Topoi (2007) 26:115–132
DOI 10.1007/s11245-006-9003-7

Beliefs in conditionals vs. conditional beliefs

Hannes Leitgeb

Published online: 20 March 2007

Abstract On the basis of impossibility results on probability, belief revision, and conditionals, it is argued that conditional beliefs differ from beliefs in conditionals *qua* mental states. Once this is established, it will be pointed out in what sense conditional beliefs are still conditional, even though they may lack conditional contents, and why it is permissible to still regard them as beliefs, although they are not beliefs in conditionals. Along the way, the main logical, dispositional, representational, and normative properties of conditional beliefs are studied, and it is explained how the failure of not distinguishing conditional beliefs from beliefs in conditionals can lead philosophical and empirical theories astray.

Keywords Belief · Conditional · Belief revision · Conditionalization · Ramsey

1 Introduction—the role of logic in the philosophy of cognition

In our cognitive lives, beliefs occupy a very peculiar systematic position. They can neither be characterized as belonging exclusively to the causal world of nature, nor do they belong exclusively to the normative domain of rationality; they rather embrace both of them.

On the one hand, beliefs are states that can cause, or can be caused by, other states or events, much in the same way as the impact of a body may cause a planet to abandon its orbit, a weird dream may be caused by the combination of too much coffee and an obsession with an unsolved mathematical problem, and a colour sensation may be caused by the presence of a coloured object. On the other hand, it seems essential for beliefs also to have propositional contents that are intended—by the agent who has these beliefs—to be true descriptions of the world. This allows us to describe beliefs, and methods of acquiring and changing beliefs, as either rational or irrational, in contrast with planetary impacts, dreams, and colour sensations, which could only be classified sensibly as rational or irrational by some huge stretch of imagination.

Since it is constitutive of beliefs and belief-forming processes that they have certain normative properties, *logic*, insofar as it is a discipline that yields partial information about what we are permitted to believe, may actually improve our understanding of the *ontology* of beliefs, i.e. what kind of entities beliefs are.

In this essay we shall draw especially on insights into logical systems for belief revision, probability, and conditionals, in order to argue that there are two different types of beliefs of ''conditional character'': *beliefs in conditionals* and *conditional beliefs*. It will turn out to be logically possible to have a conditional belief in B given A without having a corresponding belief in the conditional $A \Rightarrow B$. Or formulated differently: a conditional belief in B given A differs from a belief in (B given A). Once this is established, we will try to clarify in what sense conditional beliefs are still conditional, even though they usually lack conditional contents, and why they are still beliefs, although they are not beliefs in conditionals. Along the way we will track the cognitive traits of the two sorts of beliefs, and we will show how failure to distinguish them can easily lead philosophical and empirical theories astray.

H. Leitgeb (✉)
Department of Philosophy, University of Bristol, 9 Woodland Road, Bristol BS8 1TB, UK
e-mail: Hannes.Leitgeb@bristol.ac.uk

We do not claim these observations about beliefs in conditionals vs. conditional beliefs to be in any sense original: in fact, various authors have put forward such a distinction or something closely related to it, including Adams' (1965, 1975) seminal probabilistic account of conditionals, and more specifically Stalnaker (1984, chap 6), Edgington (1986, 1995), Gärdenfors (1988, sect 7.7), Levi (1988, 1996), and Mellor (1993) on conditional belief. However, it seems that the topic has not been studied intensively so far at all, and that the existence of (conditional) beliefs, which are neither true nor false and which have two propositional contents rather than just one, certainly has not made it into the common knowledge of philosophers of mind or cognitive scientists (if it is not disputed anyway).[1] In the literature on belief revision, the notion of conditional belief is indeed sometimes used, but often without distinguishing it properly from the notion of belief in a conditional. Furthermore, while the philosophers cited above are primarily interested in conditional beliefs as far as they are relevant for the semantics and pragmatics of conditional *sentences*, we want to focus on beliefs in conditionals and conditional beliefs *qua beliefs*, i.e. insofar as they differ as cognitive states, and we want to see what different roles they occupy in our mental architecture. As we will try to show, it is to some extent possible to deal with the latter topic independently of, and prior to, the former; the much more complex and difficult problem of mapping our linguistic practices of asserting conditionals (of various types, in various sorts of contexts) onto the mental structure of beliefs in conditionals/conditional beliefs that we are going to concentrate on will be left to be taken up by someone else.

2 Rational belief change

We are in the fortunate position to build on two well-established accounts of how to describe rational belief change: the one uses subjective conditional probability measures, the other one refers to so-called belief revision operators.[2] Since for the purposes of this paper their differences prove to be unimportant—indeed, the latter can be seen as the qualitative version of the former—we will focus mainly on belief revision. But it is instructive to start by reconsidering both of them.

According to the probabilistic or Bayesian approach (cf. Howson and Urbach 1993), an agent's belief state is represented by a subjective probability measure P which is defined on the possible propositional contents of belief; for every proposition B, the probability $P(B)$ is interpreted as the agent's degree of belief in B. Given some new evidence A, it is claimed to be rational for the agent to revise P by conditionalizing on A, i.e. for every proposition B her new degree of belief is the conditional probability

$$\mathrm{P(B/A)} \tag{1}$$

The theory of belief revision (cf. Gärdenfors 1988), on the other hand, considers an agent's belief state as being determined by two components: a logically closed "belief set" K of propositions, namely the set K of propositions that the agent currently believes in, and a revision operator $*_K$ that maps propositions to logically closed sets of propositions [following common usage we shall write "$K * A$" for "$*_K(A)$"]. It is now said to be rational for the agent to believe B in light of the new evidence A if and only if

$$B \in K * A \tag{2}$$

Of course, both probability measures P and revision functions * are defined to conform to various rationality postulates. We will not discuss these postulates in any detail, as it is not our goal to analyse or question them; it suffices to say that these postulates demand changes of belief to be "logically coherent" and that we take their validity for granted. What we are actually interested in is how these principles interfere with the doxastic representation of conditionals, where at this point we want to leave open the type of conditionals (indicative, subjunctive, otherwise?) that we are going to deal with; hence, "⇒" will be used ambiguously, its logical meaning being dependent on the context.

3 Belief change and conditionals

Given the intended reading of conditional probabilities and revision operators—"given A, the agent's degree of belief in B is so and so," "given A, the agent believes B"—it is almost inevitable that an interpretation of belief change in terms of conditionals is put forward.

[1] For example in the forthcoming entry on beliefs in the *Stanford Encyclopedia of Philosophy*, conditional beliefs do not receive any attention at all.

[2] Spohn (1988) gives another account of rational belief change which lies somewhere "in between" the probabilistic and the belief revision approach. Points very similar to the ones developed below could also be made in terms of Spohn's framework of ranking functions.

In the probabilistic case, the most famous such suggestion is known as the *conditional construal of the conditional probability (CCCP)*,[3] according to which the probability of a conditional $A \Rightarrow B$ is simply identical to the conditional probability of B given A, i.e.

$$P(A \Rightarrow B) = P(B/A)^4 \quad (3)$$

The counterpart proposal on belief revision is called the *Ramsey test*[5] for conditionals and says that a conditional $A \Rightarrow B$ is believed rationally if and only if the revision by A yields a belief in B, i.e.

$$A \Rightarrow B \in K \text{ iff } B \in K * A \quad (4)$$

Let us, tentatively, call $P(A \Rightarrow B)$ the *degree of belief in the conditional* $A \Rightarrow B$, $P(B|A)$ the *degree of conditional belief in B given A*, the mental state that is represented by the membership of $A \Rightarrow B$ in K the *belief in the conditional* $A \Rightarrow B$, and finally the mental state that is represented by the membership of B in $K * A$ *the conditional belief in B given A*. What the two theses above express can thus be interpreted in terms of the following two slogans: degrees of belief in conditionals are degrees of conditional beliefs, and beliefs in conditionals are conditional beliefs, where these identities may be assumed to hold either extensionally or necessarily, depending on the assumed status of Eqs. 3 and 4. Even if these slogans are not taken to *circumscribe* what Eqs. 3 and 4 amount to, we might still adopt the identifications they express as working assumptions from which Eqs. 3 and 4 can be derived (though perhaps not vice versa). While such identifications would leave us with a particularly clean and perspicuous doxastic account of conditionals, it turns out to be rationally impossible to follow any such approach: David Lewis (1976) showed that under certain non-triviality assumptions, CCCP cannot hold for a probability measure and all propositions A and B (where the interpretation of $\Rightarrow$ is regarded as fixed and thus independent of the chosen probability measure). Analogously, Peter Gärdenfors (1986, 1988) proved that, given certain non-triviality assumptions, there does not exists a belief revision operator such that for all belief sets K and all propositions A and B the Ramsey test for conditionals is satisfied (the interpretation of $\Rightarrow$ is again considered to be fixed). In both cases, ''non-triviality'' means that there are at least three possible but mutually incompatible propositions that are contained in the domain of the probability measures/that are consistent with a belief set for the belief revision operators. As further elaboration on Lewis' result made clear[6] —and similar refinements apply to the belief revision side—even stronger impossibility theorems can be derived which show that CCCP/the Ramsey test is bound to fail even for various ''naturally'' restricted classes of probability measures/belief sets and for various ''naturally'' restricted classes of propositions.

We conclude that for any rational agent, it *cannot* generally be true that degrees of beliefs in conditionals are degrees of conditional beliefs, and it *cannot* generally be true that beliefs in conditionals are conditional beliefs, by proof by contradiction.[7]

How about agents that are not completely rational with respect to their (degrees of) beliefs and their ways of revising (degrees of) beliefs? Perhaps those agents' degrees of beliefs in conditionals can coincide with degrees of conditional beliefs, and their beliefs in conditionals can coincide with conditional beliefs? Maybe—Eqs. 3 and 4 might hold contingently for such agents. But even if this were so, the argument from above would still show that beliefs in conditionals and conditional beliefs are ontologically different: as the existence of rational agents is certainly *possible*, it is at least *possible* that some agent's degrees of beliefs in conditionals are not generally identical to degrees of conditional beliefs, and her beliefs in conditionals are not generally identical to conditional beliefs. This possibility is sufficient to show that beliefs in conditionals and conditional beliefs are distinct *qua states*, even if we happen to live in a world in which there is no rational agent at all, and where by some lucky coincidence every agent with a belief in the conditional

[3] See Hájek and Hall (1994). One benefit of CCCP would be to yield an explanation of Adams' thesis (Adams 1965), which is usually interpreted as the claim that the degree of assertability of an indicative conditional $A \Rightarrow B$ is measured by the conditional probability $P(B|A)$. Under the additional assumption that the degree of assertability of $A \Rightarrow B$ is identified with the probability $P(A \Rightarrow B)$, Adams' thesis would even collapse into CCCP.

[4] In order to keep things as simple as possible, we suppress all obvious, though necessary, qualifications of formal statements, such as in this case: *for all A with* $P(A) > 0$ *it holds that*....

[5] See Ramsey (1929), and sect 7.1 of Gärdenfors (1988). Lindström and Rabinowicz (1995) give an excellent survey of the Ramsey test and of the literature about it. We leave open whether what is now understood as ''the'' Ramsey test in the relevant literature actually conforms Ramsey's own view about conditionals.

[6] See, e.g.: the second triviality result in Lewis (1976); Lewis (1986); Hájek and Hall (1994); Hájek (1994); Hall (1994); Milne (1997); Bradley (2000).

[7] Since a conditional belief is not a belief in a conditional, it is a fortiori also not an *implicit* belief in a conditional. So the pair *belief in a conditional* vs. *conditional belief* does not match the *explicit belief* vs. *implicit belief* distinction, if the latter are assumed to be beliefs with the same contents.

$A \Rightarrow B$ also has the corresponding conditional belief in B given A, and vice versa.

4 Beliefs and conditionals: looking back

Before we start developing our theory of conditional beliefs, thereby arguing that conditional beliefs are not misnamed as such, i.e. that they are indeed *beliefs* in the proper sense of the word, we will turn briefly to the literature on this topic so far.

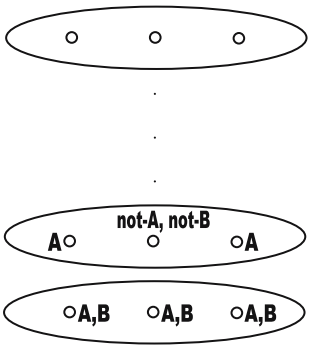
Gärdenfors (1988), on the two concluding pages of chap 7 (pp 165–166), is very explicit about the implications of the Ramsey test with respect to beliefs: ''The Ramsey test has been a leitmotiv for the analysis of conditionals in this chapter. It can be interpreted as saying that beliefs in conditionals are nothing but conditional beliefs.'' He regards his theorem as evidence against the right-to-left direction of Eq. 4, but does not continue to explore the properties of conditional beliefs any further.

Interestingly, he also states a counterexample (borrowed from Stalnaker 1984, p 105) to the left-to-right direction of Eq. 4:

> Suppose I accept that if Hitler had decided to invade England in 1940, Germany would have won the war. Then suppose I discover, to my surprise, that Hitler did in fact decide to invade England in 1940 (although he never carried out the plan). Am I now disposed to accept that Germany won the war? No, instead I will give up my belief in the conditional.

This example gives us another reason for not identifying beliefs in conditionals with conditional beliefs.

On the other hand, it would be wrong to think that there are no entailment relations whatsoever between having beliefs in conditionals and having corresponding conditional beliefs. Let us examine this in terms of Grove's (1988) sphere models, by which belief revision operators can be represented correctly and completely on the basis of ranked possible worlds. Any such model consists of a sequence of layers or spheres of worlds; by convention, we assume the layer of lowest rank to consist of the worlds with the highest subjective plausibility of being the actual world, seen from the viewpoint of the agent. The agent believes A according to such a spheres model if and only if all worlds of least rank satisfy A. The so-called belief set K as being given by a spheres model includes all sentences the agent believes in with respect to the model. Finally, it holds that $B \in K * A$, where the revision operator * is defined by the spheres model, if and only if all minimal A-worlds are B-worlds, i.e. among all the worlds that satisfy A the worlds with minimal rank (''highest plausibility'') satisfy B. In our terminology, the latter is the case if and only if the agent has the conditional belief in B given A, relative to the spheres model in question. For example in the following model the agent believes that $A \vee B$ while at the same time believing $\neg B$ given $\neg A$ (this is independent of whether the A-worlds in the second layer from below are B-worlds or not):



By considering such models, we can easily derive that if all minimal A-worlds are B-worlds, then the material implication $A \rightarrow B$ must be true in all worlds of lowest rank, i.e. if a (rational) agent has a conditional belief in B given A, she also believes in the truth of the material implication $A \rightarrow B$. On the other hand, if the agent believes in $A \rightarrow B$, *and does not believe* $\neg A$, then the A-worlds of minimal rank must be of lowest rank and thus must be B-worlds, i.e. the agent has the conditional belief in B given A (it is easy to see that it is necessary to add the italicized conjunct). Hence, there are indeed certain implication relations between conditional beliefs and beliefs in material conditionals. Furthermore, as we will see later, the externalist account of *justification* of conditional beliefs ties their justificatory status to the *truth* of their corresponding material implications.

Stalnaker (1984, chap 6) intends to show why ''...conditional belief must diverge from belief in conditional propositions. The upshot is that the explanation of one in terms of the other cannot be as simple and straightforward as one might wish'' (p 103). However, while he does not assume that conditional beliefs are necessarily beliefs that some proposition is true, he also does not exclude this possibility. In his view, ''differences in dispositions to change beliefs are always grounded in differences in beliefs about the way the world is...but this does not imply that conditional beliefs must always themselves be beliefs about the way the world is'' (p 106). In turn, conditional

propositions are regarded as projections of epistemic strategies onto the world.[8] Once again, a theory of conditional beliefs *qua* beliefs is not worked out. We will follow Stalnaker's suggestion of distinguishing conditional beliefs from beliefs in conditionals, but it will be an integral part of our theory to claim that a conditional belief is not a belief in a unique proposition of whatever sort.

Edgington (1986, 1995) defends the view that conditionals do not express propositions and do not have truth conditions. As far as beliefs and conditionals are concerned, she seems to oscillate between two positions (as pointed out by Lowe 1996): (a) conditional beliefs are beliefs in conditionals, but the latter do not express propositions; (b) conditional beliefs are not beliefs in conditionals, but we rather have a (degree of) belief in B conditional upon A. What we are going to deal with below corresponds to the latter option. According to the former, the belief that if A, B, is considered to be the belief that ''if A, B'' is true given that it has a truth value at all (see p 290 of Edgington 1995); we will not follow this way of rendering conditional beliefs since it amounts to turning them into ''half way'' beliefs in propositions. The theory that we will develop gives no indication that this position is the right account of conditional beliefs, or that it is even stable, as it might turn out to finally collapse either into option (b) or into an account of conditional beliefs as beliefs in conditional entities of a very particular sort. Among other things, Lowe (1996) attacks (b) on the following grounds: he claims (p 606) that

> we deceive ourselves if we think that we have articulated here a genuine notion of *belief*, which bears more than a nominal relationship to the standard notion, as expressed by the verb ''believe.'' The ''b'' in ''$b(B|A)$'' bears a mere orthographic similarity to the ''b'' in ''$b(A)$'', where the latter is taken to denote degree of belief in a proposition A. In short, theorists like Edgington have done nothing whatever to show that they are entitled to use the expressions ''degree of belief in A'' and ''degree of belief in B, conditional upon A'' (or ''degree of conditional belief in B, given A'') with any expectation that the words ''degree of belief'', as they figure in those expressions, bear any significant semantic relationship to one another whatsoever.

One intention of this essay is to accept Lowe's challenge by showing that conditional ''beliefs'' have to be taken seriously as beliefs. Lowe's main argument against conditional beliefs is that, by the probabilistic identity $p(A) = p(A|T)$ (where T is a tautology), the fact that conditional beliefs lack truth values would have the devastating consequence that also simple beliefs lack truth values. But one can in fact accept this probabilistic identity claim *without* regarding the belief in A as being identical to the conditional belief in A given T. The former has a unique propositional content, the latter two of them; the former is true or false, the latter is neither; the former involves a multi-track disposition, the latter does not; etc. (see the later sections). While according to all theories of justification for conditional beliefs that we will deal with later, the belief in A will be justified if and only if the conditional belief in A given a tautology T is justified, this does not imply that the two are identical as mental states, and indeed there are good reasons to believe they are not.[9]

Mellor (1993) defends the thesis that to accept a conditional ''if A, B'' is nothing else but to be disposed to infer its consequent from its antecedent, something that we will find to be a characteristic of conditional beliefs. He also claims that these dispositions do not have a content of the form $A \rightarrow B$ but rather two contents, i.e. A and B. However, at the same time Mellor regards this disposition to be *distinct* from having a belief; in view of the fact that the triviality results that we discussed before only apply to beliefs, he is thus ready to conclude that these results would not entail that conditionals necessarily lack truth conditions. In a somewhat similar spirit, Levi (1988, 1996) considers conditionals to be accepted by the agent without being believed; since conditionals are not members of belief sets, their being accepted is not in any sense affected by the Ramsey test. Conditionals do not express truth-value bearing propositions, which is why they cannot be statements of belief. Instead, Levi claims that conditionals are appraisals of beliefs with respect to their epistemic possibility.

As both Mellor and Levi do not countenance conditional *beliefs*, there is at least on the surface a

[8] This makes us wonder whether the epistemic operation of belief *update*, which was described by Grahne (1991) and Katsuno and Mendelzon (1992), might match Stalnaker's idea of projection better: Since belief update provably satisfies the Ramsey test, and since it is designed to reflect changes in the environment doxastically, it would be more plausible to regard objective conditionals as the result of projecting belief updates, rather than belief revisions, onto the world. On the probabilistic side, something similar can be said about ''projecting'' the operation of probabilistic imaging (cf. Lewis 1976) as opposed to probabilistic conditionalization.

[9] Edgington (1996), in her reply to this point of Lowe's, concedes that ''it is artificial and unhelpful to treat all probabilities as conditional, by defining $p(A)$ as $p(A$ given $T)$'' (p 622).

clear difference between our theory and theirs. On the other hand, we wonder to what extent this could be a merely terminological dispute: as we will argue later, there is a principle called ''UNIQUENESS OF CONTENT and TRUTH-APTNESS'' which is non-necessary of beliefs, while we assume Mellor and Levi would consider this principle beyond discussion and would rather turn to a terminology of ''accepting a conditional,'' ''having an inferential disposition with two propositional contents,'' ''making conditional judgements concerning the possibility of a proposition relative to transforming the current belief set by means of a further proposition'' (cf. Levi 1988, p 61), and so forth, instead of giving up UNIQUENESS OF CONTENT and TRUTH-APTNESS. We leave the question of how serious this difference is for further discussion.

Whether what we call ''conditional beliefs'' should actually be named otherwise is ultimately of only very minor importance. It *is* important, however, to realize that there is a class X of mental states the members of which (a) have been called ''beliefs'' in the past, (b) have been treated theoretically quite uniformly, (c) share important cognitive properties (INTENTIONALITY, ACTION, REPRESENTATIONAL STRUCTURE—see below), (d) are linked causally to certain types of linguistic behaviour in human agents, but where (e) X is nevertheless divided into two disjoint subclasses the members of which differ in terms of the number of propositional contents they have, their truth-aptness, their dispositional properties, and the manner in which they can be justified. In view of (a)–(d), we suggest calling the members of X, i.e. the members of both subclasses, ''beliefs''; on the basis of (e), we recommend applying the qualifications ''simple'' and ''conditional'' in order to distinguish the subclasses. However, the actual point of this paper is not this terminological stipulation but rather to clarify the distinction between the two classes and to study the cognitive characteristics of the states in the second subclass, whether these states are called ''beliefs'' or not.

We will refrain from discussing the relationship between conditional beliefs—which will be seen to be dispositional states that do not necessarily involve suppositional acts—and *beliefs in the context of a supposition*, except for noting that a formal account of the latter is urgently needed in order to facilitate a proper comparison between the two notions (Barnett's (2006) theory of conditionals, according to which conditionals express suppositions but nevertheless have truth conditions, could turn out to be relevant in this respect). Both conditional probabilities and belief revision operators are usually meant to reflect what is going to happen doxastically given some new evidence A is *believed* by a rational agent. Whether conditional belief in this sense is derivative from a more fundamental concept such as *suppositional belief*, or whether the former is conceptually prior to the latter, or whether both notions can be reduced to some joint underlying state, has to be left for further discussion (but see footnote 10 below).

5 Are conditional beliefs actually conditional?

For the sake of simplicity, we will mainly concentrate on belief revision in the following few sections (though similar points could be made for probabilistic update). We will return to a more general viewpoint at the end of the paper. Moreover, whenever we speak of propositions A, B, C,... we will restrict ourselves to propositions that ''speak'' about the world rather than about the agent's inner states (such as belief states). In particular, neither of the sentences that we will use in order to express propositions is going to include a belief operator.[10] The more general case would demand a much more intricate theoretical treatment, but surely there should be enough interest in starting with the primary case of beliefs that have strictly factual contents.

As we have seen, the identification of conditional beliefs with beliefs in conditionals contradicts our norms of rationality for belief and belief change. Since conditional beliefs therefore cannot be conditional in the sense that their contents are conditional propositions, the obvious question arises: are conditional beliefs ''conditional'' at all, and if so in what sense?

Let us return to what we said above about the conditional beliefs of rational agents in the context of the Ramsey test: while the left hand side of Eq. 4 ascribes the belief in the conditional $A \Rightarrow B$ to the agent whose

[10] For that reason, beliefs as the one in Mellor's deceiving-wife conditional (Mellor 1993, p 243), or belief updates as they are described in dynamic epistemic logic (cf. van Benthem 2006), are beyond the aims of this paper. If the belief operator ''Bel'' is allowed to occur in the sentences that express propositions such as A and B, then the theory of conditional beliefs becomes more complex in various respects. For example, when revising a belief set by means of A, one would still look for the most similar belief sets in which A is indeed believed, but one should not assume that when doing so a rational agent also is closing her beliefs on the basis of standard laws governing positive or negative introspection. From the viewpoint of the current belief set, A will rather only be *hypothetically* believed or supposed, where from the viewpoint of the revised belief set supposing is like believing except that such laws of introspection should not be taken for granted.

current belief set is K, the right hand side expresses that the same agent has a conditional belief in B given A. In terms of the standard language of doxastic logic (see Hintikka 1962), the left hand side of the Ramsey test can obviously be represented by the formula

$$Bel(A \Rightarrow B) \tag{5}$$

where ''*Bel*'' is a standard belief operator that is indexed tacitly by a name of the agent that we deal with. The question that we are interested in is now: Is it possible to express the right hand side of the Ramsey test accordingly, i.e. on the basis of ''*Bel*'' and a conditional sign ''$\Rightarrow$''?

In order to answer this question we have to delve deeper into the semantics of ''*.'' What the belief revision operator is supposed to denote is a *minimal mutilation* of the present belief set K in light of the new evidence A (see sect 3.3 of Gärdenfors 1988), i.e. $B \in K * A$ is true if and only if * determines a *minimal* change of K by A that leads to a subsequent belief in B.[11] This interpretation of ''$B \in K * A$'' can be made more precise as follows:

Let w be a possible world in which the agent's actual belief set is K. Now consider the set W' of worlds w' in which our agent believes A: then ''$B \in K * A$'' is true in w if and only if all those worlds w'' among the members of W' that are maximally similar to w are worlds in which the agent also believes B. The similarity relation in question is a doxastic one, in the sense that this similarity of worlds supervenes just on what the agent believes in these worlds, rather than on all the facts whatsoever. However, given an additional assumption on ''observer independency,'' we can presume that this doxastic similarity relation ''piggybacks'' on a standard Lewis-style similarity relation of worlds (cf. Lewis 1973) by which all possible respects of similarity are taken into account and balanced.

For let us suppose that minimal doxastic changes do not affect any mind-external facts in a world (where for simplicity we restrict ourselves to worlds in which one and the same cognitive agent exists in an otherwise agent-free universe). The worlds w'' that are most similar to w in Lewis' sense, among those worlds in which our agent believes A, can thus be determined in the following way: first, every such world w'' may be assumed to be identical to w with respect to non-mental facts; secondly, what is believed by the agent in such a world w'' is as similar as possible to what is believed by the agent in w as long as it is presupposed that A is believed in w''; by ''observer independency,'' the latter property of such worlds w'' can be considered independently of the former. It follows that the worlds that are most similar to w *in Lewis' sense* among the worlds in which our agent believes A will simply coincide with the worlds that are most similar to w *doxastically* among the worlds in which our agent believes A, since given the identity of mind-external facts a Lewis-style similarity relation simply collapses into doxastic similarity. Of course, the assumption of ''observer independency'' is strong and ultimately unrealistic; however it is still useful methodologically, if only to simplify matters while leaving much of the actual circumstances at least approximately intact (which is why it is used for analogous purposes in various areas of empirical research). Under this proviso, we can thus reformulate the semantics of the revision operator as follows: ''$B \in K * A$'' is true in w if and only if all those worlds w'' among the members of W' that are maximally similar to w in Lewis' sense are worlds in which the agent believes B. If we finally accept Lewis' semantic analysis of subjunctive conditionals, and if we presuppose that ''$\Rightarrow$'' is a subjunctive conditional sign as it is used in counterfactual claims, then we can finally express the right hand side of the Ramsey test by means of the subjunctive conditional

$$Bel(A) \Rightarrow Bel(B) \tag{6}$$

We conclude that, given our auxiliary assumptions, a rational agent has a conditional belief in B given A if and only if: *if she believed A, then she would believe B*. Hence, rather than having conditional contents, conditional beliefs turn out to be *conditional states*. If our background assumptions from above are only approximately true, there is still hope to show that conditional beliefs *approximate* conditional states. Indeed, as we will see later, recent findings on the conditional analysis of dispositions yield independent evidence that conditional beliefs are not actually conditional in the sense of Eq. 6 but that they can rather be described in terms of conditionals which result from adding ceteris-paribus clauses to the antecedents of conditionals as in Eq. 6, i.e. conditionals of the form

$$CP(Bel(A)) \Rightarrow Bel(B) \tag{7}$$

or alternatively in terms of conditionals where a corresponding ceteris-paribus component is assumed to be part of the meaning of a new subjunctive implication sign ''$\Rightarrow^{+}$,'' i.e. conditionals such as

$$Bel(A) \Rightarrow^{+} Bel(B) \tag{8}$$

[11] We speak of *a* minimal mutilation and *a* minimal change of K by A, since the rationality postulates for revision operators do not generally determine the result of revising K by A uniquely.

However, we will continue for the moment to discuss conditional beliefs according to the simpler format of Eq. 6; we will return to the more accurate ones later.[12]

6 Are conditional beliefs proper beliefs?

What we found in the last section was that an agent has the conditional belief in B given A if and only if she satisfies Eq. 6, where "$\Rightarrow$" is a subjunctive conditional sign. Obviously, Eq. 6 describes a conditional state. But what reasons do we have to regard such states as *beliefs*?

After all, not every logical composition of "*Bel*" with other expressions can be reasonably said to describe a belief. For example

$$\neg Bel(A) \tag{9}$$

is true of an agent if and only if she does not believe A. If we used a restrictive notion of (mental) state as being some kind of natural property, it would not even be clear whether a condition such as Eq. 9 characterized a state at all, let alone a belief state. This is because the space of possibilities in which an agent is *without* the belief in A is so enormously vast and heterogeneous: the agent could have strong reasons to believe in $\neg A$; she could both lack reasons to believe in A and to believe in $\neg A$; she could fail to have the conceptual resources to believe either of A and $\neg A$; or she could simply be brain-dead and thus lacking a belief in A for rather unfortunate reasons.

But whatever our ontology of states is like, it is evident that the absence of a belief, as expressed by Eq. 9, is certainly not itself a *belief* state. For similar reasons, a disjunction of the form

$$Bel(A) \vee Bel(B) \tag{10}$$

can hardly be taken to describe a belief; etc.[13]

Let us now take a liberal concept of state for granted. Hence the question is not whether what is characterized in terms of Eq. 6 is a state but rather whether it is actually a *belief* state. But what is a belief state after all?

Of course, any attempt of giving a "comprehensive" theory of beliefs in order to address this question would lead us far beyond the intended limits of this paper (not that we were able to state such a theory otherwise). Instead, we will restrict ourselves to the following list of necessary, though not even collectively sufficient, conditions on beliefs:[14]

- INTENTIONALITY: Beliefs are mental states of a cognitive agent that are intentional, i.e. which have some content. The contents of beliefs are propositions, i.e. abstract entities that may be expressed by sentences.
- ACTION: Beliefs have an action-determining, or, more generally, a behaviour-determining function for the agent ("they are mental states apt for selective behaviour towards the environment," Armstrong 1968, p 339).
- REPRESENTATIONAL STRUCTURE: Beliefs have an internal representational structure that includes representations of their contents.

The constraint that is expressed by INTENTIONALITY is, e.g., not satisfied by purely phenomenal states, which lack propositional content. ACTION would not be the case if beliefs were only caused by those mental states that cause actions, without causing such states—or actions directly—themselves. REPRESENTATIONAL STRUCTURE would be false if the ascription of believed propositional contents were merely "in the eye of the observer," i.e. if there were no representations whatsoever in the agent's cognitive system that would correspond to the propositions that are said to be believed by the agent.

Conditional beliefs, as we have introduced them before, are mental states with determinate propositional contents, so they definitely satisfy the first of these conditions. In the next two sections, we are going to deal with the other two conditions, and we will show that conditional beliefs are likely to satisfy them as well.

But before we do so, let us consider a further condition which beliefs are usually assumed to satisfy:

- UNIQUENESS OF CONTENT and TRUTH-APTNESS: Every belief has one and only one

[12] One reason for adding CP clauses to sentences such as Eq. 6 is to leave open the possibility that the agent herself interferes intentionally with the cognitive transition from $Bel(A)$ to $Bel(B)$ that is described by Eq. 6 (e.g. by means of conscious reflection on her mental state). Since we understand "belief" in a way that allows also simple "low-level" agents, who do not have conscious access to her mental states at all, to have beliefs, we will not deal with this possibility in further detail. CP clauses also take care of cases where agents simply fail to make the connection between A and B, or where an agent is insufficiently interested in forming the belief in B.

[13] The case of conjunctions is different, at least for completely rational agents, as such agents satisfy $Bel(A)$ and $Bel(B)$ if and only if they satisfy $Bel(A \text{ and } B)$ (*pace* Kyburg on the Lottery paradox, which we allow ourselves to ignore for the sake of simplicity).

[14] These are among the conditions on beliefs stated in Leitgeb (2004, chap 3). Of course, there are authors who would not even regard these assumptions on beliefs as correct, but the great majority would in fact do so.

proposition as its content. A belief is true if and only if its (unique) content is true.

What this condition adds to the first one from above is that a belief does not only have content, it actually has one *and only one* content. In contrast with beliefs in conditionals, which by definition have a unique conditional content, conditional beliefs do not conform to this uniqueness condition: as is stated clearly by Eq. 6, the conditional belief in B given A has *two* propositional contents, i.e. A and B. As the impossibility theorems that we started with show, there is no general way of amalgamating the two contents of a conditional belief such that the latter could finally be understood again as a belief *simpliciter* or simple belief, i.e. a belief that can be described by a sentence of the form ''$Bel(A)$.''

So we are facing the following dilemma: either we give up the standard assumption that beliefs have unique contents and that a belief is true or false according to whether its content is, or we are forced to conclude that what we called conditional *belief* before is not a kind of belief after all. A way out of this dilemma becomes clearer once we examine the reasons why UNIQUENESS OF CONTENT and TRUTH-APTNESS is generally assumed. The answer we suggest is twofold:

First of all, all the examples of beliefs that we are usually aware of satisfy this condition. Although UNIQUENESS OF CONTENT and TRUTH-APTNESS does not seem to be implied or presupposed by any of the other conditions on beliefs that were mentioned above, beliefs such as the belief that I am sitting here, the belief that Wittgenstein was a philosopher, the belief that $F = ma$, and the belief that $2 + 2 = 4$ all have unique contents. Hence, it might be that we have accepted UNIQUENESS OF CONTENT and TRUTH-APTNESS for the very same reason for which we regarded conditional beliefs as beliefs in conditionals in the first place: CCCP and the Ramsey test would thus turn out to be the formal expressions of a presupposition that became part of our concept of belief in view of the fact that it seemed to lack any obvious counterexamples.

If this diagnosis is right, then it should be possible to show that the mental states that we referred to above as conditional beliefs share all the core properties of beliefs while *lacking* UNIQUENESS OF CONTENT and TRUTH-APTNESS. If so, this would leave us with a reason to regard UNIQUENESS OF CONTENT and TRUTH-APTNESS as not being essential for beliefs after all and to accept conditionals beliefs as beliefs proper. What we are going to do in the next two sections is to investigate whether conditional belief states actually conform to the other principles from above, i.e. ACTION and REPRESENTATIONAL STRUCTURE. At the same time, this will give us opportunity to consider some of the differences that conditionals beliefs and beliefs in conditionals, or beliefs *simpliciter* in general, do or do not exhibit.

The second reason why UNIQUENESS OF CONTENT and TRUTH-APTNESS has at least prima facie plausibility is an epistemological one: the fact that it allows us to assign truth values to beliefs yields an explanation of why beliefs are rational or irrational, justified or unjustified. If UNIQUENESS OF CONTENT and TRUTH-APTNESS fails, it will not only be impossible to determine the truth value of certain beliefs on the basis of their uniquely specified contents, we will actually have to accept that certain beliefs are *not truth-apt*, i.e. do not have truth conditions at all. This is because every conditional belief will be regarded to involve two propositional contents, and although each of these propositions can be said to correspond to, or deviate from, reality if taken in isolation, there is no method of fusing these two contents into a unique proposition which could finally be compared to reality in order to make conditional beliefs come out as true or false. Hence, it is also no longer clear how such ''beliefs'' could become justified and thus to aim at truth, since the justification of a belief is usually taken to involve some kind of evidence in favour of its truth. Conditional beliefs might therefore be in danger to be expelled from the space of reasons, which in itself would be a strong reason for not counting them as proper beliefs.

We are going to address this worry after having dealt with ACTION and REPRESENTATIONAL STRUCTURE in the next two sections.

7 Conditional beliefs, dispositions to act, and Ramsey's account of belief

When we speak of conditional beliefs in the following three sections, we do so in view of the fact that it will follow that conditional beliefs are actually *beliefs*; as long as this has not been established yet, one might think of ''conditional belief'' to refer to any mental state that is described by the right hand side of Eq. 4, be it a state of belief or a different sort of state.

There are several theories of what having a belief consists in. But independent of what theory of beliefs one is inclined to defend, it is generally regarded constitutive of the belief in A to involve various dispositions, most notably the disposition of acting as if A were the case (see Braithwaite 1932). For a representationalist, these dispositions will be effected by some

mental representation that plays a similar role in our cognitive systems as the sentences that express the contents of beliefs do in our linguistic practices. For a functionalist, such dispositions will be part of the functional role of the belief in A. For a dispositionalist, having these dispositions will simply be exhaustive of what it is to believe A; and so forth.

So what are the dispositions, if any, that a conditional belief in B given A would consist in? The answer should be obvious by now: the disposition of acquiring the belief in B given the circumstances that A is believed. For example: an agent has the conditional belief that Gabriel is mortal given that Gabriel is a man if and only if the agent is disposed to believe the former when she believes the latter. In many cases, we will express the triggering of this disposition in a somewhat different way: e.g. by saying that the agent *infers* that Gabriel is mortal from the fact that Gabriel is a man; or, if she had believed Gabriel to be immortal before, that she would have *revised* her belief that Gabriel is immortal, such that she starts to believe Gabriel to be mortal once she has evidence that Gabriel is a man (and not an angel or an number as she might have believed before); or that the agent *learns* that Gabriel is mortal on the basis of learning that Gabriel is a man etc. But the underlying dispositional character is always the same, namely the one that is displayed by the subjunctive conditional (6) from above. Conditional beliefs even have dispositional properties of a very distinguished kind: while we found a conditional belief to entail a so-called single-track disposition, i.e. a disposition to show one particular type of manifestation in one particular type of circumstance, a belief such as the belief that Gabriel is mortal, or the belief that ice is dangerously thin,[15] involves a multi-track disposition, i.e. a manifold of circumstance–manifestation pairs. A further characteristic of the conditional belief in B given A is that its dispositional properties involve beliefs—the belief in A and the belief in B—rather than actions or mental states different from beliefs; since the belief in A and the belief in B involve dispositions themselves, conditional beliefs can actually be understood as higher-order single-track dispositions.

It follows that conditional beliefs partially determine an agent's acquisition and change of simple beliefs; since the latter in turn co-determine the agent's overall activities, the former do as well, where these activities could include the agent's intentional actions, which are directed towards the outer world, but also various sorts of patterns of unintentional behaviour as well as mental episodes that bring about new inner states of the agent. We conclude that conditional beliefs have at least indirectly an action-determining, or, more generally, a behaviour-determining function for the agent, and thus that principle ACTION is satisfied.[16]

There are two slight complications that affect this dispositional account of conditional beliefs. First of all: Does a conditional belief in B given A necessarily lead to a belief in B in *all possible circumstances* in which A is believed? *No*. The reason for this is, in the terminology of the proponents of non-monotonic logic (cf. Brewka et al. 1997): conditional beliefs do not necessarily obey the monotonicity principle "if x has a conditional belief in B given A, x also has the conditional belief in B given A and C" (whatever the C). Or, from the viewpoint of conditional logic: since subjunctive conditionals do not obey a monotonicity rule by which antecedents could be strengthened arbitrarily, Eq. 6 from above does not entail

$$Bel(A)\&Bel(C) \Rightarrow Bel(B) \tag{11}$$

which, by the logical equivalence of $Bel(A)$ and $Bel(C)$ and $Bel(A$ and $C)$ in standard doxastic logic, means that Eq. 6 does not imply

$$Bel(A\&C) \Rightarrow Bel(B) \tag{12}$$

either. So it is consistent that both $Bel(A) \Rightarrow Bel(B)$ and $\neg[(Bel(A)$ and $Bel(C)) \Rightarrow Bel(B)]$ are the case, and this logically possibility can indeed be found realized in the actual world quite frequently.

For example reconsidering the obligatory and omnipresent paradigm case of non-monotonic reasoning: an agent might have the conditional belief that Harry is able to fly given Harry is a bird, without having the conditional belief that Harry is able to fly given both that Harry is a bird *and Harry is a penguin* (or dead or having his feet set in concrete or...). The agent's corresponding disposition will thus not be to acquire the belief that Harry can fly in every possible circumstance in which she believes that Harry is a bird, but rather she will have the disposition to acquire the belief that Harry can fly in every circumstance in which she believes that Harry is a bird *and this is all that is relevantly believed by the agent at the time*. This latter

[15] This is an example of Ryle, who introduced the terminology of "multi-track dispositions" in his study of belief; see Ryle (1949, pp 134f).

[16] It should be noted that if ACTION were strengthened by demanding that beliefs ought to determine actions *directly*, in particular, without a detour through other states of beliefs, then it is likely that iterated simple beliefs of the form described by "$Bel(Bel(A))$" would fail to count as beliefs, which would be highly counter-intuitive.

disposition obviously does not entail the conditional belief that Harry is able to fly given both that Harry is a bird and Harry is a penguin.

''Relevantly'' is to be understood in this context as ''*causally* and *cognitively* relevantly'': every other belief at the time that was causally active and played a role in the agent's cognitive procedures was already contained (explicitly or implicitly) in the belief that Harry is a bird. Note that this refined analysis does not rule out conditional beliefs such as the first one from above where the belief in Gabriel being a man indeed leads to the belief that Gabriel is mortal *whatever else* is believed by the agent at the time; while the disposition to believe B given the total relevant belief in A does not necessarily include the disposition to believe B given a total relevant belief in A and C, it also does not necessarily exclude it. In this sense, the ''if all that is believed is that...'' format of ascribing conditional belief dispositions is simply weaker, and in this sense more general than the naïve material ''if it is believed that...'' format.[17]

The second problem that our dispositional analysis of conditional beliefs faces is given by well-known objections to the conditional analysis of dispositions in general. We have claimed that an agent has the conditional belief in B given A if and only if she is disposed to believe B given all that is relevantly believed by her originally is A. But secondly we have tacitly presupposed the latter to be the case *if and only if* the agent satisfies a subjunctive conditional of the form: if all the agent relevantly believed at the time were A, then she would start to believe B. It is this latter equivalence claim that came under attack in recent years in view of the possible presence of dispositional *finks* and *antidotes*: According to Martin (1994), an object's disposition is finkish when the object loses the disposition after the occurrence of the stimulus but before the expected manifestation can occur, such that consequently that manifestation does not occur at all. Bird (1998) discusses the effects of antidotes, where an antidote to a disposition would be something which, when applied before the time of the expected manifestation of the disposition, has the effect of breaking the causal chain that leads to this manifestation, such that again the manifestation does not occur. Since the presence of finks or antidotes cannot be excluded in the cases of conditional beliefs either—though they may be expected to be exceptional—Eq. 6 has to be refined in one of the ways indicated by Eq. 7 or 8 from above. But this does not affect any of our conclusions on conditional beliefs so far and it will do not do so in the following when we continue pretending that Eq. 6 is actually the case, for the sake of simplification.

Now that we have quite a clear picture of the cognitive role of conditional beliefs, let us turn to a class of beliefs which, at first glance, have very similar properties but which nevertheless differ from conditional beliefs in various crucial aspects:

It seems conditional beliefs fit nicely into Ramsey's (1929) classical account of beliefs (which was worked out later by Armstrong 1973), according to which there are actually two different kinds of beliefs. As Stalnaker (1984, p 101) summarizes the view: ''there are on the one hand beliefs about particular matters of fact, which are pictures or maps of reality by which we guide our actions. On the other hand there are general beliefs, which are dispositions to extend or change our maps of reality''; the latter beliefs being described by Armstrong (1973, p 5) as '''habits of inference' which dispose us to move from a belief about some particular matter of fact to a further belief about some particular matter of fact.''

As mentioned before, the source of these thoughts is Ramsey (1929) where we find the corresponding distinction between beliefs of the ''primary sort'' and beliefs in ''variable hypotheticals'':

> A belief of the primary sort is a map of neighbouring space by which we steer (p 146).
> Many sentences express cognitive attitudes without being propositions; and the difference between saying yes or no to them is not the difference between saying yes or no to a proposition (pp 147f).
> To believe that all men are mortal—what is it? Partly to say so, partly to believe in regard to any x that turns up that if he is a man he is mortal. The general belief consists in (a) A general enunciation, (b) A habit of singular belief (p 148).
> Variable hypotheticals are not judgements but rules for judging 'If I meet a ϕ, I shall regard it as a ψ'. This cannot be *negated* but it can be *disagreed* with by one who does not adopt it (p 149).

It is tempting to read our previous distinction between beliefs *simpliciter* and conditional beliefs into these quotations. However, while there would be a kernel of truth in doing so, there are also reasons to be cautious and, in the end, to reject any such identification.

This is because Ramsey actually runs together the following types of beliefs: (a) *singular beliefs*, i.e. beliefs about a single object, situation, or fragment of environment, with *beliefs simpliciter*, i.e. beliefs with a single propositional content; (b) *general beliefs*, i.e.

[17] Levesque (1990) is the primary source on the logic of total belief or knowledge ascriptions of the ''all that is believed is that...'' form. Carnap's and Hempel's classical accounts of statistical explanation and inductive reasoning in terms of principles of ''total evidence'' and ''maximum specificity'' use a very similar format.

beliefs about an open-ended and potentially infinite collection of objects, situations, or about an open-ended and potentially infinite environment, with *conditional beliefs*, i.e. beliefs of the form described by Eq. 6 from above.

The background of these identifications is given by Ramsey's pragmatic-instrumentalist account of general statements as enunciating habits or rules rather than expressing propositions (a view shared by the later Wittgenstein and to some extent also by the Vienna Circle). But from the modern semantic point of view, universal statements (''for all *x*...''), statistical statements (''the statistical probability of *x*...is 0.8''), normality statements (''for normal *x*...''),[18] and other general statements have clear truth conditions by which they express certain constraints on what the world has to be like in order for them to be true. In contrast with rules, there is also no problem whatsoever about negating such statements or about applying to them any other sort of compositional procedure. Hence, general beliefs are simply special cases of beliefs *simpliciter*, namely, *beliefs in general propositions*; consequently, just as all other beliefs in propositions, also general belief states are not to be mixed up with conditional belief states.

One might think that Gärdenfors' impossibility result does not actually contradict a ''Ramsey test'' of the form

$$A[x] \Rightarrow_x B[x] \in K \text{ iff } B[a] \in K * A[a] \tag{13}$$

with a variable-binding operator ''$\Rightarrow_x$'' and an arbitrary individual constant ''*a*,'' but in fact one can show that a similar impossibility theorem can be derived in such a context as well: for assume Eq. 13 and suppose that $K \subseteq K'$; if $B[a] \in K * A[a]$, it follows that $A[x] \Rightarrow_x B[x] \in K$, hence $A[x] \Rightarrow_x B[x] \in K'$, and thus $B[a] \in K * A[a]$. So Eq. 13 entails the following monotonicity principle: if $K \subseteq K'$ then $K * A[a] \subseteq K' * A[a]$. But this is all that is needed in order to derive Gärdenfors' theorem.

This having been said, the following relationship between general beliefs and conditional beliefs does indeed hold: if a rational agent has the general belief that $A[x] \Rightarrow_x B[x]$, where ''$\Rightarrow_x$'' expresses any of the quantifiers mentioned above, then the same agent must have a conditional belief in $B[a]$ given $A[a]$, i.e.: as every rational agent is bound to respect the validity of detachments, a general belief in one of Ramsey's ''variable hypotheticals'' implies every conditional belief in an instantiation of the consequent of this hypothetical, given the corresponding instantiation of the antecedent of this hypothetical. So

$$Bel(A[x] \Rightarrow_x B[x]) \tag{14}$$

implies

$$Bel(A[a]) \Rightarrow Bel(B[a]) \tag{15}$$

whereas the other direction of this implication does not hold necessarily, which confirms that general beliefs and conditional beliefs do not coincide—not to mention the differences we have outlined above: Eq. 14 describes a belief in a unique proposition, which Eq. 15 does not; Eq. 14 describes a state which is true or false, while Eq. 15 does not; Eq. 14 involves a quantifier binding a variable in the two open formulas ''$A[x]$'' and ''$B[x]$,'' which is not so in the case of Eq. 15.

8 Conditional beliefs, non-monotonic reasoning, and connectionist representation

Now that we have found ACTION to be satisfied by conditional beliefs, let us turn to the question of whether conditional beliefs also have a representational structure that includes representations of their contents, i.e. whether REPRESENTATIONAL STRUCTURE is true of conditional beliefs as well.[19]

As it stands, this seems to be completely underdetermined by the way conditional beliefs were introduced above: without any further specification, ''the'' mental state that is represented by the membership of B in $K * A$ could be anything whatsoever as long as it subserves the dispositional state that was analysed in the last section. However, given it is the mental states of actual cognitive agents that we are referring to—whether human, animal, or artificial—it is certainly unlikely that such dispositional states are embodied in such agents without an underlying system of representations the function of which is to encode the propositions which figure as the contents of these conditional beliefs. For the very same reason, we think that in order to have simple beliefs, such as the belief that Gabriel is a man, a human agent must rely on representations of those propositions that are the contents of these beliefs—e.g. a representation of the proposition that Gabriel is a man. Whether or not REPRESENTATIONAL STRUCTURE is only satisfied by the types of cognitive agents that we are

[18] See Schurz (2001) for a theory of such ''normic'' laws.

[19] We will disregard any question of the sort to what extent, and in what sense, idealisations such as the usual logical closure properties of probability measures and belief sets apply to real cognitive agents.

acquainted with, or whether it is necessary of every cognitive agent whatsoever for principal reasons, we would expect conditional beliefs not to differ from beliefs *simpliciter* with respect to REPRESENTATIONAL STRUCTURE if conditional beliefs are to count as beliefs proper. As we are going to argue in the following, whatever system of representations is used to provide for the representational structure of beliefs *simpliciter*, the very same system can be employed in order to effect a representational system that would be adequate for conditional beliefs. We will exemplify this claim by focusing on the two kinds of representation that are currently regarded to be the most plausible candidates for what mental representations in human cognition are like: symbolic representations in rule-based computation systems on the one hand and distributed representations in neural networks on the other.

From the viewpoint of the former, beliefs *simpliciter* are usually regarded to correspond to an agent's having a sentential representation stored in the agent's ''belief box,'' or, as computer scientists would say, in her factual knowledge base. Let us assume an agent indeed possesses the capabilities of saving, manipulating, and retrieving mental sentences A and B[20] in the way such that putting A and B into the factual knowledge base means the agent acquires both the belief in A and the belief in B; furthermore, having stored $\neg A$ in the belief box is equivalent to the agent's having the belief that $\neg A$ is the case; similarly for $A \wedge B$, and so forth. Clearly, this is exactly the sort of architecture that is presupposed by much of today's cognitive psychology and philosophy of mind, and which is used and applied by most of the classical artificial intelligence applications in computer science (see, e.g. Newell 1990).

Extending this type of representational system to a corresponding one for conditional beliefs is fairly straightforward; indeed, a good part of the theoretical and practical work on non-monotonic reasoning in theoretical computer science deals with exactly this problem: additional to factual knowledge bases of the type mentioned above, it turned out to be useful to equip computers with conditional knowledge bases consisting of so-called *defaults* with a

$$\text{premise}(A)\text{—conclusion}(B)$$

or antecedent-consequent structure, where possible defeating propositions (C) that would make the default inferentially inapplicable are expressed in either of two ways: (a) explicitly, as in ''if (A) a is a bird and (C) for all you know it is consistent to assume a is able to fly, then (B) conclude that a is able to fly,'' such that the defaults are of the form

$$\frac{A : C}{B} \tag{16}$$

or (b) implicitly, i.e. they are accounted for by the inferential patterns that are permissible to be applied by the agent: ''If (A) a is a bird, then (B) conclude that a is able to fly [you may apply this rule if all you relevantly know factually is that a is able to fly, but you are no longer allowed to apply it if there is an additional statement C in your factual knowledge base].'' In the latter case, the default has the ''conditional'' form

$$A|\sim B \tag{17}$$

(''$|\sim$'' is the sign that would be typically used in such a context in order to express defaults; cf. Kraus et al. 1990).

In both cases the representations that are put into the ''conditional belief box'' are composed of the sort of representations that are allowed to be put into the belief box for simple beliefs. As computer implementations of non-monotonic reasoning procedures show, agents are very well able to generate the characteristic dispositional properties of conditional beliefs on the basis of such conditional knowledge bases by means of applying rules of inference to them (see Kraus et al. 1990, for various rule-based systems of non-monotonic logic). Thus, if REPRESENTATIONAL STRUCTURE holds for beliefs *simpliciter*, and if the reason for this being so is that the contents of these beliefs are represented by mental sentences such as A and B, then there is no reason why the same should not be true of conditional beliefs based on conditional default representations involving A and B such as the ones mentioned above.

There is one additional point about defaults that is worth being noticed: In the literature on non-monotonic reasoning, one can find all sorts of readings and interpretations of these conditional default expressions.[21] Often they are explained as descriptive statements of either of the following forms: (a) *if A then normally B*; (b) *if A then typically B*; (c) *if A then it is very likely that B*; (d) *normal As are Bs*; (e) *(proto-)typical As are Bs*; (f) *by far most of the As are Bs*. While it would of course be possible to store sentences of any of these kinds in a knowledge base and let a reasoning mechanism be applied to it, one can see from the way the implemen-

[20] Or, more strictly: the *mental counterparts* of the sentences A and B that we use in the language by which we ascribe these beliefs, i.e. usually, some natural language.

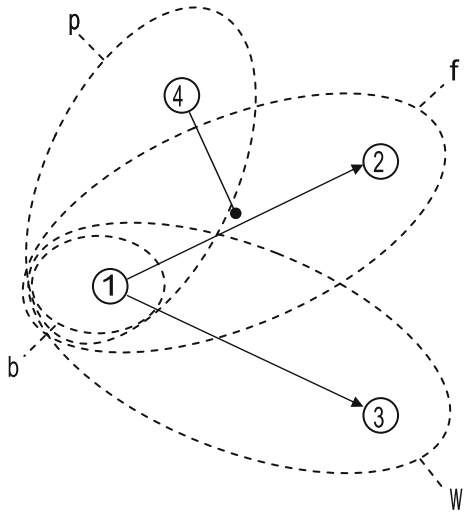
[21] Most of these interpretations can be found instantiated in Ginsberg (1987) and Brewka (1996)

tation of these conditional defaults is usually set up in non-monotonic reasoning that their whole function is to serve an inferential mechanism by which B is inferred given that all that is believed is A; defaults are signs in the agent's database the sole purpose of which is to bring about the conditional belief in B given A, which is why we chose to explain them in terms of conditional imperatives (''if...*conclude* that...'') rather than as plain descriptive conditionals. Unlike descriptive sentences, defaults are usually not negated, they are not parts of disjunctions, and they cannot be nested.[22] In fact, in many cases conditional defaults are either taken to be rules (as in default logic; cf. Reiter 1980) or to be metalinguistic statements about two object level sentences A and B which are claimed to stand in a binary non-monotonic consequence relation $|\sim$ (see, e.g. Makinson 1994). Hence, the obvious way of interpreting the presence of a default in a conditional knowledge base is neither in terms of a belief in a singular proposition, as in (a)–(c) above, nor in terms of a belief in a general proposition, as suggested by (d)–(f), but rather straightforwardly as a conditional belief. The reason why computer scientists feed their computer's conditional belief boxes with these defaults may be on grounds of believing in the truth of any of (a)–(f), but the beliefs that are ''created'' by putting these defaults in their machines' conditional knowledge bases—given these machines were sufficiently agent-like otherwise—are most likely conditional beliefs, not beliefs *simpliciter*. The failure of distinguishing the former from the latter can thus lead to a misunderstanding of what mental states and processes a non-monotonic reasoning mechanism should be assumed to instantiate if it were employed within the cognitive architecture of a sufficiently complex computer agent. Similarly, if human cognition indeed turned out to be based on symbolic computation, then conditional defaults—symbols in the mind directing the transition between symbolic representations in the box of simple beliefs—would be plausible candidates for being the representational substrata of conditional beliefs as described by Eqs. 6 and 15, rather than beliefs of the kind described by Eq. 5 or 14.

According to connectionism, the second main paradigm in cognitive science, mental representations in human and animal cognitive systems are *not* of the localized and symbolic form discussed so far; instead, they are distributed over the units and connections of biological or artificial neural networks, such that the states of a single unit or connection in such a network are ''subsymbolic,'' i.e. below the representational level (cf. Smolensky 1988). Although the viability and need of folk-psychological concepts such as *belief* in the context of the connectionist paradigm has been questioned (cf. Churchland 1989), several ways of representing propositions in artificial neural networks have been suggested, which might nevertheless enable connectionists to account for propositional attitudes (see Macdonald and Macdonald 1995).

In the following we will consider a little toy example that is intended to show that simple occurrent beliefs can be implemented as the occurrence of activation patterns in neural networks, and that once this is established one gets a corresponding implementation of conditional beliefs in these networks for free. For the sake of simplicity, we will deal with very ''coarse-grained'' neural networks that consist of nodes, excitatory connections between nodes, and inhibitory connections leading from nodes to excitatory connections; no thresholds or weights are attached to any of these components. The activity of nodes (0 or 1) is assumed to spread along excitatory lines, but if an inhibitory connection leads from a node to such an excitatory line, then the activity of that node is able to stop the flow of energy in the excitatory connection.

For example, such a network might consist of four nodes (1, 2, 3, 4), two excitatory connections, and one inhibitory connection, as shown in the following diagram:



We assume that each of the four nodes can be activated by some external input that is fed to the nodes for some amount of time. For example, if node 1 is the only node that receives an input, it gets activated and remains so for this amount of time. Its activation spreads to nodes 2 and 3 along the excitatory connections; the activation pattern 1, 2, 3 stays stable until the input changes. But if both node 1 and node 4 receive

[22] Autoepistemic logic, being an extension of standard modal logic, is an exception to this pattern; see Levesque (1990).

the external input, then the activity of 4 prohibits the transfer of activation from node 1 to node 2, while node 3 is still activated via the uninhibited excitatory line that leads from 1 to 3. Thus, in this case, the pattern 1, 3, 4 will be the stable state of the network.

Now let us add a simple interpretation of the corresponding activation patterns in terms of beliefs *simpliciter*: we regard the single node pattern 1 as representing B, pattern 1, 2 as representing F, pattern 1, 3 as representing W, and pattern 1, 4 as representing P, where the expressions ''B,'' ''F,'' ''W,'' ''P'' can be regarded as short-hands for the propositions expressed by ''*a* is a bird,'' ''*a* is able to fly,'' ''*a* has wings,'' and ''*a* is a penguin,'' respectively. Furthermore, we may extend this assignment of statements to patterns also to the case of complex formulas: let the pattern which belongs to the negation of a sentence be the set of nodes which are not contained in the pattern of the formula itself; e.g. since 1, 2 is the pattern that belongs to F, it follows that 3, 4 is the pattern of $\neg F$; accordingly, $\neg P$ has the pattern 2, 3, because the pattern of P is 1, 4; and so forth. Finally, we define the pattern of a conjunctive sentence to be the union of the patterns that are assigned to the sentences which are joined by ''$\wedge$''; therefore, $F \wedge W$ is represented by pattern 1, 2, 3; $B \wedge F$ is represented by 1, 2; and the pattern of $B \wedge P$ is 1, 4. But also more complex formulas have patterns assigned to them: e.g. the pattern that corresponds to $\neg F \wedge W$ is 1, 3, 4—the union of the pattern of $\neg F(3, 4)$ and the pattern of $W(1, 3)$. In this way, both the inputs and the stable states of our simple network can be regarded as representing the contents of simple (occurrent) beliefs. Whenever the pattern of a proposition is active in the network, i.e. all of its nodes are active (and perhaps further ones), let us suppose that the network believes this proposition to be the case. As the underlying representations are defined by patterns of activity that may be spread over the whole network, they conform to the intended format of connectionist representations.

How can we extend such a system of representations to one for conditional beliefs? In a sense, we do not even have to do so: determining the representations of simple beliefs as patterns of activation is already sufficient for giving us the representations that we need in order to implement conditional beliefs in our little toy network. For example, the fact that the sole activation of node 1 leads to the pattern 1, 2, 3 can be interpreted in the way that the network infers from the total information B (bird) the conclusion $F \wedge W$ (is able to fly and has wings); in other words: the network has the conditional belief in $F \wedge W$ given B. If, on the other hand, the input to the network is 1, 4, i.e. the nodes which belong to the pattern of $B \wedge P$ are activated by the input, then the final output pattern is 1, 3, 4, which is the pattern of $\neg F \wedge W$: so the network concludes from having the information that there is a penguin bird as her total input the derived information that there is something which is not able to fly but which still has wings. In other words: given the interpretation of activation patterns that we have presupposed above, the network conditionally believes in *a* being able to fly and having wings given *a* is a bird, while at the same time believing conditionally that *a* is not able to fly but still has wings given that *a* is a *penguin* bird. Hence, an input pattern represents the proposition that is ''given'' in a conditional belief, the stable state it causes represents the proposition that can be concluded from the given propositions on the basis of a conditional belief, and the conditional belief itself is embodied by the dynamics of the network that connects the one pattern to the other. This way of implementing a conditional belief state in a network does not rely on any localized representations at all: the propositional contents of the beliefs *simpliciter* involved are represented by distributed patterns of units; the dispositional properties of conditional beliefs are ''coded'' by the topology of the network connections as a whole. Despite being grossly oversimplified in several respects, this example should be sufficient to show that if REPRESENTATIONAL STRUCTURE holds for beliefs *simpliciter* in view of their contents being represented by neural patterns of activation, then there is no reason why REPRESENTATIONAL STRUCTURE should not also hold for conditional beliefs on the basis of the dispositional properties of the network if considered as a dynamical system. Indeed, one can even prove general soundness and completeness theorems which express that each of the systems of cumulative non-monotonic logic that are introduced in Kraus et al. (1990) is sound and complete with respect to proper artificial neural network semantics very much like the one sketched before.[23] It is also well-known that these systems of non-monotonic logic are inter-translatable with corresponding systems of belief revision. The fact that the example above lacked any sort of representation of a conditional connective that could figure in the neural representation of compositions or nestings of conditionals should not come as a surprise anymore; after all, as we have seen before, a

[23] This account of belief and reasoning in neural networks, together with the mentioned soundness and completeness theorems, is worked out in detail by Leitgeb (2004, 2005). While the notion of conditional belief is not used there, Leitgeb's notion of *an agent being disposed to infer B given all she believes is A* corresponds to it. There are of course many other suggestions of ''neural networks semantics'' to be found in the literature; however, the cited approach is among the few ones that are justified on the basis of general model-theoretic adequacy theorems; see Blutner (2004) for a similarly justified approach.

conditional belief in B given A is to be distinguished from the belief in the conditional "if A then B," as it does not have a unique conditional content that results from combining A and B by a conditional connective.

9 The justification of conditional beliefs

Since we found to have good reasons to believe conditional beliefs to satisfy ACTION and REPRESENTATIONAL STRUCTURE, it remains to be seen whether they can also be regarded sensibly as *justified* mental states, despite the fact that they are not truth-apt.

Obviously, there is nothing like a conditional belief that is *knowledge* in the standard sense, as conditional beliefs are neither true nor false whilst knowledge entails truth.[24] But we will show that there is hope for both externalist and internalist notions of *justification* being applicable to conditional beliefs.

Let us start with an externalist view on justification: according to one of the standard accounts (Goldman 1986), a belief is justified in the sense of process-reliabilism if and only if it was generated by a reliable process (where processes are regarded as types that can be instantiated repeatedly, and where reliability is understood as some sort of truth-conduciveness). But what does "reliable" mean if the output of such a process is supposed to be a cognitive state that does not have a truth value, namely, a state of conditional belief? In order to answer this question we have to reconsider what an agent would regard as an epistemically failed manifestation of a conditional belief in B given A. The situation to be avoided by the agent is obviously of the following form: the agent believes A; she acquires the belief in B on the basis of the conditional belief in B given A; her subsequent belief in B turns out to be false. However, the occurrence of any such circumstance can only be blamed on the conditional belief if the original simple belief in A was true, since otherwise the conditional belief might merely have "truthfully" preserved a falsity that was part of the agent cognitive system from the start. Thus we are led to the following process-reliabilist account of justification for conditional beliefs: (a) a conditional belief is justified if and only if it was produced by a reliable process; (b) consider any process that only produces conditional beliefs:[25] such a process is reliable if and only if most of the conditional beliefs it produces, or the great majority thereof, do not lead from truths to falsehoods. So the contribution that a conditional belief in B given A makes to the reliability of the process by which it is formed turns out to be measured by the truth value of the corresponding material implication $A \rightarrow B$. If the latter is true, then the generation of the conditional belief increases the reliability of the conditional-belief-producing process, and the reliability gets decreased otherwise.[26] Of course, this does not mean that the conditional belief in B given A is true if and only if $A \rightarrow B$ is true, since, as we have seen, conditional beliefs do not have truth values at all. But the *justification* of a conditional belief in B given A in terms of reliable processes is indeed tied closely to the truth of the material implication $A \rightarrow B$.

What might an internalist account of justification, or rationality, for conditional beliefs look like? Fortunately, here we can rely on theories of justification that have already been developed for such, or very similar, purposes (though perhaps not using the same terminology): On the *coherentist* picture, a conditional belief will be justified (with respect to a system S of beliefs) if and only if (a) it belongs to S, (b) S is coherent, and (c) the coherence of S is sufficiently accessible to the agent to enable her to put forward arguments in favour of it on the basis of the coherence of S. The coherence of a belief system can be explained in terms of rationality constraints on beliefs and on the way these beliefs get updated (the axioms for subjective probability spaces and those for belief revision operators are standard examples); the constraints themselves are justified on the basis of Dutch book arguments.[27]

In the case of *foundationalist* internalism, justification is supposed to be transferred from basic to non-basic beliefs. Whatever the notion of deductive or inductive justification-transfer—the reason-for relation—looks like that is used to explain in what sense beliefs can be epistemically founded on, and argued for, on the basis of more primitive beliefs, it must automatically yield a notion of justification for condi-

[24] There may be *conditional knowledge*, though we will not develop this thought here any further.

[25] This definition could easily be generalized to mixed processes, i.e. processes that generate both beliefs *simpliciter* and conditional beliefs, but let us restrict ourselves to the pure case for simplicity.

[26] Alternatively, we might say that conditional beliefs in B given A with false A should not be counted at all—rather than counting positively—as they are irrelevant with respect to justification. However, the difference between these two accounts is largely negligible.

[27] While it seems that no attempts have been made to give Dutch book arguments in favour of the axioms of belief revision, there is no principal reason why one should not do so: by well-known translations between ranked models and non-standard probability measures (see, e.g. the appendix of Lehmann and Magidor 1992), it would be possible to give Dutch book arguments for belief revision operators on the basis of bets with non-standard (e.g. infinitesimal) betting ratios.

tional beliefs: call a conditional belief in B given A justified if and only if the transition from the belief in A to the belief in B is justification-transmitting, or A is a reason for B. The fact that conditional beliefs are not beliefs *simpliciter* themselves and thus are not in need of the same sort of justification as simple beliefs, might even save the foundationalist from an epistemological third man or regress argument.

An internalist justification of conditional beliefs is therefore not only possible, it might actually be an integral component of any successful internalist treatment of justification.

We conclude that conditional beliefs can still be reasonably called *justified* or *unjustified*. By having a conditional belief an agent still aims at the truth, but she does so indirectly and instrumentally—by believing B given she believes in A, the agent aims at the truth of B given that her belief in A is true.

10 Summary—repercussions for philosophy and cognitive science

Summing up: On the basis of theorems by Lewis and Gärdenfors, we have argued that there is a class of mental states called ''conditional beliefs'' which are distinct from beliefs in conditionals and which are characterized by the following facts:

- Conditional beliefs have two propositional contents.
- Conditional beliefs are neither true nor false.
- Conditional beliefs are not beliefs in a material implication, though having a belief in a material implication entails having a corresponding conditional belief.
- Conditional beliefs are conditional states that can be described in terms of subjunctive conditionals (with ceteris-paribus clauses added).
- The conditional belief in B given A involves the disposition of acquiring the belief in B given all that the agent relevantly believes at the time is A; these dispositions are higher-order single-track dispositions which have a behaviour-determining function for the agent.
- Conditional beliefs are not beliefs in general propositions, though having a general belief entails having certain conditional beliefs.
- In cognitive systems that are based on symbolic computation, conditional beliefs can be implemented by means of non-descriptive default expressions, the syntactic components of which represent propositions.
- In connectionist systems, conditional beliefs can be embodied by the dynamics of neural pattern transformations, where patterns of activated units represent propositions.
- Conditional beliefs may be justified externalistically in terms of the reliable processes that generate them; the contribution that a conditional belief in B given A makes to the reliability of such a process is determined by the truth value of the material implication $A \rightarrow B$.
- Conditional beliefs can also be justified internalistically, or rather: standard coherentist and foundationalist theories of internalist justification rely on an account of justification for conditional beliefs.
- The property of having a unique propositional content, and of being true or false, is not essential for beliefs, though it is essential for beliefs *simpliciter*.

This leaves us with the following open questions, which address philosophical logicians and philosophers of mind, philosophers of language, and cognitive psychologists, respectively:

1. How can this theory be generalized to a theory for conditional beliefs with introspective propositional contents? [For example: What does the proper logical representation of conditional beliefs with introspective contents look like in view of limitative results such as the one by Fuhrmann (1989)?]
2. In which ways does the distinction between conditional beliefs and beliefs in conditionals map onto our linguistic practices involving conditionals? (For example: Is there any correspondence between indicative conditionals and conditional beliefs on the one hand and between subjunctive conditionals and beliefs in conditionals on the other?)
3. How can conditional beliefs be distinguished from beliefs in conditionals empirically, i.e. in psychological experiments? (For example: Did the celebrated Wason selection task examine reasoning on the basis of the supposition counterpart—''assume that if a card has...on one side, then it has...on the other side''—of a conditional belief, or rather reasoning on the basis of the supposition counterpart of a belief in a conditional?)

We have come a long way, from logical results on conditionals to empirical questions on human cognition. *Pace* Frege and Husserl, logic—with a little help from philosophy and cognitive science—might help us understand our mental capacities after all.

References

Adams EW (1965) The logic of conditionals. Inquiry 8:166–197
Adams EW (1975) The logic of conditionals. Reidel, Dordrecht

Armstrong DM (1968) A materialist theory of mind. Routledge, London
Armstrong DM (1973) Belief, truth and knowledge. Cambridge University Press, Cambridge
Barnett D (2006) Zif is if. Mind 115:519–565
van Benthem J (2006) Dynamic logic for belief revision. J Appl Non-Class Logic (in press)
Bird A (1998) Dispositions and antidotes. Philos Q 48:227–234
Blutner R (2004) Nonmonotonic inferences and neural networks. Synthese 142:143–174
Bradley R (2000) A preservation condition for conditionals. Analysis 60(3):219–222
Braithwaite RB (1932) The nature of believing. Proc Aristotelian Soc 33:129–146
Brewka G (ed) (1996) Principles of knowledge representation. CSLI Publications, Stanford
Brewka G, Dix J, Konolige K (1997) Nonmonotonic reasoning. An overview. CSLI Lecture Notes 73. CSLI Publications, Stanford
Churchland PM (1989) A neurocomputational perspective: the nature of mind and the structure of science. MIT, Cambridge, MA
Edgington D (1986) Do conditionals have truth conditions? Critica 18:3–30
Edgington D (1995) On conditionals. Mind 104:235–329
Edgington D (1996) Lowe on conditional probability. Mind 105:617–630
Eells E, Skyrms B (1994) Probability and conditionals. Belief revision and rational decision. Cambridge University Press, Cambridge
Fuhrmann A (1989) Reflective modalities and theory change. Synthese 81:115–134
Gärdenfors P (1986) Belief revisions and the Ramsey test for conditionals. Philos Rev 95:81–93
Gärdenfors P (1988) Knowledge in flux: modeling the dynamics of epistemic states. Bradford Book/MIT, Cambridge, MA
Ginsberg ML (ed) (1987) Readings in nonmonotonic reasoning. Morgan Kaufmann, Los Altos
Goldman AI (1986) Epistemology and cognition. Harvard University Press, Cambridge
Grahne G (1991) Updates and counterfactuals. In: Allen JA, Fikes R, Sandewall E (eds) Principles of knowledge representation and reasoning. Morgan Kaufmann, San Mateo, pp 269–276
Grove A (1988) Two modelings for theory change. J Philos Logic 17:157–170
Hájek A (1994) Triviality on the cheap? In: Eells E, Skyrms B (eds) Probability and conditionals. Belief revision and rational decision. Cambridge University Press, Cambridge, pp 113–140
Hájek A, Hall N (1994) The hypothesis of the conditional construal of conditional probability. In: Eells E, Skyrms B (eds) Probability and conditionals. Belief revision and rational decision. Cambridge University Press, Cambridge, pp 75–111
Hall N (1994) Back in the CCCP. In: Eells E, Skyrms B (eds) Probability and conditionals. Belief revision and rational decision. Cambridge University Press, Cambridge, pp 141–160
Hintikka JKK (1962) Knowledge and belief. Cornell University Press, Ithaca, NY
Howson C, Urbach P (1993) Scientific reasoning: the Bayesian approach, 2nd edn. Open Court, Chicago
Katsuno H, Mendelzon A (1992) On the difference between updating a knowledge base and revising it. In: Gärdenfors P (ed) Belief revision. Cambridge University Press, Cambridge, pp 183–203
Kraus S, Lehmann D, Magidor M (1990) Nonmonotonic reasoning, preferential models and cumulative logics. Artif Intell 44:167–207
Lehmann D, Magidor M (1992) What does a conditional knowledge base entail? Artif Intell 55:1–60
Leitgeb H (2004) Inference on the low level. An investigation into deduction, nonmonotonic reasoning, and the philosophy of cognition. Kluwer, Dordrecht
Leitgeb H (2005) Interpreted dynamical systems and qualitative laws: from inhibition networks to evolutionary systems. Synthese 146:189–202
Levesque H (1990) All I know: a study in autoepistemic logic. Artif Intell 42:263–309
Levi I (1988) Iteration of conditionals and the Ramsey test. Synthese 76:49–81
Levi I (1996) For the sake of the argument. Ramsey test conditionals, inductive inference, and nonmonotonic reasoning. Cambridge University Press, Cambridge
Lewis D (1973) Counterfactuals. Harvard University Press, Harvard
Lewis D (1976) Probabilities of conditionals and conditional probabilities. Philos Rev 85:297–315. Reprinted in Lewis D, Philosophical papers, vol II. Oxford University Press, Oxford, pp 133–152
Lewis D (1986) Probabilities of conditionals and conditional probabilities II. Philos Rev 95:581–589. Reprinted in Lewis D Papers in philosophical logic. Cambridge University Press, Cambridge, pp 57–65
Lindström S, Rabinowicz W (1995) The Ramsey test revisited. In: Crocco G, Farinas del Cerro L, Herzig A (eds) Conditionals from philosophy to computer science. Oxford University Press, Oxford, pp 157–202
Lowe EJ (1996) Conditional probability and conditional beliefs. Mind 105:603–615
Macdonald C, Macdonald G (eds) (1995) Connectionism: debates on psychological explanation, vol 2. Basil Blackwell, Oxford
Makinson D (1994) General patterns in nonmonotonic reasoning. In: Gabbay DM, Hogger CJ, Robinson JA (eds) Handbook of logic in artificial intelligence and logic programming, vol 3. Clarendon, Oxford, pp 35–110
Martin CB (1994) Dispositions and conditionals. Philos Q 44:1–8
Mellor DH (1993) How to believe a conditional. J Philos 90(5):233–248
Milne P (1997) Quick triviality proofs for probabilities of conditionals. Analysis 57(1):75–80
Newell A (1990) Unified theories of cognition. Harvard University Press, Cambridge
Ramsey FP (1929) General propositions and causality. In: Mellor DH (ed) F.P. Ramsey, Philosophical papers. Cambridge University Press, Cambridge, 1990
Reiter R (1980) A logic for default reasoning. Artif Intell 13:81–132
Ryle G (1949) The concept of mind. Hutchinson, London
Schurz G (2001) What is ''normal''? An evolution-theoretic foundation of normic laws and their relation to statistical normality. Philos Sci 68:476–497
Smolensky P (1988) On the proper treatment of connectionism. Behav Brain Sci 11:1–23
Spohn W (1988) Ordinal conditional functions. A dynamic theory of epistemic states. In: Harper WL, Skyrms B (eds) Causation in decision, belief change, and statistics, vol II. Kluwer, Dordrecht, pp 105–134
Stalnaker RC (1984) Inquiry. The MIT Press, Cambridge, MA