ultrafilters: See e.g.

<http://mathworld.wolfram.com/Ultrafilter.html> for the definition.

You might say: what do I care about these funny names that the mathematicians like to use? And that's absolutely right: the names are totally obsolete really. But it is truly remarkable that the same mathematical structures that we found to be exemplified by the logical space of propositions, by the set of propositions believed by an inferentially perfectly rational person, and by the set of true propositions, are also investigated heavily by mathematicians for their own purposes. We discovered filters by studying rational belief, while mathematicians discovered filters by studying the set of neighbourhoods of a point in a topological space, for example. They do not speak of propositions, rational belief, or truth, but some of their fundamental and most fruitful concepts – Boolean algebra, filter, and ultrafilter – apply to objects of their own mathematical interest just as much as they apply to objects of our philosophical interest.

Quiz 24:

(1): Let $W = \{w_1, \dots, w_8\}$ and take $B_W = \{w_2, w_5\}$ to be the least proposition again that is believed by an inferentially perfectly rational person p . For each of the following three propositions, please determine whether it is believed by p , or whether it is disbelieved by p (that is, its negation is believed), or whether p suspends judgement on it (that is, p neither believes it nor disbelieves it):

- (i) $\{w_1, w_5, w_7\}$,
- (ii) $\{w_1, w_3\}$,
- (iii) $\{w_1, w_2, w_4, w_5\}$.

(2): How many propositions (that is, subsets of $\{w_1, \dots, w_8\}$) does p believe, given that $\{w_2, w_5\}$ is a subset of all, and only, believed propositions?

[Solution](#)

Another example concerning B_W and belief: Say, you believe that Conny is presently either in the living room or in the kitchen, but you do not know at all in which of the two; in principle, there is also the quite remote possibility that she is in the neighbor's house, but at that point of time you would rule that out (however, given new evidence, you might be willing to reconsider that possibility). And, say, her location is also what you presently care about (since you want to ask her something or the like).

Now, if you want to model that situation in terms of possible worlds, you might choose a set W of precisely three possible worlds w_1, w_2, w_3 : Conny is in the living room at world w_1 , she is in the kitchen at world w_2 , and she is at the neighbor's house at world w_3 . B_W would be $\{w_1, w_2\}$ (since you believe that she is in the living room or in the kitchen), and you believe every proposition that is a superset of B_W (in this case this only means that, other than B_W itself, you also believe the less specific proposition W , that is, that she is in the living room or in the kitchen or in the neighbor's house). The point of all of that is to introduce a simple formal manner of expressing what your beliefs are like in that situation. The fact that there are three possible worlds in W , and that there are two worlds in B_W , reflects the fact that you do not have complete information in the situation that we are describing: your beliefs are incomplete. E.g., you neither believe that Conny is in the living room (after all, for all that you know, she might be in the kitchen), nor do you believe that she is in the kitchen (after all, for all that you know, she might be in the living room). Accordingly, neither $\{w_1\}$ nor $\{w_2\}$ is believed by you, as $\{w_1, w_2\}$ is not a subset of either of them.

Speaking of different possible worlds in our lectures will merely serve as a nice way of allowing one to say in formal terms that one's beliefs are incomplete.

3.10 Rational Degrees of Belief (07:50)

The view that we have developed so far concerning belief is incomplete in many ways, of course, but in one very salient respect in particular: as I said before, I believe that Socrates is a philosopher, I believe that the weather will be nice tomorrow, I believe that $2 + 2 = 4$; say, all of that is indeed the case. But additionally, it seems, I might believe each of these three propositions – that Socrates is a philosopher, that the weather will be nice tomorrow, that $2 + 2 = 4$ – with different strength. Presumably, I am most firmly convinced of $2 + 2 = 4$: maybe I even believe this simple mathematical law with maximal possible strength. I am very sure that Socrates is a philosopher, even though I admit that historical sources can get things wrong; which is why I might not be absolutely certain that Socrates is a philosopher, I am just very sure about it. Finally, my strength of belief in the weather being fine tomorrow might be even less than that: after all, I find the reliability of tv weather forecasts to be limited, and the same holds of my own powers of weather prediction. The upshot is: I believe each of the three propositions, but I do not believe

each of them with equal strength. On the one hand, there is all-or-nothing belief, which is a binary matter: either one believes a proposition, or one does not. In that sense, I believe each of the three propositions. But on the other hand, there are also strengths of belief: degrees of firmness with which one believes a proposition. These are not just an all-or-nothing affair: these strengths of belief can be greater or smaller.

But what are these strengths of belief really? In many ways they should be like all-or-nothing beliefs: they should dispose us to act given certain conditions; they should allow us to reason about the world; and so on. But how do they do so?

Let us first try to picture strengths of belief somehow:

Reconsider our diagram of the eight possible worlds from before, with the three propositions X , Y , Z , and all of their logically possible combinations (see Figure 3.18).

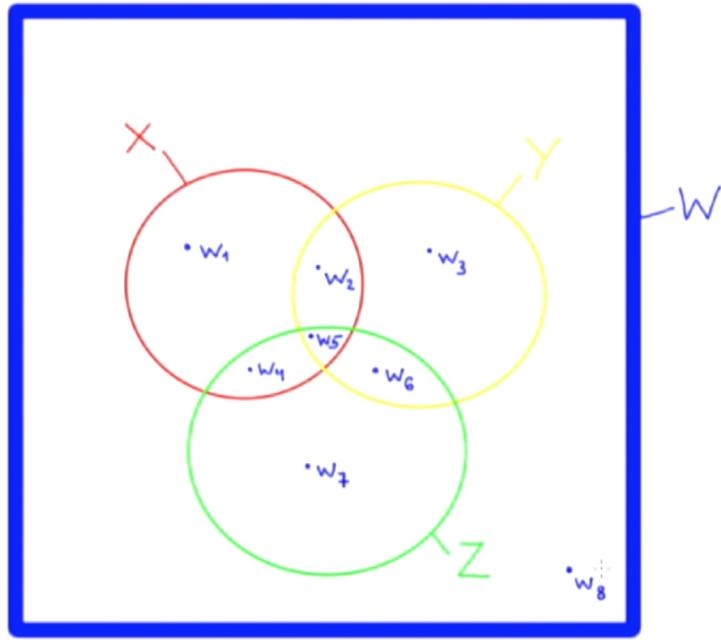


Figure 3.18: Distribution of w_1, \dots, w_8

There is a feature of that diagram that we haven't exploited so far: as pictured in the diagram, all of the various propositions, the various regions in the diagram, can be measured in terms of their areas: the areas of some of these regions are large, while the areas of other regions are small.

In order to make it easier to determine such areas, let us assume that the side length of the total set W of possible worlds, the outer square, is 1 (see Figure 3.19). And instead of regions in the form of circles, let us consider propositions of the shape of little squares.

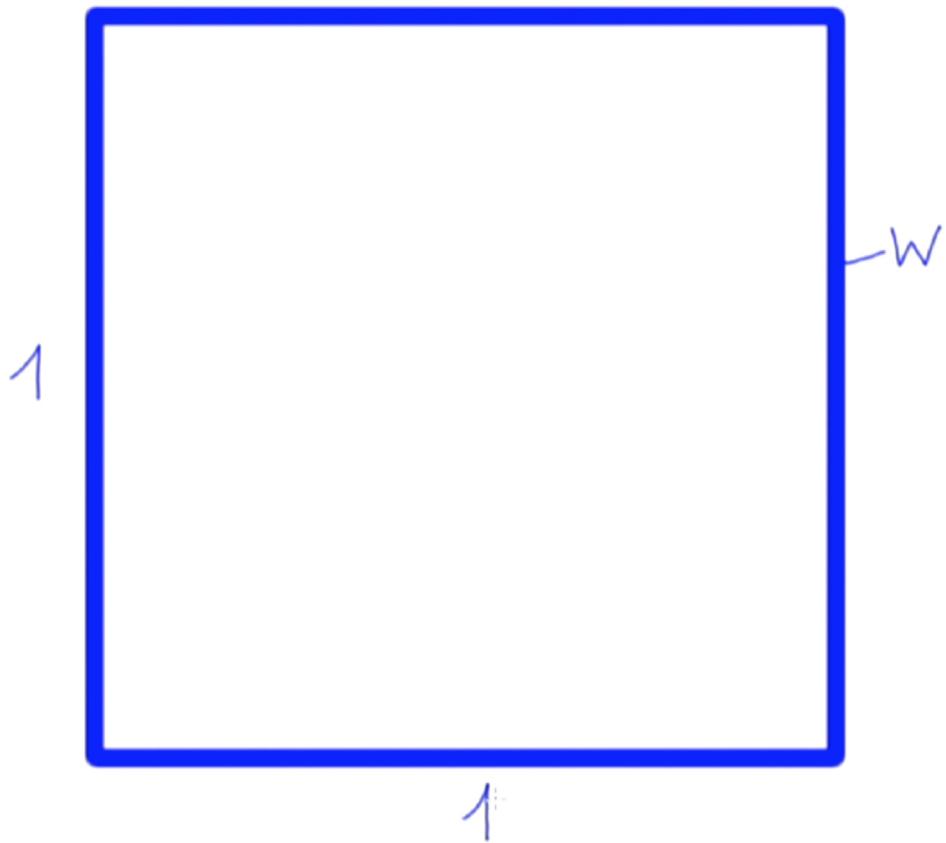


Figure 3.19: The side length of W equals 1

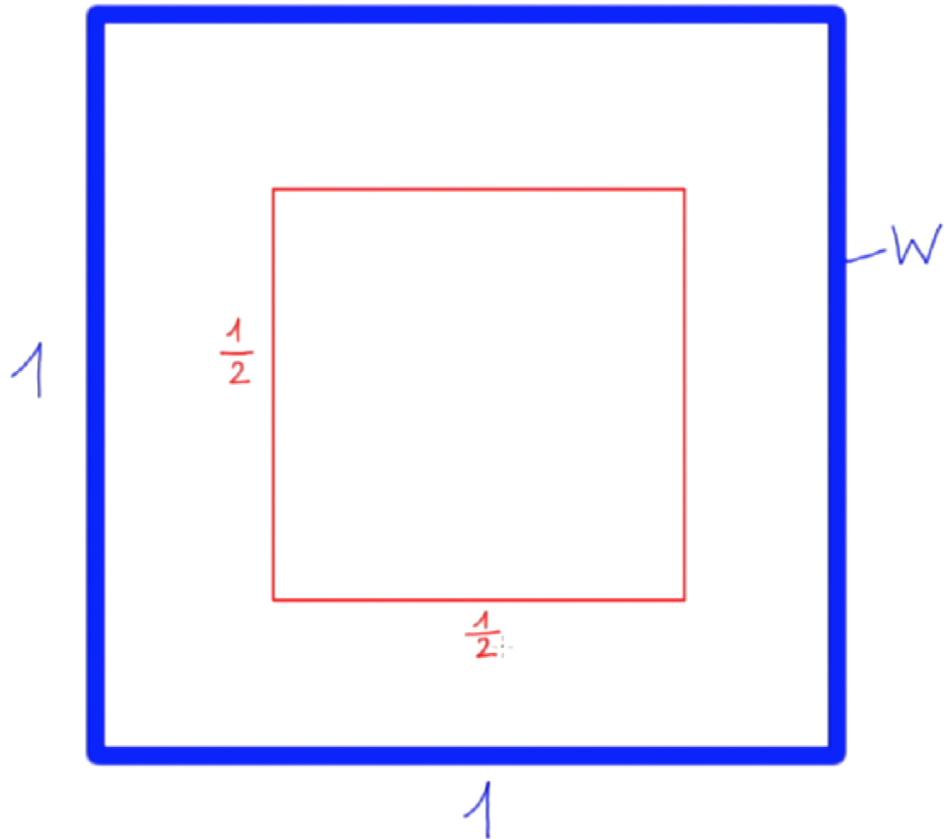


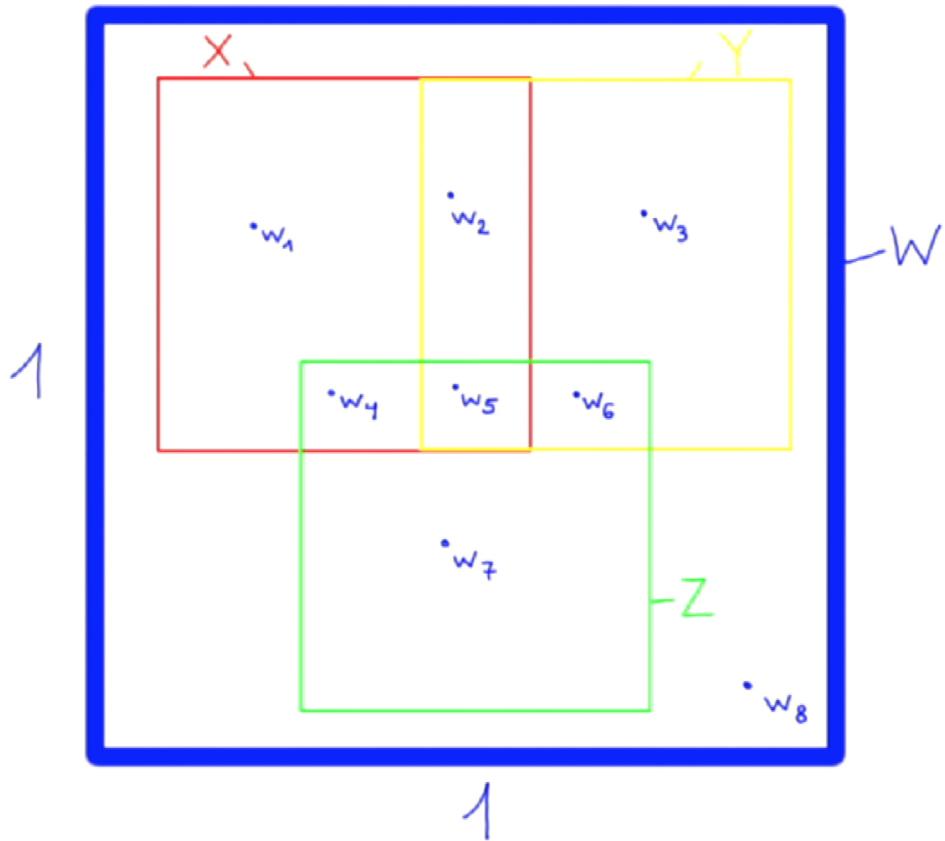
Figure 3.20: Propositions as squares

For instance, this proposition (see Figure 3.20) occupies an area of size

$$\frac{1}{4}$$

The area of the little square is $\frac{1}{4}$ of the total area of the whole square.

And this (Figure 3.21) is our previous diagram again, but now X , Y , and Z have the shape of squares:

Figure 3.21: X , Y , and Z as squares

As far as our previous considerations on propositions and belief are concerned, we could have worked with that diagram before – nothing would have changed. But in the present diagram it is easy to see that each of X , Y , Z occupies an area of $\frac{1}{4}$ each: $\frac{1}{4}$ of the total set W of all possible worlds, which of course occupies a maximal area of 1.

Here is another diagram of our eight possible worlds and our by now familiar propositions on these worlds (see Figure 3.22):

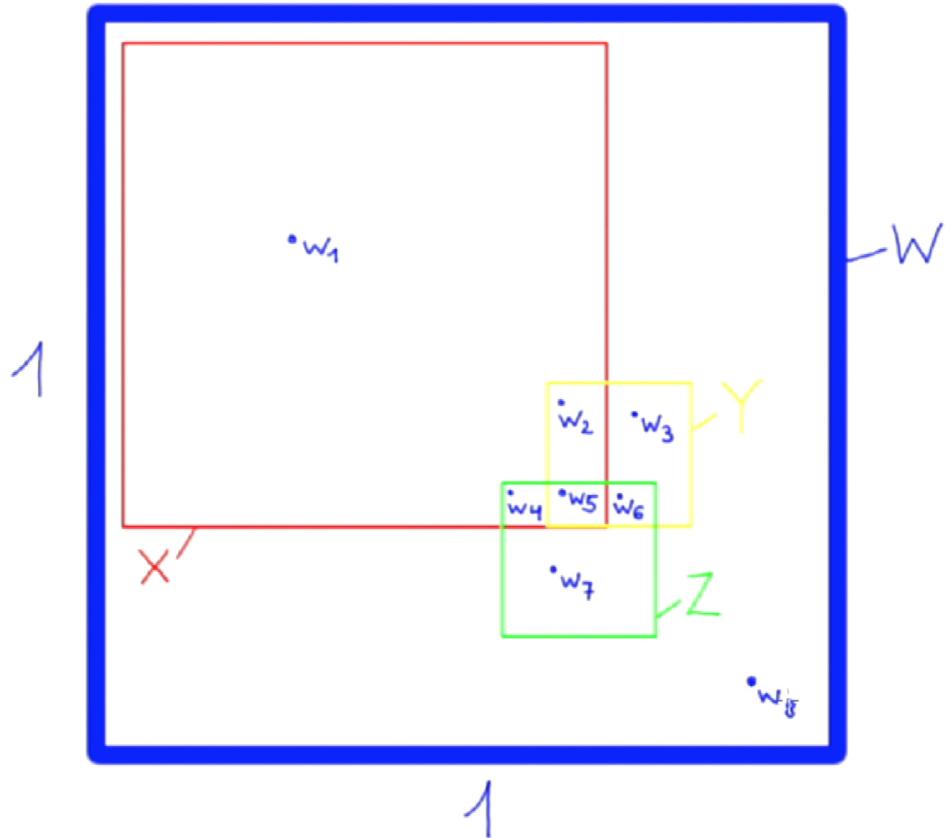


Figure 3.22: X , Y , and Z as squares of different size

As you can see, the region of proposition X is now much larger than before, and the areas of Y and Z are much smaller than that of X . Once again, we can build their logical combinations in the same way as before, it is just that some of the areas of the resulting regions have changed. For instance, consider the proposition $X \wedge \neg Y \wedge \neg Z$, that is, the intersection of X with the complement of Y and the complement of Z : that is the set $\{w_1\}$, as it would have been in all of the other diagrams of X , Y , Z that we had considered before. But now the area of the proposition $\{w_1\}$ is pretty big, unlike the areas for $\{w_1\}$ in any of the previous diagrams.

That is one way of visualizing strengths of belief in certain propositions – such strengths of belief are very much like the areas of the regions that depict the propositions. If the proposition $\{w_1\}$ occupies a large area, then this is like assigning a relatively great strength of belief to it; one regards it as likely that w_1 is the actual world. If the same proposition occupies a small area in a diagram, then this is like assigning a relatively weaker strength

of belief to it: one regards it as less likely that w_1 is the actual world.

Obviously, this is just a picture, a metaphor, if you like: actually, believing a proposition with a certain strength is a mental state that has not much to do with drawing large or small blobs in a diagram. But the picture might still help us develop a clearer conception of what strengths of belief are all about.

Consider the following facts about the areas of propositional regions in our present kind of diagrams:

(Slide 67)

- Area 1: $A(W) = 1$.

By our choice of the overall set W of possible worlds being depicted by a square of side length 1, the area of W is 1, of course.

(Slide 68)

- Area 2: For all propositions (regions) X , $0 \leq A(X) \leq 1$.

As depicted in our present type of diagram, all propositions occupy an area that is at most the area of the full square: so the area of whatever proposition depicted in our diagrams is less than or equal to 1. Since it does not make any sense in our diagrams to speak of negative areas, all areas of propositions or regions must also be greater than or equal to 0.

(Slide 69)

- Area 3: For all propositions (regions) X , for all propositions (regions) Y , if $X \cap Y = \{\}$, then $A(X \cup Y) = A(X) + A(Y)$.

This is just a simple law of area composition: if two propositions or regions do not overlap, that is, they have empty intersection, then the area occupied by their union is the sum of the areas of the single propositions. For example:

$X \wedge Y$ does not overlap with $\neg X \wedge Y$ (see Figure 3.23). Accordingly, the area of $X \wedge Y$ taken together with $\neg X \wedge Y$, in other words, the area of all of Y , is nothing but the sum of the area of $X \wedge Y$ and the area of $\neg X \wedge Y$.

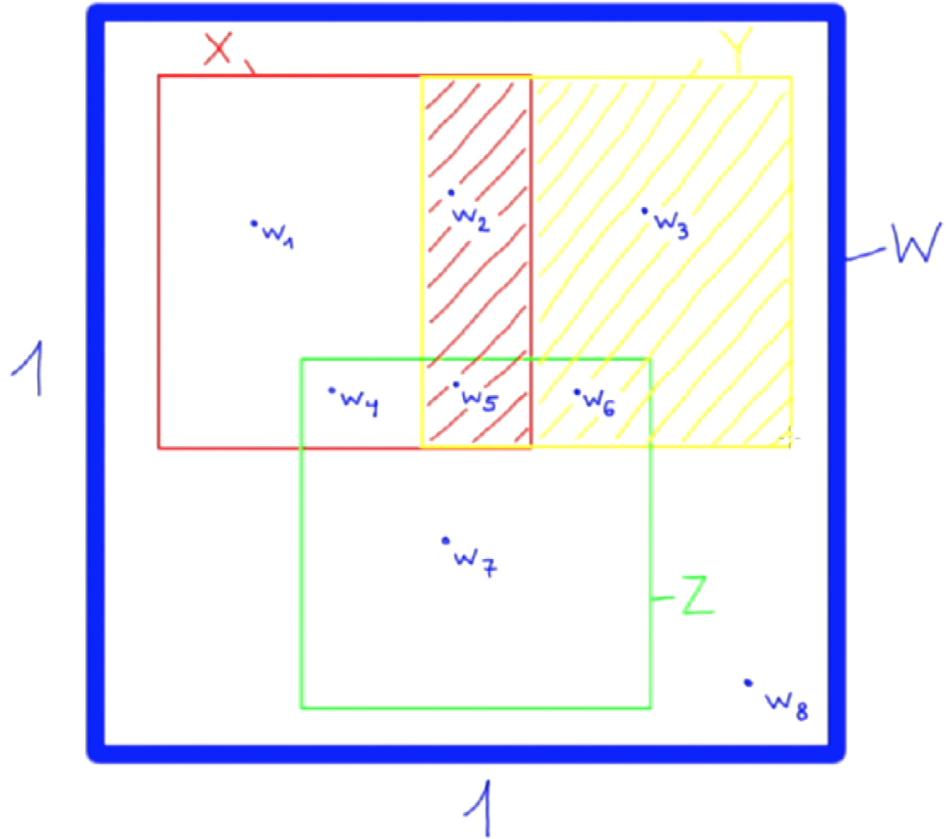


Figure 3.23: $X \wedge Y$ does not overlap with $\neg X \wedge Y \dots$

Quiz 25:

Draw a square of side length 1 (representing W) as in the lecture. Try out various different ways of drawing within the square two blobs representing X and Y , respectively. Please check for yourself that the following two principles on areas (two-dimensional geometrical sizes) are satisfied independently of how you draw the blobs:

- (i) $A(\neg X) = 1 - A(X)$ (“The area of the complement of X is equal to 1 minus the area of X ”).
- (ii) If X is a subregion of Y , then $A(X)$ is less than or equal to $A(Y)$.
(There is no need to prove anything at this point. For the moment, please just draw and look.)

3.11 Postulates on Rational Degrees of Belief (12:54)

O.k. Let us now forget about the function A measuring areas of geometrical regions in a square of side length 1. Let us think about A instead as measuring a person's strengths of belief, her degrees of belief, in propositions, where we regard 1 as the maximal possible degree of belief, and 0 as the minimal possible degree of belief. One might also say: 1 represents 100% belief, while 0 represents 0% of belief. In any case, even when bearing this revised interpretation of the function ' A ' in mind, each of the principles Area 1-3 look plausible.

(Slide 70)

- Area 1: $A(W) = 1$.
- Area 2: For all propositions (regions) X , $0 \leq A(X) \leq 1$.

Take Area 1: this means now that a person believes the proposition W , which is expressed by, e.g., a sentence of the form $A \vee \neg A$, with maximal strength 1; if the person is rational, then this is certainly what she should do; after all, a logical truth cannot be false. Area 2 simply follows from our interpretation of 0 and 1 as minimal and maximal strengths of belief, respectively.

(Slide 71)

- Area 3: For all propositions (regions) X , for all propositions (regions) Y , if $X \cap Y = \{\}$, then $A(X \cup Y) = A(X) + A(Y)$.

$$X \cap \neg X = \{\}.$$

$$X \cup \neg X = W.$$

$$A(X \cup \neg X) = A(W) = 1.$$

$$\text{By Area 3: } A(X \cup \neg X) = A(X) + A(\neg X).$$

$$\text{Therefore: } 1 = A(X) + A(\neg X).$$

$$\text{Equivalently: } A(\neg X) = 1 - A(X).$$

Finally, Area 3 is plausible, too: consider e.g. the case of two propositions X and $\neg X$: these two propositions have empty intersection; no possible world could make both of these propositions true simultaneously; so $\neg X$ is indeed one of the sets Y that have empty intersection with X . We know that $X \cup \neg X$ is W , and the strength of belief in W , $A(W)$, equals 1. So whatever the strength of belief in X is, it summed up with the strength of belief in $\neg X$ must yield 1, according to principle Area 3. Or equivalently: the degree of belief in $\neg X$ must be 1 minus the degree of belief in X .

This implies that if the degree of belief in X is very high, close to the degree of belief in W , the degree of belief in its negation $\neg X$ must be low, and the other way around, which is exactly what we would expect of a person's degrees of belief, if the person is rational. For instance, it would be irrational to believe X almost as firmly as a logical truth, and also to believe the negation of X almost as firmly as a logical truth: the truth of X rules out the truth of $\neg X$, and vice versa, and that should show up in the strengths of belief that one invests into X and $\neg X$, respectively. The degree of belief in X represents a certain percentage of belief, the degree of belief in $\neg X$ does so, too, and principle Area 3 with its new interpretation makes the two of them sum up to 100%, that is, to the degree of belief in $X \cup \neg X$, the degree for the total set of possible worlds. All of that sounds very reasonable indeed. And Area 3 is just the generalization of the special case where Y is identical to $\neg X$ to the more general case where Y might not be the complement of X , but where Y has still empty intersection with X .

Here is thus what seem to be plausible postulates on rational degrees of belief:

(Slide 72)

- Rational Degree of Belief 1: If a person is inferentially perfectly rational (with W as her set of entertainable possible worlds), then her degree of belief function P is such that: $P(W) = 1$.
- Rational Degree of Belief 2: If a person is inferentially perfectly rational (with W as her set of entertainable possible worlds), then her degree of belief function P is such that:

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

(Slide 73)

- Rational Degree of Belief 3: If a person is inferentially perfectly rational (with W as her set of entertainable possible worlds), then her degree of belief function P is such that:

For all propositions X (for all subsets X of W), for all propositions Y (for all subsets Y of W):

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

We speak of perfectly rational persons here again, because people in general can mess things up: in the case of degrees of belief just as much as in the case of all-or-nothing belief that we had considered before. But perfectly rational people do not. Or alternatively: our postulates Rational Degree of Belief 1-3 do not express empirical descriptions of the psychology of degrees of belief, but normative constraints on what the degrees of belief of a person ought to be like. Moreover, we also speak of inferentially perfectly rational person's

strengths of belief: as in the case of all-or-nothing belief before, the persons that we are considering right now do not have to be perfectly rational in each and every respect; as long as they are inferentially perfectly rational, they should take care that their degrees of belief distribute over the various possible logical or set-theoretic combinations of propositions in the way that gets expressed by the three postulates. For instance, as mentioned before, if the degree of belief in X is very high, close to the degree of belief in W , then the degree of belief in its negation $\neg X$ must be low, and vice versa: this is like a graded, a numerical kind of inference to the effect that support of X decreases support of $\neg X$, and vice versa. That is why we use the term ‘inferentially perfectly rational’ again.

Why do we use the symbol ‘ P ’ now instead of ‘ A ’? The reason is that our postulates Rational Degree of Belief 1-3 correspond to what mathematicians, statisticians, philosophers, psychologists, computer scientists, physicists, and so on call the ‘axioms of probability’. P above is nothing but a probability measure: hence the ‘ P ’. Given that in the present context ‘ P ’ denotes a person’s degree of belief function, her strength of belief function, we might also call P a person’s subjective probability measure.

(Slide 74)

Different kinds of probability:

- Statistical probability.
 - Objective single-case probability.
 - Geometrical probability.
- ⋮

This is important in so far as there are different kinds of probability around: e.g., in medical statistics, one might say that the probability of getting lung cancer for a male heavy smoker in Munich of age between 30 and 50 years is so-and-so. In this case, the probability is not assigned to a proposition at all but rather to a property – the property of getting lung cancer – and this is done so relative to a class of individuals – the male heavy smokers in Munich of age between 30 and 50 years. At least at first glance, this statistical probability is very different from the degrees of belief that we are currently interested in. Or take a quantum physicist who is telling us that the probability that this particular atom will decay in that particular period of time is so-and-so: this seems to be a different kind of probability again. It is the objective, worldly probability of a particular physical event to happen – the atom decaying in that period of time – and at least at first sight this seems to be independent again of what beliefs someone might have about this. Accordingly, the areas of geometrical regions in a square of side length 1 that we had considered before are purely mathematical probabilities: geometrical ones, if you like. We used these geometrical probabilities to picture degrees of belief somehow, but in the

meantime we are dealing with the real stuff: with subjective probabilities – a person's degrees of belief in propositions.

(Slide 75)

Different kinds of probability:

- Statistical probability.
- Objective single-case probability.
- Geometrical probability.
- ⋮
- **Subjective probability.**

Remark on probability: The mathematical axioms of probability are due to the Russian mathematician Andrey Nikolaevich Kolmogorov. As pointed out in the lecture, the purely mathematical structure of probability spaces is open to various different interpretations about which you can learn more at

<http://plato.stanford.edu/entries/probability-interpret/>.

For our own purposes in this lecture, only the interpretation of probabilities in terms of degrees of belief will be relevant.

Let us see what conclusions we can draw from our postulates on rational degrees of belief:

Let W be a finite, non-empty set of possible worlds. If the Rational Degree of Belief postulates 1-3 are the case (for that given set W of possible worlds), then for every inferentially perfectly rational person's degree of belief function P (on W as her set of entertainable possible worlds), the following holds for all propositions X, Y :

(Slide 76)

- (i) $P(\neg X) = 1 - P(X).$
- (ii) $P(\{\}) = 0.$
- (iii) $P(Y) = P(X \cap Y) + P(\neg X \cap Y).$
- (iv) If X is a subset of Y , then $P(X) \leq P(Y).$

O.K. Lets have a look at that in detail.

- (i) we have already determined before:

(Slide 77)

- (i) Show: $P(\neg X) = 1 - P(X)$.

By Rational Degree of Belief 3: $P(W) = P(X) + P(\neg X)$.

By Rational Degree of Belief 1: $P(W) = 1$.

Hence: $1 = P(X) + P(\neg X)$.

That is: $P(\neg X) = 1 - P(X)$.

- (ii) follows by plugging in W for ‘ X ’ in (i):

(Slide 78)

- (ii) Show: $P(\{\}) = 0$.

Plug in ‘ W ’ for ‘ X ’ in (i): $P(\neg X) = 1 - P(X)$.

$P(\neg W) = 1 - P(W)$.

Since $\neg W = \{\}$, and $P(W) = 1$ by Rational Degree of Belief 1:

$$P(\{\}) = 1 - 1 = 0.$$

- (iii) follows immediately from Rational Degree of Belief postulate 3:

(Slide 79)

- (iii) Show: $P(Y) = P(X \cap Y) + P(\neg X \cap Y)$.

Since $(X \cap Y) \cup (\neg X \cap Y) = \{\}$, it follows from Rational Degree of Belief 3:

$$P((X \cap Y) \cup (\neg X \cap Y)) = P(X \cap Y) + P(\neg X \cap Y).$$

$$P(Y) = P(X \cap Y) + P(\neg X \cap Y).$$

The proposition $(X \cap Y) \cup (\neg X \cap Y)$ whose degree of belief we are taking on the left side, is actually identical to Y : it includes all the members of Y that are also members of X , taken together with all the members of Y that are members of $\neg X$; since every member of Y must either be a member of X or a member of $\neg X$, this is precisely the set of all members of Y , that is, the set Y . So it follows that

$$P(Y) = P(X \cap Y) + P(\neg X \cap Y),$$

as claimed in (iii).

Finally, (iv) follows from (iii):

(Slide 80)

(iv) Show: If X is a subset of Y , then $P(X) \leq P(Y)$.

By (iii): $P(Y) = P(\textcolor{red}{X} \cap Y) + P(\neg X \cap Y)$.

If X is a subset of Y , then $\textcolor{red}{X} \cap Y = \textcolor{red}{X}$.

So: $P(Y) = P(\textcolor{red}{X}) + P(\neg X \cap Y)$.

By Rational Degree of Belief 2: $0 \leq P(\neg X \cap Y) \leq 1$.

Hence: $P(Y) = P(X) + P(\neg X \cap Y) \geq P(X)$.

That is: $P(X) \leq P(Y)$.

Remark on axioms of probability and what follows from them: Some of the statements that we have just *proven* correspond to those that we *inspected* to hold in the quiz before when we interpreted ‘ P ’ in terms of areas of regions. Within our proof of these statements on the basis of the axioms of probability, the precise interpretation of ‘ P ’ was actually irrelevant, only its mathematical structure was important.

Quiz 26:

Assume our Rational Degree of Belief postulates 1-3 to be the case (for a given set W of possible worlds). Let P be the degree of belief function of an inferentially perfectly rational person.

Show that for all propositions X, Y : $P(X \cap Y)$ is less than or equal to $P(X)$.

[Solution](#)

3.12 A Theorem on Rational Degrees of Belief (07:06)

There is more to derive: one can actually show something like the probabilistic counterpart to the little theorem on all-or-nothing belief that we had dealt with earlier today. Back then we showed that given the Rational Belief Postulates and W being finite, the belief system of an inferentially perfectly rational person can always be generated from a particular proposition, B_W .

The counterpart in the present context is: given the Rational Degree of Belief Postulates, and W being finite again, the degree of belief function of an inferentially perfectly rational person can always be generated from an assignment of non-negative real numbers to worlds, where the numbers that are assigned in that way sum up to 1.

Here is the theorem:

(Slide 81)

Theorem:

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

If the Rational Degree of Belief postulates 1–3 are the case (for the given set W of possible worlds), then for every inferentially perfectly rational person's degree of belief function P (on W as her set of entertainable possible worlds), there is a function B , such that the following holds:

- B assigns to each world w in W a non-negative real number $B(w)$.
- The sum of all the values of B on worlds in W is 1: that is, if $W = \{w_1, \dots, w_n\}$, then $B(w_1) + \dots + B(w_n) = 1$.
- For all propositions X (over W), $P(X)$ is the sum of the values of B on worlds in X ; that is:

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

Let me illustrate what this little theorem says in terms of our set W of eight possible worlds again. We are facing 2^8 , that is, 256 propositions. Given all of the assumptions in the theorem, the degree of belief function P of an inferentially perfectly rational person assigns to each of these 256 propositions a degree of belief, such that the axioms of probability are satisfied. One might think, therefore, that in order to determine what the degree of belief function of such a person is like, one would have to determine 256 real numbers: one for each proposition. But according to the theorem above, the situation is much simpler than that: once again rationality reduces complexity, amongst other things. The theorem says that our rational person's degree of belief function P is already determined by just 8 real numbers, one for each world in W ; the function B maps the worlds in W to precisely these numbers.

For instance, B could be the following function:

(Slide 82)

E.g., assume P to be generated by:

$$B(w_1) = 1/15$$

$$B(w_2) = 1/3$$

$$B(w_3) = 1/15$$

$$B(w_4) = 1/15$$

$$B(w_5) = 1/3$$

$$B(w_6) = 1/15$$

$$B(w_7) = 1/15$$

$$B(w_8) = 0$$

(See Figure 3.24.)

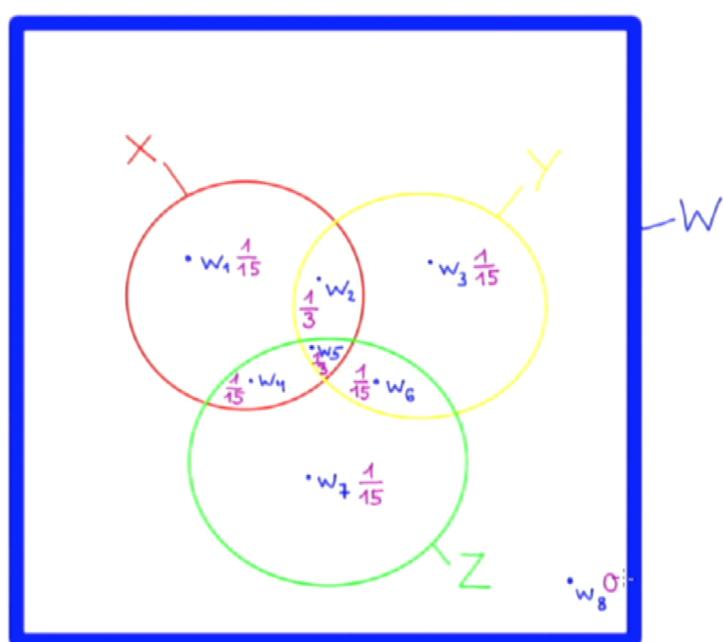


Figure 3.24: Values of B

If one sums up all the values of B ,

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

then that sum is indeed equal to 1.

If, for our given inferentially perfectly rational person's function P , this very function B is the one claimed to exist by the theorem, then we know precisely how to determine P from B :

For example, $P(\{w_1\}) = P$ applied to the proposition that includes only w_1 as a member – is the sum of values of B for the worlds in that proposition; but the only member of the proposition in question is w_1 : hence, $P(\{w_1\}) = B(w_1) = 1/15$.

More generally, we see that for each world w , $B(w)$ is nothing but $P(\{w\})$.

Now, what is our rational person's degree of belief in the proposition X , that is, what is $P(\{w_1, w_2, w_4, w_5\})$? That degree of belief is the sum of the values of B for worlds that are members of that proposition: the sum $B(w_1) + B(w_2) + B(w_4) + B(w_5)$. And that sum is $1/15 + 1/3 + 1/15 + 1/3 = 2/3 + 2/15$, which is precisely 0.8.

(Slide 84)

$$\begin{aligned} P(\underbrace{\{w_1, w_2, w_4, w_5\}}_X) &= B(w_1) + B(w_2) + B(w_4) + B(w_5) = \\ &= 2/3 + 2/15 = 0.8 \end{aligned}$$

The person that we are describing right now believes that Socrates is a philosopher with strength 0.8, or 80%. Accordingly, $P(\neg X)$ is the sum of the rest of the B -values, those for the worlds outside of X , and of course these sum up to 0.2: after all, we know already that $P(X) + P(\neg X)$ must sum up to 1. Similarly, $P(X \cap Y)$ is equal to $2/3 : B(w_2) + B(w_5)$, which is $1/3 + 1/3$, or 0.6666..., or 66%, whatever. And so on. By the theorem, every degree of belief function P of an inferentially perfectly rational person can be generated in such manner. And if one thinks about it, it follows again that for given P its corresponding generating function B is uniquely determined, just as it had been the case for rational belief and B_W previously.

I won't state the proof of this theorem here. But it is not difficult: basically one shows that by the laws of probability, which are captured by our Rational Degree of Belief postulates 1-3, the probability of a set can always be decomposed into the sum of the probabilities of its single-member subsets; and the probabilities of these single-member sets give us precisely the values of the B -function; just as in the previous example, $B(w_1)$ had been identical to $P(\{w_1\})$.

Quiz 27:

Just as in the lecture, assume that an inferentially perfectly rational person's degree of belief function P is determined by the following function B (in the manner explained by the last theorem):

$$B(w_1) = 1/15$$

$$B(w_2) = 1/3$$

$$B(w_3) = 1/15$$

$$B(w_4) = 1/15$$

$$B(w_5) = 1/3$$

$$B(w_6) = 1/15$$

$$B(w_7) = 1/15$$

$$B(w_8) = 0$$

Determine the following degrees of belief:

- (i) $P(\{w_8\})$,
- (ii) $P(\{w_4, w_7, w_8\})$,
- (iii) $P(\{w_1, w_2, w_3, w_5, w_6\})$.

[Solution](#)

3.13 Rational Degrees of Belief and Bets (15:16)

Let me add a bit more to the justification for our Rational Degree of Belief postulates instead: o.k., so we motivated them by the manner in which areas of geometrical regions are determined, and we found them plausible as constraints on an inferentially perfectly rational person's degree of belief function independently. But can we argue for them also in some yet more convincing manner?

Indeed, that is possible. In order to do so we need to take a step back and reconsider the role that beliefs play in the determination of action. Just as discussed before for all-or-nothing belief, also degrees of belief dispose ourselves to act in a particular way given that certain circumstances are the case.

A salient type of such circumstances are given by betting situations, presupposing that one has the corresponding desire to bet. That is where degrees of belief in the truth of propositions present themselves most nicely in terms of action: in this case, in terms of buying or selling bets.

Let me explain what I mean by a bet on the proposition X :

First of all, I pay a certain amount of money for the bet, I buy the bet from a so-called bookie in order to be in, to be able to make the bet with him. Say, what I pay for the bet is an amount of money of the form $q \cdot S$ Euro.

If ultimately X turns out to be false, so I was wrong, then that is also the amount of money that I will lose by betting on X : it's my net loss. But if X turns out to be true, and hence I was right, I win S Euro: since I paid $q \cdot S$ Euro initially, and I win S Euro afterwards, I will make $S - q \cdot S$ Euro overall by being right, that is, by betting on X and X indeed being the case. $S - q \cdot S$ is the net profit that I make then.

(Slide 85)

I buy a bet on the proposition X :

- First of all, I pay $q \cdot S$ Euro.
- If X is false, my net loss is $q \cdot S$ Euro.
- If X is true, my net profit is $S - q \cdot S$ Euro.

S is often called the ‘stake’ of the bet; it is the sum of the net amount of money that I win in case X is true together with the net amount of money that I lose in case X is false, so

(Slide 86)

S : *stake* of the bet; $S = (S - q \cdot S) + q \cdot S$.

q is called the ‘betting quotient’ of this bet on X ; q is identical to my net loss in case X is false, $q \cdot S$, divided by the stake, S :

(Slide 87)

q : *betting quotient* of the bet; $q = \frac{q \cdot S}{S}$

The idea is now that

(Slide 88)

I regard a bet on X as fair if and only if my degree of belief $P(X)$ in X equals the betting quotient q of the bet:

$$q = P(X)$$

If I regard a bet on X as fair, then I should always be happy buying the bet on X , if I am gambling person, which we are presupposing now.

For example; let P be my degree of belief function.

(Slide 89)

E.g., $P(X) = 1/2$, $q = 1/2$:

- First of all, I pay $q \cdot S = S/2$ Euro.
- If X is true, my net profit is $S - q \cdot S = S/2$ Euro.
- If X is false, my net loss is $q \cdot S = S/2$ Euro.

Assume $P(X) = 1/2$: by what I said before, I regard a bet on X as fair if and only if the betting q of the bet is $1/2$, that is, equal to my degree of belief in X . I pay $q \cdot S$ Euro for any such bet, that is, $S/2$ Euro; if X is the case, then I will win S . The net profit is $S/2$ Euro. If X is not the case, then I will simply lose the $S/2$ Euro that I had paid for the bet. The net loss is then $S/2$ Euro again. My belief in X is equally strong as my belief in $\neg X$, and accordingly I gain just as much in case X is true as I lose in case $\neg X$ is true. Which is why such a bet should appear fair to me. If my degree of belief in X were greater than the betting quotient of $1/2$, then I should be even happier buying the bet: for I would regard it as more likely to end up with a net profit of $S/2$ Euro than to end up with a net loss of $S/2$ Euro; buying the bet would seem advantageous to me. If my degree of belief in X were less than the betting quotient of $1/2$, then I should be unhappy to buy the bet: for I would regard it as less likely to end up with a net profit of $S/2$ Euro than to end up with a net loss of $S/2$ Euro; I would feel disadvantaged. In either of these two cases the bet should not seem fair to me.

Or consider the extreme case $P(X) = 1$:

(Slide 90)

E.g., $P(X) = 1$, $q = 1$:

- First of all, I pay $q \cdot S = S$ Euro.
- If X is true, my net profit is $S - q \cdot S = 0$ Euro.
- If X is false, my net loss is $q \cdot S = S$ Euro.

By what I said before, I regard a bet on X as fair if and only if the betting quotient q of the bet is 1, that is, equal to my degree of belief in X . I pay S Euro for the bet; if X is the case, then I win S Euro. My net profit is 0 Euro. Nothing lost, nothing gained. If X is not the case, then I lose all of S . But I do not care: for I am certain that X is the case anyway, since I assign the maximal degree of belief to X . If $P(X) = 1$, I will simply regard every bet on X as fair in which I do not gain anything in net value if X is the case, and in which I lose money if X were not the case. That I should still regard such a bet as fair is due to my extreme certainty about X .

And so on. Let me also put on record that in a case in which the stake S is precisely 1 Euro, I regard a bet as fair if and only if I pay precisely my degree of belief $P(X)$ in Euro in order to buy the bet. For example, if $P(X) = 0.4$, then I pay 0.4 Euro, and this will seem fair to me.

I should add that regarding a bet as fair is a two-edged sword: if I regard a bet as fair, I should be happy being on the other side of the bet, too, that is, selling the bet on X rather than buying it. In that case I will become the bookie, and someone else might buy the bet on X from me. In the case of a fair bet, this should be fine with me again. In that reversed case, if X turns out to be false, my opponent's net loss will be $q \cdot S$ Euro – her bet was on X , and X is false – and thus my net profit will be $q \cdot S$ Euro. But if X turns out to be true, my opponents net profit will be $S - q \cdot S$ Euro, and accordingly my net loss will be $S - q \cdot S$ Euro. Everything is simply reversed then:

(Slide 91)

If I sell a bet on the proposition X , things are reversed:

- First of all, my opponent pays $q \cdot S$ Euro.
- If X is true, my net loss is $S - q \cdot S$ Euro.
- If X is false, my net profit is $q \cdot S$ Euro.

In all of that, we are assuming a very idealized setting: for instance, the extra fun of the process of betting itself (if there is one) is not taken into account at all, it is not measured in terms of money in any way; and the like. Anyway.

Here is why I am bringing all of this up: Let us assume now that I do not distribute my beliefs according to the laws of probability. As it turns out, it is possible then to come up with a sequence of bets all of which I regard as fair, but whatever is going to happen – whatever the actual world is like – I am going to lose money, if I enter all of these bets. In other words: if my degrees of belief do not conform with the standard of subjective probability, then someone can make a sequence of bets with me in which I am logically guaranteed to lose money overall, come what may. Although all of the bets seem fair to me. This shows that not distributing degrees of belief according to the norms of subjective probability theory is irrational. And hence our Rational Degree of Belief postulates 1-3 from before are justified.

All of this can be proven by mathematical means in the form of a precise theorem, but I will restrict myself just to an example here:

(Slide 92)

Assume two propositions X and Y to have empty intersection, and:

$$P(X) = 0.4, P(Y) = 0.4, \text{ but } P(X \cup Y) = 0.7 (!)$$

According to the laws of probability, that degree of belief in $X \cup Y$ should be 0.8, but let us assume that I am violating these laws by assigning a degree of belief of 0.7 to $X \cup Y$.

Here are three bets that we are going to consider; I will make things simple by assuming the stakes in question always to be 1 Euro. Here is the bet:

(Slide 93/1)

- Bet on X : $P(X) = 0.4, S_X = 1$ Euro, $q_X = 0.4$.

So I regard this bet on X as fair, since its betting quotient equals my degree of belief in X .

(Slide 93/2)

- Bet on X : $P(X) = 0.4, S_X = 1$ Euro, $q_X = 0.4$.

X is true: my net profit is $S_X - q_X \cdot S_X = 1 - q_X = 0.6$ Euro.

$\neg X$ is true: my net loss is $q_X \cdot S_X = q_X = 0.4$ Euro.

Then here is the bet on Y :

(Slide 94/1)

- Bet on Y : $P(Y) = 0.4, S_Y = 1$ Euro, $q_Y = 0.4$.

Thus I consider this bet on Y to be fair again.

(Slide 94/2)

- Bet on Y : $P(Y) = 0.4, S_Y = 1$ Euro, $q_Y = 0.4$.

Y is true: my net profit is $S_Y - q_Y \cdot S_Y = 1 - q_Y = 0.6$ Euro.

$\neg Y$ is true: my net loss is $q_Y \cdot S_Y = q_Y = 0.4$ Euro.

Finally, there is also a

(Slide 95/1)

- Bet on $X \cup Y$: $P(X \cup Y) = 0.7, S_{X \cup Y} = 1$ Euro, $q_{X \cup Y} = 0.7$.

Once again the bet is fair in my eyes. Because of that, I should be happy to sell that bet in this case, too. Let's assume someone else is betting on $X \cup Y$ against me. So if $X \cup Y$ is true, the net profit of my opponent will be $S_{X \cup Y} - q_{X \cup Y} \cdot S_{X \cup Y} = 1 - q_{X \cup Y} = 0.3$ Euro.

And if $\neg(X \cup Y)$ is true, then the net loss of my opponent will be $q_{X \cup Y} \cdot S_{X \cup Y} = q_{X \cup Y} = 0.7$ Euro. That means for me:

(Slide 95/2)

- Bet on $X \cup Y$: $P(X \cup Y) = 0.7$, $S_{X \cup Y} = 1$ Euro, $q_{X \cup Y} = 0.7$.

$X \cup Y$ is true: my net loss is $S_{X \cup Y} - q_{X \cup Y} \cdot S_{X \cup Y} = 1 - q_{X \cup Y} = 0.3$ Euro.

$\neg(X \cup Y)$ is true: my net profit is $q_{X \cup Y} \cdot S_{X \cup Y} = q_{X \cup Y} = 0.7$ Euro.

(Slide 96/1)

Say, I bet on X , I bet on Y , and someone else bets on $X \cup Y$ against me; my degrees of belief tell me that all of these bets are fair.

So I am happy to be part of each of them. But independently of what is going to happen, I am bound to lose money overall. Here is why:

(Slide 96/2)

- Case 1: X is true and Y is true, hence $X \cup Y$ is true:

logically impossible, since X and Y have empty intersection!

- Case 2: X is true and Y is false, and hence $X \cup Y$ is true:

$0.6 - 0.4 - 0.3 = -0.1$. Overall I lose 0.1 Euro.

- Case 3: X is false and Y is true, and hence $X \cup Y$ is true:

$-0.4 + 0.6 - 0.3 = -0.1$. Overall I lose 0.1 Euro.

- Case 4: X is false and Y is false, and hence $X \cup Y$ is false:

$-0.4 - 0.4 + 0.7 = -0.1$. Overall I lose 0.1 Euro.

Whatever case is the actual one, I will lose money overall. My degrees of belief dispose me to act in this irrational manner, they dispose me to enter a sequence of bets that look fair to me according to these degrees of belief but in which I am logically guaranteed to lose money. This cannot be rational. Therefore, a degree of belief function that is not a subjective probability measure cannot be a rational degree of belief function.

This line of argument in favour of degrees of belief having to be subjective probabilities derives from the famous work by the British philosopher, logician, mathematician, and economist Frank Plumpton Ramsey, as well as from that of the Italian mathematician and statistician Bruno de Finetti – work done in the first half of the 19th century. One can also prove a converse of this result: if one's degrees of belief conform with the laws of probabilities, then one can show there cannot be any sequence of bets in which one is guaranteed to lose money come what may. One may of course lose money in certain

possible cases, but not in all possible cases. In this sense, one is always better off by having a degree of belief function that is a probability measure than by having one that is not. Our Rational Degree of Belief postulates 1-3 from above are vindicated.

Remark on degrees of belief and betting: These betting arguments in favor of the axioms of subjective probability are called ‘Dutch book arguments’ in the relevant literature. If you want to know about them in more detail, please take a look at

<http://plato.stanford.edu/entries/dutch-book/>.

Or you might want to check out the shorter summary on Dutch books that is part of

<http://plato.stanford.edu/entries/epistemology-bayesian/>;

that Stanford Encyclopedia entry also gives a nice overview of how to do epistemology on the basis of the fundamental concept of rational degree of belief more generally, but please note that to some of the topics that it discusses we will only turn in subsequent lectures (e.g., we will deal with the topic of conditionalization in Lecture 4, Stephan will deal with confirmation in Lecture 5, and so on).

A terminological remark: doing epistemology in terms of subjective probabilities is sometimes called ‘Bayesian epistemology’. More generally, ‘Bayesianism’ refers to the program of applying subjective probability theory in various different areas of research; accordingly, there is Bayesian philosophy of science, Bayesianism in cognitive psychology, and the like. One final thing: the Dutch book arguments are both philosophically and mathematically non-trivial, which is why it is perfectly fine for you not to understand all the details at this point. What you should memorize, though, is this basic insight: if one does not distribute one’s degrees of belief according to the axioms of probability, then one is subject to a “Dutch book”: that is, there is a sequence of bets that will look fair given one’s degrees of belief, but where one is guaranteed to lose money overall if one makes all the bets in the sequence.

Quiz 28:

Say, someone believes $X \wedge \neg X$ with a degree of 0.3. (Obviously, this is not in line with the axioms of subjective probability, as they demand believing the contradictory proposition $X \wedge \neg X$ with the minimal degree of 0.) Can you offer a bet to that person so that the person is guaranteed to lose money?

[Solution](#)

3.14 The Lottery Paradox (10:33)

There is something that might have been on your mind already for quite some time now. We have studied all-or-nothing belief, and indeed rational all-or-nothing belief first; after which we turned to degrees of belief, and indeed rational degrees of belief. But are we

supposed to have beliefs on both kinds of scale: all-or-nothing beliefs and numerical beliefs simultaneously?

At least at first glance, it seems so, yes. In our original example on belief, our inferentially perfectly rational person believed $X \cap Y$, that Socrates is a philosopher and Plato is a teacher of Aristotle. And then we considered an example concerning degrees of belief, in which an inferentially perfectly rational person believed $X \cap Y$ with a degree of $2/3$ or $0.66666\dots$ on a numerical scale from 0 to 1. There does not seem to be any reason why the one inferentially perfectly rational believer could not be identical to the other inferentially perfectly rational believer. Indeed it does seem to be the case that I believe certain propositions – ask me, and you will see that I will assert them – and I also assign certain degrees of belief to the same propositions, strengths of belief that will show up e.g. in my betting behaviour, if you offer me the corresponding bets. I seem to exemplify all-or-nothing beliefs and degrees of belief in the very same propositions simultaneously.

But am I still rational in doing so? What do my beliefs and degrees of belief need to be like in order for me to be rational on both sides at the same time?

This question is more difficult than one might think. Consider the following argument – a paradox again, which is called the Lottery Paradox in the relevant literature and which goes back to the US-American philosopher and computer scientist Henry Kyburg who passed away just a couple of years ago in 2007:

Consider a lottery with 1,000,000 tickets that is going to take place soon. Assume that I am perfectly aware of this being so.

The first premise of the argument simply says:

(P1) I am absolutely certain that some ticket will win in the lottery; hence, my degree of belief in the proposition that ticket 1 will not win in the lottery and ticket 2 will not win in the lottery and ... and ticket 1,000,000 will not win in the lottery is 0:

(Slide 98)

Consider a lottery with 1000000 tickets:

(P1)

$$P(\text{not ticket 1 wins} \wedge \text{not ticket 2 wins} \wedge \dots \wedge \text{not ticket 1000000 wins}) = 0.$$

I exclude the possibility that no ticket will win in the lottery; for some ticket will be drawn.

The second premise expresses that I regard the lottery as fair; as far as I can tell, there is no positive or negative bias towards any of the tickets:

(P2) My degree of belief in the proposition that ticket 1 will win is $1/1000000$. My degree of belief in the proposition that ticket 2 will win is $1/100000$ My degree of belief in the proposition that ticket 1000000 will win is $1/100000$.

Or slightly more formally:

(Slide 99)

(P2)

$$P(\text{ticket 1 wins}) = 1/1000000$$

$$P(\text{ticket 2 wins}) = 1/1000000$$

:

$$P(\text{ticket 1000000 wins}) = 1/1000000$$

I consider each ticket to be equally likely to win. And indeed, for each of them it is very unlikely that it will win.

Which brings me to the next premise:

(P3) For every proposition X : I believe X if and only if my degree of belief in X is greater than or equal to some threshold, say, 0.9.

That is,

(Slide 100)

(P3) For every proposition X :

I believe X if and only if $P(X) \geq 0.9$.

This sounds very plausible again: if I believe something in the all-or-nothing sense, then if I am rational, my degree of belief in it should be high; and the other way around, too. The threshold 0.9 might seem a bit arbitrary, but at least something in that ballpark should do.

I also assume:

(Slide 101)

(P4) For every proposition X , for every proposition Y :

if I believe X and I believe Y , then (if I am perfectly rational) I also believe $X \wedge Y$.

P4 corresponds to what our Rational Belief postulate 4 says about inferentially perfectly rational believers: their beliefs are closed under taking conjunctions or intersections.

Finally, we have:

(P5) For every proposition X , my degree of belief in X is 1 minus the degree of belief in $\neg X$.

More formally:

(Slide 102)

(P5) For every proposition X :

$$P(\neg X) = 1 - P(X).$$

$$(\text{And: } P(X) = 1 - P(\neg X).)$$

This is what we know to hold for every inferentially perfectly rational person's degree of belief function, according to our Rational Degree of Belief postulates 1 and 3.

So these are the premises.

From them we can conclude, by logic:

(Slide 103)

(C)

$$P(\text{not ticket 1 wins} \wedge \text{not ticket 2 wins} \wedge \dots \wedge \text{not ticket 1000000 wins}) \geq 0.9, \text{ and}$$

$$P(\text{not ticket 1 wins} \wedge \text{not ticket 2 wins} \wedge \dots \wedge \text{not ticket 1000000 wins}) = 0.$$

which is absurd: my degree of belief in one and the same proposition cannot be greater than or equal to 0.9 and equal to 0 at the same time. Once again we have encountered a logically valid argument that leads to absurdity, given only premises that look plausible if taken just by themselves.

How do the premises imply the conclusion logically?

First of all:

(Slide 104)

The two premises

(P2)

$$P(\text{ticket 1 wins}) = 1/1000000$$

:

$$P(\text{ticket 1000000 wins}) = 1/1000000$$

(P5) For every proposition X :

$$P(\neg X) = 1 - P(X).$$

taken together entail:

for each i , $P(\text{not ticket } i \text{ wins}) = 999999/1000000$.

The probability of any particular ticket not winning is huge.

Secondly:

(Slide 105)

Therefore, by premise

(P3) For every proposition X :

I believe X if and only if $P(X) \geq 0.9$.

it follows:

I believe that not ticket 1 wins.

:

I believe that not ticket 1000000 wins.

(Slide 106)

So, by applying premise

(P4) For every proposition X , for every proposition Y :

if I believe X and I believe Y , then (if I am perfectly rational) I also believe $X \wedge Y$.

multiple times, it follows:

- I believe that (not ticket 1 wins \wedge not ticket 2 wins $\wedge \dots \wedge$ not ticket 1000000 wins).

I have the corresponding conjunctive belief, since my beliefs are rationally closed under conjunction, under the logical ‘and’ or the set-theoretic intersection operation.

But by applying P3 again, this time applying in the left-to-right direction, this would mean that the probability that ticket 1 does not win and ticket 2 does not win and ... and ticket 1000000 does not win is greater than or equal to the threshold, in this case, 0.9. However, this contradicts P1, which says that the probability that no ticket will win is 0. So we end up with a contradictory conclusion.

(Slide 107)

- I believe that (not ticket 1 wins \wedge not ticket 2 wins $\wedge \dots \wedge$ not ticket 1000000 wins).

with the two premises

(P1)

$$P(\text{not ticket 1 wins} \wedge \text{not ticket 2 wins} \wedge \dots \wedge \text{not ticket 1000000 wins}) = 0.$$

(P3) For every proposition X :

I believe X if and only if $P(X) \geq 0.9$.

leads to the contradiction C.

Which premise is the bad guy?

(Slide 108)

(P1) $P(\text{not ticket 1 wins} \wedge \text{not ticket 2 wins} \wedge \dots \wedge \text{not ticket 1000000 wins}) = 0$.

(P2)

$$P(\text{ticket 1 wins}) = 1/1000000$$

⋮

$$P(\text{ticket 1000000 wins}) = 1/1000000$$

(P3) For every proposition X :

I believe X if and only if $P(X) \geq 0.9$.

(P4) For every proposition X , for every proposition Y :

if I believe X and I believe Y , then I believe $X \wedge Y$.

(P5) For every proposition X :

$$P(\neg X) = 1 - P(X).$$

P1, P2, and P5 look unproblematic, as long as one buys subjective probabilities at all. So we should concentrate on P3 and P4: Maybe P3 is wrong – maybe it is not the case that one believes rationally a proposition if and only if one's rational degree of belief in that proposition is high enough? Or perhaps that is fine, in principle, it is just that the threshold in this case ought to be 1: one believes all and only what has the maximal degree of belief of 1? But that cannot be right generally, at least, since it does seem to be the case that I believe the weather to be fine tomorrow and also believe that proposition with the same strength with which I believe that $2 + 2 = 4$. So what makes me determine the threshold to be 1 in the lottery case but less than 1 in the weather case? Or is really P4 the culprit: perhaps rationality does not commit inferentially perfectly rational persons to close their beliefs under conjunction? But what kind of attitude is all-or-nothing belief then, if believing that the actual world is in X , and believing that the actual world is in Y , does not commit us rationally to believing that the actual is both in X and in Y , that is, in $X \cap Y$? Or should we have simply thrown away the whole concept of all-or-nothing belief from the start, at the very moment in which we discovered the merits of the concept of degree of belief? But wouldn't that move seem a bit drastic, as in this way we would be forced to throw great junks of more than 2000 years of traditional epistemology into the garbage can – of one of the central areas of philosophy in which the concept of rational all-or-nothing belief has traditionally played a major role.

We will not, and cannot, answer these questions here. The questions are hotly debated, as we speak, in an area of philosophy that is called formal epistemology and in which mathematical methods are applied to epistemological questions and problems concerning knowledge, belief, justification, rationality, and reasoning – problems and questions such as the ones that we were dealing with in this lecture. At this point it is very much an open question how rational all-or-nothing belief and rational degrees of belief relate to each other. Perhaps you will find the right answer?

Quiz 29:

Think about it yourself – which of the premises of the Lottery Paradox argument would you suggest to give up?

Remark on the Lottery Paradox: Here you can read more about the Lottery Paradox:

<http://plato.stanford.edu/entries/epistemic-paradoxes/#LotLotPar>.

There is a closely related paradox, called the ‘Preface Paradox’, which you might also like to check out:

<http://plato.stanford.edu/entries/epistemic-paradoxes/#PrePar>.

Finally, at

<http://plato.stanford.edu/entries/formal-belief/>

you will find more about how belief can be represented formally, whether in terms of subjective probabilities of propositions, all-or-nothing (qualitative) beliefs in propositions, or in some other way.

3.15 Conclusions (07:51)

In today’s lecture we first developed an account of propositions as sets of possible worlds: we were able to derive some plausible principles on propositions from that account together with the principles of set theory, and we reconstructed various important concepts concerning propositions, such as negation, conjunction, disjunction, and logical consequence for propositions in set-theoretic terms. Then we analyzed the rationality of belief by means of some principles of rational belief, we found that these principles entailed a reduction of complexity vis-a-vis believed propositions, and we saw that the formal structures that emerged from the principles are actually well-known also from a part of modern mathematics, from algebra. Finally, we did the same for the rationality of degrees of belief, we postulated rational degrees of belief to be probabilities, and we sketched a justification of this from the rationality of betting behaviour. In various of these parts, we drew conclusions on rational belief or degrees of belief my means of mathematical theorems. Finally, we discussed the lottery paradox which led to the open question whether, and if so, how, belief and degrees of belief can exist rationally side by side.

As usual, let me conclude this lecture by addressing some worries that you might have.

Firstly, you might be a bit disappointed by our conception of propositions as sets of possible worlds: “I still do not know what a proposition is”, you might say, “because you did not tell me what exactly a possible world is”. If you think so, then you are absolutely right. Once we have determined what a possible world is, what properties and relations apply to possible worlds and which do not, then our account of propositions will be complete, in some sense: propositions are sets of possible worlds, and if we understand possible worlds to a reasonable extent, then, since we already understand sets to a reasonable extent, we would also understand propositions to a reasonable extent. But I did not give you that first part, a theory of possible worlds. The task of supplying us with such a theory of possible worlds belongs to metaphysics, traditionally, another one of the classic areas of philosophy.

Once the metaphysicians have figured out what possible worlds are, we can simply use what they give us as an input to our set-theoretic understanding of propositions and get as an output a rich theory of propositions. Unfortunately, there is no widely accepted theory of possible worlds around so far in metaphysics: there are just various such theories. Each of these theories could be fed into what we said above, and we would end up with propositions as sets of possible worlds in the sense of that particular theory of possible worlds. For the time being, we simply have to live with the fact that propositions in our sense are only determined up to the notion of possible world on which that account of propositions is based.

But let me give this a slightly more positive twist: our initial principles ($\text{Id}_{\text{Worlds}}$) and (Id_{Prop}), the concepts of the negation of a proposition and of the conjunction and disjunction of propositions, the relationship of logical implication between propositions – all of that only concerned the formal structure of propositions. Our account of propositions as sets of possible worlds constituted one way of delivering that intended formal structure in a particularly nice way. Many metaphysicians, who do argue with each other about what possible worlds are, would still subscribe to that formal structure: so whatever conclusions we were able to draw about propositions, and about attitudes towards propositions, as long as they were derived just from that formal structure, these metaphysicians should be willing to embrace them. By proceeding just on the presumption that propositions are sets of possible worlds, while leaving open the exact nature of possible worlds, we gained a lot of generality, and we should have managed not to lose too many of the philosophers who work on the topic of possible worlds, even when they do not agree with each other on the metaphysical nature of possible worlds, and in fact even when they never will. Reaching that level of generality is another advantage of applying mathematical, that is, structural methods, in philosophy.

A second issue: we have talked a lot about perfectly rational persons in this lecture. But that is not us: none of us is perfectly rational. So why should we be interested at all in this ethereal beings? The answer is that, amongst others, philosophy is a normative business: philosophers want to determine what we should aspire to. This does not mean that we should not work out also in what way we, real-world flawed persons, can at least approximate perfectly rational persons, and what means we can take in order to make sure that we do. Today I was only concerned with the goal, and not so much with how we can reach that goal. But both topics are interesting and important.

Thirdly, and obviously, I have put a lot of important aspects of the rationality of belief and degrees of belief to one side in this lecture. Basically, we were only dealing with inferential rationality here: the rationality of belief or degrees of belief in propositions that result from applying logical operations to other propositions – propositions that might be believed as well or which in any case have certain degrees of belief themselves. As interesting as that might be, obviously this does not exhaust the topic of rationality for belief: not by far. But

that had not been the claim anyway: this is just a start. There is more to come on this matter, even within the tight boundaries of these introductory lectures on mathematical philosophy. Other topics, unfortunately, we will not be able to cover within our lectures at all. For instance: how about rational belief about rational belief? Sometimes I believe a proposition, and I also believe that I believe that proposition. Or I do not believe a proposition, and I believe that I do not believe it. Is this always so, if one is perfectly rational? What are the laws of rational belief in belief, of rational introspection? This is an important topic not just in epistemology but also in the areas of doxastic and epistemic logic: the logics of belief and knowledge. In these areas it is often more helpful again to deal with sentences or formulas rather than with propositions or sets. Formal methods play a crucial role again, but the methods belong much more to philosophical logic than to algebra. Similarly, by the way, one can also assign degrees of belief, subjective probabilities, to sentences or formulas instead of propositions; as you will see, Stephan will be doing so in his lectures. Sometimes this is preferable, but most of the time it does not matter that much.

I believe that this is the end of Lecture 3. And I also assign a high degree of belief to the proposition that this is the end. No Lottery Paradox shall stop me from doing so.

Here are some references to the relevant background literature:

On modal logic:

Hughes, G. and Cresswell, M., *A New Introduction to Modal Logic*, London: Routledge, 1996.

Blackburn, P., de Rijke, M. and Venema, Y., *Modal Logic*, Cambridge: Cambridge University Press, 2001. (This book is quite advanced mathematically!)

On doxastic and epistemic logic:

Hintikka, J., Knowledge and Belief. *An Introduction to the Logic of the Two Notions*, Ithaca, NY: Cornell University Press, 1962. The book was reprinted again more recently as: Hintikka, J., *Knowledge and Belief. An Introduction to the Logic of the Two Notions*, prepared by V.F. Hendricks & J. Symons, London: King's College Publications, 2005.

Hendricks, V. F., *Mainstream and Formal Epistemology*, New York: Cambridge University Press, 2006.

Van Ditmarsch, H., van der Hoek, W. and Kooi, B., *Dynamic Epistemic Logic*, Synthese Library Vol. 337, 2007.

On subjective probability theory/Bayesianism:

Skyrms, B., *Choice and chance*, 4th edition, Belmont, CA: Wadsworth/Thompson, 2000.

Jeffrey, R., *Subjective Probability. The Real Thing*, Cambridge: Cambridge University Press, 2004.

Howson, C. and Urbach, P., *Scientific Reasoning: The Bayesian Approach*, second edition, La Salle, Illinois: Open Court, 1993.

Earman, J., *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, Mass.: The MIT Press, 1992.

On all-or-nothing belief vs degrees of belief:

Kyburg, H.E.Jr., *Probability and the Logic of Rational Belief*, Middletown, CT: Wesleyan University Press, 1961.

Levi, I., *Gambling With Truth. An Essay on Induction and the Aims of Science*, New York: Knopf, 1967.

Makinson, D., "The Paradox of the Preface", *Analysis* 25 (1965), 205?207.

Finally, concerning the view on the topic from theoretical computer science:

Halpern, J.Y., *Reasoning about Uncertainty*. Cambridge, Mass.: The MIT Press, 2003.

Appendix A

Quiz Solutions Week 3: Rational Belief

Quiz 19:

Take our principles (Id_{Worlds}) and (Id_{Prop}) from the lecture as given.

(1): Assume that a proposition X is true at world w but it is not true at world w' . Can we determine from this whether w is identical to/distinct from w' ?

(2): Assume that propositions X and Y are true at precisely the same possible worlds. Can we determine from this whether X is identical to/distinct from Y ?

SOLUTION Quiz 19:

Solution (1): Yes – w must be distinct from w' . In fact, we do not actually need (Id_{Worlds}) and (Id_{Prop}) for this: the logical principle of the indiscernibility of identicals is sufficient. By the indiscernibility of identicals, if w were identical to w' , then the two of them would need to have the same properties. But w has a property that w' does not have, as X is true in w but not in w' , which is why w must be different from w' .

Solution (2): Yes – X must be identical to Y . In order to derive this, one needs to rely on our principle (Id_{Prop}) .

[Back to quiz](#)

Quiz 20:

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 ; finally, determine (and color) graphically (i) the proposition $W \setminus (X \cup Y)$, and also (ii) the proposition $(W \setminus X) \cap (W \setminus Y)$. What does the graphical representation in (i) look like in comparison with that in (ii)?

SOLUTION Quiz 20:

They look the same. The proposition $W \setminus (X \cup Y)$ is identical to the proposition $(W \setminus X) \cap (W \setminus Y)$ (see Figure A.1). The corresponding law $W \setminus (X \cup Y) = (W \setminus X) \cap (W \setminus Y)$ (or in logical terms: the negation of the disjunction of X with Y is logically equivalent to the conjunction of the negation of X with the negation of Y) is the set-theoretic version of one of the two so-called De Morgan laws, which are named after the British mathematician Augustus De Morgan.

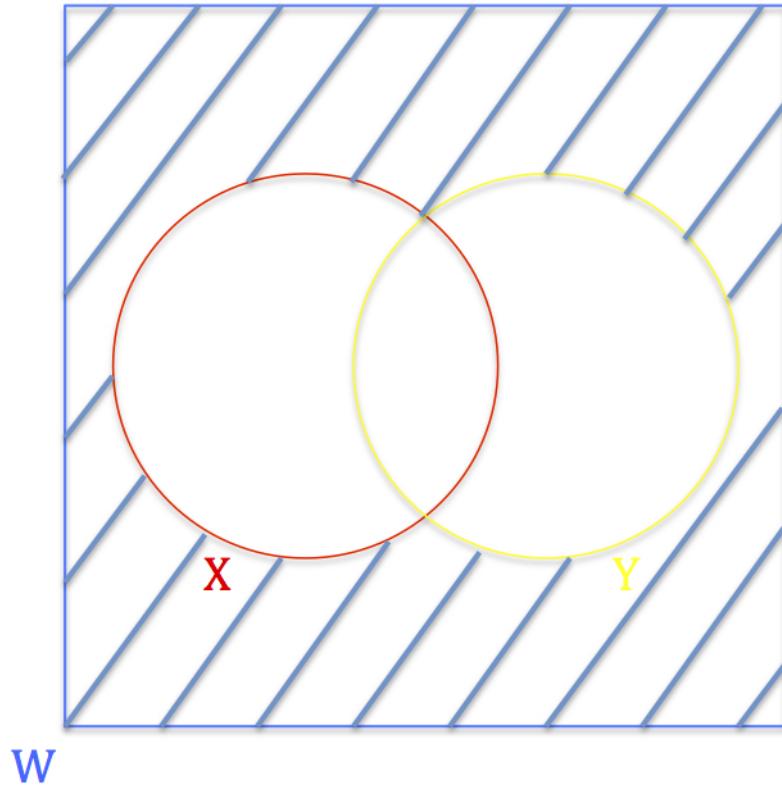


Figure A.1: $W \setminus (X \cup Y) = (W \setminus X) \cap (W \setminus Y)$

[Back to quiz](#)

Quiz 21:

Just as in the lecture, let $W = \{w_1, \dots, w_8\}$, $X = \{w_1, w_2, w_4, w_5\}$, $Y = \{w_2, w_3, w_5, w_6\}$, $Z = \{w_4, w_5, w_6, w_7\}$.

- (1): Determine $(X \cap Y) \cap \neg Z$.
- (2): Determine $\neg(X \cap Z) \cup \neg\neg Y$

SOLUTION Quiz 21:

Solution (1): $\{w_2\}$.

Solution (2): $\{w_1, w_2, w_3, w_5, w_6, w_7, w_8\} (= W \setminus \{w_4\})$.

[Back to quiz](#)

Quiz 22:

Show that if X is a subset of Y and X is a subset of Z , then X is also a subset of $Y \cap Z$.

SOLUTION Quiz 22:

Suppose that X is a subset of Y and X is also a subset of Z . By the definition of ‘subset of’, we need to prove that every member of X is also a member of $Y \cap Z$.

So let x be an arbitrary member of X . Since X is a subset of Y , it follows that x is also a member of Y . Accordingly, since X is a subset of Z , it also follows that x is a member of Z . So x is a member of Y and a member of Z . This implies that x is a member of $Y \cap Z$ (since that intersection is precisely the set of all objects that are members of Y and members of Z). But that is what we needed to show.

[Back to quiz](#)

Quiz 23:

Assume our postulates Rational Belief 1-4.

Show that the following is the case: If a person is inferentially perfectly rational, then if she believes a proposition X , she does not also believe the negation of X .

SOLUTION Quiz 23:

Suppose a person is inferentially perfectly rational, and suppose furthermore that she believes X . Now assume for reductio that the person also believes $\neg X$: By postulate Rational Belief 4, she believes then also $X \cap \neg X$. But $X \cap \neg X$ is identical to the empty set, which means that the person believes the empty set. But that contradicts postulate Rational Belief 2. Hence, by reductio, the person does not believe $\neg X$, which is what we wanted to show.

[Back to quiz](#)

Quiz 24:

Our theorem says (I suppress some details, such as concerning p being inferentially perfectly rational and W being the set of all worlds): if Rational Belief postulates 1-4 hold, then there is a non-empty proposition B_W , such that for all propositions X , person p believes X if and only if B_W is a subset of X .

But actually one can also prove the converse of this: if there is a non-empty proposition B_W , such that for all propositions X , person p believes X if and only if B_W is a subset of X , then Rational Belief postulates 1-4 hold. Please prove this converse statement.

SOLUTION Quiz 24:

Assume that there is a non-empty proposition B_W , such that for all propositions X , person p believes X if and only if B_W is a subset of X .

Now we derive from this our postulates Rational Belief 1-4:

Ad Rational Belief 1: B_W is a subset of W , which is why p indeed believes W , by the assumption above.

Ad Rational Belief 2: B_W is non-empty, by assumption, which is why B_W is not a subset of $\{\}$. Therefore, p does not believe $\{\}$, by the assumption above again.

Ad Rational Belief 3: Suppose p believes X , and X is a subset of Y . Since p believes X , B_W must be a subset of X , by the assumption above. But because X is a subset of Y , it follows that B_W is also a subset of Y , which entails that p also believes Y , by the assumption from above again.

Ad Rational Belief 4: Suppose p believes X , and p believes Y . By the assumption from above, this implies that B_W is a subset of X , and B_W is a subset of Y . But that entails (as we have seen in a previous quiz) that B_W is also a subset of $X \cap Y$. By the assumption from above, therefore, p also believes $X \cap Y$.

Done!

[Back to quiz](#)

Quiz 25:

(1): Let $W = \{w_1, \dots, w_8\}$ and take $B_W = \{w_2, w_5\}$ to be the least proposition again that is believed by an inferentially perfectly rational person p . For each of the following three propositions, please determine whether it is believed by p , or whether it is disbelieved by p (that is, its negation is believed), or whether p suspends judgement on it (that is, p neither believes it nor disbelieves it):

- (i) $\{w_1, w_5, w_7\}$,
- (ii) $\{w_1, w_3\}$,
- (iii) $\{w_1, w_2, w_4, w_5\}$.

(2): How many propositions (that is, subsets of $\{w_1, \dots, w_8\}$) does p believe, given that $\{w_2, w_5\}$ is a subset of all, and only, believed propositions?

SOLUTION Quiz 25:

Solution (1):

- (i) p suspends judgement on $\{w_1, w_5, w_7\}$,
- (ii) p disbelieves $\{w_1, w_3\}$,
- (iii) p believes $\{w_1, w_2, w_4, w_5\}$.

Solution (2): p believes $2^6 = 64$ propositions. Here is why: w_2 and w_5 are members of every believed proposition. w_1 is either a member of a believed proposition or it is not (2 possibilities), similarly w_3 is either a member of a believed proposition or it is not (2 possibilities), w_4 is either a member of a believed proposition or it is not (2 possibilities), w_6 is either a member of a believed proposition or it is not (2 possibilities), w_7 is either a member of a believed proposition or it is not (2 possibilities), and w_8 is either a member of a believed proposition or it is not (2 possibilities). In total this yields $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 64$ possibilities of constructing a superset of $\{w_2, w_5\}$.

[Back to quiz](#)

Quiz 27:

Assume our Rational Degree of Belief postulates 1-3 to be the case (for a given set W of possible worlds). Let P be the degree of belief function of an inferentially perfectly rational person.

Show that for all propositions X, Y : $P(X \cap Y)$ is less than or equal to $P(X)$.

SOLUTION Quiz 27:

We know already that the set $X \cap Y$ is a subset of X . And we have derived before from our Rational Degree of Belief postulates 1-3 that for all propositions X, Y , if X is a subset of Y , then $P(X)$ is less than or equal to $P(Y)$. Putting the two together (where now $X \cap Y$ is the relevant subset of X), it follows that $P(X \cap Y)$ is less than or equal to $P(X)$.

[Back to quiz](#)

Quiz 28:

Just as in the lecture, assume that an inferentially perfectly rational person's degree of belief function P is determined by the following function B (in the manner explained by the last theorem):

$$B(w_1) = 1/15$$

$$B(w_2) = 1/3$$

$$B(w_3) = 1/15$$

$$B(w_4) = 1/15$$

$$B(w_5) = 1/3$$

$$B(w_6) = 1/15$$

$$B(w_7) = 1/15$$

$$B(w_8) = 0$$

Determine the following degrees of belief:

- (i) $P(\{w_8\})$,
- (ii) $P(\{w_4, w_7, w_8\})$,
- (iii) $P(\{w_1, w_2, w_3, w_5, w_6\})$.

SOLUTION Quiz 28:

(i) 0.

(ii) $2/15$.

(iii) $13/15$ ($= 3/15 + 2/3 = 3/15 + 10/15$; or using (ii), as $\{w_4, w_7, w_8\}$ is the complement of $\{w_1, w_2, w_3, w_5, w_6\}$: $13/15 = 1 - 2/15$).

[Back to quiz](#)

Quiz 29:

Say, someone believes $X \wedge \neg X$ with a degree of 0.3. (Obviously, this is not in line with the axioms of subjective probability, as they demand believing the contradictory proposition $X \wedge \neg X$ with the minimal degree of 0.) Can you offer a bet to that person so that the person is guaranteed to lose money?

SOLUTION Quiz 29:

Yes. Offer the following bet on $X \wedge \neg X$ to the person: We know that the person's degree of belief $P(X \wedge \neg X)$ in $X \wedge \neg X$ is 0.3. Let $S_{X \wedge \neg X} = 1$ Euro, $q_{X \wedge \neg X} = 0.3$. The person will regard this bet on $X \wedge \neg X$ as fair, since the betting quotient equals the person's degree of belief in $X \wedge \neg X$. If $X \wedge \neg X$ is true, then the person's net profit will be $S_{X \wedge \neg X} - q_{X \wedge \neg X} \cdot S_{X \wedge \neg X} = 1 - q_{X \wedge \neg X} = 0.7$ Euro. If the negation of $X \wedge \neg X$ is true, then the person's net loss will be $q_{X \wedge \neg X} \cdot S_{X \wedge \neg X} = q_{X \wedge \neg X} = 0.3$ Euro. But of course $X \wedge \neg X$ is false (indeed, logically false), which is why the negation of $X \wedge \neg X$ is true come what may; therefore, the person is guaranteed to lose 0.3 Euro.

[Back to quiz](#)

List of Figures

3.1	X is a subset of W and w is a member of X	8
3.2	$\neg X$	11
3.3	$X \cap Y$	13
3.4	$X \vee Y$	14
3.5	Distribution of w_1, \dots, w_8	17
3.6	$\neg \neg X$	19
3.7	$X \wedge Y$	21
3.8	$X \vee \neg Y$	22
3.9	$\neg X \wedge Z$	23
3.10	$(X \wedge Y) \wedge Z$	24
3.11	$\neg X \wedge \neg Y \wedge \neg Z$	25
3.12	$W = X \vee \neg X$	26
3.13	$X \wedge \neg X = \{\}$	27
3.14	X	30
3.15	$X \cup Y$	31
3.16	$X \cap Y \subseteq X$	33
3.17	$B_W = X \cap Y$	46
3.18	Distribution of w_1, \dots, w_8	52
3.19	The side length of W equals 1	53
3.20	Propositions as squares	54
3.21	X , Y , and Z as squares	55
3.22	X , Y , and Z as squares of different size	56
3.23	$X \wedge Y$ does not overlap with $\neg X \wedge Y$	58
3.24	Values of B	66
A.1	$W \setminus (X \cup Y) = (W \setminus X) \cap (W \setminus Y)$	86