Revamping Hypothetico-Deductivism: A Dialectic Account of Confirmation

Gregor Betz*

October 2, 2011

Abstract

We use recently developed approaches in argumentation theory in order to revamp the hypothetico-deductive model of confirmation, thus alleviating the well-known paradoxes the H-D account faces. More specifically, we introduce the concept of dialectic confirmation on the background of the so-called theory of dialectical structures [Betz, 2010, 2011]. Dialectic confirmation generalises hypothetico-deductive confirmation and mitigates the raven paradox, the grue paradox, the tacking paradox, the paradox from conceptual difference, and the problem of novelty.

1 Introduction

According to the hypothetico-deductive account of confirmation, a scientific hypothesis is confirmed by its true deductive implications, in particular by its true empirical consequences. Thus, a successful prediction lends, for example, evidential support to the theory that entailed it. While the H-D model was spelled out and championed by logical positivists, it had figured prominently in the history of the philosophy of science before, being advocated, e.g., by scholars such as John Herschel and Stanley Jevons [c.f. Losee, 2001]. Yet, as we will review below, the H-D account of confirmation faces systematic objections to the effect that it is nowadays largely dismissed as indefensible. Notwithstanding this criticism, the H-D model seems to provide an adequate description of historical as well as current scientific practice, as its opponents admit [e.g. Glymour, 1980, p. 48]. Moreover, today's mainstream theory of

*Karlsruhe Institute of Technology, Germany, email: gregor.betz@kit.edu.

confirmation, i.e. Bayesianism, cites in its favour its ability to give a justification of hypothetico-deductive methods [Talbott, 2008, Howson and Urbach, 1989, p. 82]. Thus, a defence of the H-D model may still be worthwhile.

This paper takes on the challenge, arguing that hypothetico-deductivism provides yet an attractive outlook on empirical confirmation. In particular, it shows how recent developments in argumentation theory can be used to refine the simple H-D model so as to avoid the diverse objections which have been put forward against it.

The overall argument this paper unfolds proceeds as follows. The view that a hypothesis is confirmed by its true deductive implications gives rise to a bunch of well-known problems. Section 2 briefly reports these objections hypothetico-deductivism faces. The revamp of the H-D account that aims at circumventing those problems is couched within a specific argumentation-theoretic framework: The so-called theory of dialectical structures, being introduced in section 3. On the background of that framework, section 4 puts forward a dialectic notion of confirmation which modifies the simple H-D account. The concept of dialectic confirmation, we argue in section 5, allows one to solve the paradoxes outlined in section 2.

2 Problems of hypothetico-deductivism

Raven paradox. Consider the hypothesis H that all ravens are black and let a be an individual raven.[1] Clearly, H implies that a is black. According to the hypothetico-deductive account of confirmation, observing that a is black therefore confirms H. Likewise, the hypothesis H' that all non-black things are non-ravens is confirmed by verifying that some non-black object b is not a raven, but, e.g., a green leaf. Yet, the hypothesis H' is logically equivalent with H. Given that any two logically equivalent hypotheses are supported by the very same evidence, verifying that the green object b is a leaf should also confirm H. But that, or so it seems, is not the case: H does not speak of green things at all, let alone leaves, so how could

[1] Originally discussed by Hempel [1945, p. 11] as a problem of the instantial model of confirmation, the raven paradox represents an objection to the H-D account in general, because every instantial confirmation amounts to a hypothetico-deductive one [Lipton, 2004, p. 16]. See also Fitelson and Hawthorne [2010] for a recent review and a Bayesian response to the problem.

it be confirmed by observing a green leaf?

Grue paradox. As Goodman [1983, pp. 73ff.] has famously argued, the hypothesis that all emeralds are green implies exactly the same available empirical data (as of today) as the hypothesis which states that all emeralds are grue—where grue is the property of being green and being observed before some future point in time t or being blue and not being observed before t. According to the hypothetico-deductive account, both hypotheses are equally confirmed by our available evidence. Yet this is apparently not the case: We all (evidently rightly) infer, from our occasional encounters with jewellery, that emeralds are green, not grue.

Tacking paradox. Whatever is implied by some hypothesis H is also implied by the conjunction $H\&C$, where C represents an arbitrary, possibly irrelevant or even silly, claim. Thus, whatever confirms H, confirms, according to hypothetico-deductivism, $H\&C$ as well.[2] But this is, again, an absurd consequence. The observation of interference patterns confirms, for example, the wave-theory of light, whereas it does not confirm the hypothesis conjoining the wave-theory of light and the claim that humans are descended from kangaroos.

Paradox from conceptual difference. Consider two theoretical hypotheses, H_1 and H_2, that are empirically equivalent but conceptually distinct. Hypothetico-deductivism has it, then, that the evidential support for H_1 is the same as for H_2. But now assume: "H_1, but not H_2, is derivable from a more general theory T, which also entails another hypothesis H. An empirical consequence e of H is obtained. e supports H and thereby T. Thus e provides indirect evidential warrant for H_1, of which it is not a consequence, without affecting the credentials of H_2." [Laudan, 1991, p. 464] In contradiction to hypothetico-deductivism, H_1 and H_2 are not confirmed by the same items of evidence.

[2]This paradox is also referred to as the "problem of irrelevant conjunction" [cf. Glymour, 1980, p. 31]. Provided the highly plausible views that (i) a statement confirms its deductive consequences and that (ii) the confirmation relation is transitive, the tacking paradox gives rise to an additional *reductio*: It follows that any piece of evidence, which confirms some statement at all, confirms every statement whatsoever [e.g. Goodman, 1983, pp. 67-68].

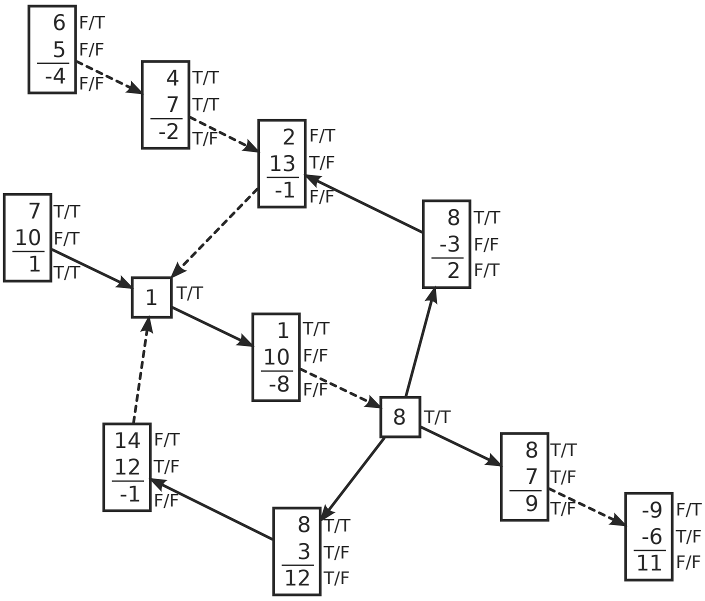


Figure 1: A dialectical structure with two complete positions attached. Truth values are symbolised by "T" (true) and "F" (false).

Problem of novelty. According to the hypothetico-deductive account of confirmation, a hypothesis is confirmed by entailed items of evidence no matter when the respective observations were made, or whether they were unexpected or not. As a result, hypothetico-deductivism collides with the highly plausible view that new and surprising predictions which turn out to be correct confirm a hypothesis H to a higher degree than old, well-known evidence which is implied by H.[3]

3 Fundamentals of the theory of dialectical structures

A **dialectical structure** $\tau = \langle T, A, U \rangle$ is a set of deductively valid arguments (premiss-conclusion structure), T, on which an attack relation, A, and a support relation, U, are defined as follows ($a, b \in T$):

- $A(a, b) :\iff$ a's conclusion is contradictory to one of b's premisses;
- $U(a, b) :\iff$ a's conclusion is equivalent to one of b's premisses.[4]

[3]It is, in other words, not clear whether hypothetico-deductive accounts of confirmation can avoid accomodationism.

[4]The theory of dialectical structures is more thoroughly developed in Betz [2010].

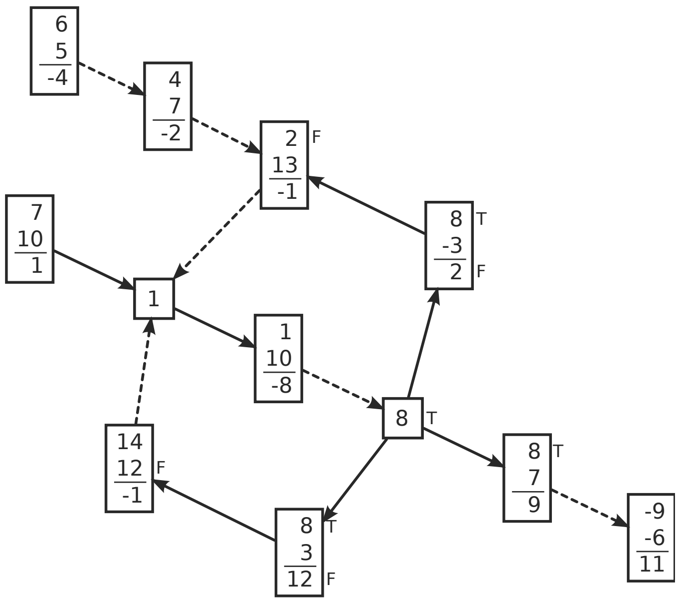


Figure 2: A dialectical structure with one partial positions attached.

Complex debates—philosophical, political or scientific ones—can be reconstructed as dialectical structures. Figure 1 depicts a purely formal example of a dialectical structure. Numbers stand for sentences, and a negative number denotes the negation of the sentence which is designated by the corresponding positive integer. Each box represents an argument or a thesis.[5] Continuous and dashed arrows indicate the support and attack relationship, respectively.

Relative to a dialectical structure τ, which in a sense depicts the state of a debate, one can specify the positions of different proponents. We may, generally, distinguish complete and partial positions. A **complete position** $\mathcal{Q}$ (a proponent can adopt) on τ is a truth value assignment to all sentences which figure in arguments in T, i.e. $\mathcal{Q} : S \to \{t, f\}$, where S is the set of all sentences in τ. A **partial position** $\mathcal{P}$ (a proponent can adopt) on τ is a truth value assignment to some sentences which figure in arguments in T, i.e. $\mathcal{P} : S' \to \{t, f\}$, where $S' \subseteq S$. Whereas figure 1 shows two complete positions defined on a dialectical structure, figure 2 gives an example for a partial position defined on the very same debate.

[5]Introducing theses into the argument maps does not require one to modify the definition of a dialectical structure. Formally, theses can simply be understood as arguments containing exactly one premiss and a conclusion identical with that premiss.

Partial positions can be combined. Let $\mathcal{P}_1 : S_1 \rightarrow \{t, f\}$ and $\mathcal{P}_2 : S_2 \rightarrow \{t, f\}$ be two partial positions which agree on $S_1 \cap S_2$. The **conjunction** of these positions, $(\mathcal{P}_1 \& \mathcal{P}_2) : S_1 \cup S_2 \rightarrow \{t, f\}$, can be defined by,

$$p \mapsto \begin{cases} \mathcal{P}_1(p) & \text{if } p \in S_1 \\ \mathcal{P}_2(p) & \text{if } p \in S_2 \setminus S_1 \end{cases} .$$

Obviously, the arguments that make up a dialectical structure impose certain constraints on what a proponent can reasonably assert. Not every complete or partial position can rationally be adopted. Thus, a complete position $\mathcal{Q}$ on τ is **(dialectically) coherent** if and only if

1. equivalent sentences are assigned the same truth value;
2. contradictory sentences are assigned complementary truth values;
3. if every premiss p of some argument $a \in T$ is assigned the value "true", then a's conclusion is assigned the value "true", too.

A partial position $\mathcal{P} : S' \rightarrow \{t, f\}$ on τ is **(dialectically) coherent** if and only if it can be extended to a complete position $\mathcal{Q}$ on τ ($\mathcal{P} = \mathcal{Q}|_{S'}$) which is coherent.

Returning to the complete positions depicted in figure 1, we may note:

- The left-hand-side complete position in that example is coherent. It complies with the coherence conditions set up above.
- The right-hand-side complete position is, however, not coherent. Yet, that position does not merely defy one coherence constraint. Actually, all three conditions are violated. Thus, the right-hand-side position violates constraint (1) because the tokens of sentence 10 are assigned different truth values; it violates constraint (2) because the contradictory sentences 3/-3 are both considered true; and it violates constraint (3) because the conclusion of argument (4,7;-2) is false despite its premisses being true.

Moreover, the partial position shown in figure 2 is not coherent, either, since it cannot be extended to a complete, coherent position. To see this, consider, first of all, the argument (8,-3;2). Since premiss (8) is true and the conclusion (2) is false, the remaining premiss (-3) has to be false. Otherwise the complete position would violate the third constraint. Secondly, in regard of argument (8,3;12), sentence (3) must be false for analogous reasons. Yet (3) and (-3) cannot be false in the same time because of the second constraint. Hence, the partial position cannot be extended to a coherent complete position.

Based on these primitive notions, we can now introduce the concepts of dialectic entailment and degree of partial entailment, which will play a major role in our attempt to rephrase hypothetico-deductivism. Thus, a partial position $\mathcal{P}_2$ **dialectically entails** a partial position $\mathcal{P}_1$, iff all coherent and complete positions which extend $\mathcal{P}_2$ equally extend $\mathcal{P}_1$. The concept of dialectic entailment may be generalised by following Wittgenstein's basic idea in the *Tractatus* (and identifying cases with complete and coherent positions on τ): The **degree of partial entailment** of a partial position $\mathcal{P}_1$ by a dialectically coherent partial position $\mathcal{P}_2$, can be defined as,

$$\begin{aligned} \textsc{Doj}(\mathcal{P}_1|\mathcal{P}_2) &:= \frac{\text{number of cases with } \mathcal{P}_1 \text{ and } \mathcal{P}_2}{\text{number of cases with } \mathcal{P}_2} \\ &= \frac{\begin{array}{c}\text{number of complete \& coherent positions that}\\ \text{extend } \mathcal{P}_1 \text{ and } \mathcal{P}_2\end{array}}{\begin{array}{c}\text{number of complete \& coherent positions that}\\ \text{extend } \mathcal{P}_2\end{array}}. \end{aligned} \quad (1)$$

As a consequence, $\textsc{Doj}(\mathcal{P}_1|\mathcal{P}_2) = 1$ if and only if $\mathcal{P}_2$ dialectically entails $\mathcal{P}_1$. Degrees of partial entailment satisfy, under certain conditions which we shall assume to hold[6], the **axioms of probability theory**.[7]

[6]This is the problem: For every probability measure over a set of statements, it holds that $P(A \vee B) = P(A) + P(B)$ for contrary A, B. Now assume that the three sentences $A \vee B$, A and B figure in some τ and that there is no dialectically coherent position according to which both A and B are true. Still, this does not guarantee that the (unconditional) degrees of partial entailment of A and B add up to the (unconditional) degree of partial entailment of $A \vee B$. This is because not every coherent complete position according to which A is true assigns $A \vee B$ the value "true"—unless an argument like (A;$A \vee B$) is included in τ. Thus, degrees of partial entailment satisfy the probability axioms only if the respective dialectical structure is suitably augmented by simple arguments as indicated.

[7]This is, however, not to say that the concept of degree of partial entailment represents a theoretical explication of our pre-theoretical notion of probability (as in "It is very improbable that she wins the national lottery"). The theory of dialectical structures is therefore not committed to the assumption that all cases (i.e. positions) are equally distributed in the sense of being equally likely.

Finally, the **degree of justification** of a partial position $\mathcal{P}$ can be defined as its degree of partial entailment from the empty set,

$$\textsc{Doj}(\mathcal{P}) := \textsc{Doj}(\mathcal{P}|\emptyset) \tag{2}$$

$$= \frac{\text{number of complete \& coherent positions that extend } \mathcal{P}}{\text{number of complete \& coherent positions}}. \tag{3}$$

It can be shown that degrees of justification possess the following properties [cf. Betz, 2011]:

- Introducing an independent[8] argument that supports (attacks) some thesis t increases (decreases) t's degree of justification.
- Introducing an independent argument that supports (attacks) a supporting argument for some thesis t increases (decreases) t's degree of justification.
- Introducing an independent argument that supports (attacks) some argument which attacks thesis t decreases (increases) t's degree of justification.
- Incorporating premisses of independent arguments that support (attack) some thesis t into the background knowledge increases (decreases) t's degree of justification.

As the degree justification is an indicator of a partial position's robustness, and as it is reasonable to maximise the robustness of one's position, a high degree of justification points to the belief-worthiness of a position [cf. Betz, 2011].[9]

Degrees of justification, as defined above, are sensitive to changes in the dialectical structure which represents the reconstruction of a debate, e.g. a scientific controversy. Far from being a disadvantage, it is precisely this dependency which allows one to understand why the belief-worthiness, and the overall confirmation, of a hypothesis might change in the course of an ongoing argumentation.

[8] i.e. an argument whose premisses are neither equivalent nor contradictory to any sentences already contained in the debate.

[9] A formal investigation of the concept of degree of justification in Betz [forthcoming], based on multi-agent simulations of debate dynamics, suggests that degrees of justification may also serve as an indicator of truthlikeness. These results, if actually correct, could substantially strengthen this paper's dialectic defence of the H-D account.

4 A dialectic account of confirmation

The degree of justification of some hypothesis H relative to a dialectical structure and a given body of background knowledge largely determines H's overall belief-worthiness. If $\text{DOJ}(H)$ equals 1, H is true.[10] If $\text{DOJ}(H) = 0$, H is false. In addition, the higher $\text{DOJ}(H)$, the more arguments count in favour of H, and the less objections H faces. The overall belief-worthiness of H *given* some evidence E is nothing but the *conditional* degree of justification $\text{DOJ}(H|E)$, i.e. H's degree of justification provided E is incorporated into the background knowledge. It is straightforward to assess the confirmatory impact of E on H by comparing $\text{DOJ}(H)$ and $\text{DOJ}(H|E)$. If, for instance, establishing E as true increases H's degree of justification, E inductively supports H. Thus,

$$\begin{aligned} E \text{ confirms } H &\iff \frac{\text{DOJ}(H|E)}{\text{DOJ}(H)} > 1, \\ E \text{ disconfirms } H &\iff \frac{\text{DOJ}(H|E)}{\text{DOJ}(H)} < 1, \\ E \text{ is neutral reg. } H &\iff \frac{\text{DOJ}(H|E)}{\text{DOJ}(H)} = 1. \end{aligned}$$

The ratio $\text{DOJ}(H|E)/\text{DOJ}(H)$ can be defined as the **degree of confirmation** of H by E.[11] It corresponds to the probabilistic ratio measure of confirmation and gauges the relative increase of H's degree of justification which results from incorporating E into the body of background knowledge.[12] The overall degree of justification of H once the evidence E is verified, i.e. the belief-worthiness of H given E and the state of the debate, trivially satisfies,

$$\text{DOJ}(H|E) = \frac{\text{DOJ}(H|E)}{\text{DOJ}(H)} \times \text{DOJ}(H), \tag{4}$$

[10]It is tacitly understood that all degrees of justification are determined relative to some body of background assumptions, B, without making this explicit.

[11]This ratio is undefined if $\text{DOJ}(H) = 0$. But note that $\text{DOJ}(H) = 0$ implies $\text{DOJ}(H|E) = 0$. The degree of justification of a hypothesis whose falsehood is implied by the background knowledge cannot be increased by augmenting the background knowledge.

[12]This said, the ratio measure represents, obviously, only one among many alternative ways for defining quantitative degrees of confirmation in a probabilistic way [see, e.g., Fitelson, 1999]. This paper, however, is content by resolving the paradoxes of the H-D account based on the ratio measure. It remains to be studied how different measures of confirmation behave in the dialectic framework.

and thus depends on (i) the degree of confirmation of H by E, and (ii) the prior degree of justification of H irrespective of E. These are the two directions from where support in favour of some hypothesis might generally stem. Evidential support, on the one side, as well as theoretical support "from above" [cf. Hempel, 1966, pp. 38-40], on the other side, both increase a hypothesis' overall degree of justification. Consequently, the degree of confirmation of H by some evidence is not the only indicator for whether H is highly justified and belief-worthy.[13]

Note that degrees of justification as well as degrees of confirmation can be calculated, given some dialectical structure, for arbitrary pairs of hypotheses and items of evidence. As a consequence, the dialectic account allows one

- to distinguish the confirmatory relevance of different pieces of evidence for a given hypothesis, as well as
- to determine the specific evidential support some empirical data bestows on different hypotheses, e.g. different parts of a complex theory.

This account thus settles a major short-coming of the simple H-D model of confirmation which Glymour [1980, pp. 30ff.] rightly identifies.

In terms of mathematical structure, the dialectic account set forth so far closely resembles Bayesian Confirmation Theory. So, in which sense does it represent a revamp of the H-D account at all? Besides being based on a fully deductive argumentation framework, the dialectic account generalises the key tenet of the H-D model, namely that a hypothesis is confirmed by its deductive implications.[14] For according to the dialectic model, a hypothesis is confirmed by those statements it

[13] Thus, we clearly distinguish (a) the overall degree of belief-worthiness of a hypothesis and (b) the degree of confirmation of that hypothesis by some evidence. Larry Laudan once remarked that "'possessing the same positive instances' and 'being equally well confirmed' boil down to the same thing only in the logician's never-never land. (It was Whewell, Peirce and Popper who taught us all that theories sharing the same positive instances need not be regarded as equally well tested or equally belief-worthy.)" [Laudan, 1990, 282-283]. Whilst we don't consider the theory exposed in this section a Neverland's account of confirmation in the first place, the desideratum Laudan articulates, namely to differentiate empirical consequences on the one hand and overall belief-worthiness on the other hand, is clearly met.

[14] It concurs, in this respect, with previous, non-Bayesian attempts to defend the H-D account against the paradoxes (reviewed in improved upon by Sprenger [2011a,b]), which replaced the classic concept of logical entailment—originally used to articulate hypothetico-deductivism—by a (weaker) notion of relevant implication [e.g. Schurz, 1991, Gemes, 1993, 1994, 2005]. Now, instead of developing a formal notion of relevant entailment, this paper employs the concept of partial dialectic entailment as a substitute for logical entailment.

partially entails (to a certain, sufficient degree). More precisely, the degree of partial entailment of E by H fully determines the degree of confirmation of H by E, as follows directly from Bayes' theorem.[15]

5 Resolving the problems

Raven paradox. We may begin our dialectical analysis of the raven paradox with the obvious observation that the hypothesis H does not entail that there exists a black raven. H does entail—provided the auxiliary assumption that a is a raven—that a is black. Likewise, hypothesis H' does not entail that there exists a green leaf. But it does entail—provided the auxiliary assumptions that (i) b is green and that (ii) nothing is green and black in the same time—that b is not a raven. So, as a first thing to note, the situation is not fully symmetrical. To observe that some (given) raven is black confirms H. To observe that some (given) leaf is green does, however, not confirm H'. What confirms H', and thus H, is the observation that some (given) non-black entity (e.g. something green) is not a raven (but a leaf). This apparent asymmetry is also revealed by a more detailed analysis. Setting

Key	Proposition
E_1	Object a is black.
A_1	Object a is a raven.
E_2	Object b is a leaf.
E_3	Object b is not a raven.
A_4	Object b is green.
A_5	Nothing is black and green in the same time.
A_6	Nothing is a leaf and a raven in the same time.

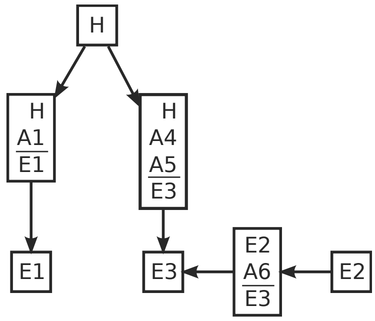
the dialectical structure in figure 3 depicts the inferential relations that hold between the various pieces of evidence and the hypothesis H. Specifically, H entails the items

[15] Bayes' theorem, in its simplest variant, reads,

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}.$$

We thus have,

$$\frac{P(H|E)}{P(H)} \propto P(E|H).$$



	$\text{DOJ}(H\|E_i)$	$=$	$\frac{\text{DOJ}(H\|E_i)}{\text{DOJ}(H)}$	$\times$	$\text{DOJ}(H)$
$i = 1$	0.47		1.18		0.40
$i = 2$	0.41		1.02		0.40

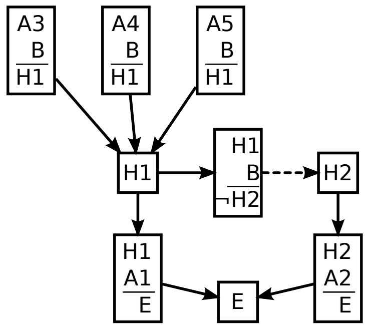
Figure 3: Argument map depicting the dialectical structure of the raven paradox. The table gives the corresponding degrees of partial entailment and degrees of confirmation.

E_1 and E_3—namely via A_1, and via A_4 and A_5, respectively. E_3, in turn, is backed by the argument (E_2,A_6;E_3). Relative to this dialectical structure, degrees of partial entailment and degrees of confirmation can be calculated[16]; results are displayed in the table in figure 3. The hypothesis H is entailed by E_1 to a higher degree than by E_2. Likewise, the degree of confirmation of H by E_1 is significantly higher than by E_2. Still, E_2 does positively confirm H, albeit to a very small degree. Thus, the raven paradox is solved: Observing that some raven is black confirms the raven hypothesis H to a higher degree than observing that some green object is a leaf—notwithstanding that the latter lends some minimal support to H, too.[17]

Grue paradox. A resolution of the grue paradox, understood as an objection to the hypothetico-deductive account of confirmation, should begin by stressing that theory-confirmation, as well as theory-choice, is always carried out (a) against some body of (theoretical) background knowledge and (b) with regard to a finite, contingently composed set of alternatives. The purpose of a philosophical theory

[16] For this as well as the following numerical calculations see the supplementary material.

[17] It might be objected that A_5 and A_6 should count as background knowledge. That would, indeed, make the dialectic situation symmetrical and E_2 (resp. E_3) confirmed H to the same degree as E_1. Note, however, that even in this case, H is not at all confirmed by the observation that some object x, which is a leaf (non-raven), is actually green (non-black); but only by the observation that some green (non-black) entity is not a raven and hence no counter-example to H. In the end, this doesn't seem to be highly paradoxical.

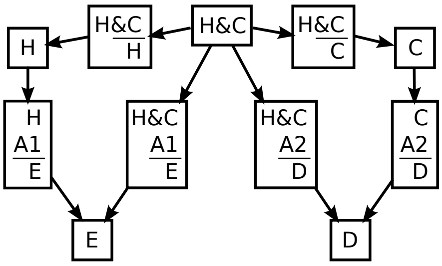


	$\mathrm{DoJ}(H_i \vert E)$	=	$\frac{\mathrm{DoJ}(H_i \vert E)}{\mathrm{DoJ}(H_i)}$	×	$\mathrm{DoJ}(H_i)$
$i = 1$	0.80		1.03		0.77
$i = 2$	0.10		1.03		0.097

Figure 4: Argument map depicting the dialectical structure of the grue paradox. The table gives the corresponding degrees of partial entailment and degrees of confirmation.

of confirmation, we take it, is not to explain and justify how empirical data—and empirical data *alone*—may support theoretical hypotheses. On the contrary, such a theory must clarify how empirical data, given some background knowledge, may lent support to one hypothesis, compared to a set of rivals.

This said, the grue paradox can be resolved by explicitly including the theoretical arguments for (or against) the rival hypotheses, i.e. the support 'from above', into the analysis. In contrast to a similar strategy discussed by Goodman [1983, p. 77] himself—namely to determine the *lawlikeness* of a general statement in terms of its overall theoretical support so as to set apart the lawlike statement that all emeralds are green (H_1) from the merely accidental claim that they are grue (H_2)—we suggest to take the theoretical support directly into account in the dialectical analysis, without touching the concept of lawlikeness. Let us consider, to spell out this idea, Goodman's example in some more detail. The relevant background knowledge which determines the overall belief-worthiness of H_1 and H_2 comprises, for instance, the claim that the optic properties of materials supervene on their molecular or atomic structure. The colour of a stone therefore simply doesn't depend on when we observe it. In a dialectical analysis, this general background knowledge translates into a theoretical argument in favour of H_1 (or, alternatively, an argument against H_2). Apart from the theoretical support for H_1, the rival hypotheses entail, by con-



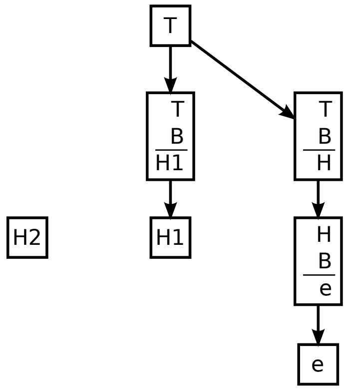
	$\textsc{DoJ}(H_i \vert E)$	=	$\frac{\textsc{DoJ}(H_i \vert E)}{\textsc{DoJ}(H_i)}$	×	$\textsc{DoJ}(H_i)$
$i = 1$	0.59		1.14		0.52
$i = 2$	0.18		1.14		0.16
$i = 3$	0.53		1.02		0.52

Figure 5: Argument map depicting the dialectical structure of the tacking paradox. The table gives the corresponding degrees of partial entailment and degrees of confirmation with $H_1 := H$, $H_2 := H\&C$ and $H_3 := C$.

struction, the same available evidence E. Figure 4 depicts the respective dialectical structure. The theoretical arguments which back H_1 rely on auxiliary assumption $A_3, \ldots, A_5$ as well as some background knowledge abbreviated by "B". The arguments (H_1,A_1;E) and (H_2,A_2;E) state that both hypotheses entail the available evidence. As the corresponding table reports, the evidence E confirms H_1 as much as H_2. The degree of justification of H_1 given E, however, is significantly higher than of H_2—a fact entirely due to H_1's greater prior degree of justification. So even though the empirical evidence confirms both hypotheses to the same degree, H_1 is, generally, much better justified. Refined hypothetico-deductivism gives by no means rise to the contrary-to-fact implication that H_1 and H_2 are, all things considered, equally well supported.[18]

Tacking paradox. Consider a situation as described in the tacking paradox above. We shall assume that claim C, conjoined to the original hypothesis H, is neither

[18] Along this general line of reasoning, we may equally solve the additional illustrative riddle Goodman mentions to exemplify his paradox [Goodman, 1983, p. 73]. So, why does observing that a given piece of copper conducts electricity (E_1) confirm the hypothesis that copper is conductive (H_1), whereas observing that a man in this room is a third son (E_2) does not confirm, or so it seems, the hypothesis that all men in this room are third sons (H_2)? The answer is that, contrary to one's first impression, both pieces of evidence do lend support to the respective hypothesis. In spite of this fact, however, H_1 is, all things considered, more belief-worthy than H_2 because of additional theoretical support for H_1 (all metals are conductors) as well as theoretical objections to H_2 (less than x% of all parents have 3 sons) and, possibly, direct evidential disconfirmation of H_2 (this man over there pretends to have no siblings).



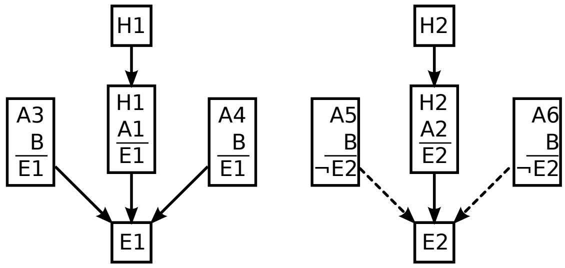
	$\text{Doj}(H_i\|e)$	$=$	$\frac{\text{Doj}(H_i\|e)}{\text{Doj}(H_i)}$	$\times$	$\text{Doj}(H_i)$
$i=1$	0.6		1.05		0.57
$i=2$	0.5		1.0		0.5

Figure 6: Argument map depicting the dialectical structure of the paradox from conceptual difference. The table gives the corresponding degrees of partial entailment and degrees of confirmation.

empirically empty nor empirically equivalent with H, i.e. C has observational implications, say D, which are not implied by H. Consequently, the conjoined hypothesis $H\&C$ implies observational statements D as well as E whereas H implies E only. Figure 5 represents these inferential relations by a dialectical structure, presuming that the derivation of E and D relies on some auxiliary assumptions. As can be seen from the corresponding table, E increases, as a matter of fact, the prior degree of justification of the simple hypothesis H (that is: confirms H) to the same degree as the conjoined hypothesis $H\&C$. Yet, the individual hypotheses C receives much less empirical confirmation from E than H or $H\&C$. Moreover, the simpler hypothesis displays, because of ample prior justification, the greatest conditional degree of justification given E and is, in particular, far better justified than the conjoined hypothesis.[19]

Paradox from conceptual difference. In order to solve the paradox from conceptual difference, we will show that, according to the dialectic account, H_1 is confirmed by the piece of evidence e although H_1 doesn't imply e. Figure 6 contains

[19]This solution resembles the Bayesian response to the tacking paradox inasmuch as it is accepted that even irrelevant conjunctions receive some (albeit marginal) empirical support [cf. Hawthorne and Fitelson, 2004].



	$\text{DOJ}(H_i\|E_i)$	=	$\frac{\text{DOJ}(H_i\|E_i)}{\text{DOJ}(H_i)}$	×	$\text{DOJ}(H_i)$
$i = 1$	0.5		1.06		0.47
$i = 2$	0.5		1.33		0.38

Figure 7: Argument map depicting the dialectical structure of the problem of novelty. Note that the dialectical structure consists of two independent sub-structures, and that the hypotheses H_1 and H_2 don't represent rivals. The table gives the corresponding degrees of partial entailment and degrees of confirmation.

a reconstruction of the dialectical situation pictured in the problem set-up. Since H_1 and H_2 possess the same empirical implications, E, and since we're not going to investigate how any hypothesis is confirmed by that very evidence, we may disregard E for the sake of simplicity in our analysis. Again, "B" refers to background assumptions which are taken for granted. Note that because H_1 and H_2 are "conceptually distinct", the general theory T implies H_1 without entailing H_2. As the respective table details, the individual item of evidence e imparts indeed some positive backing to the hypothesis H_1, increasing its prior degree of justification by 5%.

Problem of novelty. We shall demonstrate that refined hypothetico-deductivism can take novelty into account when assessing the confirmatory impact of some evidence. In order to do so, we consider two independent situations. In our first case, some hypothesis H_1 implies a true, albeit expected item of evidence E_1. In the second case, however, hypothesis H_2 entails the novel datum E_2, which—surprisingly, and unexpectedly—turns out to be true. Generally spoken, the fact that some evidence E is expected, or even well-known, corresponds, in terms of the dialectical analysis, to the fact that E is supported by arguments which largely rely on given background assumptions. A piece of evidence E is, on the contrary, unexpected and surprising, if there are arguments against E being the case. In line with this general

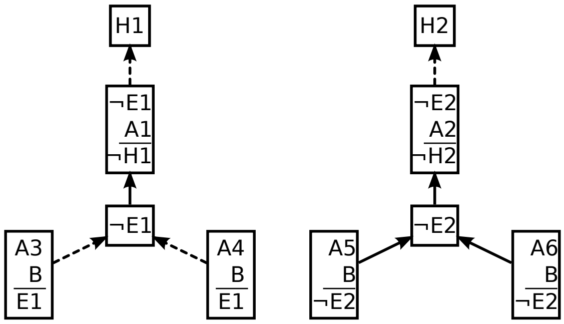


Figure 8: Transformed dialectical structure which results from argument-wise contraposition of figure 7.

reasoning, we take it that E_1 is supported by two arguments, while E_2 faces two objections which render it prima facie unexpected. Figure 7 shows the dialectical structure that maps these two cases. Note, as a caveat, that H_1 and H_2 must not be thought of as rival hypotheses. Rather, we have depicted two independent situations which can be compared in order to gauge the effect of novelty on confirmation. As set out in the respective table, the unexpected and surprising evidence E_2 increases H_2's degree of justification by 34% against merely 6% for H_1 and the previously expected empirical datum E_1. Thus, novelty boosts confirmatory strength. Note that the superior impact of the novel evidence results from the lower prior degree of justification of H_2. Once the evidence is verified, both hypothesis actually possess, all things considered, the same degree of justification.

These quantitative results can be given a qualitative explanation within the theory of dialectical structures. As a preliminary step, we observe that different arguments may encode the very same inferential relations. Consider for instance $(P_1,P_2;P_3)$, $(\neg P_3,P_2;\neg P_1)$, and $(P_1,\neg P_3;\neg P_2)$—all three arguments basically say that P_1, P_2 and $\neg P_3$ cannot be true in the same time. Because such an inferential relation can be represented by different arguments, different dialectical structures, too, may encode the very same inferential relations. Specifically, replacing one argument in τ by a contrapositive variant yields an equivalent dialectical structure τ' such that a position is dialectically coherent in τ iff it is coherent in τ'. Given these general, preparatory remarks, it is plain that the dialectical structure depicted in figure 8 is equivalent to the one shown in figure 7. In the transformed dialectical structure (figure 8), each hypothesis is directly attacked by an argument. But while the argument against H_1 is, in turn, attacked by two counter-arguments, thus weakening the direct objection, the argument attacking H_2 is backed, and thereby strengthened, by two additional arguments. The two direct objections against H_1 and H_2 rely on the assumption that the corresponding evidence (E_1 and E_2, respectively) is false. Thus, verifying these items of evidence eclipses the objections: Establishing E_1 blocks the comparatively weak objection to H_1, establishing E_2 shuts off the relatively strong argument against H_2. Comprehensibly, validating E_2 has therefore a higher confirmatory impact than validating E_1. Moreover, once both direct counter-arguments are eclipsed, the two hypotheses, facing no further objections, are, in terms of degree of justification, on a par.

6 Conclusion

This paper argued that modifying the H-D account of confirmation in the light of recent advances in argumentation theory allows one to resolve the paradoxes which beset hypothetico-deductivism. We may identify, more specifically, two points which have proven crucial in alleviating the problems: This is, firstly, the introduction of the concept of degree of partial entailment which, in turn, enabled us to set forth a quantitative notion of degree of confirmation; and, secondly, the neat conceptual distinction between (i) the evidential support some observation bestows on a hypothesis and (ii) the overall belief-worthiness (all things considered) of a hypothesis.

The credibility of the formal analysis put forward in this paper could be greatly enhanced by concrete case-studies which consist in reconstructions of scientific debates and which show the abstract confirmatory mechanism at work. That research could, e.g., address the interplay between inductive and deductive argumentation in real scientific controversies.

References

Gregor Betz. *Theorie dialektischer Strukturen.* Klostermann, Frankfurt am Main, 2010.

Gregor Betz. On degrees of justification. *Erkenntnis*, forthcoming, 2011.

Gregor Betz. *Debate Dynamics: How Controversy Improves Our Beliefs*. Synthese Library. Springer, Dordrecht, forthcoming.

Branden Fitelson. The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, 66(3):362–378, 1999.

Branden Fitelson and James Hawthorne. How Bayesian Confirmation Theory handles the paradox of the ravens. In Ellery Eels and J.H. Fetzer, editors, *The Place of Probability in Science*, Boston Studies in the Philosophy of Science, pages 247–276. Springer, Dordrecht, 2010.

Ken Gemes. Hypothetico-deductivism, content and the natural axiomitization of theories. *Philosophy of Science*, 60(3):477–487, 1993.

Ken Gemes. A new theory of content I: Basic content. *Journal of Philosophical Logic*, 23(6):595–620, 1994.

Ken Gemes. Hypothetico-deductivism: Incomplete but not hopeless. *Erkenntnis*, 63(1):139–147, 2005.

Clark N. Glymour. *Theory and Evidence*. Princeton University Press, Princeton, N.J., 1980.

Nelson Goodman. *Fact, Fiction, and Forecast*. Harvard University Press, Cambridge, Mass., 4th edition, 1983.

James Hawthorne and Branden Fitelson. Discussion: Re-solving irrelevant conjunction with probabilistic independence. *Philosophy of Science*, 71(4):505–514, 2004.

Carl G. Hempel. Studies in the logic of confirmation (I.). *Mind*, 54(213):1–26, 1945.

Carl G. Hempel. *Philosophy of Natural Science*. Prentice-Hall Foundations of Philosophy Series. Prentice-Hall, Englewood Cliffs, N.J.,, 1966.

Colin Howson and Peter Urbach. *Scientific Reasoning : The Bayesian Approach*. Open Court, La Salle, Ill., 1989.

Larry Laudan. Demystifying underdetermination. In C. Wade Savage, editor, *Scientific Theories*, volume 14 of *Minnesota Studies in the Philosophy of Science*, pages 267–297. University of Minnesota Press, Minneapolis, 1990.

Larry Laudan. Empirical equivalence and underdetermination. *The Journal of Philosophy*, 88(9):449–472, 1991.

Peter Lipton. *Inference to the Best Explanation*. Routledge, London, 2004.

John Losee. *A Historical Introduction to the Philosophy of Science*. Oxford University Press, Oxford ; New York, 4th edition, 2001.

Gerhard Schurz. Relevant deduction. *Erkenntnis*, 35(1):391–437, 1991.

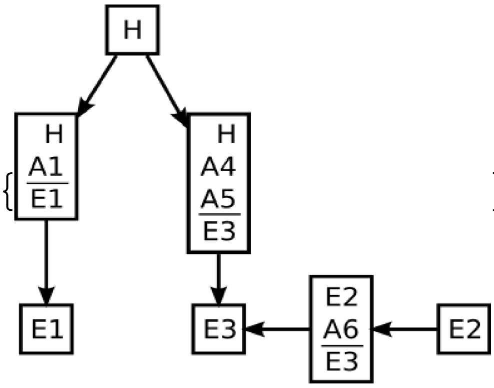
Jan Sprenger. Hypothetico-deductive confirmation. *Philosophy Compass*, 6(7):497–508, 2011a.

Jan Sprenger. A synthesis of Hempelian and hypothetico-deductive confirmation. *Manuscript*, 2011b. URL http://www.laeuferpaar.de/Papers/H-DHempel.pdf.

William Talbott. Bayesian epistemology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition, 2008. URL http://plato.stanford.edu/archives/fall2008/entries/epistemology-bayesian/.

```
(*************General functions***************)
(*doj of p in tau*)
DOJ[p_, tau_] :=
  SatisfiabilityCount[tau && p] / SatisfiabilityCount[tau];
(*conditional doj of p given q in tau*)
DOJ[p_, q_, tau_] := DOJ[p, tau && q];
```

```
(**********Calculation of DOJs for RavenParadox**********)
```

{  };


```
(*Set up the boolean constraints*)
ClearAll[tau, a1, a2, a3, a4, a5, a6];
a1 = (H && A1) ⇒ E1;
a2 = (H && A4 && A5) ⇒ E3;
a3 = (E2 && A6) ⇒ E3;
tau = a1 && a2 && a3;
```

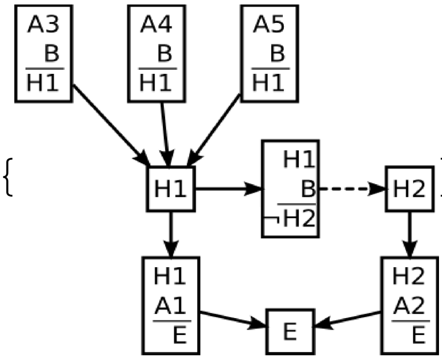
```
(*DOJs as in figure 3*)
TableForm[ DOJ[H, E1, tau]   DOJ[H, E1, tau] / DOJ[H, tau]   DOJ[H, tau]
           DOJ[H, E2, tau]   DOJ[H, E2, tau] / DOJ[H, tau]   DOJ[H, tau] ,
 TableHeadings → {None, {"P(H|Ei)=", "P(H|Ei)/P(H)*", "P(H)"}}]
```

P(H|Ei)=	P(H|Ei)/P(H)*	P(H)
$\frac{25}{53}$	$\frac{187}{159}$	$\frac{75}{187}$
$\frac{11}{27}$	$\frac{2057}{2025}$	$\frac{75}{187}$

```
TableForm[N[%], TableHeadings → {None, {"P(H|Ei)=", "P(H|Ei)/P(H)*", "P(H)"}}]
```

P(H|Ei)=	P(H|Ei)/P(H)*	P(H)
0.471698	1.1761	0.40107
0.407407	1.0158	0.40107

```
(**********Calculation of DOJs for Grue Paradox**********)
```

{  };

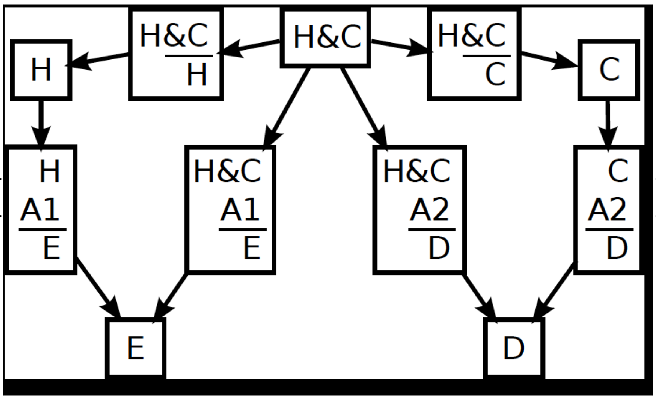

```
(*Set up the boolean constraints*)
ClearAll[tau, a1, a2, a3, a4, a5, a6];
a1 = (H1 && B) ⇒ ! H2;
a2 = (H1 && A1) ⇒ E;
a3 = (H2 && A2) ⇒ E;
a4 = (A3 && B) ⇒ H1;
a5 = (A4 && B) ⇒ H1;
a6 = (A5 && B) ⇒ H1;
tau = a1 && a2 && a3 && a4 && a5 && a6 ;
(*DOJs as in figure 4*)
TableForm[ ( DOJ[H1, E && B, tau]   DOJ[H1, E && B, tau] / DOJ[H1, B, tau]   DOJ[H1, B, tau]
             DOJ[H2, E && B, tau]   DOJ[H2, E && B, tau] / DOJ[H2, B, tau]   DOJ[H2, B, tau] ),
 TableHeadings → {None, {"P(Hi|E)=", "P(Hi|E)/P(Hi)*", "P(Hi)"}}]
```

P(Hi|E)=	P(Hi|E)/P(Hi)*	P(Hi)
$\frac{4}{5}$	$\frac{31}{30}$	$\frac{24}{31}$
$\frac{1}{10}$	$\frac{31}{30}$	$\frac{3}{31}$

```
TableForm[N[%], TableHeadings → {None, {"P(Hi|E)=", "P(Hi|E)/P(Hi)*", "P(Hi)"}}]
```

P(Hi|E)=	P(Hi|E)/P(Hi)*	P(Hi)
0.8	1.03333	0.774194
0.1	1.03333	0.0967742

```
(**********Calculation of DOJs for tacking paradox 09/2011**********)
```

{

};

```
(*Set up the boolean constraints*)
(*H1 = H; H2 = H&C*)
ClearAll[tau, a1, a2, a3, a4, a5, a6];
a1 = (H1 && A1) ⇒ E;
a2 = (H2 && A1) ⇒ E;
a3 = (H2 && A2) ⇒ D;
a4 = (H3 && A2) ⇒ D;
a5 = H2 ⇒ H3;
a6 = H2 ⇒ H1;
tau = a1 && a2 && a3 && a4 && a5 && a6;
```

```
(*DOJs as in figure 5*)
```

TableForm[

DOJ[H1, E, tau]	DOJ[H1, E, tau] / DOJ[H1, tau]	DOJ[H1, tau]
DOJ[H2, E, tau]	DOJ[H2, E, tau] / DOJ[H2, tau]	DOJ[H2, tau]
DOJ[H3, E, tau]	DOJ[H3, E, tau] / DOJ[H3, tau]	DOJ[H3, tau]

,

TableHeadings → {None, {"P(Hi|E)=", "P(Hi|E)/P(Hi)*", "P(Hi)"}}]

P(Hi|E)=	P(Hi|E)/P(Hi)*	P(Hi)
$\frac{10}{17}$	$\frac{58}{51}$	$\frac{15}{29}$
$\frac{3}{17}$	$\frac{58}{51}$	$\frac{9}{58}$
$\frac{9}{17}$	$\frac{87}{85}$	$\frac{15}{29}$

```
TableForm[N[%], TableHeadings → {None, {"P(Hi|E)=", "P(Hi|E)/P(Hi)*", "P(Hi)"}}]
```

P(Hi|E)=	P(Hi|E)/P(Hi)*	P(Hi)
0.588235	1.13725	0.517241
0.176471	1.13725	0.155172
0.529412	1.02353	0.517241

```
(**********Calculation of DOJs for footnote 12**********)
```

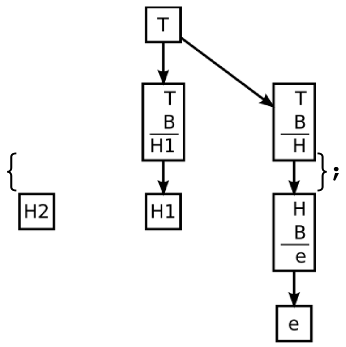
```
(*Set up the boolean constraints*)
(*H1 = H; H2 = H&C*)
ClearAll[tau, a1, a2, a3, a4, a5, a6];
a1 = (H1 && B) ⇒ E;
a2 = H2 ⇒ H1;
a3 = H2 ⇒ C;
tau = a1 && a2 && a3;
Print["P(H|E)/P(H)=", N[DOJ[H1, E && B, tau] / DOJ[H1, B, tau]]];
Print["P(H&C|E)/P(H&C)=", N[DOJ[H2, E && B, tau] / DOJ[H2, B, tau]]];
Print["P(C|E)/P(C)=", N[DOJ[C, E && B, tau] / DOJ[C, B, tau]]];
```

P(H|E)/P(H)=1.4

P(H&C|E)/P(H&C)=1.4

P(C|E)/P(C)=1.05

```
(**********Calculation of DOJs for paradox from concept. diff.**********)
```




```
(*Set up the boolean constraints*)
(*H1 = H; H2 = H&C*)
ClearAll[tau, a1, a2, a3, a4, a5, a6];
a1 = (T && B) ⇒ H1;
a2 = (T && B) ⇒ H;
a3 = (H && B) ⇒ E;
tau = a1 && a2 && a3;
```

```
(*DOJs as in figure 6*)
```

TableForm[

DOJ[H1, E && B, tau]	DOJ[H1, E && B, tau] / DOJ[H1, B, tau]	DOJ[H1, B, tau]
1 / 2	1	1 / 2

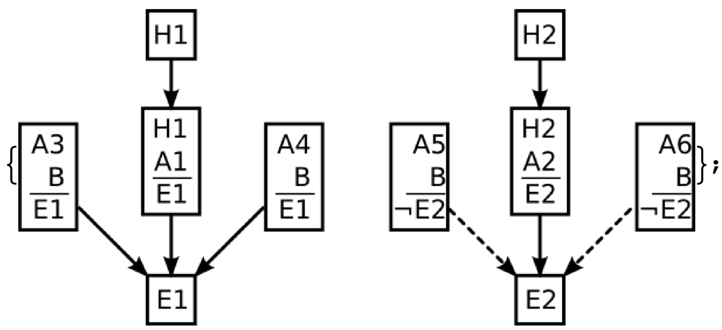
,
TableHeadings → {None, {"P(Hi|e)=", "P(Hi|e)/P(Hi)*", "P(Hi)"}}]

P(Hi|e)=	P(Hi|e)/P(Hi)*	P(Hi)
$\frac{3}{5}$	$\frac{21}{20}$	$\frac{4}{7}$
$\frac{1}{2}$	1	$\frac{1}{2}$

```
TableForm[N[%], TableHeadings → {None, {"P(Hi|e)=", "P(Hi|e)/P(Hi)*", "P(Hi)"}}]
```

P(Hi|e)=	P(Hi|e)/P(Hi)*	P(Hi)
0.6	1.05	0.571429
0.5	1.	0.5

```
(**********Calculation of DOJs for the problem of novelty**********)
```




```
(*Set up the boolean constraints*)
ClearAll[tau, a1, a2, a3, a4, a5, a6];
a1 = (H1 && A1) ⇒ E1;
a2 = (A3 && B) ⇒ E1;
a3 = (A4 && B) ⇒ E1;
a4 = (H2 && A2) ⇒ E2;
a5 = (A5 && B) ⇒ ! E2;
a6 = (A6 && B) ⇒ ! E2;
tau = a1 && a2 && a3 && a4 && a5 && a6 ;
```

```
(*DOJs as in figure 7*)
TableForm[{{DOJ[H1, E1 && B, tau], DOJ[H1, E1 && B, tau] / DOJ[H1, B, tau], DOJ[H1, B, tau]},
           {DOJ[H2, E2 && B, tau], DOJ[H2, E2 && B, tau] / DOJ[H2, B, tau], DOJ[H2, B, tau]}},
 TableHeadings → {None, {"P(Hi|Ei)=", "P(Hi|Ei)/P(Hi)*", "P(Hi)"}}]
```

P(Hi|Ei)=	P(Hi|Ei)/P(Hi)*	P(Hi)
$\frac{1}{2}$	$\frac{19}{18}$	$\frac{9}{19}$
$\frac{1}{2}$	$\frac{4}{3}$	$\frac{3}{8}$

```
TableForm[N[%], TableHeadings → {None, {"P(Hi|Ei)=", "P(Hi|Ei)/P(Hi)*", "P(Hi)"}}]
```

P(Hi|Ei)=	P(Hi|Ei)/P(Hi)*	P(Hi)
0.5	1.05556	0.473684
0.5	1.33333	0.375

■ Backup