Ludwig-Maximilians-Universität München

Munich Center for Mathematical Philosophy

and

Mathematisches Institut

Probabilistic Conditionals

Bachelor Thesis in **Mathematics**

Author: Alexandra Surdina
Advisors: Prof. DDr. Hannes Leitgeb (MCMP),
Prof. Dr. Hans-Dieter Donder (Mathematisches Institut)
Date: August 05, 2013

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisor, Hannes Leitgeb, for his guidance, patience and his infectious enthusiasm. He agreed to supervise a Bachelor's thesis in mathematics although his professorship is at the Faculty for Philosophy, and has always found the time for questions when needed. Many times, I entered his office with some problem and each time I left not only with some idea how to solve the problem I brought with me, but also with plenty of others, which I hope to get an opportunity to pursue in the future.

I would also like to thank Hans-Dieter Donder for offering a fascinating introduction into classical logic in his lectures, and for giving me the possibility to write a thesis at the Faculty for Philosophy by providing an unbureaucratic connection to the Mathematical Institute.

Last but not least, thanks to my friends for patiently listening to me whenever the party we attended had the misfortune of taking place somewhere with whiteboards. I would particularly like to thank Tosca Lechner, Linus Mattauch, Jost Migenda and Georg Sauerwein for critical questions and helpful remarks.

Contents

Chapter 1

Introduction

Take the following sentence:

> 'If you read this thesis completely, you will invent time-travel.'

If you are asked whether you would believe this statement or not, your reply will most likely be 'no'. For one thing, time travel is most likely physically impossible or at least technologically very far away; besides, there is no evident causal relationship between the two parts of the sentence.
A candidate for a formal equivalent to the sentence above is a statement of the form

$$p \to q\,.$$

In the expression $p \to q$, p is called the antecedent and q is called the consequent. The expression is read as 'if p then q' and understood as 'if p is true, then q is also true', i.e. $p \to q$ is false precisely if p is true but q is false. Now, take p to mean 'you read this thesis completely' and take q to mean 'you will invent time-travel'. If you are now asked whether you would consider the implication $p \to q$ to be true, your reply will differ depending on whether you read this thesis completely or not: in fact, you can *make* the implication true by *not reading* a part of this paper.
The logical connective $\to$ is called the *material conditional* and you just encountered one of its peculiar properties: if the antecedent is false, the conditional statement is always true. This is a property which goes against human understanding of implications in English sentences. It is one of many questionable features of the material conditional.

This thesis deals with a possible answer to the question: how can we define a conditional differently in order to avoid such features? Ramsey (1931) suggested that we evaluate a conditional statement by adding the antecedent hypothetically to our knowledge of the world and then asking ourselves, whether the consequent is true. Conditionalising on the antecedent in probability theory provides a formal framework for evaluating the consequent given the antecedent. The present work thus explores attempts to define conditionals probabilistically.
Two attempts to capture Ramsey's idea in formal logic will be presented: the first is the definition of a conditional by Adams (1966), that takes conditional probabilities of the consequent given the antecedent as a measure for acceptance of a conditional (and consecutively defines a probability of a conditional by conditional probabilities). The second one is formulated in terms of possible worlds (Stalnaker, 1980) and formalised into a logical system C2, which will later turn out to be equivalent to a special case of probabilistic conditionals.

Therefore, probabilistic conditionals constitute an alternative candidate of formalising conditional statements in Ramsey's sense and avoiding the paradoxes of material implication. One motivation for a *probabilistic* theory of conditionals is the idea to bring together the elaborate formal framework of probability theory, with a still controversial interpretation, and the area of conditional statements, where no formal theory (or too many formal theories) has been established yet.

The core of this thesis consists in an examination of the compelling hypothesis that probabilities of conditionals are not merely defined by conditional probabilities but can be considered as *elements* of the probability space themselves. It is compelling because this way, it would be possible to treat conditionals as propositions, e.g. calculate the probability of a conditional given a proposition or even *another conditional*, which is not possible in Adams' framework. This hypothesis is followed by some impossibility results by David Lewis (Lewis, 1976, 1986), and Alan Hájek and Ned Hall (Hájek, 1994, Hájek & Hall, 1994) that pose strong evidence against it.

This Bachelor's thesis is primarily a work in mathematics, not in philosophy. This implies two things:
First, it focuses on formal arguments rather than interpretation. There is a lot more to say on the semantic aspects of Adams and Stalnaker conditionals, or why the hypothesis about probabilistic conditionals is on the one

hand so compelling and, on the other hand, which (non-formal) arguments speak against it. To embrace even just the central part of these ‘softer’ results would go beyond the scope of this thesis. A more thorough account from the perspective of philosophy can be found in Alan Hájek’s PhD dissertation (Hájek, 1993).
Second, in contrast to theses in philosophy, this thesis does not defend and culminate in a particular position (which is common in philosophy) but rather presents an overview over the development of a collection of results from different authors (which is common in mathematics).

The thesis is structured as follows:

In Chapter 2, I will portray some situations that are problematic about the material conditional as a possible formalisation of conditional statements. Additionally, I will describe Grice’s theory of implicatures in defence of the material conditional and explain why it is insufficient to solve all the arising paradoxes.

In Chapter 3, I will provide the formal framework to be able to speak of the probability of a sentence given another. To do that, I will describe the axioms of probability theory and introduce the notion of the probability of sentences in a formal language. A reader familiar with these concepts may skip this chapter.

The main part of this thesis consists of Chapters 4 and 5. In Chapter 4, I will present an explanation of why a probabilistic theory of conditionals is appealing, based on Ramsey’s idea how truth values of conditionals are usually evaluated, and present two attempts to capture this idea, namely the conditionals of Adams and Stalnaker. In Chapter 5, based on the work of Lewis, Hájek and Hall, I shall present the central hypothesis regarding the probability of a conditional – ‘probabilities of conditionals are conditional probabilities’ – and arguments disproving it.

Chapter 2

The Material Conditional and its Problems

This chapter motivates introductions of alternative definitions of conditionals by highlighting some of the problems of the material conditional.
As we will see in the first part of the chapter, material implication as a candidate for a formalisation of implications of conversational sentences (i.e. English sentences of the form 'if something is the case, then something else is the case') leads to sentences that do not agree with out intuitions regarding conversational statements – such as the fact that from inconsistent premises anything follows.
In the second part of the chapter, an additional theory is introduced in an attempt to amend those conditional sentences that are true but counterintuitive, namely Grice's theory of *implicatures* that aims to determine which true material implications are conversationally appropriate and which are not.

2.1 Paradoxes of Material Implication

The material implication (sometimes called *material conditional*), here denoted by '$\to$', is a logical connective which, in classical logic, has the following truth value function:

p	q	$\neg p$	$\neg q$	$p \to q$	$\neg q \to \neg p$
T	T	F	F	T	T
F	T	T	F	T	T
T	F	F	T	F	F
F	F	T	T	T	T

The conditional $p \to q$ is equivalent to $\neg q \to \neg p$, which one can see from the fact that they have the same truth table.

Principle of Explosion

One problem of the material conditional is the principle of explosion which states that from inconsistent premises, *anything* follows:

$$\models ((\neg A \wedge A) \to B) \tag{2.1}$$

This property is counterintuitive because A and B do not need to be 'causally' related in any way. Consider the implication from A and $\neg A$ to B in the following example:

(A) *The sun is shining.*
($\neg A$) *The sun is not shining.*
(B) *Paul Erdös just turned into a zombie.*

According to the truth table, this is a valid consequence which, even though many mathematicians may wish for it to be true, few human beings would intuitively accept.

Further Paradoxes

Some other paradoxes that arise from the truth table of the material implication are

$$p \to (q \to p)\,, \tag{2.2}$$

$$p \to (q \vee \neg q)\,, \tag{2.3}$$

$$(p \to q) \vee (q \to r)\,. \tag{2.4}$$

The Expression (2.2) is true since if p is true, $q \to p$ is also true for all q. The Implication (2.3) is true since $q \vee \neg q$ is true for all q so the validity follows from (2.2). Finally, (2.4) is true since $q \vee \neg q$ is true; if q, then $p \to q$ for all p, and if $\neg q$, then $q \to r$ yields $(\neg q \wedge q) \to r$. This is an implication with a contradiction in the antecedent and true according to the principle of explosion mentioned before.

We now illustrate the paradoxical nature by three examples. For this purpose, consider a version of the duck test:

> If it looks like a duck, swims like a duck, and quacks like a duck, then it is probably a duck.

The Implication (2.2) allows for the following statement:

> Thomas Edison invented the light bulb. Therefore, if it looks like a duck, swims like a duck, and quacks like a duck, then Thomas Edison invented the light bulb.

This constitutes an example to (2.2) by taking p to be the sentence 'Thomas Edison invented the light bulb', and q to be 'it looks like a duck, swims like a duck, and quacks like a duck'. It is true because the consequent is true, Thomas Edison *did* invent the light bulb. It is also paradoxical because the implication does not represent a causal connection we would accept.

Statement (2.3) yields the following sentence in 'duck terminology':

> If it looks like a duck, then either it rains or it does not rain.

Since in (2.4) r can be p, the fact that this implication is true results in the truth of the following example sentence:

> If it is a duck, then it is a cat – or if it is a cat, then it is a duck.

This is nonsensical as in this case p and q are not only not causally related but even contradict each other.

2.2 Intuitiveness, Implications and Implicatures

In the main part of this thesis, probabilistic conditionals are presented as a way to avoid the undesirable properties of material implication. The purpose of this section[1] is to present H.P. Grice's theory of implicatures, pointing out

[1] This section is based on Davis (2013).

that there are cases where correct material implications are not conversationally appropriate; since our intuitions refer to implications in conversations, the unintuitiveness of some properties of the material conditional could be resolved this way. At the end of the section, I will explain why Grice's account is not adequate for a formal theory of 'allowed' conditionals.

In interpersonal communication, a speaker might imply something different from the meaning a sentence itself conveys. The systematic study of such cases was first undertaken by Grice (1975). Grice distinguished between *conversational* implicatures and *conventional* implicatures.

A conversational implicature relates not only to the uttered words but also to the context. Consider this example of a conversation:

Alice:	Are you going to Eve's party?
Bob:	I have to work.

The sentence Bob used does not *mean* that Bob is not going to the party and yet this is the message conveyed by the speaker due to the context (i.e. Alice expecting an answer to her question).
A conventional implicature on the other hand is part of the sentence's meaning, for instance in the following case:

Bob:	(a) It's a Monday; I have to work.
	(b) It being a Monday implies that I have to work.

Besides this categorisation of implicatures, Grice developed a theory for the description and explanation of conversational implicatures. According to this theory, human communication is based on a general principle of cooperation:

> *Cooperative Principle:* Contribute what is required by the accepted purpose of the conversation.

Grice also states several specifications of this principle, such as the rule: *be as informative as required* (by the purpose of the conversation). In the following conversation, this rule has an immediate influence on the meaning of a sentence:

Arthur:	Hello! Where's Alice?
Merlin:	She is either in the cafeteria or in Bob's office.

Due to the rule, it is implied that Merlin does not know where exactly Alice is, for then his answer would either be:

(C) She is in the cafeteria, or
(O) She is in Bob's office.

Hence this results in the fact that if C is the case[2], the implicature $C \vee O$ additionally provides the information that he does not know where she is. For if the speaker knew that C is the case, she should not assert $C \vee O$.
If we believe that Alice is in the cafeteria *we also believe that Alice is in Bob's office or the cafeteria*; similarly, we believe that if Arthur just asked a question, then he is a human being, probably male, which we also would not state since the information 'Arthur just asked a question' is sufficient; adding statements of the other beliefs would not make the description of the situation any more informative.

The main difficulty is that Grice's theory seems to be too weak: for almost each implicature, it seems that the rule used to produce it can be also used to produce inexistent implicatures (Davis, 2013). Take the following example sentence:

'John met a woman.'

An implicature of this is:

'John did not meet his wife.'

For otherwise the speaker would have been less informative. But analogously, we can go from the statement

'John broke a finger.'

to the next, stronger statement:

'John did not break his own finger.'

This is clearly not what was the sentence about John breaking his finger was intended to mean.

Grice's theory explains correct examples of the nature of human communication very well. But although his strategy to determine implicatures –

[2] And therefore, C or O is the case.

that is, asking: does the statement obey the Cooperative Principle? – delivers correct implicatures, the criteria derived from the Cooperative Principle can explain false implicatures (which carry other meanings than the intended one) just as well. Thus, the logic of material implication in combination with Grice's theory of implicatures still remains unsatisfactory.

Chapter 3

Axioms and Basics of Probability Theory

The previous chapter presented some of the ways in which the material implication is inadequate. Now, in order to be able to define conditionals in a probabilistic way, it is necessary to be able to speak of such things as 'the probability that the sky is blue', or 'the probability that the reader is going to wake up tomorrow and find herself in the year 1742'.
To do so, a basic version of a formal system of probability theory[1] is introduced in this chapter. In particular, we will define the notion of probabilities in the context of formal languages.
The last part[2] of the chapter provides a brief introduction into different ways to grasp the concept of probability and hereby establishes the context of subjective probability, in which the rest of the thesis is set.

3.1 Probability of Subsets of a Set

Definition. Let $\Omega \neq \varnothing$ be a set. An *algebra* on Ω is a set $F \subseteq \mathcal{P}(\Omega)$ of subsets of Ω with the following properties:

(I) $\Omega \in F$.

(II) $A \in F \;\Rightarrow\; \Omega \backslash A \in F$ (closed under complement with respect to Ω).

(III) $B, C \in F \;\Rightarrow\; B \cup C \in F$ (closed under finite unions).

[1]Based on Klenke (2008).
[2]See Hájek (2012).

Definition. Let $P : F \to \mathbb{R}$ be a function obeying

(A) $P(A) \geq 0$ for all $A \in F$ (non-negativity),

(B) $P(\Omega) = 1$ (normalisation),

(C) $P(A \cup B) = P(A) + P(B)$ for all $A, B \in F$ such that $A \cap B = \varnothing$ (finite additivity).[3]

Then (Ω, F, P) is called a *probability space* and P a *probability function.* Elements of F are called *events.*

Conditional Probability

We conclude Section 3.1 with the definition of conditional probability and a review of Bayes' theorem for reference in future chapters.

Definition (Conditional Probability)**.** Let (Ω, F, P) is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}\,.$$

Theorem (Bayes' Theorem)**.** *Let P be a probability function on a probability space (Ω, F, P) and let $P(A) > 0$. Then,*

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}\,.$$

Proof.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Therefore,

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)\,.$$

Dividing by $P(A)$, the result immediately follows. □

[3]Sometimes closure under *countable* instead of just finite unions is required for F; subsequently, *countable additivity* (also called *sigma additivity*) is required instead of (C): $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ if the A_i are pairwise disjunct. This is a stronger statement which is only necessary for more technical cases, such as proving the continuity of the probability function P for decreasing sequences of sets.

3.2 Formal Languages

3.2.1 The Language of Propositional Logic

Definition. A *formal language* consists of a set of symbols (the *alphabet*) and a set of rules on how to form strings of symbols (the *grammar*).

Definition. The *language of propositional logic*[4] consists of:

(i) Infinitely many propositional variables: $\{a_i | i \in \mathbb{N}\}$

(ii) Symbols for the propositional connectives[5]: $\wedge, \vee, \neg, \supset, \leftrightarrow$

(iii) Parentheses: (,)

Definition (well-formed formulae)**.**

(i) Every propositional variable is a well-formed formula.

(ii) If f is a well-formed formula, so is $\neg f$.

(iii) If f, g are well-formed formulas, so are $(f \wedge g)$, $(f \vee g)$, $(f \supset g)$ and $(f \leftrightarrow g)$.

(iv) Nothing else is a well-formed formula.

Definition (truth values and valuations)**.** Each proposition has one of the two truth values 'true' (T) or 'false' (F). The truth value of each well-formed formula is uniquely determined by the truth values of its parts; a function mapping each sentence to a truth value is called a *valuation*.

Definition (atomic formulae)**.** Each well-formed formula is either an *atomic* or a *non-atomic* formula; non-atomic formulae are characterised by the property that they can be divided into partial expressions where the truth value of the non-atomic well-formed formula depends on the truth values of the atomic formulae it consists of.

3.2.2 Probability of Sentences of a Formal Language

Analogously to the probabilities of subsets of a set, one can attach probabilities to sentences in a formal language[6]:

[4] See Goldrei (2005).

[5] The material conditional contained in $\mathcal{L}$ is denoted by $\supset$ in order to avoid mistaking the different specifications of '$\rightarrow$' in Chapters 4 and 5 for the material conditional.

[6] Based on Demey et al. (2013)

Definition. Let F be a collection of sentences in a formal language $\mathcal{L}$ that is closed under truth functional combinations. A *probability function* P on $(\mathcal{L}, F, P)$ is a function $P : F \to \mathbb{R}$ such that

(a) $P(A) \geq 0$ for all A in F (non-negativity),

(b) If T is a logical truth[7], then $P(T) = 1$ (normalisation),

(c) $P(A \vee B) = P(A) + P(B)$ for all $A, B \in F$ such that A and B are logically incompatible, i.e. in classical logic $A \wedge B \vdash \bot$ (finite additivity).

In this case, $(\mathcal{L}, F, P)$ is called a *probability space* and elements in F are called *events*.

The definition of conditional probability and Bayes' theorem also hold for formal language-based probability spaces. In the remainder of the thesis, let $\mathcal{L}$ be the language of propositional logic.

3.3 Interpretations of Probability

For the formal definitions of probability theory, the aforementioned axioms are the standard; but a completely different aspect to consider is the understanding of the *concept* of probability[8]. What is meant by 'the probability of an event'? If the probability of an outcome of an experiment is 0.5, does it mean that the event will happen close to 50 out of 100 times under the same conditions? What if the experiment cannot be repeated under exactly the same conditions, for example because the outcome changes the conditions? Interpretations of probability can be roughly divided into five types: classical, logical, subjectivist, frequentist, and propensity interpretations (Hájek, 2012). While the different types branch out even further, the discrimination that is especially relevant for the study of conditionals is the one between subjectivist and objectivist accounts.

Objective Probabilities

In interpretations of this kind, probabilities are considered as a property of the outer world. For instance, in *frequentist* interpretations a probability is seen as the relative frequency of occurences of the event in question (or, in

[7] A logical truth is a proposition that can never be false; for example, a tautology.

[8] This section is based on (Hájek, 2012).

case of infinite sequences of events, the limiting relative frequency). *Propensity* interpretations avoid the problem of single measurements (for example, one single coin toss does nott result in a frequentist probability) by relating probabilities to the causes of outcomes of a particular experiment.

Subjective Probabilities

Subjective interpretations can be characterised by the statement: 'Probability is degree of belief.' This approach identifies probabilities with degrees of confidence of different agents at particular times.
Unconstrained subjectivism, posing no requirements to the agent, is not tenable; people have been shown to 'violate the probability calculus in spectacular ways' (Hájek, 2012).
From the point of view of someone interested in coherent probability interpretations, a more interesting case results from the assumption that the agent is *rational*. In particular, a rational agent's beliefs have to be logically consistent. It can be shown that logical consistence implies agreement with the axioms of probability calculus.

Now we have the necessary tools in order to go on to the next chapter, which will deal with the question: how do we evaluate conditionals, and how can this be formalised?

Chapter 4

Conditionals

After the necessary background to discuss probabilistic conditionals has been provided in the previous chapter, this chapter is based on Ramsey's idea how one can determine whether to accept a conditional. It will explain the idea and present two alternatives to the case of the material implication that are based on this idea: first, Adams Conditionals; and second – after a very short introduction to modal logic necessary to define them – Stalnaker Conditionals. In the last part, a formal framework for Stalnaker conditionals (Stalnaker's C2 logic) will be described.

4.1 Ramsey Test

In a footnote to his article, Ramsey (1931) suggested the following acceptability condition for a conditional:

> If two people are arguing 'If p, then q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; so that in a sense 'If p, q' and 'If p, $\overline{q}$' are contradictories. We can say that they are fixing their degree of belief in q given p. If p turns out false, these degrees of belief are rendered void. If either party believes *not* p for certain, the question ceases to mean anything to him except as a question about what follows from certain laws or hypotheses.

That is: if you want to determine whether to believe $p \to q$, add p hypothetically to your set of beliefs and determine your degree of belief in q.

4.2 Adams Conditionals

According to Ernest Adams (1988), conditionals are not factual claims that have truth values but bear conditional probabilities (Skyrms, 1994). In a sense, they constitute a probabilistic formalisation of the Ramsey test: two dialogue partners evaluating $p \to q$ by 'adding p hypothetically to their stock of knowledge' in the framework of probability theory means precisely conditioning on p.

Depending on which interpretation of probability one considers (probabilities as objective chances or as degrees of belief; see Section 3.3), his original hypothesis results in two theories:

With the degree-of-belief interpretation, conditionals play a role in updating our degrees of belief: Let q be a proposition and $P(q)$ the probability which we assigned to it. Given some new evidence e, we update our degree of belief $P(q) \mapsto P(q|e)$, the conditional probability of q on e. The original Adams hypothesis is that this updated probability equals to the probability of the conditional, $P(q|e) = P(q \to e)$ (Skyrms, 1994). This second account of conditionals with probability functions 'based on the concept of knowledge rather than truth' (Stalnaker, 1980) is therefore relevant to Stalnaker's system of conditional logic and to the hypothesis. In Adams' theory of conditionals, probabilities are always understood as *subjective* probabilities.

It is also possible to build a theory of conditionals interpreting conditional probabilites as objective chances. Such an account of conditionals is closely related to the framework of rational choice (see Hájek & Hall (1994)).

4.3 A Glimpse of Modal Logic

In an article in 1912, C.I. Lewis found himself startled by two of the theorems in Russell and Whitehead's *Principia Mathematica*. He was referring to two formulas we already encountered in Chapter 2 as paradoxes of material implication:

$$\neg p \to (p \to q) \tag{2.1}$$

('From a false proposition, anything follows.')

and

$$p \to (q \to p) \tag{2.2}$$

('A true proposition follows from anything.')

According to C.I. Lewis, these paradoxes by themselves do not pose a contradiction or danger of any sort; they merely result from the particular definition of implication as the material implication, which is an inadequate representation of what we would like 'implication' to mean:

> 'In themselves, they are neither mysterious sayings, nor great discoveries, nor gross absurdities. They exhibit only, in sharp outline, the meaning of 'implies' which has been incorporated into the algebra.' (C.I. Lewis, 1912)

Take Statement (2.2), for instance. From p being true it does not logically follow that p logically follows from any particular proposition at all, let alone *any* proposition. To incorporate a better-behaved meaning of 'implies' into logic, C.I. Lewis introduced the strict conditional.

4.3.1 Possible Worlds Semantics for Modal Logic

In order to introduce C.I. Lewis' strict conditional and Stalnaker's understanding of a conditional, some background on the operators of modal logic[1] is now provided.

Recalling the language of propositional logic in Section 3.2.1, a *valuation* assigns a truth value (T or F) to a sentence. It does so by assigning a truth value to each propositional variable p; then the truth values of complex (non-atomic) sentences are calculated using truth tables.

In modal semantics, a set W of possible worlds is introduced. A *valuation* gives a truth value to each propositional variable p *for each possible world* in W. The truth value assigned to p for a world $w \in W$ may differ from the truth value assigned to p for $w' \in W$.

Notation: $v(p, w)$ denotes the truth value of the atomic sentence p at world w given by the valuation v.

Definition. The *operator of possibility* is denoted by $\Diamond$, where $\Diamond q$ is to be interpreted to be 'it is possible that q'. The *operator of necessity* is denoted by $\Box$, where $\Box p$ should be read as: 'it is necessary that p'. These operators are dual to each other in the following sense:

[1] This section is based on Ballarin (2010).

$$\Box p \leftrightarrow \neg\Diamond\neg p$$
$$\Diamond q \leftrightarrow \neg\Box\neg q$$

(Usually, one of these statements is required by definition; the other one follows.)

Definition (truth values for complex sentences)**.**

$$(\neg) \qquad v(\neg A, w) = T \text{ iff } v(A, w) = F$$
$$(\to) \qquad v(A \to B, w) = T \text{ iff } v(A, w) = F \text{ or } v(B, w) = T$$
$$(\Box) \qquad v(\Box A, w) = T \text{ iff for every world } w' \in W, v(A, w') = T$$

For $(\neg)$ and $(\to)$, this is normal truth table behaviour. Moreover, given that $\Diamond A = \neg\Box\neg A$, $(\Box)$ insures that $\Diamond A$ is true exactly in those cases where A is true in at least one possible world.

4.3.2 C.I. Lewis' Strict Conditionals

Definition. The *strict implication* $p \multimap q$ is defined as:

$$p \multimap q := \neg\Diamond(p \wedge \neg q)$$

The axioms for C.I. Lewis' logical system **S3** are:

s1 $(p \wedge q) \multimap (q \wedge p)$

s2 $(p \wedge q) \multimap p$

s3 $p \multimap (p \wedge p)$

s4 $((p \wedge q) \wedge r) \multimap (p \wedge (q \wedge r))$

s5 $((p \multimap q) \wedge (q \multimap r)) \multimap (p \multimap r)$

s6 $(p \wedge (p \multimap q)) \multimap q$

s7 $\Diamond(p \wedge q) \multimap \Diamond p$ (consistency)

s8 $(p \multimap q) \multimap (\neg\Diamond q \multimap \neg\Diamond p)$

The rules of **S3** are:

- **Uniform substitution:** A valid formula remains valid if a formula is uniformly substituted in it for a propositional variable (that is, if all occurrences of the variable are substituted by the formula while leaving all other variables fixed).

- **Substitution of Strict Equivalents:** If $\Phi \strictif \Psi$ and $\Psi \strictif \Phi$, Φ and Ψ can be substituted for each other.

- **Adjunction:** If Ψ and Φ are inferred, then $\Psi \wedge \Phi$ can also be inferred.

- **Strict Inference:** If Φ and $\Phi \strictif \Psi$ have been inferred, Ψ can also be inferred.

The strict conditional also has its problems (Egré & Cozic, 2012). For instance, inferences by contraposition and transitivity are still possible. That is,

$$p \strictif q \models \neg q \strictif \neg p$$

and

$$(p \strictif q) \wedge (q \strictif r) \models p \strictif r\,.$$

To see why inference by contraposition is undesirable, consider the following conditional sentence as an example: 'If it is Anne's birthday, she will get a present today.' By contraposition we get: 'If Anne will not get a present today, it is not Anne's birthday.' This statement is wrong; for example, it might be Christmas.

Transitivity is 'bad' since strengthening is a special case of transitivity: if transitivity *did* hold, we would have that if $A \to B$ was true, $(A \wedge C) \to B$ would be true as well. But this also does not properly capture conversational reasoning; consider the sentences 'If I buy a bottle of water in the summer, it won't freeze.' and 'If I buy a bottle of water in the summer *and* put it in the freezer, it won't freeze.'

4.4 Stalnaker Conditionals

Stalnaker considered the problem of implication in some sense the other way around than Adams: instead of thinking of how a conditional statement must look like to be true or not, his approach is based on the perspective of a selection function. Such a function constitutes a tool to determine whether a possible world is such that the conditional is true in this particular world and, if not, to select another one.

Definition. A *selection function* is a function f that maps each pair (p, w) where p is a propositional variable in $\mathcal{L}$ and w is a possible world in the set of possible worlds W, to a possible world w':

$$\begin{aligned} f : (W, \mathcal{L}) &\to W \\ (w, p) &\mapsto w' \end{aligned}$$

A conditional is said to be true in the *actual* world whenever its consequent is true in the *selected* world.

Using the corner ('>') as the conditional connective, the *assertion* which the conditional $(p > q)$ makes is: the consequent is true in the world that is selected:

> $A > B$ is true in $\alpha \in W$ if B is true in $f(A, \alpha)$;
> $A > B$ is false in α if B is false in $f(A, \alpha)$.

Thus, conditional logic is an extension of modal logic making use of possible-world semantics.

4.5 Stalnaker's C2 Logic

The class of valid formulas of Stalnaker's conditional logic portrayed in the previous section coincides with the class of theorems in Stalnaker's formal system **C2**.

C2 contains the primitive connectives $\supset$ (formalisation of the material implication), $\neg$, and $>$. The connectives $\vee$ and $\wedge$ can be constructed using these primitive connectives:

$$\begin{aligned} (A \vee B) &= (A \supset (\neg B)) \\ (A \wedge B) &= \neg(\neg A \vee \neg B) \end{aligned}$$

The modal operators of possibility and necessity are defined using the corner:

$$\begin{aligned} \Box A &:= (\neg A) > A \\ \Diamond A &:= \neg(A > \neg A) \\ A \lesseqgtr B &:= (A > (B \wedge (B > A))) \end{aligned}$$

The axioms for **C2** are:

a1 Any tautologous well-formed formula is an axiom.

a2 $\Box(A \supset B) \supset (\Box A \supset \Box B)$

a3 $\Box(A \supset P) \supset (A > B)$

a4 $\Diamond A \supset ((A > B) \supset \neg(A \supset \neg B))$

a5 $A > (B \vee C) \supset ((A > B) \vee (A > C))$

a6 $(A > B) \supset (A \supset B)$

a7 $A \lesseqgtr B \supset ((A > C) \supset (B > C))$

The rules of inference are:

- **Modus ponens:** If A and $A \supset B$ are theorems, then so is B.
- **Gödel rule of necessitation:** If A is a theorem, then $\Box A$ is a theorem.

Properties of Stalnaker conditionals

Robert Stalnaker states the conditional denoted by ‘$>$’ to be intermediate between the strict implication and the material implication. The following unsusual features are particularly remarkable (Stalnaker, 1980):

- The inference of contraposition, which is valid for both ‘$\supset$’ and ‘⥽’, is invalid for ‘$>$’. $A > B$ may be true while $\neg B > \neg A$ is false. In conversational usage, the inference of contraposition is in general also invalid (see 4.3.2).
- Transitivity does not hold. This is desirable, since strengthening is an conversationally unplausible inference (see 4.3.2), and strengthening is a special case of transitivity.

This chapter has introduced three types of conditionals: Adams conditionals as bearers of conditional probabilities, C.I. Lewis’ strict conditionals in a type of modal logic, and Stalnaker conditionals in terms of strict conditionals and possible worlds semantics. The next chapter will now deal with a fundamental hypothesis regarding the nature of conditionals.

Chapter 5

The CCCP Hypothesis

This chapter will introduce the hypothesis that conditional probability is the probability of the conditional. This differs from Adams' account of conditionals insofar as it does not only *define* the probability of a conditional by conditional probability, but hopes for a more fundamental type of equality: the hypothesis claims that conditionals can be considered as elements of the probability space, that they can be assigned probabilities, and that these probabilities coincide with the conditional probabilities of the consequent given the antecedent.

The most general version of the hypothesis is: '*Conditional probability is the probability of the conditional*'. That is,

$$P(A \to B) = P(B|A)\,, \tag{5.1}$$

where $A, B \in \Omega$ and $P(A) > 0$. This is the ***C**onditional **C**onstrual of **C**onditional **P**robability* – a term that was introduced by Hájek & Hall (1994). They differentiate between the following statements that result by quantifying over the probability functions and conditionals in four different ways. An informal account of those ways is given here, and is later made more precise in Section (5.2).

(1) There is some[1] '$\to$' such that for all probability functions P, CCCP holds.

(2) There is some '$\to$' such that for all probability functions P that could represent a rational agent's system of beliefs, CCCP holds.

[1]Please note that in this paragraph, '$\to$' does not stand for the material conditional! Here, it denotes *some* conditional that is yet to be specified. This notation is kept throughout the remainder of this chapter unless stated otherwise.

(3) For each probability function P there is some '$\rightarrow$' such that CCCP holds.

(4) For each probability function P that could represent a rational agent's system of beliefs, there is some '$\rightarrow$' such that CCCP holds.

Hájek & Hall (1994) show that the second and fourth versions, although underspecified[2], are not tenable. For the other two versions, this has been shown before by Lewis (1986).
In Section 5.1 of this chapter, a simple argument against the general formulation of the hypothesis is presented. In Section 5.2, the four different versions of the hypothesis are restated in a slightly different terminology, and it is shown that the hypothesis with few additional requirements already entails the characteristic principles for Stalnaker's C2 logic introduced in Section 4.4. This is followed by Lewis' triviality results and their proofs in Section 5.3 and Hájek and Hall's triviality results in Section 5.4. The chapter concludes by reviewing what the triviality results mean for the hypothesis in Section 5.5.

5.1 Disproving the CCCP Hypothesis

The first version of the hypothesis, being the most general formulation, is easier to disprove than the other versions. It is sufficient to find for any '$\rightarrow$' *one* probability function and a pair of events such that Equation (5.1) does not hold.

Proposition. *The version of CCCP that is given by (1) in the list above does not hold.*

Definition. P is called a *CCCP-function*, if $P(Y|X) = P(X \rightarrow Y)$ for any pair of events X, Y.

Proof. Let P_1, P_2 be two distinct CCCP-functions for a '$\rightarrow$', and let $P_3 = \frac{1}{2}P_1 + \frac{1}{2}P_2$. Let A, B be two events such that $P_i(A)$, $P_i(B)$, $P_i(A \cap B)$ (where $i = 1, 2$) are all positive[3] and $P_1(B|A) \neq P_2(B|A)$. Assume P_3 is a CCCP-function. Now,

$$P_3(B|A) = \frac{P_3(A \cap B)}{P_3(A)} = \frac{\frac{1}{2}P_1(A \cap B) + \frac{1}{2}P_2(A \cap B)}{\frac{1}{2}(P_1(A) + P_2(A))}$$

[2] Without choosing a definition of a rational agent.

[3] Since $P(X) \geq 0$ for any probability function P and event X, in the context of probability, 'positive' means 'strictly greater than zero'.

by the definition of conditional probability. Furthermore,

$$\begin{aligned}P_3(A \to B) &= \frac{1}{2}P_1(A \to B) + \frac{1}{2}P_2(A \to B) \\ &= \frac{1}{2}P_1(B|A) + \frac{1}{2}P_2(B|A) \\ &= \frac{1}{2}\frac{P_1(A \cap B)}{P_1(A)} + \frac{1}{2}\frac{P_2(A \cap B)}{P_2(A)}\,.\end{aligned}$$

This holds because P_1 and P_2 are CCCP-functions. By assumption, $P_3(B|A) = P_3(A \to B)$ and therefore, equating

$$\frac{\frac{1}{2}P_1(A \cap B) + \frac{1}{2}P_2(A \cap B)}{\frac{1}{2}(P_1(A) + P_2(A))} = \frac{1}{2}\frac{P_1(A \cap B)}{P_1(A)} + \frac{1}{2}\frac{P_2(A \cap B)}{P_2(A)}$$

we get

$$\begin{aligned}&P_1(A)P_2(A)(P_1(A \cap B) + P_2(A \cap B)) \\ &= \frac{1}{2}(P_1(A \cap B)P_2(A) + P_2(A \cap B)P_1(A)) \cdot (P_1(A) + P_2(A))\,.\end{aligned}$$

This is equivalent to

$$\begin{aligned}&P_1(A \cap B)P_1(A)P_2(A) + P_2(A \cap B)P_1(A)P_2(A) \\ =&\frac{1}{2}(P_1(A \cap B)P_1(A)P_2(A) + P_1(A \cap B)P_2(A)^2 \\ &+ P_2(A \cap B)P_1(A)P_2(A) + P_2(A \cap B)P_1(A)^2)\,.\end{aligned}$$

This results in

$$\begin{aligned}&P_1(A \cap B)P_1(A)P_2(A) + P_2(A \cap B)P_1(A)P_2(A) \\ &= P_1(A \cap B)P_2(A)^2 + P_2(A \cap B)P_1(A)^2\,. \end{aligned} \qquad (\star)$$

The last equality does not necessarily hold, as is demonstrated next in a case-by-case analysis.

(i) Let P_1, P_2 be such that $P_1(A \cap B) = P_2(A \cap B) := P(A \cap B) \neq 0$ but $P_1(A) \neq P_2(A)$. Then $(\star)$ results in:

$$2P(A \cap B)P_1(A)P_2(A) = P(A \cap B)P_2(A)^2 + P(A \cap B)P_1(A)^2$$

This is equivalent to:

$$P(A \cap B)(P_1(A) - P_2(A))^2 = 0$$

Since $P(A \cap B) \neq 0$, it must hold that $(P_1(A) - P_2(A))^2 = 0$ and therefore $P_1(A) = P_2(A)$, a contradiction.

(ii) Let $P_1(A \cap B) \neq P_2(A \cap B)$ but $P_1(A) = P_2(A)$. Then a contradiction follows in analogy to (i).

(iii) Let $P_1(A \cap B) \neq P_2(A \cap B)$ and $P_1(A) \neq P_2(A)$.

$$\begin{aligned} 0 &= P_1(A \cap B)\left(P_1(A)P_2(A) - P_2(A)^2\right) + P_2(A \cap B)\left(P_1(A)P_2(A) - P_1(A)^2\right) \\ &= (P_1(A) - P_2(A)) \cdot P_1(A \cap B)P_2(A) + (P_2(A) - P_1(A)) \cdot P_2(A \cap B)P_1(A) \\ &= (P_1(A) - P_2(A)) \cdot (P_1(A \cap B)P_2(A) - P_2(A \cap B)P_1(A)) \end{aligned}$$

Since $P_1(A) \neq P_2(A)$, it follows that $P_1(A \cap B)P_2(A) - P_2(A \cap B)P_1(A) = 0$ and thus

$$\frac{P_1(A \cap B)}{P_1(A)} = \frac{P_2(A \cap B)}{P_2(A)} .$$

This is equivalent to

$$P_1(B|A) = P_2(B|A)$$

and hereby contradicts the assumption.

□

5.2 A More Formal Approach to CCCP

In this section[4], the additional terminology of models and conditional algebras is introduced and the different versions of the CCCP hypothesis are restated in this more formal framework that is useful to describe Lewis' triviality results in Section 5.3. It is then shown in Part 5.2.1 of this section that by requiring the CCCP hypothesis and three additional axioms to hold, the principles of Stalnaker's C2 logic can be obtained, establishing the relevance of the triviality results for the C2 logic.

Definition. Let (W, F, P) be a probability space. $(W, F, P, \rightarrow)$ is called a *model* and $(W, F, \rightarrow)$ a *(conditional) algebra*.

Definition. It is said that CCCP *holds* for a model $(W, F, P, \rightarrow)$ if for all $A, B \in F$ with $P(A) > 0$ it holds that $P(A \rightarrow B) = P(B|A)$. If CCCP holds, the conditional operator denoted by $\rightarrow$ is called a *CCCP-conditional* for the probability space (W, F, P) and P is called a *CCCP-function* for this space. A model $(W, F, P, \rightarrow)$ is called *trivial* if P has at most four values of conditional probabilities. A probability space is called trivial if any model containing it is trivial.

[4] This section is based on Hájek & Hall (1994).

For a given trivial probability space, one can construct a '$\rightarrow$' that extends it to a trivial model where CCCP holds. This implies that stating that the property for a conditional to be a CCCP-conditional is only meaningful in the context of non-trivial probability spaces; this makes the triviality results tantamount to disproving the hypothesis, for they disprove the interesting cases.

To construct such a '$\rightarrow$', let (W, F, P) be a trivial probability space. Due to triviality, $P(X \cap Y) \in \{0, 1, P(X), P(Y)\}$ for all $X, Y \in F$. Define '$\rightarrow$' as follows: if $P(X \cap Y) = 0$, let $X \rightarrow Y = \varnothing$; if $P(X \cap Y) = 1$, let $X \rightarrow Y = W$; and let $X \rightarrow Y = Y$ otherwise.
Now, $X \rightarrow Y$ can only be assigned the probabilities $P(\varnothing) = 0$, $P(W) = 1$ or $P(Y)$. Thus, this '$\rightarrow$' together with (W, F, P) provides a trivial model for which CCCP holds.
For this reason, the property of being a CCCP-conditional is only meaningful when considering non-trivial probability spaces and models.

Definition. Let $\mathcal{L}_{\rightarrow}$ denote the language of propositional logic together with an additional[5] implication '$\rightarrow$' such that $\mathcal{L}_{\rightarrow}$ is closed under $\rightarrow$.

Given W and F, we can now reformulate the four versions of the CCCP hypothesis in this terminology:

(1') There is some '$\rightarrow$' such that CCCP holds for all models $(W, F, P, \rightarrow)$, where P is defined on $\mathcal{L}_{\rightarrow}$. (Universal version)

(2') There is some '$\rightarrow$' such that CCCP holds for all models $(W, F, P, \rightarrow)$, where P is defined on $\mathcal{L}_{\rightarrow}$ and could describe a rational agent's set of beliefs. (Belief function version)

(3') For each P there is some '$\rightarrow$' such that CCCP holds for the model $(W, F, P, \rightarrow)$. (Universal tailoring version)

(4') For each P that could describe a rational agent's set of beliefs, there is some '$\rightarrow$' such that CCCP holds for the model $(W, F, P, \rightarrow)$. (Belief function tailoring version)

Since '$\rightarrow$' should represent a conditional, one can pose additional requirements[6] to the model $(W, F, P, \rightarrow)$:

(L1) For all $A, B \in F$, $A \cap (A \rightarrow B) \subseteq AB$.[7] (Modus ponens)

[5] Again, '$\rightarrow$' does not denote the material conditional; the material conditional is already contained in $\mathcal{L}$.
[6] Hájek & Hall (1994)
[7] AB denotes $A \cap B$.

(L2) For all $A, B, C \in F$, $(A \to B) \cap (A \to C) \subseteq A \to (BC)$. (Entailment of consequent)

(L3) For all $A, B, C \in F$, $(A \to B) \cap (B \to A) \cap (B \to C) \subseteq A \to C$. (Weakened transitivity)

5.2.1 Probabilistic Entailment Results

Let (W, F, P) be a model obeying (L1), (L2) and CCCP, and let $A, B, C \in F$. Then the following properties hold:

(PE1) If $A \subseteq B$ and $P(A) > 0$, then $P(A \to B) = 1$.

Proof.

$$\begin{aligned} P(A \to B) &= P(B|A) && \text{(CCCP)} \\ &= \frac{P(A \cap B)}{P(A)} && (P(A) > 0) \\ &= \frac{P(A)}{P(A)} && (A \subseteq B) \\ &= 1 \end{aligned}$$

□

(PE2) If $P(AB) = 0$ and $P(A) > 0$, then $P(A \to B) = 0$.

Proof.

$$\begin{aligned} P(A \to B) &= P(B|A) && \text{(CCCP)} \\ &= \frac{P(A \cap B)}{P(A)} && (P(A) > 0) \\ &= 0 && (P(A \cap B) = 0) \end{aligned}$$

□

(PE3) Let $P(BC) = 0$. If $P(A) > 0$, then $P((A \to B) \cap (A \to C)) = 0$.

Proof. Axiom (L2) states that $(A \to B) \cap (A \to C) \subseteq A \to (BC)$. Thus,

$$\begin{aligned}
P((A \to B) \cap (A \to C)) &\leq P(A \to BC) \\
&= P(BC|A) && \text{(CCCP)} \\
&= \frac{P(ABC)}{P(A)} && (P(A) > 0) \\
&= 0\,. && (P(AB) = 0 \text{ and } ABC \subseteq AB)
\end{aligned}$$

□

(PE4) If $P(A) > 0$, then $P((A \to B) \cup (A \to \overline{B})) = 1$ (where $\overline{B} = \neg B$ denotes the complement of B).

Proof.

$$\begin{aligned}
&P((A \to B) \cup (A \to \overline{B})) \\
&= P(A \to B) + P(A \to \overline{B}) - P((A \to B) \cap (A \to \overline{B})) && \text{(additivity)} \\
&= P(A \to B) + P(A \to \overline{B}) && \text{(PE3)} \\
&= P(B|A) + P(\overline{B}|A) && \text{(CCCP)} \\
&= 1
\end{aligned}$$

□

(PE5) If $P(A) > 0$, then $P(C \cap \neg(A \to B)) = P(C \cap (A \to \overline{B}))$.

Proof.

$$\begin{aligned}
&P(C \cap \neg(A \to B)) \\
&= P(((C \cap \neg(A \to B)) \cap ((A \to B) \cup (A \to \overline{B})))) && \text{(PE4)} \\
&= P(C \cap \neg(A \to B) \cap (A \to \overline{B})) + \underbrace{P(C \cap \neg(A \to B) \cap (A \to B))}_{=0} \\
&= P(C \cap A \to \overline{B}) - P(C \cap (A \to \overline{B}) \cap (A \to B)) \\
&= P(C \cap A \to \overline{B}) && \text{(PE3)}
\end{aligned}$$

where the next-to-last equality holds because of expansion by cases:
$P((C \cap A) \to B)$
$= P(((C \cap A) \to \overline{B}) \cap (A \to B)) + P(((C \cap A) \to \overline{B}) \cap \neg(A \to B))$. □

(PE6) If $P(A) > 0$, then $P(A \to B \cap A \to C) = P(A \to BC)$.

Proof. For $P(A) > 0$,

$$\begin{aligned} &P((A \to B) \cap (A \to C) \cap \neg(A \to BC)) \\ &= P((A \to B) \cap (A \to C) \cap (A \to \overline{B})) && \text{(PE5)} \\ &= 0\,. && \text{(PE3) and (L2)} \end{aligned}$$

Therefore,

$$\begin{aligned} &P(A \to B \cap A \to C) \\ &= P(A \to B \cap A \to C \cap A \to BC) \\ &= P(A \to BC) - P(\neg(A \to B) \cap A \to BC) \\ &\quad - P(\neg(A \to C) \cap A \to B \cap A \to BC) && \text{(L2)} \\ &= P(A \to BC) - P(A \to \overline{B} \cap A \to BC) \\ &\quad - P(A \to B \cap A \to \overline{C} \cap A \to BC) && \text{(PE5)} \\ &= P(A \to BC)\,. && \text{(PE3)} \end{aligned}$$

□

(PE7) If $P(A) > 0$, then $P(A \to B \cup A \to C) = P(A \to (B \cup C))$.

Proof.

$$\begin{aligned} &P(A \to B \cup A \to C) \\ &= P(A \to B) + P(A \to C) - P(A \to B \cap A \to C) \\ &= P(A \to B) + P(A \to C) - P(A \cap BC) && \text{(PE6)} \\ &= P(B|A) + P(C|A) - P(BC|A) && \text{(CCCP)} \\ &= P(B \cup C|A) && \text{(set theory)} \\ &= P(A \to (B \cup C)) && \text{(CCCP)} \end{aligned}$$

□

(PE8) If $P(A \cup B) > 0$, then $P(((A \cup B) \to A) \cup ((A \cup B) \to B)) = 1$.

Proof.

$$\begin{aligned} &P(((A \cup B) \to A) \cup ((A \cup B) \to B)) \\ &= P((A \cup B) \to (B \cup A)) && \text{(PE7)} \\ &= 1 && \text{(PE1)} \end{aligned}$$

□

(PE9) $P(AB) = P(A \cap (A \to B))$.

Proof. If $P(A) = 0$, then it holds trivially. Let $P(A) > 0$. Then,

$$\begin{aligned} P(AB) &= P(AB \cap (A \to B)) + P(AB \cap \neg(A \to B)) \\ &= P(AB \cap (A \to B)) + P(AB \cap (A \to \overline{B})) && \text{(PE5)} \\ &= P(AB \cap (A \to B)) && \text{(PE2)} \\ &= P(A \cap (A \to B)) - P(A\overline{B} \cap (A \to B)) \\ &= P(A \cap (A \to B))\,. && \text{(PE2)} \end{aligned}$$

□

Thus, the combination of CCCP with (L1) and (L2) implies every principle describing Stalnaker's C2 logic except for (L3), which therefore has to be required separately. Hence probabilistic equivalents are sufficient to describe the triviality results which are proved in the following section.

5.3 Lewis' Triviality Results

Four different triviality results were proved by Lewis (1976, 1986). The proofs of the triviality results are based on two steps: First, assuming that the class of CCCP-functions for an algebra is closed under certain operations. Second, constructing a contradiction based on the specific type of this assumption.

Definition. Two models $(W_1, F_1, P_1, \to_1)$ and $(W_2, F_2, P_2, \to_2)$ are said to *employ the same arrow* iff their respective algebras $(W_1, F_1, \to_1)$ and $(W_2, F_2, \to_2)$ are isomorphic. Thus, it is possible to speak of 'the algebra associated with a particular set of models', that is, the models that employ the same arrow. This algebra is defined up to isomorphism.

The natural question to examine is what the nontrivial CCCP-functions of an algebra are. Concerning the answer to that question, there are two kinds of triviality results:

1. Non-existence results. Those are of the form: *There are* no *nontrivial CCCP-functions for any algebra with certain properties.*

2. Limitation results. These can be expressed as: *The set of nontrivial CCCP-functions for any given algebra with some particular features is limited in some way.*

Definition. A probability function P_C is derived from a probability function P *by conditioning* if there is some $C \in F$ such that $\forall X \in F : P_C(X) = P(X|C)$.

Intuitively, if a person assigns a probability $P(C) \in (0,1)$ to an event C and learns some new evidence that makes her change her subjective probability for C to 1, then her subjective probability for any other proposition H should be $Q(H) = P_C(H)$ (see Joyce (2008)).

Definition. P_x is derived from P by *nondegenerate two-celled Jeffrey conditioning* if there is a proposition C and an x with $0 < x < P(\neg C)$ such that for all B in F,

$$P_x(B) = P(B) + x \cdot (P(B|C) - P(B|\neg C)).$$

Intuitively, if a person assigns probability $P(C) \in (0,1)$ to an event C and in light of some new evidence changes her subjective probability for C to x, then her subjective probability for any proposition B should be: $Q(B) = xP_C(B) + (1-x)P_{\neg C}(B)$ (see Joyce (2008)). That is, her new subjective probability function should be equal to P_x, the function derived from her old subjective probability function by Jeffrey conditioning on x. Letting x take its marginal values results in two degenerate cases: for $x = 0$, Jeffrey conditioning means not conditioning at all; for $x = 1$, Jeffrey conditioning is equivalent to the usual notion of conditioning.

Definition. A class of probability functions is *closed under conditioning* if any probability function derived from a function in the class by conditioning is also in the class. A class is closed *under nondegenerate two-celled Jeffrey conditioning* if any probability function derived from a function in the class by nondegenerate two-celled Jeffrey conditioning is also in the class.

For the remainder of this section, let $(W, F, \to)$ be an algebra ($W = \mathcal{L}$), and let S denote the set of CCCP-functions for this algebra.

First Triviality Result

Theorem. *S does not contain all probability functions definable on $(W, F, \to)$.*

Proof. Suppose that $\to$ is a universal CCCP-conditional. Let P be a probability function and the sentences $A, C \in F$ such that $P(AC) > 0$ and $P(A\overline{C}) > 0$. Since $AC \subseteq A$, $AC \subseteq C$ and $A\overline{C} \subseteq \overline{C}$, it follows that $P(A)$, $P(C)$ and $P(\overline{C})$ are also positive. Since $\to$ is a universal CCCP-conditional, it holds that

$$P(A \to C) = P(C|A)\,.$$

Due to the assumed universality of the conditional, CCCP holds for all probability functions on the given algebra; therefore, $P_B(A \to C) = P_B(C|A)$ for the function P_B derived from P by conditioning on B. Therefore, taking B as C, we get

$$P(A \to C|C) = P(C|AC) = 1\,,$$
$$P(A \to \overline{C}|C) = P(C|A\overline{C}) = 0\,.$$

Expansion by cases for the sentence $A \to C$ yields

$$\begin{aligned} P(A \to C) &= P(A \to C)|C)P(C) + P(A \to C|\overline{C})P(\overline{C}) \\ &= P(C|AC)P(C) + P(C|A\overline{C})P(\overline{C}) \\ &= 1 \cdot P(C) + 0 \cdot P(\overline{C}) \\ &= P(C)\,. \end{aligned}$$

Thus, assuming the existence of a universal CCCP-conditional we have proved the following statement: *If $P(AC)$ and $P(A\overline{C})$ are both positive, then A and C are probabilistically independent under C.*

Let C, D and E be sentences in F which are possible (i.e. have positive probabilities under some probability function P) but pairwise incompatible, that is, $P(CD) = P(DE) = P(CE) = 0$. Let A denote the disjunction $C \vee D$. Then $P(AC)$ and $P(A\overline{C})$ are positive but $P(C|A) \neq P(C)$. If such sentences exist in $\mathcal{L}$, the result obtained above is a contradiction; if the language is too primitive for three possible but pairwise incompatible sentences to exist, it is merely absurd. □

Therefore, the first triviality result can be reformulated as the following proposition, justifying the term 'triviality result'.

Proposition. *Any language having a universal probability conditional is a trivial language.*

Second Triviality Result

Theorem. *If S contains any non-trivial functions, it is not closed under conditioning.*

Proof. Suppose that $\rightarrow$ is a probability conditional for a class of probability functions S which is closed under conditioning. Let $P \in S$ and $A, C \in F$ such that $P(AC)$ and $P(A\overline{C})$ are positive. Analogously to the proof of the first triviality result it follows that $P(C|A) = P(C)$.
Now, let C, D and E be three pairwise incompatible sentences such that $P(C)$, $P(D)$ and $P(E)$ are all positive. Let A be the disjunction $C \vee D$. Then $P(AC)$ and $P(A\overline{C})$ are positive but $P(C|A) \neq P(C)$. So there are no such three sentences.
Also, P takes no more than four different values: assume that P has two distinct probability values $x, y \in (0, 1)$ such that $x + y \neq 1$. Let F and G be sentences such that $P(F) = x$ and $P(G) = y$. Then it follows that at least three of $P(FG)$, $P(\overline{F}G)$, $P(F\overline{G})$ and $P(\overline{FG})$ are positive.
To show this, assume one of the values is equal to zero; without loss of generality, let $P(FG) = 0$. Then $P(\overline{FG}) = 1 - P(FG) = 1$. □

Third Triviality Result

Lewis' first two results are problematic because of the assumption that the class of belief functions is closed under conditioning. Conditioning on a specific proposition C and its (unspecific) negation $\neg C$ might yield to conditioning on the, as Lewis (1986) puts it, wrong sort of a proposition in one case or the other and might not result in a belief function. Therefore Lewis introduces a new class of propositions called evidence propositions and for his third triviality result assumes that the class of belief function is closed under conditioning on this more specified class.

Theorem. *If S contains any non-trivial functions, it is not closed under conditioning on propositions in a single finite partition.*

Before giving the proof, the following definition motivates this specific formulation of the theorem.

Definition. A class of *evidence propositions* is the class of propositions characterising the total evidence available to a particular subject at a certain time. (These propositions are mutually exclusive and jointly exhaustive.)

This suggests a new, weaker reformulation of the hypothesis: CCCP holds throughout the class of belief functions. This class is closed under conditioning on the elements of a finite partition of evidence propositions.

Proof. Let P be a probability function in the class S. Let $C, D, E, \ldots$ be all the propositions in the partition to which P assigns positive probabilities. Assume that there are at least two such propositions. Let A be a proposition such that $P(A|C)$, $P(A|D)$, $\ldots$ are all positive and $P(C|A) \neq P(C)$. (If such $P, C, D, \ldots, A$ do not exist, call the case trivial.)
By finite additivity, the definition of conditional probability and the incompatibility of $C, D, \ldots$ we can expand by cases:

$$P(A \to C) = P(A \to C)P(C) + P(A \to D)P(D) + \ldots$$

Suppose CCCP holds for this class. Then,

$$\frac{P(CA)}{P(A)} = \frac{P(CA|C)P(C)}{P(A|C)} + \frac{P(CA|D)P(D)}{P(A|D)} + \ldots$$

Since C and D are incompatible, all but the first term on the right side vanish and by simplifying we get

$$P(C|A) = P(C)$$

in contradiction to the assumption. □

Fourth Triviality Result

Theorem. *If S contains any non-trivial functions, it is not closed under nondegenerate two-celled Jeffrey conditioning.*

Proof. Let $P \in S$. Let C be a proposition such that $P(C)$ and $P(\neg C)$ are positive. Let A be a proposition such that $P(A|C)$ and $P(A|\neg C)$ are both positive and $P(A|C) \neq P(A|\neg C)$. If such P, C and A do not exist, call the case 'trivial'.
Otherwise, suppose S is closed under nondegenerate two-celled Jeffrey conditioning. Thus, P and P_x are both CCCP-functions, i.e.

$$\begin{aligned} P(A \to C) &= \frac{P(AC)}{P(A)}\,, \\ P_x(A \to C) &= \frac{P_x(AC)}{P_x(A)}\,. \end{aligned} \tag{$\star\star$}$$

By definition of Jeffrey conditioning,

$$\begin{aligned}
P_x(A \to C) &= P(A \to C) + x[P(A \to C|C) - P(A \to C|\neg C)]\,, \\
P_x(AC) &= P(AC) + x[P(AC|C) - P(AC|\neg C)] \\
&= P(AC) + xP(AC|C) \\
&= P(AC) + xP(A|C)\,, \\
P_x(A) &= P(A) + x[P(A|C) - P(A|\neg C)]\,.
\end{aligned}$$

Using these three identities, Equation ($\star\star$) expands to

$$\begin{aligned}
&P(A \to C) + x[P(A \to C|C) - P(A \to C|\neg C)] \\
&= \frac{P(CA) + xP(A|C)}{P(A) + x[P(A|C) - P(A|\neg C)]}\,.
\end{aligned}$$

This is equivalent to

$$\begin{aligned}
&\underbrace{P(C|A)P(A)}_{=P(CA)} + P(C|A)x[P(A|C) - P(A|\neg C)] \\
&+x^2[P(A \to C|C) - P(A \to C|\neg C)][P(A|C) - P(A|\neg C)] \\
&=P(AC) + xP(A|C)\,.
\end{aligned}$$

Since $x \neq 0$ by the definition of non-degenerate Jeffrey conditioning, it is equivalent to

$$\begin{aligned}
&P(C|A)[P(A|C) - P(A|\neg C)] - P(A|C) \\
&= x[P(A \to C|C) - P(A \to C|\neg C)][P(A|C) - P(A|\neg C)]\,.
\end{aligned}$$

Since this shall hold independently of the choice of x, the left-hand side of the equation as well as the coefficient on the right-hand side must vanish.

Case 1

The right-hand side is equal to zero because $P(A|C) = P(A|\neg C)$. This contradicts the choice of A.

Case 2

The right-hand side is equal to zero because $P(A \to C|C) = P(A \to C|\neg C)$. Since the left-hand side is also equal to zero, we have

$$\begin{aligned} 0 &= P(C|A)[P(A|C) - P(A|\neg C)] - P(A|C) \\ &= \underbrace{P(A|C)[1 - P(C|A)]}_{P(A|C)P(\neg C|A)} + P(C|A)P(A|\neg C)\,. \end{aligned}$$

Since P, A and C were chosen such that $P(A|\neg C) \neq P(A|C)$ and, in particular, $P(A) \neq 0$, it follows that $P(C|A) = \frac{P(AC)}{P(A)} = \frac{P(A|C)P(C)}{P(A)}$ is positive since $P(A|C)$ and $P(C)$ were chosen to be positive. Therefore, the first term, that is: $P(A|C)P(\neg C|A)$, already does not vanish. Thus, the equation above contradicts our choice of P, A and C. $\square$

5.4 Hájek and Hall's Triviality Results

We conclude this chapter by briefly summarising some stronger triviality results that were shown by Hájek and Hall. The results of David Lewis outlined in the previous section are in part a consequence of these results. As a description of all of Hájek and Hall's triviality results goes beyond the scope of this thesis, I shall present a selection of their results without proving most of them; the interested reader is referred to Hájek & Hall (1994), Hájek (1994), and Hall (1994).

Strengthened Lewis Result

A strengthened version of Lewis' first three impossibility results can be obtained using the trick of rewriting $P(A \to B|C) = P_C(A \to B) = P_C(B|A) = P(B|AC)$.

Theorem. *If $(W, F, P, \to)$ and $(W, F, P_C, \to)$ are distinct non-trivial models where P_C is derived from P by conditioning on C, then CCCP holds for at most one of the models. It follows that in the set S of nontrivial CCCP-functions for a given algebra, no function results from another CCCP-function by conditionalisation.*

Definition. A proposition $A \in F$ is *P-atom* if $P(A) > 0$ and $\forall X \in F$: $P(AX) = 0$ or $P(AX) = P(A)$.

Proof. Suppose that $(W, F, P, \to)$ and $(W, F, P_C, \to)$ are distinct non-trivial models and CCCP holds for both. Since the models are distinct, $P(C) < 1$. Since $(W, F, P_C, \to)$ is non-trivial, C is not P-atom, for otherwise P_C would only assume the values 0 and 1. Therefore, one can choose a $D \subset C$ such that $0 < P(D) < P(C)$. Let $E = D \cup \overline{C}$. Then $P(E) < 1$. Also,

$$\begin{aligned}
P(E \to \overline{C}) &= P(E \to \overline{C}|C)P(C) + P(E \to \overline{C}|\overline{C})P(\overline{C}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{(expansion by cases)} \\
&= P(\overline{C}|EC)P(C) + P(E \to \overline{C}|\overline{C})P(\overline{C}) \qquad \text{(by Lewis' trick)} \\
&= P(\overline{C}|EC)P(C) + P(\overline{C} \cap (E \to C)) \\
&\qquad\qquad\qquad\qquad \text{(by the definition of cond. prob.)} \\
&= 0 + P(\overline{C} \cap (E \to \overline{C})) \qquad\qquad (C \cap \overline{C} = \varnothing) \\
&\leq P(\overline{C})\,.
\end{aligned}$$

Since P is a CCCP-function by assumption,

$$\begin{aligned}
P(E \to \overline{C}) &= P(\overline{C}|E) \\
&= \frac{P(E \cap \overline{C})}{P(E)} \\
&= \frac{P(\overline{C})}{P(E)}\,. \qquad\qquad (\text{since } \overline{C} \subseteq E)
\end{aligned}$$

Combining these two equations, it follows that

$$P(\overline{C}) \leq P(E)P(\overline{C})\,,$$

and thus $1 \leq P(E)$ – a contradiction.

□

Hájek and Hall's Impossibility Results

Definition. Two probability functions P and P' are called *orthogonal* if for some $A \in F$ it holds that $P(A) = 1$ and $P'(A) = 0$. (Note that this property is symmetric: if $P(A) = 0$, then $P(\overline{A}) = 1$.)

Orthogonality result (Hall): For distinct P and P' in S, P and P' are orthogonal.

Note: Lewis' results follow from this, as no two probability functions where one is derived from the other by conditioning or non-degenerate Jeffrey conditioning can be orthogonal.

Although the orthogonality result seems rather inconspicuous compared to the other results mentioned here, it has the following troublesome consequence: assume an agent's belief in $A \to B$ is given by the CCCP-function $P(A \to B)$. If this belief is now being updated given some new evidence (e.g. in the context of Bayesian learning), resulting in the belief $P'(A \to B)$, the function P' is *not* CCCP since it is derived from P by a form of conditionalising. In other words, the orthogonality result implies that any conditional belief-updating framework is not compatible with an algebra $(W, F, \to)$.

Finitude result (Hájek): If W is finite, the set S of non-trivial $CCCP$-functions for the algebra $(W, F, \to)$ is empty.

Weakened transitivity result (Hájek and Hall): There are no nontrivial CCCP-functions for any algebra whose '$\to$' obeys (L1), (L2) and (L3). This entails an earlier result by Stalnaker stating that CCCP is inconsistent with the assumption that the logic of '$\to$' is C2, since C2 contains (L1)-(L3).

5.5 Impact on CCCP

We conclude the chapter by briefly outlining what the different triviality results imply for the plausibility of the CCCP hypothesis.[8]

The universal version was refuted right at the beginning in Section 5.1. Regarding the belief function version, the orthogonality result implies that the belief function version of CCCP is only true if any two belief functions are either equal or orthogonal, which is absurd. The universal tailoring version is to be discarded as well. This follows from the finitude result together with the fact that there will be functions definable on F that will have a finite range.

The version that remains is the belief function tailoring version: 'For each

[8]This section is based on Hájek & Hall (1994).

P that could describe a rational agent's set of beliefs, there is some '$\rightarrow$' such that CCCP holds for the model $(W, F, P, \rightarrow)$.' Hájek & Hall (1994) show that although it cannot be disproved, it is still untenable for reasons provided by the triviality results. For example, if a rational agent updates her belief function, in most cases the resulting belief function will not be orthogonal to her previous belief function, which discredits most belief functions.

For these reasons, the hypothesis is to a large extent refuted.

Chapter 6

Conclusion and Outlook

In this chapter, I give a summary of the various accounts of conditionals presented in this thesis, mention some philosophically relevant ways to extend their presentation, and conclude by two ideas for further work.

The thesis started out by pointing out problems of the material conditional and by presenting Grice's theory of implicatures as a possible tool for avoiding them, which turned out to be insufficient. It continued with an introduction to probability theory in the context of propositional logic.
Ramsey's suggestion that the degree of acceptance for a conditional is the degree of belief in the consequent given the antecedent was explored: Adams' account of conditionals was portrayed, following Ramsey's idea by stating that conditionals do not bear truth values but conditional probabilities, defining them to be $P(A \to B) := P(B|A)$. Next, Stalnaker's approach to characterise conditionals in the terminology of possible worlds was depicted, along with his C2 logic.
Then the CCCP ('conditional construal of conditional probability') hypothesis was introduced: 'Conditional probability is the probability of the conditional', that is: $P(A \to B) = P(B|A)$, claiming equality of probabilities of elements in the probability space $(\mathcal{L}_{\to}, F, P)$, not just a definition $P(A \to B) := P(B|A)$ where only A and B are in the probability space $(\mathcal{L}, F, P)$ as before.
The probabilistic entailment that followed showed Stalnaker's C2 to be contained in the logic resulting by combining CCCP with three additional axioms, thus showing that the following triviality results also apply to Stalnaker's C2. Lewis' triviality results were stated and proved, and some of the stronger triviality results by Hájek and Hall were mentioned.

Finally, the meaning of these results for the different versions of the CCCP hypothesis was stated: only one of these versions was not disproved by the triviality results but was argued to be untenable.

As mentioned in the introduction, the main focus of the thesis was on formal aguments, particularly Lewis' impossibility results, rather than interpretative content. Two possible philosophical extensions to this work would be: First, the thesis did not go into details on why to believe or not to believe CCCP except for formal reasons; for a thorough account on this, see Hájek & Hall (1994).
Second, the thesis only dealt with *indicative* conditionals, that is: 'if p is the case, then q is the case'-statements of the form 'if one fact holds, so does another'; it did not involve work on *subjunctive* conditionals (also called counterfactuals) that have the form: 'if it *were* the case that p, then it *would* be the case that q' (although p is not the case, thus the term 'counterfactual').

To conclude, I suggest two, admittedly vague, ideas worth pursuing for a better understanding of conditionals and state my personal view why the failure of CCCP is not the end of a probabilistic theory of conditionals.

- Based on the probabilistic calculus in the context of propositional logic, one could try to introduce a probabilistic calculus that only accepts conditionals as events.

- Alternatively, one could try to modify Grice's theory of implication in order to apply it to the strict conditional and/or Stalnaker's C2 logic. (For example, following Grice's principles, it would not be conversationally appropriate to assert $p \to q$ if q is always true.) To avoid the difficulties of Grice's theory, one could interpret the Cooperative Principle in a stronger sense, leaving out some rules Grice derived from it, such as the rule that requires to be polite (Davis, 2013). It might also be promising to formalise Grice's notion of 'informativeness' using information theory.

To me, the failure of the CCCP hypothesis seems less devastating for a probabilistic perspective on conditionals than one might think. The hypothesis itself seemed to me like a meta-level error in some way, for $p \to q$ shall describe a (yet to be specified) relation between propositions 'if p then q', not a proposition.

In particular, considering the fact that we hoped to define '→' that way, what is an expression like $P(p|q \rightarrow r)$ ('the probability of a proposition given a relation between propositions') even supposed to mean?

References

Adams, E. (1966). Probability and the logic of conditionals. In Jaakko Hintikka & Patrick Suppes (Eds.), *Aspects of inductive logic* (pp. 265–316). North-Holland.

Adams, E. (1988). Consistency and decision: variations on Ramseyan themes. In William L. Harper & Brian Skyrms (Eds.), *Causation in decision, belief change, and statistics* (pp. 49–69). Springer.

Ballarin, R. (2010). Modern origins of modal logic. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Winter 2010 ed.). `http://plato.stanford.edu/archives/win2010/entries/logic-modal-origins/`.

C.I. Lewis. (1912). Implication and the algebra of logic. *Mind*, *21*(84), 522–531.

Davis, W. (2013). Implicature. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Spring 2013 ed.). `http://plato.stanford.edu/archives/spr2013/entries/implicature/`.

Demey, L., Kooi, B., & Sack, J. (2013). Logic and probability. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Spring 2013 ed.). `http://plato.stanford.edu/archives/spr2013/entries/logic-probability/`.

Egré, P., & Cozic, M. (2012). *Conditionals.* (to appear in M. Aloni and P. Dekker eds, Handbook of semantics (revised 16/11/2012))

Goldrei, D. (2005). *Propositional and predicate calculus: a model of argument.* Springer.

Grice, H. P. (1975). Logic and conversation. In Peter Cole & Jerry L. Morgan (Eds.), *Syntax and semantics, vol.3* (pp. 41–58). Academic Press.

Hájek, A. (1993). *The conditional construal of conditional probability.* Unpublished doctoral dissertation, Princeton University.

Hájek, A. (1994). Triviality on the cheap? In E. Eells & B. Skyrms (Eds.), *Probability and conditionals: Belief revision and rational decision* (p. 113-140). Cambridge University Press.

Hájek, A. (2012). Interpretations of probability. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Winter 2012 ed.). `http://plato.stanford.edu/archives/win2012/entries/probability-interpret/`.

Hájek, A., & Hall, N. (1994). The hypothesis of the conditional construal of conditional probability. In E. Eells & B. Skyrms (Eds.), *Probability and conditionals: Belief revision and rational decision* (p. 75-111). Cambridge University Press.

Hall, N. (1994). Back in the CCCP. In E. Eells & B. Skyrms (Eds.), *Probability and conditionals: Belief revision and rational decision* (p. 141-160). Cambridge University Press.

Joyce, J. (2008). Bayes' theorem. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2008 ed.). `http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/`.

Klenke, A. (2008). *Wahrscheinlichkeitstheorie.* Springer.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, *85*(3), 297–315.

Lewis, D. (1986). Probabilities of conditionals and conditional probabilities ii. *The Philosophical Review*, *95*(4), 581–589.

Ramsey, F. P. (1931). General propositions and causality. In H. A. Mellor (Ed.), *F. Ramsey, Philosophical papers.* Cambridge University Press, 1990.

Skyrms, B. (1994). Adams conditionals. In E. Eells & B. Skyrms (Eds.), *Probability and conditionals: Belief revision and rational decision* (p. 13-26). Cambridge University Press.

Stalnaker, R. C. (1980). A theory of conditionals. In William L. Harper, Robert Stalnaker, & Glenn Pearce (Eds.), *Ifs. Conditionals, belief, decision, chance and time* (pp. 41–55). Springer.

Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

München, den

____

Alexandra Surdina

Statement of Authorship

I hereby certify that I have written this thesis on my own and only used the listed references and aids.