On degrees of justification

Gregor Betz
Universitaet Stuttgart
gregor.betz@philo.uni-stuttgart.de

September 19, 2009

Abstract

This paper gives an explication of our intuitive notion of strength of justification in a controversial debate. It defines a thesis' degree of justification within the bipolar argumentation framework of the theory of dialectical structures as the ratio of coherently adoptable positions according to which that thesis is true over all coherently adoptable positions. Broadening this definition, the notion of conditional degree of justification, i.e. degree of partial entailment, is introduced. Thus defined degrees of justification correspond to our pre-theoretic intuitions in the sense that supporting and defending a thesis t increases, whereas attacking it decreases, t's degree of justification. Moreover, it is shown that (conditional) degrees of justification are (conditional) probabilities. Eventually, the paper explains that it is rational to believe theses with a high degree of justification insofar as this strengthens the robustness of one's position.

1 Introduction

Imagine two opponents discussing whether some claim is true or not. They put forward different arguments: fortifying their own lines of reasoning, attacking their opponent's arguments, introducing counter-arguments, etc. Once the exchange of reasons has come to an end, one might ask which of two controversial claims is more strongly justified. This seems to be a sensible question. It isn't, moreover, patently absurd to picture that some thesis is better justified than a second one, which in turn is better justified than a third one, suggesting that our intuitive notion of strength of justification is a comparative concept.

Three examples, taken from three different debates, namely about the reality global warming, scientific relativism and British foreign policy, shall carve out the vague idea that justification comes in degrees. The following list contains three central claims, A, B and C, taken from each of these debates, which can be intuitively ordered according to their respective strength of justification: A is better justified than B, and B is better justified than C.

The reality of global warming[1]:

(A) Ongoing climate change is partially man-made.

(B) As a consequence of global warming, the Greenland ice sheet will melt in the course of the next centuries.

(C) There is no significant human influence on the global climate system.

Scientific relativism:

(A) Scientific theories work, i.e. enable us to control and manipulate nature.

(B) Scientific theories become ever more empirically adequate, accumulating the success of previous theories.

(C) Scientific theories are approximately true, and converge against the true description of nature.

British foreign policy:

(A) Britain should stay in the European Union.

(B) Britain should adopt the single European currency.

(C) Britain should give up its permanent seat in the security council.

Yet why would one judge that climate change being partially man made is better justified than the prediction that Greenland will ultimately melt? Why is the advice to stay in the EU better justified than the one to join the Euro? These questions do not ask for specific arguments that warrant the respective statements; they inquire generally as to the features of the corresponding debates which are responsible for our judgement regarding the strength of justification.

One way to approach the idea which underlies this paper's attempt to give a precise explication of the notion of degree of justification consists in turning these questions upside down. Instead of "Is t well justified?" consider the question "Given the debate with its inferential relations, what would have to be the case such that t were false?"! The fewer 'possibilities' we can

[1] Compare for example Rahmstorf and Schellnhuber (2006).

envisage such that t becomes false, the better justified is the thesis t. Yet what are these 'possibilities'? We may think of them, I suggest, as *positions* one can coherently adopt in the debate. Thence, if t is true from almost every coherent position one can embrace in a debate, t is well justified, whereas if t is only true from a very specific perspective, then t is not so well justified. Furthermore, t_1 is better justified than t_2 if t_1 is false in fewer coherent positions than t_2. Returning to the examples, "Scientific theories work" is better justified than "Scientific theories are approximately true" because the former statement will be true in most coherent positions in the philosophy of science controversy, whereas the latter is true only from a few specific (e.g. realist) perspectives.

A formal explication of our pre-theoretic notion of strength of justification proves its worth by capturing and systematising our intuitions and our argumentative practice. This is, however, not to say that these intuitions are infallible. A theoretic concept of degree of justification which successfully represents many of our intuitions, thus explaining current discursive practice, is not necessarily falsified when in contradiction with other intuitions. On the contrary, these might be rectified in the light of the theoretic findings.

Having outlined the basic idea underlying this paper, I shall, next, briefly comment on how the concrete implementation of that very idea within the theory of dialectical structures relates to existing approaches in the literature. The comparative notion of justification can be caught in an argumentation theory which accommodates multiple degrees of justification, and degrees of belief, e.g. Pollock's theory of defeasible reasoning with variable degrees of belief (Pollock, 2001).[2] One possible starting point for developing such a many-value evaluation procedure are subjective probabilities. "Probabilistic argumentation frameworks" have been proposed, for instance, by Laskey and Lehner (1989), Haenni et al. (2000), Haenni and Lehmann (2003), or Kohlas (2003). Betz (2008) is an attempt to apply subjective probabilities to evaluate dialectical structures. All these approaches commonly assume that belief comes in degree, and that is what enables them to represent degrees of justification. This paper, in contrast, tries to accommodate a quantitative notion of degree of justification within a bivalence framework. In order to do so, it builds on the (basic concepts of the) theory of dialectical structures as developed by Betz (2009).[3]

[2]The approach presented in this paper, however, rejects a major principle of Pollock's theory, namely the weakest link principle.

[3]A dialectical structures is a special type of bipolar argumentation framework as developed by Cayrol and Lagasquie-Schiex (2005). Cayrol and Lagasquie-Schiex extend the abstract approach of Dung (1995) by adding support-relations to Dung's framework which originally considered attack-relations between arguments only. A specific interpretation

The subject matter of a theory of dialectical structures are debates. Debates consist of arguments and theses which can be *reconstructed* as premiss-conclusion structures.[4] Moreover, it is assumed that arguments are reconstructed as deductively valid. The set of reconstructed arguments and theses is labelled T.[5] An argument $a_1 \in T$ supports (attacks) an argument $a_2 \in T$ if and only if the conclusion of a_1, briefly: $C(a_1)$, is equivalent (contradictory) to a premiss of a_2. The support- and attack-relation, U and A respectively, that are thus defined on T make up the dialectical structure of the debate $\tau = \langle T, A, U \rangle$.[6]

A two-coloured, directed graph is an appropriate mathematical model of a dialectical structure, and a helpful visualisation, too. Accordingly, arguments are the graph's nodes; a green (red) arrow between two arguments signifies that one supports (attacks) the other. While using the colour-terminology for convenience, I will visualise these two relations as curly and straight arrows, i.e. $A(a,b) \iff a \rightsquigarrow b$ and $U(a,b) \iff a \rightarrow b$.

The general idea sketched in this introduction will be unfolded in the next two sections. Section 2 introduces the key concept of a coherent position a proponent can rationally adopt in a dialectical structure. Building on this notion, section 3 defines the degree of justification, as well as related concepts such as the conditional degree of justification and the degree of partial entailment. Section 4 proves basic properties of degrees of justification, demonstrating that the concept, as it has been defined, dovetails well with our pre-theoretic intuitions. The mathematical structure of degrees of justification is further elucidated by showing, in section 5, that degrees of justification are probabilities, and that degrees of partial entailment are conditional

of Dung's abstract framework that analyses arguments as premiss-conclusion structures is carried out in Bondarenko et al. (1997).

[4]For the reconstruction of natural language arguments, in particular philosophical ones, compare Tetens (2004).

[5]Note that, unlike in approaches by Lin and Shoham (1989) or the interpretation by Prakken and Vreeswijk (2001, p 256) of Dung (1995), T is not supposed to contain arguments which can be *constructed* given the propositions put forward in a debate (or, more generally, some INPUT) but only those arguments that have been explicitly stated (though not necessarily fully). This emphasis on real reasoning as opposed to ideal reasoning seems to be more in line with the approaches of Pollock (1987, 1995), Vreeswijk (1997), or Verheij (1996).

[6]Accordingly, if two arguments conflict, i.e. possess contrary conclusions, they do not necessarily attack each other as defined here. The "assumption attack" as well as "undercutting" an argument (cf. Pollock, 1970; Prakken and Vreeswijk, 2001) can both be represented in this framework as an attack on an argument's premiss. Moreover, indirect attacks, i.e. attacks on an argument's subconclusion $c-$, can be made explicit by reconstructing the attacked argument as two arguments, a_1 and a_2, such that $c-$ is the conclusion of a_1 and a premiss of a_2, a_1 supporting a_2 and a_2 being the argument attacked.

probabilities. The degree of justification is related to other concepts of the theory of dialectical structures, in particular the notion of τ-deducibility, in section 6. This results in distinguishing two notions of dialectical entailment. Eventually, section 7 discusses whether there is a gap between rational belief on the one hand, and beliefs with high degree of justification on the other hand. It suggests that it is rational to believe a thesis with high degree of justification insofar as one aims at maximising the robustness of one's position.

2 Coherent positions in dialectical structures

In the light of what has been said so far, defining the degree of justification requires first and foremost to explicate the notion of a position which can (i) be coherently adopted in a debate and according to which (ii) some thesis is true. Betz (2009) proposed to identify coherent positions in dialectical structures with closed and complete subdebates in equilibrium. This paper, however, develops an alternative notion of a coherent position which is based on evaluating sentences rather than entire arguments. Appendix A will relate these alternative concepts to each other, demonstrating that for every sentence-based coherent position there is a corresponding closed, complete subdebate in equilibrium, and vice versa.

This said, we can define a position on some dialectical structure τ as a truth value assignment to the sentences which figure in arguments and theses in τ.

Definition 1 (Complete position) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure and S be the sentences occurring in τ. A truth value assignment $\mathcal{Q} : S \to \{\mathbf{t}, \mathbf{f}\}$ is called a complete position on τ.*

A partial position does not necessarily assign truth values to *all* sentences in τ.

Definition 2 (Partial position) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure and S be the sentences occurring in τ. A truth value assignment $\mathcal{P} : S' \to \{\mathbf{t}, \mathbf{f}\}$ with $S' \subseteq S$ is called a partial position on τ.*

Let $\mathcal{N}$ denote the empty position ($S' = \emptyset$) which assigns no truth value to any sentence.

Note that a complete position is, according to this definition, a partial position, too. A partial position on some dialectical structure can extend another partial position.

Definition 3 (Extension of a partial position) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure, S be the sentences occurring in τ, and $S_1 \subseteq S_2 \subseteq S$. A partial position $\mathcal{P}_2 : S_2 \rightarrow \{\mathbf{t}, \mathbf{f}\}$ extends a partial position $\mathcal{P}_1 : S_1 \rightarrow \{\mathbf{t}, \mathbf{f}\}$ if and only if $\mathcal{P}_2(p) = \mathcal{P}_1(p)$ for all $p \in S_1$. $\mathcal{P}_2$ is called an extension of $\mathcal{P}_1$.*

Moreover, partial positions, provided they don't assign different truth values to the very same sentence, can be combined so as to give rise to a new (partial) position.

Definition 4 (Union of positions) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure, S be the sentences occurring in τ, and $\mathcal{P}_1 : S_1 \rightarrow \{\mathbf{t}, \mathbf{f}\}$, $\mathcal{P}_2 : S_2 \rightarrow \{\mathbf{t}, \mathbf{f}\}$ represent two partial positions on τ which agree on $S_1 \cap S_2$. The* union *of these two positions, $(\mathcal{P}_1 \& \mathcal{P}_2) : S_1 \cup S_2 \rightarrow \{t, f\}$, assigns every sentence a truth value as follows,*

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

A dialectical structure encodes logical and inferential relations between the sentences its arguments are composed of: Some of these sentence are contradictory or equivalent—which gives rise to dialectical relations (i.e. support and attack) between the arguments—, moreover, some sentences imply others, namely insofar as they occur as premisses and conclusion in an individual argument. These relations impose certain constraints on what positions can be reasonably adopted given a dialectical structure τ. E.g., it would clearly be irrational to consider all the premisses of an argument true, and its conclusion false.[7] The following definition of a dialectically coherent position summarises the conditions a position must satisfy in order to be reasonably adoptable given the logical constraints encoded in τ.

Definition 5 (Coherent complete position) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure and S be the sentences occurring in τ. A complete position $\mathcal{Q}$ on τ is dialectically coherent if and only if*

1. *$\mathcal{Q}$ assigns identical truth values to equivalent sentences in S,*
2. *$\mathcal{Q}$ assigns complementary truth values to contradictory sentences in S, and*
3. *every argument $(p_1, \ldots, p_n; c) \in T$ with true premisses, $\mathcal{Q}(p_1) = \ldots = \mathcal{Q}(p_n) = \mathbf{t}$, has a true conclusion, $\mathcal{Q}(c) = \mathbf{t}$.*

[7] Recall that we assume arguments to be reconstructed as deductively valid.

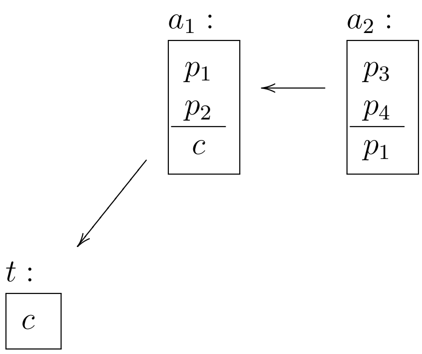
Besides the logical constraints encoded in a dialectical structure, positions might have to satisfy additional conditions imposed by a body of background knowledge. Representing background knowledge as a partial position on τ, we can define

Definition 6 (Conditionally coherent complete position) *Let* $\tau = \langle T, A, U \rangle$ *be a dialectical structure and* $\mathcal{B}$ *be a partial position on* τ*. A complete position* $\mathcal{Q}$ *on* τ *is dialectically coherent conditional to* $\mathcal{B}$ *if and only if*

1. $\mathcal{Q}$ *is dialectically coherent on* τ*, and*
2. $\mathcal{Q}$ *extends* $\mathcal{B}$*.*

Can these definitions be generalised so that they apply to partial positions, too? As a first attempt, one could simply require that a partial position had to satisfy the corresponding criteria with respect to those sentences to which it assigns truth values. That, however, would be inappropriate, as the following example shows.

Example 1



Consider the partial position $\mathcal{P}$:

	c	p_2	p_3	p_4
$\mathcal{P}$	f	t	t	t

With a view to $\{c, p_2, p_3, p_4\}$, $\mathcal{P}$ satisfies the three conditions of dialectical coherence. Still, something is wrong with $\mathcal{P}$: Although it does not assign a truth value to p_1, a proponent adopting $\mathcal{P}$ is committed to p_1, because she accepts the premisses of a_2. However, this conflicts with her view that p_2, the second premiss of a_1, is also true whereas the conclusion of a_1, c, is false. Examples of this type motivate the following definition.

Definition 7 (Coherent partial position) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure. A partial position $\mathcal{P}$ on τ is dialectically coherent if and only if there is a dialectically coherent complete position $\mathcal{Q}$ on τ which extends $\mathcal{P}$.*

$\mathcal{P}$ is dialectically coherent conditional to some partial position $\mathcal{B}$ iff there is a complete extension of $\mathcal{P}$ that is dialectically coherent conditional to $\mathcal{B}$.

We may now introduce two further concepts of which we will make use in section 5.

Definition 8 (τ-certain partial position) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure. A partial position $\mathcal{P}$ is called "τ-certain" if and only if it is extended by every coherent and complete position on τ.*

Definition 9 (τ-exclusive of a partial positions) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure. Two partial positions $\mathcal{P}_1$ and $\mathcal{P}_2$ are called "τ-exclusive" if and only if there is no coherent and complete position on τ which extends both $\mathcal{P}_1$ and $\mathcal{P}_2$.*

Let us eventually note some properties of coherent positions before we use that very concept to define degrees of justification in the next section.

1. Whether a partial position $\mathcal{P}$ is coherent or not depends on τ, in particular on the arguments in τ which impose coherence constraints (definition 5) on $\mathcal{P}$.

2. A partial position which is coherent on some τ might become incoherent if τ is extended by introducing additional arguments.

3. Dialectical incoherence, though, is monotonic in the sense that a partial position which is incoherent on τ will be incoherent on every dialectical structure that is obtained by introducing further arguments into τ.

4. Every partial position which is not dialectically coherent on some τ is logically inconsistent.

5. The reverse, however, is not true. Let τ be a dialectical structure. A partial position which is logically inconsistent is not necessarily dialectically incoherent on τ. This is because the logical relations that hold between the debate's sentences and which are responsible for the logical inconsistency are not necessarily made explicit as arguments in the dialectical structure; yet, only those inferential relations between sentences which are represented as arguments in τ are taken into account when evaluating a position for dialectical coherency. Considering

but the inferential relations that have been 'discovered' so far at a certain state of a debate, dialectical evaluation is an evaluation of real as opposed to ideal reasoning.

6. Once a partial position $\mathcal{P}$ which is dialectically coherent on τ is shown to be logically inconsistent, τ can be extended by adding further arguments (and without introducing new sentences) to τ' so that $\mathcal{P}$ becomes dialectically incoherent on τ'.

Observation 3 gives directly rise to the following lemma.

Lemma 1 (Coherent positions in expanded dialectical structures) *Let $\tau = \langle T, A, U\rangle$ and $\tau' = \langle T', A', U'\rangle$ be two dialectical structures with $T \subset T'$ and $A = A'|_T$, $U = U'|_T$. Every dialectically coherent, complete position on τ' extends a dialectically coherent, complete position on τ.*

3 The degree of justification

The preliminary analysis in the introduction suggested that the degree of justification of some thesis depends on whether it is true according to most positions one can reasonably adopt, or whether it is true only from very few, specific points of view. With the notion of a coherent position in a dialectical structure at hand, we can capture that general idea in the following definition.

Definition 10 (Degree of justification) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure and $\mathcal{P}$ be a partial position on τ. The degree of justification of $\mathcal{P}$ in τ is*

$$J_\tau(\mathcal{P}) = \frac{\sigma_\mathcal{P}}{\sigma},$$

where σ is the number of all dialectically coherent, complete positions on τ and $\sigma_\mathcal{P}$ is the number of those dialectically coherent, complete positions which extend $\mathcal{P}$.

The degree of justification of a single sentence p in τ, $J_\tau(p)$, equals the degree of justification of the corresponding partial position which merely assigns p the truth value $\mathbf{t}$.

Let us, as an illustration, calculate the degree of justification of the sentence c in example 1. First of all, this requires to identify all dialectically coherent, complete positions on τ. There exist altogether 2^5 different complete positions on τ. Not all of them are dialectically coherent. If c is true, p_1 and p_2 can take any truth value. However, if c is false, at least one of a_1's

	c	p_1	p_2	p_3	p_4
$\mathcal{P}_1$	t	t	t	t	t
$\mathcal{P}_2$	t	t	t	t	f
$\mathcal{P}_3$	t	t	t	f	t
$\mathcal{P}_4$	t	t	t	f	f
$\mathcal{P}_5$	t	t	f	t	t
$\mathcal{P}_6$	t	t	f	t	f
$\mathcal{P}_7$	t	t	f	f	t
$\mathcal{P}_8$	t	t	f	f	f
$\mathcal{P}_9$	t	f	t	t	f
$\mathcal{P}_{10}$	t	f	t	f	t
$\mathcal{P}_{11}$	t	f	t	f	f
$\mathcal{P}_{12}$	t	f	f	t	f
$\mathcal{P}_{13}$	t	f	f	f	t
$\mathcal{P}_{14}$	t	f	f	f	f
$\mathcal{P}_{15}$	f	t	f	t	t
$\mathcal{P}_{16}$	f	t	f	t	f
$\mathcal{P}_{17}$	f	t	f	f	t
$\mathcal{P}_{18}$	f	t	f	f	f
$\mathcal{P}_{19}$	f	f	t	t	f
$\mathcal{P}_{20}$	f	f	t	f	t
$\mathcal{P}_{21}$	f	f	t	f	f
$\mathcal{P}_{22}$	f	f	f	t	f
$\mathcal{P}_{23}$	f	f	f	f	t
$\mathcal{P}_{24}$	f	f	f	f	f

Table 1: Dialectically coherent, complete positions in example 1.

premisses must be false, too. A similar reasoning applies to argument a_2. Table 1 lists those positions which satisfy the coherence constraints (definition 5).

14 of these 24 coherent positions assign c the truth value $\mathbf{t}$, so

$$J_\tau(c) = 14/24 \approx 0.58.$$

Conditional degrees of justification which hold relative to a certain background knowledge can be defined as follows, making use of the concept of a conditionally coherent position.

Definition 11 (Conditional degree of justification) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure, and $\mathcal{P}, \mathcal{B}$ be two partial positions on τ. The condi-*

tional degree of justification of $\mathcal{P}$ relative to $\mathcal{B}$ in τ is

$$J_\tau(\mathcal{P}|\mathcal{B}) = \frac{\sigma_{\mathcal{PB}}}{\sigma_{\mathcal{B}}},$$

where $\sigma_{\mathcal{B}}$ is the number of all complete positions on τ which are dialectically coherent conditional to $\mathcal{B}$, and $\sigma_{\mathcal{PB}}$ is the number of those complete positions which are dialectically coherent conditional to $\mathcal{B}$ and which extend $\mathcal{P}$.

For sentences p, q, we define $J_\tau(p|q) := J_\tau(\mathcal{P}|\mathcal{Q})$ where $\mathcal{P}, \mathcal{Q}$ are partial positions which assign but p, respectively q, the truth value $\mathbf{t}$.

Assume that, in example 1, p_3 and p_4 belong to our background knowledge. How does this affect c's (conditional) degree of justification? We define a partial position $\mathcal{B}$ on $\{p_3, p_4\}$ with $\mathcal{B}(p_3) = \mathcal{B}(p_4) = \mathbf{t}$. As can be verified with table 1, there are three positions which are dialectically coherent conditional to $\mathcal{B}$, namely $\mathcal{P}_1, \mathcal{P}_5, \mathcal{P}_{15}$; c is true in two of these. So

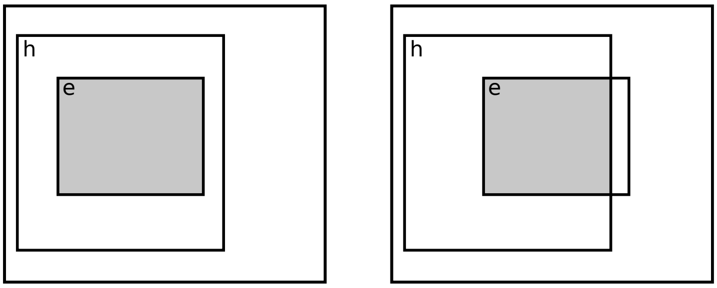
$$J_\tau(c|\mathcal{B}) = 2/3 \approx 0.67.$$

In the *Tractatus*, Wittgenstein made a suggestion how strict entailment can be generalised to the concept of degree of partial entailment (cf. proposition 5.15). The basic idea, which has later been seized by Carnap to develop his inductive logic (Carnap, 1950), is depicted in figure 1 and can be put, in abstract terms, like this. A sentence e implies a sentence h if and only if all 'cases' in which e is true are also cases in which h holds, i.e. the e-cases are included in the h-cases (panel (a) of figure 1). Given, however, that not all, but still most e-cases are h-cases, e does not imply h, but—that is essentially the proposal—e 'almost' implies h. Or: e implies h to the extent that the e-cases are included in the h-cases. And this, in turn, is expressed by the ratio of h&e-cases over all e-cases (panel (b) of figure 1). That ratio is the degree of partial entailment of h by e.

Carnap implemented this idea by interpreting 'cases' as so-called state-descriptions. In the context of the theory of dialectical structures, however, cases can be identified with dialectically coherent positions on some τ. This yields the following definition of the degree of partial entailment in a dialectical structure.

Definition 12 (Degree of partial entailment) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure and $\mathcal{P}_1, \mathcal{P}_2$ be two partial positions on τ. The degree of partial entailment of $\mathcal{P}_1$ by $\mathcal{P}_2$ in τ is*

$$I_\tau(\mathcal{P}_1|\mathcal{P}_2) = \frac{\sigma_{\mathcal{P}_1\mathcal{P}_2}}{\sigma_{\mathcal{P}_2}},$$



(a) (b)

Figure 1: Degree of partial entailment. Adapted from Carnap (1950, p. 297).

where $\sigma_{\mathcal{P}_2}$ is the number of all dialectically coherent, complete positions on τ which extend $\mathcal{P}_2$, and $\sigma_{\mathcal{P}_1\mathcal{P}_2}$ is the number of dialectically coherent, complete positions which extend $\mathcal{P}_1$ and $\mathcal{P}_2$.

Let us calculate the degree of partial entailment of c by p_4 in example 1. Thence, we have two partial positions defined on one sentence each, $\mathcal{C}(c) = \mathbf{t}$, and $\mathcal{P}(p_4) = \mathbf{t}$. There are 10 dialectically coherent, complete positions that extend $\mathcal{P}$ (namely $\mathcal{P}_1, \mathcal{P}_3, \mathcal{P}_5, \mathcal{P}_7, \mathcal{P}_{10}, \mathcal{P}_{13}, \mathcal{P}_{15}, \mathcal{P}_{17}, \mathcal{P}_{20}, \mathcal{P}_{23}$). Six of these extend $\mathcal{C}$, too (namely $\mathcal{P}_1, \mathcal{P}_3, \mathcal{P}_5, \mathcal{P}_7, \mathcal{P}_{10}, \mathcal{P}_{13}$). Thus,

$$I_\tau(\mathcal{C}|\mathcal{P}) = 6/10 = 0.6.$$

Before we investigate general properties of the degree of justification, we note that degrees of justification and degrees of partial entailment closely correspond to each other, as follows directly from the above definitions. Let τ be a dialectical structure and let $\mathcal{P}_1, \mathcal{P}_2$ be two partial positions on τ, then:

1. $J_\tau(\mathcal{P}_1) = I_\tau(\mathcal{P}_1|\mathcal{N})$, and
2. $J_\tau(\mathcal{P}_1|\mathcal{P}_2) = I_\tau(\mathcal{P}_1|\mathcal{P}_2)$.

In other words, degrees of partial entailment are nothing but conditional degrees of justification, and we shall use, in the following, $J_\tau(\ldots|\ldots)$ to refer to these.

4 General properties of degrees of justification

This section investigates some general, 'dynamic' properties of degrees of justification. It addresses the question how degrees of justification change when new arguments are introduced into a debate. As we seem to have some clear intuitions regarding this matter (e.g., introducing new supporting arguments for p does not decrease p's degree of justification), this will allow us to see whether the concept of degree of justification as defined in the previous section fits to these pre-theoretic intuitions.

The following proposition states how a thesis' degree of justification changes when new and independent arguments which directly support or attack that thesis are introduced. An argument is independent given some dialectical structure τ iff none of its premisses is equivalent or contradictory to a sentence that figures in τ.

Proposition 1 (Direct support and attack) *Let $\tau = \langle T, A, U \rangle$, $\tau' = \langle T', A', U' \rangle$ be two dialectical structures where τ' is obtained by introducing a new, independent argument a into τ, $T' = T \cup \{a\}$. Thesis $t \in T$ states sentence p.*

1. *If $a \rightarrow t$, then $J_{\tau'}(p) > J_{\tau}(p)$.*
2. *If $a \rightsquigarrow t$, then $J_{\tau'}(p) < J_{\tau}(p)$.*

Proof: Assume $a \rightarrow t$. Let σ (σ') denote the number of all dialectically coherent, complete positions on τ (τ'), and σ_p (σ'_p) the number of dialectically coherent, complete positions on τ (τ') corresponding to which p is true. Now, consider an arbitrary dialectically coherent, complete position $\mathcal{P}$ on τ corresponding to which p is true. Because argument a is assumed to be independent, and because its conclusion is true in $\mathcal{P}$, any truth value assignment to its premisses will extend $\mathcal{P}$ to a dialectically coherent, complete position on τ'. If a has n premisses, there will be 2^n dialectically coherent, complete position on τ' which extend $\mathcal{P}$. As a next step, consider an arbitrary dialectically coherent, complete position $\mathcal{Q}$ on τ corresponding to which p is false. Those and only those truth value assignments to premisses of a according to which not all premisses are true will extend $\mathcal{Q}$ to a dialectically coherent, complete position on τ'. So, there will be $2^n - 1$ dialectically coherent, complete position on τ' which extend $\mathcal{Q}$. According to lemma 1, every dialectically coherent, complete position on τ' extends a dialectically coherent, complete position on τ. Hence, we can calculate the number of positions on τ' as follows:

$$\begin{aligned} \sigma'_p &= 2^n \cdot \sigma_p \\ (\sigma' - \sigma'_p) &= (2^n - 1) \cdot (\sigma - \sigma_p). \end{aligned}$$

Therefore:

$$
\begin{aligned}
J_{\tau'}(p) &= \frac{\sigma'_p}{\sigma'} \\
&= \frac{\sigma'_p}{(\sigma' - \sigma'_p) + \sigma'_p} \\
&= \frac{2^n \cdot \sigma_p}{(2^n - 1) \cdot (\sigma - \sigma_p) + 2^n \cdot \sigma_p} \\
&= \frac{2^n \cdot \sigma_p}{2^n \cdot \sigma - \underbrace{(\sigma - \sigma_p)}_{>0}} \\
&> \frac{2^n \cdot \sigma_p}{2^n \cdot \sigma} \\
&= J_\tau(p).
\end{aligned}
$$

For symmetrical reasons, $J_{\tau'}(p) < J_\tau(p)$ if $a \rightsquigarrow t$. □

The next proposition describes the effect of indirect attacks and supports, e.g. the effect of supporting an argument which itself supports the central thesis of a debate.

Proposition 2 (Indirect support and attack) *Let $\tau = \langle T, A, U\rangle$, $\tau' = \langle T', A', U'\rangle$ be two dialectical structures. $a \in T$ is an independent argument in τ, and τ' is obtained by introducing a further independent argument b into τ, $T' = T \cup \{b\}$. Thesis $t \in T$ states sentence p.*

1. *If $a \rightarrow t$ and $b \rightarrow a$, then $J_{\tau'}(p) > J_\tau(p)$.*
2. *If $a \rightarrow t$ and $b \rightsquigarrow a$, then $J_{\tau'}(p) < J_\tau(p)$.*
3. *If $a \rightsquigarrow t$ and $b \rightarrow a$, then $J_{\tau'}(p) < J_\tau(p)$.*
4. *If $a \rightsquigarrow t$ and $b \rightsquigarrow a$, then $J_{\tau'}(p) > J_\tau(p)$.*

Proof: We calculate to how many different dialectically coherent, complete positions on τ' the respective positions on τ can be extended. (In this proof, all positions are understood to be dialectically coherent and complete.) We shall assume that a contains n premisses, and b contains m premisses. Consider the first case, i.e. $a \rightarrow t$ and $b \rightarrow a$, and let q be the conclusion of b (and therefore a premiss of a). Since a is independent in τ, the ratio of (i) positions on τ according to which p and q are true over (ii) positions on τ according to which p is true but q is false equals $2^{n-1} : 2^{n-1}$. This is because all 2^n truth value assignments to a's premisses satisfy the coherence constraint if a's conclusion, p, is true, and q is true in exactly half of these. Yet, if a's conclusion is false, there is one truth value assignment to its premisses which will not figure in a position on τ, namely the one which considers all premisses true. So, in that case, there are only $2^n - 1$ corresponding truth value assignments, 2^{n-1} of which regard q as false and $2^{n-1} - 1$ take q as true. The ratio of (i) positions on τ according to which p is false and q is true over (ii) positions on τ according to which p and q are false equals therefore $2^{n-1} - 1 : 2^{n-1}$.

Every position on τ with true q can be extended to 2^m different positions on τ'. In other words, the positions with true q are multiplied by 2^m when introducing b. Still, a position on τ with false q can only be extended to $2^m - 1$ different positions on τ'.

Given (a) the ratio of positions on τ with p and q true over positions with p true and q false, and (b) the respective multipliers, the number of positions on τ with p true is multiplied by the following factor when introducing b:

$$m_1 = \frac{2^{n-1} \cdot 2^m + 2^{n-1} \cdot (2^m - 1)}{2^n}.$$

Likewise, the number of positions on τ with p false is multiplied by the following factor when introducing b:

$$\begin{aligned}
m_2 &= \frac{(2^{n-1} - 1) \cdot 2^m + 2^{n-1} \cdot (2^m - 1)}{2^n - 1} \\
&= \frac{2^{n-1} \cdot 2^m + 2^{n-1} \cdot (2^m - 1) - \frac{2^m + 2^m}{2}}{2^n - 1} \\
&< \frac{2^{n-1} \cdot 2^m + 2^{n-1} \cdot (2^m - 1) - \frac{2^m + (2^m - 1)}{2}}{2^n - 1} \\
&= \frac{2^{n-1} \cdot 2^m + 2^{n-1} \cdot (2^m - 1)}{2^n} \\
&= m_1.
\end{aligned}$$

So the number of positions on τ with p true is multiplied by a greater factor than the number of positions with p false, and that is why p's degree of justification increases when introducing b.

We will briefly consider the second case, that is $a \rightarrow t$ and $b \rightsquigarrow a$. (Cases (3) and (4) hold for analogous reasons.) Let $\neg q$ be the conclusion of b. The ratios of positions on τ as calculated in the first case do apply. Yet, because $b \rightsquigarrow a$, a position on τ with q true can be extended to $2^m - 1$ different positions on τ'. Every position on τ with q false yields 2^m positions on τ' when introducing b. This implies for the corresponding factors m_2 and m_1:

$$m_2 = \frac{(2^{n-1} - 1) \cdot (2^m - 1) + 2^{n-1} \cdot 2^m}{2^n - 1}$$

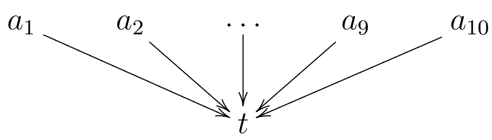
$$
\begin{aligned}
&= \frac{2^{n-1}\cdot(2^m-1)+2^{n-1}\cdot 2^m-\frac{(2^m-1)+(2^m-1)}{2}}{2^n-1}\\
&> \frac{2^{n-1}\cdot(2^m-1)+2^{n-1}\cdot 2^m-\frac{(2^m-1)+2^m}{2}}{2^n-1}\\
&= \frac{2^{n-1}\cdot(2^m-1)+2^{n-1}\cdot(2^m)}{2^n}\\
&= m_1.
\end{aligned}
$$

Thus, a position in τ with p false can be extended to, on average, more positions in τ' than a position in τ with p true. As a consequence, p's degree of justification decreases when introducing b. □

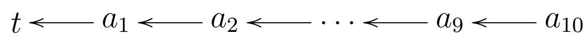
These tedious proofs have provided results that seem to correspond to our intuitions and argumentative practice: Introducing an argument that directly supports t, backing an argument which supports t, or introducing a counter-argument against an attack on t represent argumentative moves commonly carried out in order to underpin or defend t. All these strategies increase t's degree of justification. Likewise, one puts forward a direct challenge against t, backs an argument which attacks t, or introduces a counter-argument against a support for t with a view to undercutting t. Those moves actually diminish t's degree of justification.

Propositions 1 and 2 describe, qualitatively, into which direction a thesis' degree of justification changes due to a certain modification of the dialectical structure. These findings, however, didn't tell whether it is more effective, in terms of the quantitative impact on the degree of justification, to (i) support a thesis directly or to (ii) support a supporting argument, for example. In order to be able to judge the effectiveness of different argumentative moves, I carried out a small simulation study whose results will be reported in the remainder of this section. Besides investigating the quantitative effectiveness of different argumentative moves, that study allows us to analyse degrees of justification in more complex dialectical structures.

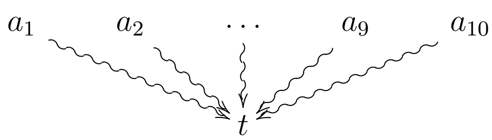
The study aimed at comparing four different argumentative strategies: (a) direct support, (b) indirect support, (c) direct attack, (d) indirect attack. I have simulated four corresponding, stylised debates which consist in the successive construction of the dialectical patterns depicted in figure 2. The ten arguments of each debate, $a_1, \ldots, a_{10}$, are introduced consecutively and at each step, the degree of justification of the thesis t is recalculated so as to measure the marginal effect of the last argument put forward. The arguments contain three premisses each while no sentence figures as premiss in two different arguments. Thence, the final states contain 31 sentences (ignoring negations), and $2^{31} \approx 10^9$ positions have to be checked for coherency, making computer help indispensable.



(a) green fan



(b) green row



(c) red fan

(d) red row

Figure 2: Final states of four stylised debates.

The left-hand panel of figure 3 depicts how t's degree of justification changes during the successive construction of the four dialectical patterns. In both fan-argumentations, the degree of justification increases (green), respectively decreases (red), continuously and monotonically in course of the argumentation process, as expected. Red and green fan-argumentation, moreover, behave symmetrically: After the final argument, a_{10}, has been introduced, we have $J_{\text{green fan}}(t) = 0.79$ and $J_{\text{red fan}}(t) = 0.21$. If the fan-argumentations were continued, degrees of justification would converge against 1 and 0 in the long run. As a consequence, the marginal effectiveness of direct supporting and attacking arguments decreases. This effect becomes already apparent after 10 steps: While a_1 changed the degree of justification by 3.3 points, a_{10} merely caused a modification of 2.3 points. The row-argumentations, however, give rise to a more complicated picture. In the green row-argumentation, t's degree of justification changes hardly any more after a few steps; from a_6 on, the first 10 decimal places remain constant. So, marginal effectiveness decreases much more drastically than in the green fan. Still, while every additional argument increases the degree of justification in the green row, the degree of justification follows a zigzag line for the red row, as a closer look reveals (right-hand panel of figure 3). The first attack, a_1, reduces t's degree of justification, the counter-attack, a_2, increases it, the counter-counter-attack decreases it, etc. etc. But here, as in the green row, the marginal effectiveness of additional arguments decreases dramatically, too. In both row-argumentations, degrees of justification have changed, after 10 steps, from an initial value of 0.5 to 0.54, respectively 0.47, only. In sum, this simulation suggests that the effectiveness of additional arguments in terms of impact on a thesis' degree of justification decreases significantly with its distance to t.

5 Degrees of justification are probabilities

This caption is in need of a qualification. It is misleading to say that conditional degrees of justification are probabilities, full stop. To see this, consider the following dialectical structure.

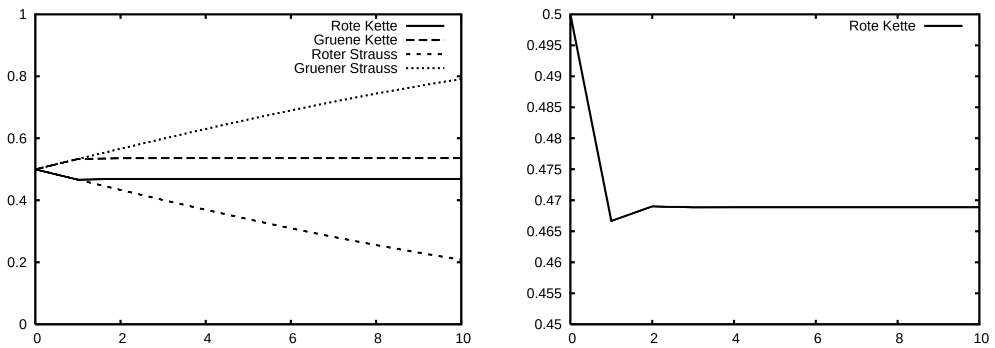
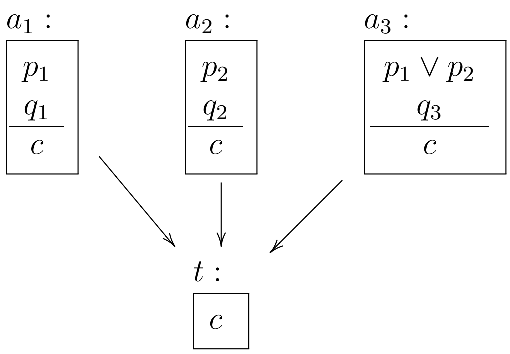


Figure 3: Degrees of justification of t (y-axis) in stylised debates as function of arguments introduced (x-axis). (a) green fan: Gruener Strauss; (b) green row: Gruene Kette; (c) red fan: Roter Strauss; (d) red row: Rote Kette.

Example 2



Assume that sentences p_1 and p_2 are logically incompatible. Consequently, according to probability calculus, $P(p_1) + P(p_2) = P(p_1 \vee p_2)$. Yet, as is obvious for reasons of symmetry, $J_\tau(p_1) = J_\tau(p_2) = J_\tau(p_1 \vee p_2)$. Therefore, it appears, degrees of justification aren't probabilities. The underlying problem which gives rise to this and similar counter-examples is that the logical relations between p_1, p_2, and $p_1 \vee p_2$ are not represented in the dialectical structure, and therefore not taken into account when calculating degrees of justification.

We will address and account for this problem as follows. First of all, we show that degrees of justification of *partial positions* are probabilities. On the basis of this result, we demonstrate that degrees of justification of single sentences can be identified with degrees of justification of partial positions under specific conditions, namely insofar as the inferential relations referred to in the probability statements are fully represented in the dialectical structure.

The concept of degree of justification hence yields a further interpretation of the formal, mathematical probability theory—additional to the logical, subjective, frequency, and propensity interpretation of probability.

In addition to the connective introduced in definition 4, which immediately allows us to calculate the degree of justification of the conjunctive combination of two positions, we have to define the degree of justification of the disjunctive combination of two positions. Although there is, generally, no single position which corresponds to the disjunction of two positions, the definition of the respective degree of justification is straightforward.

Definition 13 (Degree of justification of disjunction of positions) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure, and let $\mathcal{P}_1$, $\mathcal{P}_2$ represent two partial positions on τ. The degree of justification of the* disjunction *of these two positions is defined as,*

$$J_\tau(\mathcal{P}_1 \vee \mathcal{P}_2) = \frac{\sigma_{\mathcal{P}_1 \vee \mathcal{P}_2}}{\sigma},$$

where $\sigma_{\mathcal{P}_1 \vee \mathcal{P}_2}$ refers to the number of coherent and complete positions on τ which extend $\mathcal{P}_1$ or $\mathcal{P}_2$ (or both).

Conditional degrees of justification of such disjunctions are defined correspondingly.

With respect to this definition, we can now show,

Proposition 3 (Degrees of justification are probabilities) *Degrees of justification satisfy the Kolmogorov axioms.*

Proof: Let $\tau = \langle T, A, U \rangle$ be a dialectical structure. We have to show,

(i) $0 \leq J_\tau(\mathcal{P}) \leq 1$ for every position $\mathcal{P}$,

(ii) $J_\tau(\Omega) = 1$ for every τ-certain position Ω,

(iii) $J_\tau(\mathcal{P}_1 \vee \mathcal{P}_2) = J_\tau(\mathcal{P}_1) + J_\tau(\mathcal{P}_2)$ for two τ-exclusive positions $\mathcal{P}_1$ and $\mathcal{P}_2$.

Ad (i): Clear, because every degree of justification is by definition greater than or equal to 0 and less than or equal to 1. Ad (ii): Let Ω be a τ-certain position. Therefore, as $\sigma_\Omega = \sigma$,

$$J_\tau(\omega) = \frac{\sigma_\omega}{\sigma} = 1.$$

Ad (iii): The number of coherent and complete positions which extend $\mathcal{P}_1$ or $\mathcal{P}_2$ equals the number of coherent and complete positions which extend

$\mathcal{P}_1$ plus the number of coherent and complete positions which extend $\mathcal{P}_2$ minus—accounting for double-count—the number of coherent and complete positions which extend both $\mathcal{P}_1$ and $\mathcal{P}_2$, i.e. which extend $\mathcal{P}_1\&\mathcal{P}_2$. Thus, since $\mathcal{P}_1$ and $\mathcal{P}_2$ are τ-exclusive,

$$\begin{aligned} J_\tau(\mathcal{P}_1 \vee \mathcal{P}_2) &= \frac{\sigma_{(\mathcal{P}_1 \vee \mathcal{P}_2)}}{\sigma} = \frac{\sigma_{\mathcal{P}_1} + \sigma_{\mathcal{P}_2} - \overbrace{\sigma_{\mathcal{P}_1\mathcal{P}_2}}^{=0}}{\sigma} \\ &= \frac{\sigma_{\mathcal{P}_1}}{\sigma} + \frac{\sigma_{\mathcal{P}_2}}{\sigma} \\ &= J_\tau(\mathcal{P}_1) + J_\tau(\mathcal{P}_2). \end{aligned}$$

□

Moreover, degrees of partial entailment are conditional probabilities, as the following proposition states.

Proposition 4 (Conditional probabilities) *Degrees of partial entailment are conditional probabilities.*

Proof: Let $\tau = \langle T, A, U\rangle$ be a dialectical structure. We have to show that $J_\tau(\mathcal{P}|\mathcal{Q}) = J_\tau(\mathcal{P}\&\mathcal{Q})/J_\tau(\mathcal{Q})$ for arbitrary $\mathcal{P}$, $\mathcal{Q}$. Yet this is simple,

$$\begin{aligned} J_\tau(\mathcal{P}|\mathcal{Q}) &= \frac{\sigma_{\mathcal{P}\mathcal{Q}}}{\sigma_{\mathcal{Q}}} \\ &= \frac{\sigma_{(\mathcal{P}\&\mathcal{Q})}}{\sigma_{\mathcal{Q}}} \\ &= \frac{\sigma_{(\mathcal{P}\&\mathcal{Q})}/\sigma}{\sigma_{\mathcal{Q}}/\sigma} \\ &= \frac{J_\tau(\mathcal{P}\&\mathcal{Q})}{J_\tau(\mathcal{Q})}. \end{aligned}$$

□

Let us reconsider the problem initially raised in this section. When do degrees of justification of single sentences satisfy the Kolmogorov axioms? This would clearly be the case if degrees of justification of partial positions and single sentences suitably agreed—specifically, if $J_\tau(p \wedge q) = J_\tau(\mathcal{P}\&\mathcal{Q})$ and $J_\tau(p \vee q) = J_\tau(\mathcal{P} \vee \mathcal{Q})$, where $\mathcal{P}$ and $\mathcal{Q}$ denote partial positions which assign but p and q the value $\mathbf{t}$. This condition is obviously violated in the example considered at the beginning of this section. Thus, for instance, not every position according to which p_1 is true or p_2 is true considers $p_1 \vee p_2$ true, as well. Still, the following theorem states that every dialectical structure may be easily extended by simple arguments so that degrees of justification of single sentences agree with the degrees of justification of the corresponding partial positions.

Proposition 5 (Single sentence degrees of justification) *Let $\tau = \langle T, A, U\rangle$ be a dialectical structure and $p, q, p \vee q, p \vee q \in S$ be four sentences in τ. $\mathcal{P}$ and $\mathcal{Q}$ denote partial positions that assign exactly one truth value and according to which p, respectively q, is true. Then, it is always possible to introduce additional, valid arguments into τ without introducing new sentences so that $J_\tau(p \wedge q) = J_\tau(\mathcal{P}\&\mathcal{Q})$ and $J_\tau(p \vee q) = J_\tau(\mathcal{P} \vee \mathcal{Q})$.*

Proof: The stated equations hold if

- $\mathcal{E}$ is a coherent extension of $\mathcal{P}\&\mathcal{Q}$ iff $\mathcal{E}(p \wedge q) = \mathbf{t}$, and
- $\mathcal{E}$ is a coherent extension of $\mathcal{P}$ or of $\mathcal{Q}$ iff $\mathcal{E}(p \vee q) = \mathbf{t}$.

In order to guarantee these two conditions, it suffices to introduce the valid arguments $(p,\ q;\ p \wedge q)$, $(p \wedge q;\ p)$, $(p \wedge q;\ q)$, $(p;\ p \vee q)$, $(q;\ p \vee q)$, $(p \vee q,\ \neg p;\ q)$, $(p \vee q,\ \neg q;\ p)$. The coherence constraints imposed by these additional arguments make sure that (1) $p \wedge q$ is true in $\mathcal{E}$ iff p is true and q is true, and (2) $p \vee q$ is true in $\mathcal{E}$ iff p is true or q is true. □

So, what do these three propositions signify? First of all, we have not proven that degrees of justification represent the probability of a certain thesis being true. All we have shown is that degrees of justification realise, or are a model of, a certain mathematical structure. What makes this result significant is that the respective mathematical structure—probability theory—is so well known and that the theorems of probability theory as well as the probability calculus can all at once be applied to degrees of justification.

6 Degrees of justification and dialectical entailment

This section investigates how the concepts of degree of justification and degree of partial entailment relate to the notion of dialectical entailment, in particular to the concept of τ-deducibility. In order to express that a sentence follows from other sentences given the inferential relations encoded in a certain dialectical structure, Betz (2009) defined τ-deducibility as follows:

Definition 14 (τ-deducibility) *Let $\tau = \langle T, A, U\rangle$ be given. A statement c is deducible in τ from $P = \{p_1 \dots p_n\}$, briefly "$P \vdash_\tau c$", iff there is an argument $a \in T$ with conclusion c and there is a green, acyclic subgraph $\tau' \subseteq \tau$ such that (i) a is the only sink of τ', and (ii) $\Pi_{\tau'} \subseteq P$.*

Here, $\Pi_{\tau'}$ refers to the free premisses in the green subgraph τ', i.e. the set of premisses which are not equivalent with the conclusion of an argument in τ'.

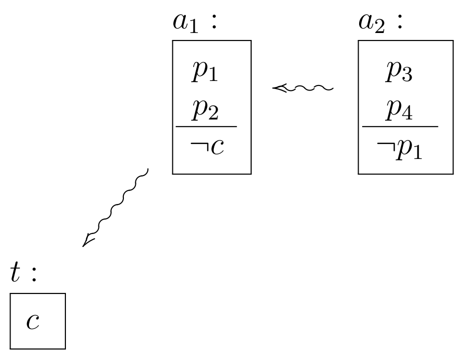
Example 1 is a green, acyclic dialectical structure whose only sink is t. Its free premisses, Π, include p_2, p_3, p_4. So, relative to example 1, c is τ-deducible from these, $\{p_2, p_3, p_4\} \vdash_\tau c$. Moreover, let $\mathcal{P}$ and $\mathcal{C}$ denote the partial positions which consider p_2, p_3, p_4, respectively c true. As there is no dialectically coherent, complete position $\mathcal{Q}$ which extends $\mathcal{P}$ and $\mathcal{C}$, we have $\sigma_{\mathcal{CP}} = \sigma_{\mathcal{P}}$, and thence the degree of partial entailment of c by p_2, p_3, p_4 is:

$$J_\tau(\mathcal{C}|\mathcal{P}) = \frac{\sigma_{\mathcal{CP}}}{\sigma_{\mathcal{P}}} = 1.$$

This observation can be generalised. If $P \vdash_\tau c$ on some dialectical structure τ and if the partial position $\mathcal{P}$ which considers sentences in P as true is coherent at all (otherwise the degree of partial entailment would not be defined), then $J_\tau(c|\mathcal{P}) = 1$. This is due to the fact that every complete position on a green, acyclic subgraph which considers its free premisses true, has to consider its conclusions true in order to be dialectically coherent.

However, the reverse does not hold in general. Consider:

Example 3



Here, $\neg p_1$ is clearly not τ-deducible from $\{c, p_2\}$. There is, however, no dialectically coherent, complete position which considers p_1, p_2, c true in the same time. In other words, $\neg p_1$ is true according to every dialectically coherent, complete position which extends the partial position $\mathcal{P}$ defined on $\{c, p_2\}$ with $\mathcal{P}(c) = \mathcal{P}(p_2) = \mathbf{t}$. Thus, we have $J_\tau(\neg p_1|\mathcal{P}) = 1$.

Cases like example 3 seem to pinpoint a disadvantage of the concept of τ-deducibility. For doesn't it make perfectly sense to say that, given the inferential relations encoded in the dialectical structure, c and p_2 imply $\neg p_1$—although there is no corresponding green subgraph? The inadequacy of the purely formal, syntactic concept of τ-deducibility motivates the introduction of a semantic notion of dialectical entailment.

Definition 15 (Dialectic-semantic entailment) *Let $\tau = \langle T, A, U \rangle$ be given. A statement c is dialectic-semantically entailed in τ by $P = \{p_1 \dots p_n\}$, briefly "$P \models_\tau c$", iff every dialectically coherent, complete position on τ which extends $\mathcal{P}$ also extends $\mathcal{C}$; where $\mathcal{P}, \mathcal{C}$ are partial positions with $\mathcal{P}(p_1) = \ldots = \mathcal{P}(p_n) = \mathbf{t}$ and $\mathcal{C}(c) = \mathbf{t}$.*

It follows immediately from this definition that $P \models_\tau c \iff J_\tau(c|\mathcal{P}) = 1$. In particular, regarding example 3, we have $\{c, p_2\} \models_\tau \neg p_1$.

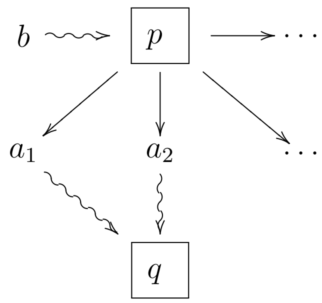
Inasmuch as the notion of dialectic-semantic entailment seems to rectify certain shortcomings of the concept of τ-deducibility, it seems to be advisable to replace, in a theory of dialectical structures, the latter by the former. This means, specifically, to reformulate discursive aims such as fulfilling burdens of proof (cf. Betz, 2009, def. 8), in terms of dialectic-semantic entailment rather than τ-deducibility.

7 Robustness

Having formally defined and explored the notion of degree of justification, we shall now, based on the precise explication, step back and ask what that notion actually signifies. What does it tell us about a debate? What precisely can we infer from a thesis' degree of justification being rather high, or low, in a certain dialectical structure?

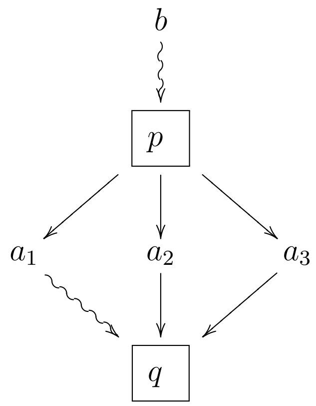
In order to answer these questions, it is helpful to recall the vague, intuitive idea the previous sections tried to explicate: The question "How well is this-and-this thesis justified?" has been rephrased as "What would have to be the case such that this thesis were false?", thence expressing the idea that the fewer positions there are in debate according to which t is false, the higher t's degree of justification. This general idea already reveals that one enters a fundamentally different mode of debate evaluation when one inspects a thesis' degree of justification as compared to asking, for example, whether a single, specific position is dialectically coherent, or who carries a burden of proof. Calculating degrees of justification presupposes considering all positions that can be coherently adopted in a debate, and presupposes, moreover, to regard them as equal (in some sense). To put it differently, calculating degrees of justification involves adopting a view from nowhere on a certain debate. This mode of evaluation starkly contrasts with the debate evaluation as set out in Betz (2009), which is basically debate evaluation from specific perspectives: The latter puts the focus on evaluating single positions in a debate, examining, in particular, whether the proponent's position is coherent, and whether the proponent carries a burden of proof. More specifically, the fact that introducing an argument a increases t's degree of justification neither indicates that doing so is an effective strategy to foster the discursive aims of a proponent who claims t, nor does a discursive strategy which helps a proponent to defend t necessarily increase t's degree of justification. In order to see that, consider the following two fictitious examples.

Example 4



In example 4, Anna claims q which is attacked by two arguments, both including p as premiss. Anna holds that p which is also a premiss of further arguments in her subdebate. Putting forward an argument b which concludes that $\neg p$ apparently increases $J_\tau(q)$, yet is not available as an option for Anna who would contradict herself.

Example 5



Barbara, in example 5, claims q. Her thesis, though, is attacked by an argument a_1 which starts from premiss p. That very premiss is also used in two arguments that support q, a_2 and a_3, but which Barbara does not adopt because she rejects some of their premisses. In this situation, Barbara effectively defends her thesis by attacking p with argument b. Yet, introducing the argument b decreases q's degree of justification.

These clarifications might make one wonder whether degrees of justification are of any significance at all. Of what avail is the view from nowhere? In the end, everybody is someone—not no one in particular; everybody is somehow, somewhere located in a debate—not nowhere. Isn't debate evaluation which considers specific perspectives and positions all we ultimately want and need?

So should we refrain from giving an interpretation of degrees of justification which says more than that they indicate the relative proportion of coherent positions in which the thesis is true? If one affirms this question, those parts of our intuitive notion of strength of justification which have been successfully captured by the theoretical concept of degree of justification would only have a minor role to play in debate evaluation. More dramatically, we would not be able to explain why one should belief in a thesis which is well justified: rational belief would have nothing to do with degree of justification. This would, at least, be puzzling. Thence, in the remainder of this section, I am going to discuss how the gap between rational belief and degree of justification might be bridged.

I see two similar explanations for why it is rational to believe sentences with high degree of justification: (a) because believing a partial position with a higher degree of justification provides one with more future alternatives regarding the extension of one's position, and (b) because a partial position with a higher degree of justification is more immune to falsification. Both explanations stress that well justified partial positions are more robust; and both take off from the same starting point: Consider a dialectically coherent partial position $\mathcal{P}$ on some τ. $\mathcal{P}$ might represent the entire position a proponent has so far adopted in a debate, or it might delineate the core position of a proponent, i.e. those truth value assignments she is resolved to hold on to, come what may. The question is: Why should a rational proponent be interested in $J_\tau(\mathcal{P})$ being high? First of all, recall that a high degree of justification signifies that there are many dialectically coherent, complete positions in τ to which $\mathcal{P}$ can be extended. In contrast, the lower the degree of justification, the less dialectically coherent, complete positions on τ extend $\mathcal{P}$. In the extreme, there will be but one dialectically coherent, complete position that extends $\mathcal{P}$; and in that case, $\mathcal{P}$, although possibly defined on a small subset of τ's sentences only, determines the truth values of all sentences in τ.

Explanation (a) urges us to picture a proponent who is constructing her position on τ step by step, and who has thus far adopted $\mathcal{P}$. Now such a proponent might be interested in $\mathcal{P}$ having a high degree of justification because that minimises her current commitment to future truth value assignments. Recent results in decision theory substantiate this point. Thus, Rosenhead

(2001) argues that "robustness" is, in general, an adequate decision principle when "there is radical uncertainty about the future, and where decisions can or must be staged sequentially" (p. 181). As the proponent builds up her position (from a certain basis) stepwise and hardly knows what to think about theses she hasn't considered in detail yet, both conditions, in a certain sense, apply in our case. If we say that "an acceptable configuration" which is "reachable" corresponds to, in our case, a complete and dialectically coherent position which extends $\mathcal{P}$, the following definition of the so-called robustness score boils down to the degree of justification:

> The robustness of any of the candidate initial commitments under consideration can then be defined as the ratio of the number of acceptable configurations that are 'reachable' from that commitment, to the total number of configurations which have been identified as having acceptable expected performance at the planning horizon. (Rosenhead, 2001, p. 189)

Why, then, is it rational to believe in theses with higher degree of justification rather than in theses with lower one? It is rational to do so when completing one's own partial position over a dialectical structure because:

1. It is rational to adopt a dialectically coherent, complete position.
2. It is rational to maximise future options when taking sequential decisions under uncertainty.

This is the first suggestion how the gap between rational belief and degree of justification can be bridged. The second explanation (b) stresses that partial positions with higher degree of justification are more immune to falsification. Assume that $\mathcal{P}$ represents the core beliefs of a proponent. Now if $J_\tau(\mathcal{P})$ is very low, truth values of most sentences in the debate are determined by $\mathcal{P}$ and, therefore, fixing a truth value of some sentence in τ, i.e. the growth or modification of the background knowledge, risks rendering the entire partial position (conditionally) incoherent. So, if $J_\tau(\mathcal{P})$ is very low, the proponent is inflexible and runs the risk that a single commitment (outside her core position) forces her to modify her core position. If, however, $J_\tau(\mathcal{P})$ is high, the proponent's core position can be coherently combined with many different truth value assignments, the position is flexible and robust against future growth and modification of the background knowledge, it is immune to falsification. This is the second reason why a rational proponent is interested in adopting a partial position with high degree of justification.

Although this reasoning explains to which extent it is rational to adopt a position with high degree of justification, it should be noted, as a caveat, that it is not unreasonable to adopt a position with low degree of justification in the same sense as it would be unreasonable to adopt a dialectically incoherent position. A proponent who adopts an incoherent position contradicts herself. This is not true for a proponent who adopts a position with low degree of justification. Such a position can still be coherent (unless $J_\tau(\mathcal{P}) = 0$). And it might be perfectly rational—e.g. because the proponent is eager to maintain certain theses—to adopt it.

8 Conclusion

This paper defined and investigated the quantitative concept of a thesis' degree of justification. That concept explicates our intuitive, comparative notion of strength of justification—a notion we tried to trace by the following, paradigmatic question: What would have to be the case, given the corresponding dialectical structure and body of background knowledge, such that the thesis were actually false? The theoretic concept turned out to fit to our pre-theoretic intuitions, since what typically counts as a support and defence of (attack on) a thesis, actually increases (decreases) its degree of justification. Moreover, the concept of degree of justification has been related to other notions of the theory of dialectical structures, in particular to τ-deducibility, and it has been shown to satisfy the axioms of probability theory.

The last section discussed the significance of degrees of justification. I suggested that the apparent gap between rational belief and degree of justification can be bridged with a view to the property of being a robust position. High degree of justification implies high robustness, and this, in turn, ensures that a partial position (a) can be flexibly extended in many different ways when constructing a complete position, and (b) is highly immune to falsification.

A Appendix: Dialectically coherent positions and complete, closed subdebates in equilibrium

Before we relate the notion of a dialectically coherent position to the concept of a complete, closed subdebate in equilibrium, I repeat, without comment, the relevant definitions from Betz (2009).

Definition 16 (Validity-function) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure. A function $v : T \to \{0, 1\}$ is called a validity-function on τ iff for all $a \in T$: $(v(a) = 0 \leftrightarrow \exists b \in T : b \rightsquigarrow a \wedge v(b) = 1))$.*

If the validity-function exists on τ and is unique, it is labelled "ϑ" and an argument $a \in T$ is called "τ-valid" iff $\vartheta(a) = 1$, "τ-invalid" otherwise.

Definition 17 (Free premiss) *Let $\tau = \langle T, A, U \rangle$ be given. A premiss p of an argument in τ is called "bound in τ" iff*

$$\exists a \in T : \Big[\vartheta(a) = 1 \wedge \Big((p \Leftrightarrow C(a)) \vee (p \Leftrightarrow \neg C(a))\Big)\Big].$$

If and only if a premiss is not bound in τ, it is "free in τ". The set of all free premisses of τ is called Π_τ.

Definition 18 (Equilibrium) *A dialectical structure $\tau = \langle T, A, U \rangle$ is said to be in equilibrium iff not*

$$(p \in \Pi_\tau \vee \Pi_\tau \vdash_\tau p) \wedge (\neg p \in \Pi_\tau \vee \Pi_\tau \vdash_\tau \neg p)$$

for some sentence p.

Definition 19 (Stance-attribution) *Let $\tau = \langle T, A, U \rangle$, and $O = \{o_1, \ldots, o_k\}$ be a set of proponents. A function $S : O \to \mathbf{P}(T)$ which assigns each proponent a subset $T_i \subseteq T$ is called a stance-attribution on τ. $\tau_i = \langle S(o_i), A|_{S(o_i)}, U|_{S(o_i)} \rangle$ is the subdebate accepted by o_i. A proponent o_i claims that*

- *All $p \in \Pi_{\tau_i}$ are true.*
- *All $C(a)$ (with $a \in S(o_i)$ is τ_i-valid) are true.*

Definition 20 (Closed subdebates) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure and $S : O \to \mathbf{P}(T)$ a stance-attribution on τ. A subdebate τ_i induced by S is called "closed" iff there is no $a \in (T \setminus T_i)$ such that $\Pi_{\tau_i} = \Pi_{\tau'}$, $\tau' = \langle S(o_i) \cup \{a\}, A|_{S(o_i) \cup \{a\}}, U|_{S(o_i) \cup \{a\}} \rangle$.*

Betz (2009) stipulated that a subdebate has to be complete in order to represent a position a proponent can rationally adopt in a debate, for otherwise the status assignment might not even exist on her subdebate. The following attempt to relate the concept of a coherent position (as truth value assignment) and the notion of a closed, complete subdebate in equilibrium will show that subdebates have to satisfy an additional condition in order to represent rational positions: for each sentence whose negation occurs in the debate as well, the proponent has to assert exactly one of both in a thesis. As the completeness condition already required that specific theses exist in a dialectical structure, I propose to modify and extend the definition of a complete stance-attribution as follows instead of introducing a further condition.

Definition 21 (Complete stance-attribution) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure. The stance-attribution $S : \{o_1, \ldots, o_k\} \to \mathbf{P}(T)$ is called "complete" iff for every induced subdebate τ_i $(i = 1 \ldots k)$ there is a τ_i-valid thesis $t \in T_i$ stating either p or $\neg p$*

1. *for every pair of contradictory sentences $p, \neg p$ which both occur in τ while neither p nor $\neg p$ occurs in τ_i,*
2. *for every conclusion p of a τ_i-invalid argument which neither attacks nor supports another argument in τ_i, and*
3. *for every red circle C in τ_i such that*
   (a) *t attacks one of C's arguments,*
   (b) *t is neither part of a red circle itself nor connected to a red circle via a red directed path from that circle to a_C, and*
   (c) *t is assigned the validity value 1 according to a partial evaluation of τ_i, $\vartheta_{\text{partial}}$, which excludes all arguments in red circles.*

As a final preliminary concept, we introduce

Definition 22 (Generated v-function) *Let $\tau = \langle T, A, U \rangle$ be a dialectical structure and $\mathcal{Q}$ be a complete position on τ. A function $v : T \to \{0, 1\}$ is generated by $\mathcal{Q}$ iff for every argument $a \in T$ with premisses $p_1, \ldots, p_n$:*

$$v(a) = 1 \iff \mathcal{Q}(p_1) = \ldots \mathcal{Q}(p_n) = \mathbf{t}.$$

v is called a v-function.

With these definitions at hand, we can now proof

Proposition 6 (Construction of dialectically coherent position) *Let $S : \{o_1, \ldots, o_k\} \to \mathbf{P}(T)$ be a complete stance-attribution on the dialectical structure $\tau = \langle T, A, U \rangle$. If the induced subdebate τ_i is closed, in equilibrium, and the validity function ϑ exists on τ_i, then there is a dialectically coherent, complete position $\mathcal{Q}$ on τ such that for the v-function v which is generated by $\mathcal{Q}$:*

$$\forall a \in T : v(a) = 1 \iff \vartheta(a) = 1.$$

Proof: Let Π_{τ_i} denote the set of free premisses of τ_i. We construct $\mathcal{Q}$ on τ_i first, show that $\mathcal{Q}|_{\tau_i}$ satisfies the coherence constraints on τ_i, and then proceed by extending $\mathcal{Q}$ to those sentences that do not occur in τ_i.

Step 1: We set for every sentence p that occurs in τ_i

$$\mathcal{Q}(p) := \begin{cases} \mathbf{t} & p \in \Pi_{\tau_i} \vee \exists a \in T_i : [(p \leftrightarrow C(a)) \wedge \vartheta(a) = 1] \\ \mathbf{f} & \text{otherwise} \end{cases}$$

To see that $\mathcal{Q}|_{\tau_i}$ assigns complementary truth values to contradictory sentences (constraint 2 in definition 5), consider p, q with $q \leftrightarrow \neg p$ occurring in τ_i. If p is a τ_i-free premiss or the conclusion of a τ_i-valid argument, then q is not because τ_i is in equilibrium and thence $\mathcal{Q}(p) = \mathbf{t}, \mathcal{Q}(q) = \mathbf{f}$. If, in contrast, both p and q are neither τ_i-free premisses nor conclusions of τ_i-valid arguments, then both are conclusions of τ_i-invalid arguments only. Yet as τ_i is complete, there are no pairs of contradictory sentences which are but conclusions τ_i-invalid arguments. So the second case does not arise. We still have to show that $\mathcal{Q}|_{\tau_i}$ assigns conclusions the value $\mathbf{t}$ if the corresponding premisses are true (constraint 3 in definition 5): If for all premisses $p_1 \dots p_n$ of some $a \in T_i$ it holds that $\mathcal{Q}(p_1) = \dots = \mathcal{Q}(p_n) = \mathbf{t}$, then a is by construction not attacked by any τ_i-valid argument—τ_i would not be in equilibrium otherwise—, and therefore $\mathcal{Q}(C(a)) = \mathbf{t}$.

Step 2: We extend $\mathcal{Q}$ to $\tau \setminus \tau_i$ as follows (note that we consider sentences that do occur in τ but not in τ_i): Every sentence p whose negation occurs in τ_i is assigned the complementary truth value to $\mathcal{Q}(\neg p)$. Every remaining sentence is set to $\mathbf{f}$. Now, let us complete the check for dialectical coherency. Let p, q be two contradictory sentences, not both in τ_i (if one of both is in τ_i the construction ensures that they are assigned complementary truth values). But by completeness of S, there is a thesis in τ_i that states either p or q, and therefore the construction guarantees $\mathcal{Q}(p)$ and $\mathcal{Q}(q)$ are complementary. Next, does $\mathcal{Q}$ satisfy the 'deduction constraint' (constraint 3 in definition 5)? The first thing to note is that every argument $a \in \tau \setminus \tau_i$ contains at least one premiss which is false. For otherwise, every premiss p of a were either (i) a τ_i-free premiss in τ_i, (ii) a conclusion of a τ_i-valid argument, or (iii) the negation of a sentence in τ_i that is neither (i) nor (ii). Yet, since by completeness of S the only sentences in τ_i that are neither (i) nor (ii) are negations of conclusions of τ_i-valid arguments, (iii) amounts to being the conclusion of a τ_i-valid argument, i.e. (ii). Hence τ_i would not be a closed subdebate. Now because every argument a in $\tau \setminus \tau_i$ has at least one false premiss, $\mathcal{Q}$ satisfies the deduction constraint. Also, this fact guarantees that $v(a) = 0$. □

The final proposition tells us how to construct a closed subdebate in equilibrium which corresponds to a given dialectically coherent, complete position.

Proposition 7 (Construction of stance-attribution) *Let* $\tau = \langle T, A, U \rangle$ *be a dialectical structure and* v *a v-function that is generated by a dialectically coherent, complete position* $\mathcal{Q}$ *on* τ*. There exists a stance-attribution* $S : \{o\} \to \mathbf{P}(T)$ *inducing the subdebate* τ_o *such that*

1. v *is a validity function on* τ_o*,*
2. τ_o *is in equilibrium,*
3. τ_o *is closed.*

Proof: First, we construct τ_o iteratively. Let $T_0 = \emptyset$ and apply the following rule provided T_n is given

(R) Let T^* be the set of all arguments $a \in T \setminus T_n$ such that for every premiss p of a: $\mathcal{Q}(p) = \mathbf{t}$ or p negates the conclusion of an argument $b \in T_n$ with $v(b) = 1$. If $T^* = \emptyset$ then $T_o = T_n$, STOP. Otherwise $T_{n+1} = T_n \cup T^*$.

Ad 1): We show that $\vartheta : T_o \to \{0, 1\}$ with $a \mapsto v(a)$ is a validity function on τ_o. By construction an argument $a \in T_o$ has a premiss p with $\mathcal{Q}(p) = \mathbf{f}$ if and only if there is an argument b which is τ_o-valid and attacks a. Thus, ϑ does satisfy the recursive definition of a validity function.

Ad 2): Assume that τ_o were not in equilibrium, that is there were a sentence p such that both p and $\neg p$ are (i) a τ_o-free premiss or (ii) a conclusion of a τ_o-valid argument. If p were a τ_o-free premiss in argument a, then (by definition of "free premiss") $\neg p$ couldn't be the conclusion of a τ_o-valid argument. So $\neg p$ would be a τ_o-free premiss in some argument b, too. Because of dialectical coherency, $\mathcal{Q}(p)$ is complementary to $\mathcal{Q}(\neg p)$, and thence the algorithm would not have picked a and b. Yet if p were the conclusion of a τ_o-valid argument, $\neg p$ would not be τ_o-free and would thus be the conclusion of a τ_o-valid argument, too. Still, this contradicts the assumption that $\mathcal{Q}$ is dialectically coherent.

Ad 3): Assume there were an argument $a \in T \setminus T_o$ such that adding a to τ_o would not increase the set of τ_o-free premisses. Then every premiss of a would either be (i) a τ_o-free premiss of some argument in T_o (and thus be true), (ii) the conclusion of a τ_o-valid argument's conclusion (and thus be true), or (iii) the negation of a τ_o-valid argument's conclusion. Therefore, the rule (R) would have picked a and would not have stopped. □

So not only can we construct dialectically coherent, complete positions from stance-attributions, but, inversely, every coherent position corresponds to a closed subdebate in equilibrium. Note that such a subdebate is not necessarily complete since the dialectical structure τ might simply not contain enough theses.

References

Gregor Betz. Evaluating dialectical structures. *Journal of Philosophical Logic*, 38(3):283–312, 2009.

Gregor Betz. Evaluating dialectical structures with Bayesian methods. *Synthese*, 163:25–44, 2008.

Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93(1-2):63–101, June 1997.

Rudolf Carnap. *Logical Foundations of Probability.* University of Chicago Press, Chicago, 1950.

Claudette Cayrol and Marie-Christine Lagasquie-Schiex. On the acceptability of arguments in bipolar argumentation frameworks. In Lluis Godo, editor, *ECSQARU*, volume 3571 of *Lecture Notes in Computer Science*, pages 378–389. Springer, 2005.

Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.

Rolf Haenni and Norbert Lehmann. Probabilistic argumentation systems: A new perspective on the dempster-shafer theory. *International Journal of Intelligent Systems*, 18(1):93–106, 2003.

Rolf Haenni, Jürg Kohlas, and Norbert Lehmann. Probabilistic argumentation systems. In Jürg Kohlas and Serafín Moral, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 5: Algorithms for Uncertainty and Defeasible Argumentation. Kluwer, Dordrecht, 2000.

Jürg Kohlas. Probabilistic argumentation systems: A new way to combine logic with probability. *Journal of Applied Logic*, 1(3-4):225–253, June 2003.

Kathryn B. Laskey and Paul E. Lehner. Assumptions, beliefs and probabilities. *Artificial Intelligence*, 41(1):65–77, 1989.

Fangzhen Lin and Yoav Shoham. Argument systems: A uniform basis for nonmonotonic reasoning. In Ronald J. Brachman, Hector J. Levesque, and Raymond Reiter, editors, *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 245–255. Morgan Kaufmann, 1989.

John L. Pollock. Defeasible reasoning. *Cognitive Science*, 11(4):481–518, 1987.

John L. Pollock. The structure of epistemic justification. *American Philosophical Quarterly*, 4:62–78, 1970.

John L. Pollock. *Cognitive Carpentry. A Blueprint for How to Build A Person.* MIT Press, 1995.

John L. Pollock. Defeasible reasoning with variable degrees of justification. *Artificial Intelligence*, 133(1-2):233–282, December 2001.

Henry Prakken and Gerard Vreeswijk. Logics for defeasible argumentation. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 4, pages 219–318. Kluwer, Dordrecht, second edition edition, 2001.

Stefan Rahmstorf and Hans J. Schellnhuber. *Der Klimawandel.* C.H. Beck, München, 2006.

Jonathan Rosenhead. Robustness analysis. In *Rational Analysis for a Problematic World Revisited*, pages 181–207. Wiley, Chichester, second edition edition, 2001.

Holm Tetens. *Philosophisches Argumentieren.* C.H. Beck, München, 2004.

Bart Verheij. *Rules, Reasons, Arguments. Formal studies of argumentation and defeat.* Dissertation Universiteit Maastricht, 1996.

Gerard A. W. Vreeswijk. Abstract argumentation systems. *Artificial Intelligence*, 90(1-2):225–279, February 1997.