BALKAN JOURNAL OF PHILOSOPHY. Vol. 7, Issue 2, 2015

WHAT IS LEARNED FROM CONDITIONALS?

Borut Trpin
University of Ljubljana, Faculty of Arts,
Ljubljana, Slovenia
e-mail: borut.trpin@gmail.com

Abstract:
Some of the information that we learn comes to us in a conditional form. This has proven to be a problem for philosophers, who try to explain how probabilistic beliefs change when one learns from conditional sentences. The problem is that a straight-forward solution is not possible: the partial belief in the antecedent and the partial belief in the consequent either increase, decrease, or remain the same. Two existing approaches to learning from indicative conditionals are considered: an explanatory one, and another that builds on relative information minimizing with regard to the causal structure. A novel method based on epistemic entrenchment is proposed to overcome the drawbacks of the competing approaches. The method solves all the standard examples and some other examples for which existing approaches have failed to provide adequate solutions.

Key words: indicative conditionals, formal epistemology, epistemic entrenchment, Bayesian networks, explanation.

1. Introduction

Some of the information that we learn comes to us in form of conditionals. For instance, we may learn that a picnic in the park will be cancelled if it rains. We learn, at the same time, that the picnic will also probably be cancelled if the wind will be very strong, or if it will rain heavily just before the picnic. Although we have no problems learning from such conditionals, it is very hard to describe how learning from conditionals proceeds in general.

In this paper, I propose an approach to the problem of learning from conditionals. My proposal may be summed up in three main ideas:

(1) When an agent learns something from a conditional, her posterior degree of belief in either the antecedent or the consequent remains the same as her prior degree of belief (*). Posterior degree of belief in the other does, however, increase or decrease. (*) is thus epistemically more entrenched.

(2) A conditional may convey implicit information which needs to be made explicit during the analysis.

(3) After taking (2) and supporting beliefs into consideration, a simple entrenchment test can determine:

a) whether the posterior degree of belief in the antecedent or the posterior degree of belief in the consequent remains unchanged, and

b) whether the posterior degree of belief in the other increases or decreases.

The most pleasing analysis of learning from conditionals would be to apply the standard Bayesian procedure for belief updating—conditionalization. Let me formally sketch the idea behind standard conditionalization (SC):

An agent starts with some degree of belief in A and a conditional degree of belief in A given B, represented by conditional probability $Pr(A|B)$. After learning B, i.e. $Pr'(B) = 1$, her posterior belief in A changes to her (prior) conditional probability $Pr(A \mid B)$:

$$Pr'(A) = Pr(A|B)^{1} \qquad \text{(SC)}$$

Jeffrey's conditionalization (JC; also known as probability kinematics) provides a generalization of (SC) for cases when new information is not (necessarily) learned with full certainty. Formally:

$$Pr'(A) = Pr(A|B)Pr'(B) + Pr(A|\neg B)\, Pr'(\neg B) \qquad \text{(JC)}$$

It is obvious from the equation that (SC) is just a special case of (JC) when $Pr'(B)$=1.[2]

To paraphrase David Lewis (1976), alas, the most pleasing solution cannot be right. I will discuss why (JC)—and hence (SC)—cannot provide the

[1] Prior degrees of belief are denoted by Pr(-), and posterior degrees of belief by Pr'(-).

[2] Suppose $Pr'(B) = 1$. Then $Pr'(\neg B) = 0$ and: $Pr'(A) = Pr(A|B)\, Pr'(B) + Pr(A|\neg B)\, Pr'(\neg B) = Pr(A|B)*1 + Pr(A|\neg B)*0 = Pr(A|$B), which is exactly the (SC).

correct results (at least not generally) in the next sections. But let me first clarify the main concepts that will be used throughout the paper.

2. Theoretical background

2.1 Types of conditionals

When I refer to *conditionals*, I refer to standard non-nested indicative conditionals. The indicative conditional operator will be represented by an arrow (→). In the case of other conditionals, i.e. material, subjunctive, left- or right-nested conditionals, the distinction will be explicitly mentioned.

But what exactly are indicative conditionals and how do they differ from other types of conditionals?

Broadly speaking, an indicative conditional is a conditional that has both the antecedent (the if-clause) and the consequent (the main clause) in indicative mood. As noted, I only consider standard indicative conditionals for present purposes. The (indicative) mood is a grammatical feature**,** and as such also includes some non-standard types of conditionals. I will not consider the following non-standard types of indicative conditionals:

- speech act conditionals
- biscuit conditionals
- Dutchman conditionals
- even if conditionals
- independent conditionals
- counterfactual conditionals

Speech act conditionals are the cases in which the antecedent conditionally modifies not the contents of the consequent, but the speech act that the consequent carries out (Dancygier and Sweetser 2005). The typical examples are conditional tips, conditional threats, conditional questions, conditional promises, and conditional commands (Krzyżanowska 2015, p. 61). Let me list a few examples which will clear up why speech-act conditionals are excluded for present purposes.

Conditional threat: “If you don't shut your mouth, I will smack you.”

Conditional question: “Does the light turn on if I press this button?”

As the dependence between the antecedent and the consequent is non-inferential, there is nothing to be learned from them. When Tom says to Sarah: "If you don't shut your mouth, I will smack you," Sarah learns that Tom is angry, but she doesn't learn anything about her continuing with her talking or about her being smacked, at least not in the standard way. The mechanics of learning from speech-act conditionals thus remain an open issue.

Biscuit conditionals are another non-inferential type of indicative conditionals. They are called “biscuit conditionals” after a typical example from Bennett (2003): “There are biscuits on the table if you are hungry.” The biscuits are on the table regardless of the recipient's hunger. Biscuit conditionals are therefore non-standard, as they could easily be analysed in an equivalent non-conditional form, i.e. “Feel free to eat the biscuits,” and as such are not important for our analysis of learning from conditionals.

Dutchman conditionals (Hájek 2012, p. 146) are named after their typical example: “If you manage to do this, I am a Dutchman.” They are used to express a high degree of doubt in the antecedent and use an even more doubtful (or impossible) consequent as the means to convey the doubt in the antecedent. They are not conditional in their nature as they express non-conditional content, i.e. disbelief or doubt in the antecedent, and not the dependence of the consequent on the antecedent.

Even if conditionals are a special type of conditional which is easily recognized by the adverb “even.” They behave in very specific ways and will therefore be excluded.[3]

Independent conditionals (Bennett 2003, p. 149) are conditionals in which there is no real dependence between the antecedent and the consequent, e.g. “If I go outside, Africa is a continent.” Nothing may be learned from such examples, although strange conditionals in which we do not notice any connection between the clauses may suggest that there is some unexpected connection between them (Douven 2008, p. 25).

The last type of indicative conditional that will be excluded is the counterfactual conditional. Although counterfactuals are often thought to coincide with the subjunctive mood, it may be better to think of them as of the conditionals that have an antecedent that is assumed to be false. As (Krzyżanowska 2015, p. 10) pointed out, one could think Denmark is likely not ruled by a king and still assert: “If Denmark is ruled by a king, then it is a kingdom.” Learning from counterfactuals is a very interesting problem, but it is too complex for this paper.

[3] The adverb *even* is not always present in the even-if conditionals, which can lead to some puzzling consequences. For example, „(Even) If she dyed her hair, she didn't dye it blue.” Suppose the consequent is false. Modus tollens then does not hold:
She dyed her hair blue.
∴ She didn't dye her hair.

2.2. Semantics of conditionals

But what is it that makes indicative conditionals so special? One of the main problems is that it is very hard to pin down the exact semantics of indicative conditionals. But the problems do not end even when a system of semantics is clearly described, because all systems of semantics for conditionals lead to epistemological problems (Douven 2013, p.4).

There are three main candidates for a semantics of indicative conditionals: the material analysis (MA), the Stalnaker-style possible worlds analysis (PW), and the non-propositional (or, in the weaker form, the tripartite) analysis (NPA or TPA).

(MA) is the traditional approach to conditionals.[4] It is historically attributed to ancient Greek philosopher Philo of Megara (Sanford 1989, p. 14–25), who defined the semantics of (MA) in 4–3 BC.

The semantics of (MA) are very simple: a conditional ($p \supset q$) is false if and only if the antecedent (p) is true and the consequent (q) is false.

This view was prevalent throughout most of the history (in classical logic) and is still prevalent in mathematics and even in some traditions of psychology of reasoning (e.g. mental logic and mental models). While (MA) is very useful in deductive proofs in mathematics, it has some very unintuitive consequences if we try to asses everyday conditionals by it. The conditional is equivalent to the following disjunction:

$$p \supset q \leftrightarrow \neg p \vee q$$

This means that the conditional is true if its antecedent is false or the consequent is true, which leads to paradoxical examples, e.g.:

1) If Europe is not a continent, then Europe is a continent.
2) If Italy is a country, then 2+2=4.
3) If Berlusconi never existed, then Italy is not a country.

1) is true because both the antecedent is false and the consequent is true.
2) is true because both clauses are true, although there is no connection between them.
3) is true because the antecedent is false (Berlusconi exists), although it is, again, highly unassertible.

A Stalnaker-style possible world semantic of conditionals (PW) is a much better attempt to capture the epistemological relation between the antecedent and the consequent. The main idea of (PW) is that the conditional (regardless of being indicative or subjunctive) is true if and only if the consequent is true in the closest possible world to the one in which the antecedent is true *and* given that there is a possible world in which the antecedent is true. Although (PW) has some major advantages over (MA)—i.e. it escapes the typical paradoxes of (MA)—it has its problems.

For instance, Gibbard's river-boat paradox (Gibbard 1981) demonstrates that two conditionals with the same antecedent and opposite consequents ("If p, then q" and "If p, then ¬q") could both be true in the closest possible *p*-world. Further, the conditionals in which both clauses are true—despite there being no connection between them—are also considered true in (PW). For instance, "If 2+2=4, then Italy is a country" is true, because Italy is a country in the closest world where 2+2=4 is true (the actual world).

Let me now turn to the non-propositional analysis of conditionals (NPA). (NPA) is the most popular semantics amongst theorists of conditionals. According to this view a conditional is true if both the antecedent and the consequent are true, and false if the antecedent is true and the consequent false. In a case in which the antecedent is false, the truth of the conditional is undecided (Table 1).

	p	***q***	***p → q***
1	1	1	1
2	1	0	0
3	0	1	1 or 0 (undecided)
4	0	0	1 or 0 (undecided)

Table 1: Truth table for conditionals according to (NPA)

Although (NPA) captures many of our intuitions, the non-propositional nature of conditionals has a few undesired consequences. If conditionals do

[4] The material conditional operator will be represented by the horse-shoe (⊃).

not express propositions, how can we believe[5] them? I very strongly believe that salt will dilute *if* I mix it with water. Furthermore, if the semi-truth functionality (table 1) of conditionals is accepted, then conditionals could not occur in Boolean combinations, as their truth cannot always be determined (Douven 2013).

I embrace (NPA) in this paper regardless of its limitations. The proposed analysis of learning from conditionals does not require any Boolean combinations. Although I will inspect partial beliefs,[6] (NPA)'s semi-truth functionality still provides valuable guidelines. With some appropriate simplification, if an agents thinks a proposition is more likely to be true (false) than false (true), her degree of belief in this proposition is over (below) .5.

Supposing the newly learned conditional is not misleading, an agent should thus either have degrees of belief in both the antecedent and the consequent over .5 (probabilistically generalised line 1 of Table 1) or both under .5 (to prevent probabilistic cases of line 2).

3. Learning from conditionals

It may be clear from the previous sections that the problems with learning from conditionals are at last partially triggered by the very complex nature of conditionals. There are almost no common assumptions among philosophers researching conditionals. As Krzyżanowska (2015, p.2) noted, "it seems that any time a new account arises, a refutation arises sooner or later."

The situation is even worse when it comes to learning from conditionals. The problem is not just that all approaches have counterarguments, but also that learning from conditionals has not been given much attention. Only a few (somewhat general) approaches to (uncertain) learning from conditionals exist.[7]

3.1. Updating on standard material conditionals

Let us start with updating on standard material conditionals (UMC). One of its main strengths—the fact that it is very straightforward—is also its main limitation. After all, according to (UMC), after the agent learns a conditional her posterior belief in the antecedent should always *decrease* if she is uncertain about the antecedent and the consequent:

(UMC PAR)

Suppose $Pr(antecedent) > 0$ and $Pr(consequent) > 0$. According to (UMC) it then follows that $Pr'(antecedent|\ antecedent \supset consequent) \leq Pr(antecedent)$.
If additionally $Pr(antecedent) < 1$ and $Pr(consequent) < 1$, it follows that $Pr'(antecedent|\ antecedent \supset consequent) < Pr(antecedent)$.[8]

It is very hard to believe that my degree of belief in tomorrow's rain should decrease after learning: "If it rains tomorrow, the picnic in the park will be cancelled." (UMC) is thus not adequate to explain learning from conditionals.

3.2. Adams conditioning and explanatory analysis of learning from conditionals

Another approach is based on the so-called Adams conditioning, as proposed in Douven and Romeijn (2011). They developed their view in response to the so-called Judy Benjamin problem (originally in van Fraassen 1981). They started with a general rule for updating conditional degrees of belief—dubbed Adams' conditioning (AC)—and adapted it for conditionals. As (AC) is just a special case of Jeffrey's conditioning (JC), their approach is actually just a special case of (JC), with which we can analyse learning from conditionals of the form: "If A, then the odds for $B_1, \ldots, B_n$ are $c_1, \ldots, c_n$".

One of the biggest drawbacks of this account is that the posterior degree of belief in the antecedent always remains the same. While this seems fine for specific cases like the Judy Benjamin problem (which will be discussed later), it is very limited and gives the impression that it is tailored for specific examples.[9]

This is also the reason why Douven later (2012) provided a more general account of learning from conditionals based on the explanatory relation between the antecedent and the consequent (EXP). The main idea is quite simple. When an agent learns a

[5] The standard definition of "belief" is an attitude towards a proposition.

[6] It should be noted that the exact relation between truth and probability is not clear. Rational agents believe tautologies fully (with $Pr(\top) = 1$) and disbelieve contradictions ($Pr(\bot)=0$), but it is not clear how to assess partial beliefs.

[7] More research has been done in the realm of AGM theory and computer science. These analyses are, however, only concerned with categorical beliefs (Douven 2012, p. 239).

[8] For proof, see Douven and Romeijn (2011), p. 645.

[9] Refer to Lukits (2015) for further refutation of their approach.

conditional $p \rightarrow q$, she sets her conditional probability close to 1, i.e. $Pr'(q|p) \approx 1$. If there is no change in the explanatory status of the antecedent, Adams conditioning is applied. If the explanatory status goes up, $Pr'(q) \gtrsim Pr'(p) > Pr(p)$, and if the explanatory status goes down, $Pr'(q) \gtrsim Pr'(p) < Pr(p)$.

One of the main drawbacks of (EXP) is that is not clear how to determine whether the explanatory status of the antecedent goes up or down (if at all). Research of explanatory status in Bayesian epistemology is, however, increasing, and it is reasonable to expect that (EXP) will have firmer ground once more research on explanatory status is done.

There are, however, some other crucial drawbacks of (EXP). There are some cases where (EXP) excludes results which are intuitively acceptable. The same problems occur for relative information minimizers (INFOMIN). I will next examine how (INFOMIN) approaches the problem of learning from conditionals.

3.3. Minimising Kullback-Leibler divergence and causal representation

When van Fraassen introduced the Judy Benjamin problem in 1981, he demonstrated that (INFOMIN) leads to intuitively strange results.[10] I will discuss the Judy Benjamin problem later, but first I will examine what (INFOMIN) actually is and how a combination of (INFOMIN) and Bayesian causal nets provides a very robust approach (INFOMIN+CAU) (Hartmann and Rad, unpublished). (INFOMIN) is based on Kullback-Leibler divergence $D_{KL}(P'||P)$, which measures how much information is lost when moving from a prior probability distribution to a posterior probability distribution. Updating then proceeds through finding the minimum of divergence.[11]

Although (INFOMIN) yields correct results in many cases, it leads to very strange, although not necessarily wrong (Lukits 2015) results in some cases—for instance, in the Judy Benjamin problem. Hartmann and Rad combined (INFOMIN) with Bayesian causal nets and developed a technically complex but very robust model of updating on conditionals (INFOMIN+CAU). For instance, they provide scenarios with additional disabling conditions and demonstrate how their model still gets the expected results.

Their approach, however, has at least two problems:

1) Their application of (INFOMIN) depends on properly represented Bayesian causal nets. They do not provide any rules about how to determine when a representation is proper. As I will demonstrate later, a slight (and intuitively correct) change of structural representation may result in wrong results.

2) Similarly to (EXP), (INFOMIN+CAU) runs into problems in cases where both a change in degree of belief in the antecedent and the consequent appar feasible.

4. The examples

Let me first list the standard examples[12] of learning from conditionals. This will show the strength of both (EXP) and (INFOMIN+CAU). I will then show how both accounts fail to provide (all) possible results if the examples are slightly modified.

(1) The Sundowners Example

Sarah and her sister Marian are planning to have sundowners at the Westcliff hotel tomorrow. Sarah thinks there is some chance it may rain tomorrow, but thinks they could still enjoy the view from inside. To make sure, Marian makes a telephone call to the staff at the hotel. They tell her that in the event of rain, a wedding party will occupy the inside area. So she tells her sister: “If it rains tomorrow, we can't have sundowners at the Westcliff hotel.” Sarah now thinks there is no chance for sundowners at the hotel in rain. Her probability for tomorrow's rain, however, remains the same. Learning the conditional thus leaves the probability of the antecedent unchanged.

(2) The Judy Benjamin Problem

Judy Benjamin, a soldier in training, and her platoon are dropped in a swampy area which they have to patrol. The war games area is divided into the region of the Blue Army, to which Judy Benjamin and her fellow soldiers belong, and that of the Red Army. Each of these regions is further divided into Headquarters Company Area and Second Company Area. The patrol has a map, which none of

[10] Van Fraassen's conclusion is ambiguous. it is not clear whether he embraces or ridicules the unintuitive consequences of (INFOMIN).

[11] $D_{KL}(P'||P) := \sum_i \frac{P'(i)\ln(P'(i))}{P(i)}$. For further technical details of (INFOMIN), refer to Hartmann and Rad (unpublished) or van Fraassen (1981).

[12] Examples are adapted from: example 1 from Douven and Romeijn (2011), example 2 is from van Fraassen (1981), examples 3 from Douven and Dietz (2011), example 4 from Douven (2012) and example 5 from Douven and Romeijn (2011).

them understands, and they are soon hopelessly lost. Using their radio they are at one point able to contact their own headquarters. They are told by the duty officer "I don't know whether or not you have strayed into Red Army territory. But if you have, the probability is 3/4 that you are in their Headquarters Company Area." At this point the radio gives out. Supposing Judy accepts this message, how should she adjust her degrees of belief?

It seems intuitively that her degree of belief that she is in the Red Army Territory (the antecedent) should not change.

(3) The Ski Vacation Example

Harry sees his friend Sue buying a skiing suit. He is surprised because he did not know she had plans to go skiing. He knows she had an important exam a few days ago and thinks it is unlikely she passed. Then he meets Tom, his best friend and also a friend of Sue, who is just on the way to Sue to ask whether she passed the exam. Tom tells Harry: "If Sue passed the exam, her father will take her on a ski vacation." Harry recalls that Sue was buying a skiing suit, and now comes to think that it is more likely that she passed the exam. Learning the conditional thus increased his probability of the antecedent (Sue passed the exam).

(4) The Driving Test Example

Betty knows that Kevin, the son of her neighbours, was to take a driving test yesterday. She has no idea of his driving skills; she thinks it's as likely as not that he passed the test. She notices his parents have just started spading their garden. Sue's mother, who is friends with Kevin's parents, later tells her: "If Kevin passes the driving test, his parents will throw him a garden party." Sue recalls that spading has just begun and comes to think that it is very unlikely (although not fully excluded) that they will have a garden party. As a result, she thinks it is less likely that Kevin passed the test. Learning the conditional thus decreased her probability of the antecedent (Kevin passed the driving test).

(5) The Deadly Robber Example

A jeweller has been shot in his store and one of his golden watches was stolen. However, it is not clear at this point whether the two events are related. It is possible that somebody first shot the jeweller and another person took the opportunity to steal the golden watch. Kate knows her friend Henry has financial troubles and frequently walks past the store. She thinks it is likely that he is the thief, but she highly doubts he would shoot the jeweller. After the investigation of the scene, policemen conclude that a single person committed the crime. A detective comes to Kate and tells her: "If Henry stole the watch, then he also shot the owner." As a result, Kate now thinks it is very unlikely he stole the watch. Learning the conditional thus decreased her probability of the antecedent.

The examples thus show that learning a conditional may influence the antecedent in three different ways: one's degree of belief in the antecedent may (i.) increase, (ii.) decrease, or (iii.) remain the same.

I will first quickly examine whether the mentioned approaches to learning from conditionals get the intuitively expected results.

The Sundowners example

According to (EXP), after learning the conditional, i.e. setting the conditional probability $Pr(\neg sundowners|rain)\approx 1$, the explanatory status of the antecedent does not change, and Adams' conditioning is applied. Applying (AC) always results in unchanged probability of the antecedent, so (EXP) gets the intuitively expected results (Douven 2012).

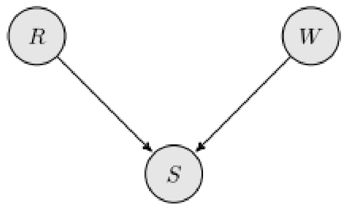

Figure 1: The Bayesian network for the sundowners example (Hartmann and Rad)[13]

(INFOMIN+CAU) also leads to expected results, although in a slightly more complex way. Hartmann and Rad, for instance, propose the following Bayesian network (Figure 1), which in their words "properly represents the causal representation [of the variables]". They then make sure that all the events are probabilistically dependent with regards to the network, adapt probabilities to the story from the example, and minimise the Kullback-Leibler divergence to get the correct results.

One of the objections I hold against their analysis of the Sundowners example is that in their collider

[13] R denotes rain, W the wedding party and S the sundowners at the hotel.

model (Figure 1), the variable *W* is independent of *R*, although Marian is told from the hotel staff that the wedding taking place inside the hotel depends on the rain. To be fair, Hartmann and Rad actually define *W* as a binary variable with two values, W: “There is a wedding party” and ¬W: “There is no wedding party” Thus defined *W* is obviously independent of *R*.

But is the correct result still achieved if we define *W* as a variable with two values, W: “There is a wedding party inside the hotel” and ¬W: “There is no wedding party inside the hotel” and add the dependence of *W* on *R* (Figure 2)?

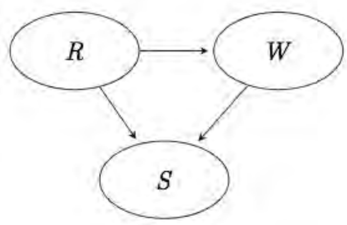

Figure 2: Alternative Bayesian network for the sundowners example

It turns out that if the Bayesian network is represented in this way, and if the Kullback-Leibler divergence is minimized, Sarah's posterior probability of tomorrow's rain decreases after learning the conditional “If it rains tomorrow, we can't have the sundowners at the hotel.”[14] This does not refute (INFOMIN+CAU), but rather shows that more detailed criteria on how to properly represent situations is needed.

The Judy Benjamin Problem

(EXP) proceeds similarly in the Judy Benjamin Problem, although the probability of the conditional is not close to 1, but .75. The reason for a similar procedure is that learning the conditional “If you are in the Red Army area (R), the odds are 3:1 that you are in their Headquarters Area (RH),” i.e. setting $Pr(RH|R)=0.75$, does not change the explanatory status of the antecedent. This means that Adams' conditioning is applied and the posterior probability of being in the red area remains the same, .5.

The results of (INFOMIN+CAU) are similar, although the procedure is, again, more complex. According to Hartmann and Rad, there is no direct causal connection between being in the Red Army area and being in their Headquarters Area. The only possibility is, therefore, that both events have a common cause *X*. It is not necessary to determine what the common cause is, as it does not have any influence on their procedure, which after minimising the Kullback-Leibler divergence, again it yields the expected results.

The Ski Vacation and The Driving Test examples

The Ski Vacation (SV) and The Driving Test (DT) examples intuitively require the agents to adjust their degrees of belief in the antecedent in two opposite ways: either to increase it (SV) or decrease it (DT). Both examples, however, turn out to be deeply related on a structural level.

According to (EXP), learning the conditional makes the explanatory status of the antecedent (Sue passed the exam) go up in (SV), while it goes down (exactly the opposite) in (DT). Posterior degrees of belief in the antecedent in the two examples thus adjust in two opposite directions: it increases in (SV) and decreases in (DT).

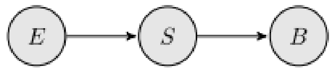

Figure 3: The Bayesian network for the ski vacation example

Structural similarity in the two examples is also noted in (INFOMIN+CAU). Hartmann and Rad thus propose structurally identical models: the event from the consequent depends on the antecedent, and the additional observed event further depends on the consequent. They get the expected results after minimising the Kullback-Leibler divergence. While the proposed network for (SV) (Figure 3) seems obviously appropriate, the proposed network for (DT), on the other hand, does not. It is not hard to imagine that spading the garden prevents the garden party, especially considering it temporally precedes the garden party. Such representation of the situation would then result in a collider network (like in Figure 1) and, likely, in wrong results. An examination of the consequences of such a Bayesian network is omitted because of its complexity.

The Deadly Robber Example

The Deadly Robber Example (DR) was originally provided as an example of epistemic entrenchment, which was not further developed into a general approach to belief updating on conditionals. It is, however, not hard to envision how (EXP)

[14] Proof omitted because of its complexity and length.

would solve (DR) and get the correct results. After learning that "If Henry stole the watch, he also shot the owner," the explanatory status of the antecedent goes down, and as a result, Kate finds it less likely that Henry was the culprit.

It is not clear what the appropriate representation of the problem would be in (INFOMIN+CAU), although it seems reasonable that both clauses would have a common cause. The problem would then be similar to the Judy Benjamin Problem, although it is intuitively not acceptable that the degree of belief in the antecedent remains unchanged.

As I have shown, both (EXP) and (INFOMIN+CAU) provide intuitively expected results in all the examples. But look at what happens in the modified driving test example:

(6) The Modified Driving Test Example

Betty knows that Kevin, the son of her neighbours, was to take a driving test yesterday. She knows he has been illegally driving since he was 12 and <u>thinks it is very likely that he passed the test</u>. She notices his parents have just started spading their garden. Sue's mother, who is friends with Kevin's parents, later tells her: "If Kevin passed the driving test, his parents will throw him a garden party." Sue has two possibilities, depending on the strength of her belief that Kevin passed the test: she could either decrease her degree of belief in the antecedent (like in the original example), or keep her belief in the antecedent unchanged and increase her degree of belief in the consequent, i.e. come to think that the garden party will take place despite spading.

According to both (EXP) and (INFOMIN+CAU) it is, however, only possible that Sue decreases her degree of belief in the antecedent. Problems of this nature motivated my proposed approach, which seems to overcome them.

5. Learning from conditionals revised

As is clear from the previous cases, when one learns a conditional and actually learns something new, then she keeps, increases, or decreases her prior degree of belief in the antecedent. But the antecedent only provides half of the story. In the case that the probability of the antecedent increases or decreases, the probability of the consequent remains unchanged. In the case that the probability of the antecedent remains the same, the probability of the consequent increases or decreases. There are therefore not just 3 possibilities (ant. unchanged, increased, or decreased), but 4.

After learning "If P, then Q," the degrees of belief in the antecedent and the consequent may adjust in the following 4 ways:

Learning "If m, then n."	The degree of belief in the antecedent m	The degree of belief in the consequent n
a)	does not change: $Pr'(m) = Pr(m)$	increases: $Pr'(n) > Pr(n)$
b)	does not change: $Pr'(m) = Pr(m)$	decreases: $Pr'(n) < Pr(n)$
c)	increases: $Pr'(m) > Pr(m)$	does not change: $Pr'(n) = Pr(n)$
d)	decreases: $Pr'(m) < Pr(m)$	does not change: $Pr'(n) = Pr(n)$

Table 2: Possible cases after learning from conditionals

As is clear from Table 2, strength of belief in one of the clauses always remains unchanged (antecedent in lines a) and b); consequent in c) and d)). The clause which retains the same strength of belief is epistemically more entrenched. My aim is to provide guidelines that would enable a determination of which of the clauses is more entrenched.

5.1 Epistemic entrenchment

But what is epistemic entrenchment (EE) after all? The term originates from AGM belief revision theory, although it can also be applied to partial beliefs. Douven and Romeijn (2011), for instance, claim that Kate from the Deadly Robber Example finds it easier to give up her partial belief in Henry's theft because her (partial) belief that he is not a killer is more entrenched.

I claim that one of the clauses is always more entrenched than the other, i.e. degree of belief in this clause remains unchanged. But why and when is one clause more entrenched than the other? I will try to answer this in the next section.

5.2 What can we learn from the examples?

When one is told something in the conditional form, she does not just (i) *explicitly* learn that the consequent will occur in the event of the antecedent, but (ii) (in some cases) also *implicitly* learns something about other events.

The Sundowners Example serves as a good illustration of implicit learning from conditionals: when Marian tells Sarah that they cannot have sundowners if it rains tomorrow, Sarah implicitly infers that the inside area of the hotel will be occupied. Otherwise Marian's conditional would not make sense, as it is clear from the story that she initially thinks they could have their sundowners inside in the case of rain. This observation then plays an important role in the actual learning from the conditional.

Another observation that can be taken from the examples is that learning from conditionals depends on how coherent the newly learned conditional is with not just previous (partial) beliefs in its clauses, but also with other beliefs that support them.

For instance, Betty from the driving test example has strong *support* for her belief that the garden party is unlikely (she observed the spading of the garden). Her belief about Kevin's (non)success with the test is, on the other hand, not supported at all, because she has no knowledge about his driving skills.

Finally, which belief is more entrenched depends on its strength. Strength of belief does not coincide with the degree of belief, although it is determined by it. I define strength in a simple and straightforward manner: beliefs of degree .5 (it is as likely as not) are very weak. Strength of beliefs with degrees lower or higher than .5, on the other hand, symmetrically increases.

This, again, can easily be illustrated with the driving test example. Betty's prior degree of belief in Kevin's success was exactly .5 (she thought it was as likely as not that he passed the test), and thus had no strength, while her belief that it is unlikely a garden party will take place was very strong, and hence more entrenched.

5.3. The method

These three observations are also at the core of the presently proposed model of learning from conditionals.

Step 1: When one learns "If *M*, then *N*," she explicitly learns that *N* follows from *M*, i.e. *N* can be inferred from *M*.

Step 2: Check if the conditional conveyed any implicit information. The simplest way to do so is to answer the following question from the point of view of the learning agent: Could the consequent follow from the antecedent alone? If not, what other conditions would also need to be fulfilled? These implicit additional conditions are thus also learned.

Step 3: How (if at all) are beliefs in the antecedent and the consequent supported? Further, if the supporting beliefs of the consequent are in conflict with the antecedent, its supporting beliefs, or additional implicitly learned conditions, then this casts doubt on them (in the context of the conditional), and their strength is reduced to zero.

If the implicitly learned additional conditions are, however, in conflict with the consequent or its supporting beliefs, this similarly casts doubt on them and reduces their strength to zero.

Step 4: Determine the degree of belief in both clauses with regards to support and then determine their strength, with the following simple entrenchment function:[15]

$$S(m) = \begin{cases} -2Pr(m) + 1, Pr(m) \in [0, .5) \\ 2Pr(m) - 1, Pr(m) \in [.5, 1] \end{cases}$$

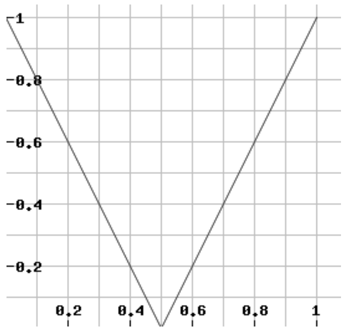

Figure 4: Graph of the simple entrenchment function

Step 5: Compare the strength of the antecedent and the consequent. This leaves three possibilities:

i.) $S(ant) > S(cons)$
ii.) $S(ant) < S(cons)$
iii.) $S(ant) = S(cons)$

In (i) the antecedent is stronger (hence, more entrenched), and thus remains unchanged after learning the conditional, in (ii) the consequent is more entrenched, and in the rare event of (iii) the antecedent is more entrenched (*).

Step 6: I now know which belief remains unchanged. Whether the other increases or decreases depends on their prior degrees of belief:

If $S(ant) > S(cons)$,
then if $Pr(ant) > .5$,
$Pr(ant) \geq Pr'(cons) > Pr(cons)$
Else, i.e. $Pr(ant) < .5$,
$Pr(ant) \leq Pr'(cons) < Pr(cons)$

[15] The range of strengths, i.e. interval [0,1], was motivated by probabilities, which are spread out in the same range. The function could be more generally defined as $S(m) = \begin{cases} -x\, Pr(m) + \frac{x}{2}, Pr(m) \in [0, .5), x \in \mathbb{R} \\ x\, Pr(m) - \frac{x}{2}, Pr(m) \in [.5, 1], x \in \mathbb{R} \end{cases}$

If $S(ant)<S(cons)$,

then if $Pr(cons)>.5$,

$$Pr(cons)\geq Pr'(ant)>Pr(ant)$$

Else, i.e. $Pr(cons)<.5$,

$$Pr(cons)\leq Pr'(ant)<Pr(ant)$$

In the case in which beliefs in both the antecedent and the consequent are of the same strength (*), the antecedent is entrenched, as the agent explicitly learns that the consequent depends on the antecedent.

5.4. The Examples Revisited

Let me now demonstrate how the proposed method works "in action." I will show that it leads to intuitively expected results not just in the examples from the literature, but also in modified examples where other approaches fail, e.g. the modified driving test example.

The Sundowners Example

What does Sarah learn when her sister tells her "If it rains tomorrow (*R*), we can't have sundowners at the hotel. (¬*Su*)"?

She first explicitly learns that (¬*Su*) can be inferred from (*R*). She then considers whether (*R*) entails (¬*Su*). It does not: Sarah thinks they could still have sundowners inside the hotel if it rains. Sarah thus abductively infers that the inside area of hotel will not be available.

Further, her low degree of belief in (¬*Su*) is supported by her background knowledge that they can have the sundowners either outside or inside the hotel *and* that both areas are very rarely occupied. But Sarah also implicitly learned that the inside area of the hotel will not be available if it rains (or at least one area otherwise). This is in conflict with her support for low degree of belief in (¬*Su*): she now thinks they can only have the sundowners outside. Strength of her belief in (¬*Su*) is reduced to 0 and her $Pr(\neg Su) = .5$

As $S(R)\geq S(\neg Su) = 0$, the posterior degree of belief in (*R*) remains unchanged. If she thinks rain is unlikely (although not impossible), her posterior $Pr'(\neg Su)$ lies in this range:

$$Pr(R)=Pr'(R)\leq Pr'(\neg Su)<Pr(\neg Su)=.5$$

If, however, Sarah thinks it's likely that it will rain, she increases $Pr'(\neg Su)$:

$$Pr(\neg Su)<Pr'(\neg Su)\leq Pr(R)= Pr'(R).$$

The results are intuitively correct.

The Judy Benjamin example

When Judy learns that "If you are in Red Army area, the odds are 3:1 that you are in their Headquarters Area," she does not learn anything implicitly. Her beliefs for being either in the Red Army area or being in the Headquarters Area (or Red Headquarters area, given that she is in the Red Army area) have no support, and hence no strength: Judy and her platoon are lost and think it's as likely as not that they are in any quadrant of the equally divided area (Figure 5).

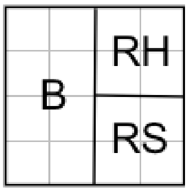

Figure 5: Judy Benjamin's initial mental map

As both clauses have the same strength, the antecedent is more entrenched. If the conditional information was provided with odds 1:0, i.e. "If you are in the Red Army area (*RH* v *RS*), then you are in their Headquarters Area (*RH*)," she would increase $Pr(RH)=.25$ to $Pr'(RH)=.5$. As the conditional is only partial (with odds 3:1 or, equivalently, probability of .75), this needs to be taken into account, so: $Pr'(RH)=.5*.75=.375$. As (*R*) is entrenched, $Pr'(R)=Pr(R)=.5$. This is also what one would expect.

It must be noted, though, that this may not be the only possibility. As Lukits (2014) pointed out, intuitions are not always correct. Judy Benjamin has more possibilities, depending on which of her prior beliefs is entrenched.

She starts with prior beliefs $Pr(R)=Pr(B)=.5$ and $Pr(RH)=Pr(RS)=.25$.

If she entrenches her belief in (*R*), then the above results are correct. But she could also entrench her belief in (*RH*) and still update according to the conditional:

$$Pr(RH)=3/12,\ Pr(RS)=1/12, Pr(B)=8/12.$$

In this case, her degree of belief in the antecedent would decrease from .5 to 1/3.

She could also entrench her belief in (*RS*):

$$Pr(RS)=1/4,\ Pr(RH)=3/4, Pr(B)=0.$$

In this case the probability of the antecedent (*R*) would increase to 1.

The odds for being in (*RH*) are 3:1 in comparison to (*RS*) in all of the above possibilities, and it is not impossible to imagine cases when this could be true.[16] For example, the last possibility could make sense if the blue area is fully occupied by Red Army, and thus does not exist any more. It is not clear why the radio operator would then use a conditional, although it is a possibility.

All of the alternative possibilities, however, depend on supporting beliefs which are not available to Judy, but rather to the radio operator. Judy thus correctly entrenches (R). But if she had different supporting beliefs, her belief updating wouldn't proceed in the same way. My proposed account is thus resistant to Lukits (2014) counterexamples.

Ski vacation example

When Tom tells Harry "If Sue passed the exam (E), her father will take her skiing. (Sk)," Harry doesn't learn anything implicitly, as (Sk) could follow from (E). His degree of belief in the success of her exam is low, which is supported by her past results (known to Harry), her testimony, or by other sources. It is important to note that it is not in conflict with Harry's support for partial belief in (Sk).

Which belief is stronger—(Sk) or (E)—depends on how "extreme" they are. Does Harry think it is more likely that Sue's father will take her skiing because she was buying a skiing outfit, or that it is more *un*likely that she passed the exam? Both options seem possible:

1.) If Harry thinks Sue had almost no chance of passing the exam, and thinks that it is somewhat likely that she was just wishfully buying her skiing outfit, then $S(E)>S(Sk)$ and as $Pr(E)<.5$:

$$Pr(E)=Pr'(E)\leq Pr'(Sk)<Pr(Sk)$$

2.) If Harry initially thinks it's very likely that Sue's father will take her skiing since she was buying the skiing outfit, and thinks it's only somewhat likely that she didn't pass the exam, then $S(E)<S(Sk)$ and as $Pr(Sk)>.5$:

$$Pr(E)<Pr'(E)\leq Pr'(Sk)=Pr(Sk)$$

The results of our method are thus in line with intuitions, but also cover the special case, which is not covered either by (EXP) or (INFOMIN+CAU).

Driving test example

When Betty learns "If Kevin passed his driving test (T), his parents will throw him a garden party (G)," she doesn't learn any implicit information—passing the driving test would suffice for Kevin's parents to throw him a garden party.

Betty initially doesn't have any support to believe he actually passed the test, so her $Pr(T)=.5$ and $S(T)=0$. She has strong support to think a garden party is unlikely (she noticed the spading of the garden had just begun), so her $Pr(G)<.5$, and thus $S(T)<S(G)$. As $Pr(G)<.5$:

$$Pr(G)=Pr'(G)\leq Pr'(T)<Pr(T) \qquad \text{(DT-1)}$$

As we expected, she learns that it is unlikely that Kevin passed the test.

If Betty, on the other hand, has strong support for (T), and hence $Pr(T)>.5$, as in the modified example, then everything would depend on which belief is stronger, i.e. whether (i) $Pr(T)$ is closer to 0 or (ii) $Pr(G)$ to 1.

In case (i) $S(T)>S(G)$ and as $Pr(T)<.5$:

$$Pr(G)<Pr'(G)\leq Pr(T)=Pr'(T)$$

Betty now comes to think a garden party is likely despite the spading.

In case (ii), the results are the same as in (DT-1) above. Both results seem to be in line with intuitions.

The Deadly Robber Example

I conclude the demonstration of the proposed method with the deadly robber example. When Kate learns "If Henry stole the watch (W), he also killed the owner (K)," she implicitly learns that the culprit was a single person (SP). Otherwise she couldn't infer (K) from (W).

Kate thinks (W) is likely. This is supported by her belief that Henry had financial troubles and that the thief was not necessarily also the killer. The additional condition, (SP), is in conflict with her supporting belief for (W), which reduces the strength of (W) to 0. Her belief in (K) is very low, which is supported by her knowledge of her friend Henry and his character. Either way, $S(K)>S(W)=0$.

As $Pr(K)<.5$:

$$Pr(K)=Pr'(K)\leq Pr'(W)<Pr'(W)$$

The result is again in line with intuitions. This example is interesting, because the implicitly learned information undermines the strength of the antecedent. Exactly the reverse was the case in The Sundowners Example.

6. Conclusion

My proposed approach is far from being the final answer to the problem of learning from conditionals. After all, I limited myself to simple conditionals, and it is not clear what would happen in the

[16] For further discussion refer to Lukits (2014).

case of nested conditionals. This needs to be considered in the future.

Another aspect that deserves more attention and is usually overlooked in the discussion of updating on conditionals is their social nature. One can learn something from her perceptions, reasoning, introspection, or testimony.[17] However, all of the conditionals from the examples were learned through testimony, and as such all the examples also touch on questions of social epistemology. For instance, why should the recipient of the conditional trust the sender?

Does this reveal something about the epistemological nature of conditionals? Do I learn anything if I come to know about conditionals from my own reasoning? One could explain this phenomenon (away) by claiming that all the conditionals from the examples were testimonial because otherwise the stories would be, mildly said, modernistic streams of thoughts, or simply boring. This seems a bit rash. I will leave this question open—as it is not of direct relevance—and simply emphasize that (based on the examples) learning from conditionals seems to be essentially a product of testimony.

Another important question that was excluded is how learning from counterfactuals proceeds. After all, counterfactuals are important in philosophy of science and science in general. For instance, causal relations are often described with counterfactuals: "If *A* had not occurred, *B* would still occur." One learns that B is causally independent of A, but the procedure is obviously much different from our proposed approach. After all, neither *A* nor *B* are more likely after we get to know this counterfactual.

It would also be insightful to see how people psychologically learn from conditionals, especially in cases where multiple outcomes seem possible (for instance, in the "normal" and modified driving test example and in the ski vacation example). This would allow for a check of whether the empirical data supports this method.

It is thus clear that my proposal is still in the early phase and has a lot of space for growth in different directions, but still provides (intuitively) correct results not just in the cases already addressed by other approaches, but also in some cases where other accounts do not provide the expected results.

References

Bennett J (2003) *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press.

Dancygier B, Sweetser E (2005) *Mental Spaces in Grammar: Conditional Constructions*. Cambridge: Cambridge University Press.

Douven I (2008) The evidential support theory of conditionals. *Synthese* 164(1): 19–44

Douven I (2012) Learning conditional information. *Mind and Language* 27(3): 239–263

Douven I (2013) The epistemology of conditionals. In *Oxford Studies in Epistemology* eds T S Gendler, J Hawthorne. Oxford: Oxford University Press.

Douven I, Romeijn J (2011) A new resolution of the Judy Benjamin problem. *Mind* 120: 637–670

Douven I, Dietz R (2011) A Puzzle About Stalnaker's Hypothesis. *Topoi* 30:31–37

Gibbard A (1981) Two recent theories of conditionals. In eds W L Harper, R Stalnaker, G Pearce, 211–247. Dordrecht: Ifs. Reidel.

Hájek A (2012) The fall of "Adams' thesis"?. *Journal of Logic, Language and Information* 21(2): 145–161

Krzyżanowska K (2015) Between if and then. Dissertation, University of Groningen.

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

Lukits, S (2015) Maximum Entropy and Probability Kinematics Constrained by Conditionals. *Entropy* (17): 1690–1700

Sanford D H (1989) If P, then Q. New York: Routledge.

Van Fraassen, B C (1981) A Problem for Relative Information Minimizers in Probability Kinematics. *The British Journal for the Philosophy of Science* 32(4): 375–379

[17] Memory is omitted as it merely retains knowledge. Although there are counterarguments to this, the epistemological nature of memory does not interest us for present purposes.