Shangin, V. O. 2023. “A Supraclassical Probabilistic Entailment Relation” [in English]. *Filosofiya. Zhurnal Vysshey shkoly ekonomiki [Philosophy. Journal of the Higher School of Economics]* 7 (4), 215–238.

Vasily Shangin*

A Supraclassical Probabilistic Entailment Relation**

Submitted: June 07, 2023. **Reviewed:** Sept. 20, 2023. **Accepted:** Sept. 30, 2023.

Abstract: The paper presents an original supraclassical nontrivial plausible entailment relation $\approx$ that employs Kolmogorov’s probability theory. Its crucial feature is the primitiveness of a conditional probability, which one calculates with the help of the method of truth tables for classical propositional logic. I study the properties of the entailment relation in question. In particular, I show that while being supraclassical, i. e., all classical entailments and valid formulas are $\approx$-valid, but not vice versa, it is not trivial and enjoys the same form of inconsistency as classical entailment $\models$ does. I specify the place of the proposed probability entailment relation in certain classifications of nonclassical entailment relations. In particular, I use Douven’s analysis of some probabilistic entailment relations that contains dozens of properties that are crucial for any probabilistic entailment relation, as well as Hlobil’s *choosing your nonmonotonic logic: shopper’s guide*, due to the fact that $\approx$ is not monotonic, and Cobreros, Egré, Ripley, van Rooij’s entailment relations for tolerant reasoning. At last, I perform a comparative analysis of classical, the proposed, and some other entailment relations closely related to the latter: those introduced by Bocharov, Markin, Voishvillo, Degtyarev, Ivlev, where the last two entailment relations are based on the so-called principle of reverse deduction, which is an intuitively acceptable way to connect classical and probabilistic entailment relations.

Keywords: Classical Logic, Probabilistic Logic, Bayes, Evidence, Entailment, Reverse Deduction.

DOI: 10.17323/2587-8719-2023-4-215-238.

*Vasily Shangin, PhD in Philosophy; Associate Professor at the Lomonosov Moscow State University (Moscow, Russia); Institute of Logic Cognitive Science and Personality Development (Moscow, Russia), shangin@philos.msu.ru, ORCID: 0000–0002–7380–3213.

Acknowledgements: Some topics of the present paper were discussed with A. Weir and N. L. Arkhiereev, to whom the author expresses his infinite gratitude. The author is thankful to the anonymous referee, whose invaluable comments helped improve the embryonic draft of this paper. The author’s cordial thanks go to the marvelous Ksyu. Many thanks are due to the developers of the QuillBot software (https://quillbot.com/) for the priceless help in checking and paraphrasing the English of this paper. Last but not least, the author wishes to express thanks to the linguistic editor of this journal for the final invaluable corrections. The study was funded by RSCF, grant № 22–18–00450, https://rscf.ru/project/23-28-00801/.

1. INTRODUCTION

The key motivation of the present paper is to propose some other original plausible entailment relation that employs conditional probability as a primitive and is a nontrivial generalization of the classical entailment relation. Due to the fact that each plausible entailment discussed in this paper employs probability theory, I use the terms of "plausible entailment" and "probabilistic entailment" as synonyms. I want my probabilistic entailment relation to be defined simply as classical entailment relation is. Hence, the proof-theoretic method of truth tables allows me to propose a variant of the Bayesian account which defines a rational probability measure of a propositional formula A via its truth table within the framework of **CPL**.[1] While choosing between a conditional (binary) or an unconditional (unary) probability as a primitive, I prefer the former due to certain relevant shortcomings of the latter.

> Binary probability functions, also known as *conditional probability functions*, are often defined in terms of singulary ones. For example, in (Carnap, 1950, Kolmogorov, 1956), etc., $P(A/B)$ is set at $P(A \wedge B)/P(B)$ when $P(B) \neq 0$, but otherwise is left without a value. [...] Partial functions, however, are unwieldy and—perforce—of limited service. An alternative approach, favoured by Keynes as early as 1921 and, hence, antedating Kolmogorov's, has now gained wide currency, thanks to such diverse writers as Reichenbach, Jeffreys, Von Wright, Renyi, Carnap (in post-1950 publications), Popper, etc. Handling binary probability functions as you would singulary ones, you adopt constraints suiting your understanding of $P(A/B)$ and then own as your binary probability functions all functions meeting these constraints.

This representative list of writers lacks Wittgenstein, whose *Tractatus Logico-Philosophicus*, published in 1922, defines a binary probability as follows (a truth-ground is his term for an entry of a truth table, where a formula is true; the notation is original):

> If T_r is the number of the truth-grounds of the proposition "r", T_{rs} the number of those truth-grounds of the proposition "s" which are at the same time truth-grounds of "r", then I call the ratio $T_{rs} : T_r$ the measure of the *probability* which the proposition "r" gives to the proposition "s" (Wittgenstein, 1922: 5.15).

I refer to von Wright's exposition of the evolution of Wittgenstein's accounts of probability and note that the exposition contains a list of

[1] **CPL** stands for classical propositional logic.

19th-century thinkers, including Bolzano, who favor the primitiveness of conditional probability (von Wright, 1969: 263–265).

At last, while studying *the question of extending propositional logic to a logic of plausible reasoning* and positing *four requirements that any such extension should satisfy*, van Horn proposes the following argument pro conditional probability (*the italic is mine*) (Horn, 2017: 313):

> We see therefore that, although it is the conditional probabilities $c(h, e)$ that most interest Carnap, unconditional probabilities are for him more fundamental. *In contrast, we take conditional plausibilities as the fundamental concept* and, rather than imposing the laws of probability, seek to derive them.

As a result, I obtain a supraclassical probabilistic entailment $\approx$ relation, where all classical tautologies and entailments are valid, too, whilst the opposite is wrong. The obtained probabilistic entailment is anticipated to lack certain properties in order to avoid the textbook triviality argument that any supraclassical consequence relation holds. A variant of it is Exercise 1.50 in Mendelson's classic textbook (Mendelson, 1997: 43). Roughly speaking, it could be shown that the argument is not $\approx$-valid, due to its lack of transitivity.

The price that is paid in order to preserve the nontrivial supraclassicality is an indefinite position of $\approx$ in terms of the properties that it has. To clarify the position, I use some nomenclatures of nonclassical entailment relations as well as some analysis of their properties. Due to its monotonicity-free, Hlobil's *shopper's guide* is of invaluable help (Hlobil, 2018). In Section 3.2, I use his classification in order to specify the position of $\approx$ with the help of Douven's extensive analysis of dozens of properties, which he finds in the literature devoted to probabilistic entailments (see Section 3.1). Speaking briefly, it turns out that $\approx$ is not Hlobil's favorite nonmonotonic logic, whilst the number of properties that $\approx$ holds, according to the Douven analysis, is quite standard, which makes it a rather weak logic.

In the end, let me address the following referee's suggestion (the translation from Russian is mine).

> The paper leaves an ambivalent impression. On the one hand, in studying the properties of plausible (probabilistic) entailment, the author bases his conclusions on quite important logical results [...], i. e., the methodology he employs is well-grounded. On the other hand, the author chooses the probabilistic entailment as the object of his analysis, which is based on the method of calculating logical probability with the help of truth tables that Wittgenstein proposed in *Tractatus Logico-Philosophicus*. And though Wittgenstein's idea is attractive with its

> simplicity and clarity, it is commonplace that one cannot formulate an adequate probabilistic logic on the basis of this idea because I face some difficulties here. [...] it is not occasionally, therefore, that the method in question is employed in pedagogy as a rule. However, in the conclusive part of his paper, the author himself provides a comparative analysis of the original probabilistic entailment and its alternatives, which belong to certain textbooks on logic rather than scientific literature. One comes to the conclusion that, speaking metaphorically, the author *uses a sledgehammer to crack a nut.*

I express my consent to the essence of the quote above. It was one of my motivations to write this paper to find out whether the scientific literature contains any explanation of the textbook approaches mentioned here. It was a real surprise for me to find out that teaching logical introductory courses and doing the science of logic do not go hand in hand in this aspect. Hence, the problem of searching for the well-foundedness of pedagogical approaches arises. And its solution certainly needs employing modern logical methods from the scientific literature, i. e., it needs *using a sledgehammer to crack a nut.* As a result, as the readers find out below, it is not the case that each pedagogical approach under discussion is well-founded, i. e., it was not completely worthless to crack the pedagogical nut with the scientific sledgehammer. On the other hand, while solving the problem, I came up with the idea to extend this impractical pedagogical approach as much as possible. By not straying far from the standard propositional language, I want to go beyond classical logic without falling into inconsistency. To this end, the main result of this paper — some supraclassical probabilistic entailment relation — is proposed.

The paper is structured as follows. In Section 2, I expose a supraclassical probabilistic entailment relation, where a conditional probability is a primitive. In Section 3, this relation is classified on certain nomenclatures found in the literature. Section 4 discusses some closely related alternatives to my approach. Section 5 summarizes the paper and outlines future research.

2. A SUPRACLASSICAL PROBABILISTIC ENTAILMENT RELATION

In this section, I first list Leblanc's probability axiomatization (Leblanc, 1983) and then define an original probabilistic entailment relation $\approx$, where a primitive is a binary probability and a unary probability is therefore definable via the former. To this effect, I employ a mechanical method of truth tables and an approach to plausible reasoning that relies upon Kolmogorov's probability theory (Keynes, 1921; Kolmogorov, 1956; Lorenz et al., 2019, Carnap, 1962).

Unless specified otherwise, henceforth I fix some standard language $\mathcal{L}$ of **CPL** over the standard alphabet with the conventional connectives and the notion of a formula. The letters A, B, C run over formulae and Γ runs over sets of formulae as usual. **CPL** is defined with Tarski's T, F-semantics, the standard definitions of satisfiable, valid, contradictory formulae and entailment relation $\models$. By default, $\top$ ($\bot$) denotes a fixed valid (contradictory) formula rather than the "verum" ("falsum") constant, as usual. What is unusual is the following

Definition 2.1. A is said to be *plausible* iff it is satisfiable and invalid.

In what follows, I employ propositional parts of certain axiomatizations of the unary and binary probability measures P whose provisions are based on (Popper, 1955) and thoroughly discussed in (Leblanc, 1983: 87–88, 107–109, accordingly).

Definition 2.2. (Popper-Leblanc's unary probability) The probability measure of B (denoted by $P(B)$) is a one-place total function which satisfies the following provisions:

1. $0 \leq P(B)$,
2. $P(\neg(B \wedge \neg B)) = 1$,
3. $P(B) = P(B \wedge A) + P(B \wedge \neg A)$,
4. $P(B) \leq P(B \wedge B)$,
5. $P(B \wedge A) \leq P(A \wedge B)$,
6. $P(B \wedge (A \wedge C)) \leq P((B \wedge A) \wedge C)$.

Textbook knowledge has the following

Remark 2.3. $P(\neg B) = 1 - P(B)$ follows from the axioms in Definition 2.2.

Definition 2.4. (Leblanc's binary probability) The probability measure of B given A (denoted by $P(B/A)$) is a total two-place function which satisfies the following provisions:

1. There are a statement B and a statement A such that $P(B/A) \neq 1$,
2. $0 \leq P(B/A)$,
3. $P(B/B) = 1$,
4. If there is a statement C such that $P(C/A) \neq 1$, then $P(\neg B/A) = 1 - P(B/A)$,
5. $P(B \wedge A/C) = P(B/A \wedge C) \cdot P(A/C)$,
6. $P(B \wedge A/C) = P(A \wedge B/C)$,
7. $P(B/A \wedge C) = P(B/C \wedge A)$.

Throughout the paper, I employ Γ to be $\{A_1, \ldots, A_k\}$, $k \neq 0$. Hence, $\bigwedge_{i=1}^{k} A_i$ denotes any conjunction consisting of $A_1, \ldots, A_k$, where the conjuncts are ordered and associated arbitrarily.

Definition 2.5. A binary probability of B given $A_1, \ldots, A_k$, $k \neq 0$, is determined via a *joint* truth table of $A_1, \ldots, A_k, B$ as follows, where n is the number of its rows simultaneously containing T, for each $A_1, \ldots, A_k$, and m, $m \leq n$, is the number of its rows simultaneously containing T, for each $A_1, \ldots, A_k$ and B:

- ⋄ $P(B/\bigwedge_{i=1}^{k} A_i) = 1$, if $n = 0$;
- ⋄ $P(B/\bigwedge_{i=1}^{k} A_i) = \dfrac{m}{n}$, otherwise.

The readers easily determine that $P(p/q) = \frac{1}{2} = P(p/\top)$ and $P(p/\bot) = 1$. Note that $P(\bot/\bot) = 1$.

Remark 2.6. The readers might not have found the first clause in Definition 2.5 anticipated because it leads straightforwardly to an informal and counter-intuitive fact that an impossible event makes any event certain. I refer the readers to (the discussion about) Definition 2.12 below: in short, I purportedly want to have a probability measure of this kind because it could serve as the basis of a supraclassical nontrivial probabilistic entailment, i. e., an entailment that generalizes **CPL**-entailment at any triviality-free cost.

The following lemma justifies Definition 2.5.

Lemma 2.7. Definition 2.5 meets the provisions of Definition 2.4.

Proof. Let me employ the notation $r(A)$ meaning that in a truth table for A, r is a number of rows containing T for A.

$P(p/q) = \frac{1}{2}$ proves Provision 1. Provision 2 follows from $P(\bot/\top) = 0$, the totality of P and the non-negativity of the numerator and denominator in the respective fraction. $P(p/p) = P(\top/\top) = P(\bot/\bot) = 1$ proves Provision 3. Under the *if*-clause in Provision 4, A is not $\bot$.[2] Hence, $\frac{r(\neg B,A)}{r(A)} + \frac{r(B,A)}{r(A)} = \frac{r(\neg B,A)+r(B,A)}{r(A)}$. Note that for any row in a truth table, where A is true, either $\neg B$, or B is true. Hence, $r(\neg B, A) + r(B, A) = r(A)$. In order to prove Proposition 5, note that $\frac{r(A\wedge C,B)}{r(A\wedge C)} \cdot \frac{r(A,C)}{r(C)} = \frac{r(A\wedge C,B)}{r(C)}$ because $r(A \wedge C) = r(A, C)$. Hence, $\frac{r(A\wedge C,B)}{r(C)} = \frac{r(B\wedge A,C)}{r(C)}$ because $r(A \wedge C, B) = r(B \wedge A, C)$. Provisions 6 and 7 are provable because $r(B \wedge A) = r(A \wedge B)$ and $r(C \wedge A) = r(A \wedge C)$, accordingly. □

[2]In this case, $P(B/\bot) = P(\neg B/\bot) = 1$.

A unary probability is traditionally defined via the binary probability.

Definition 2.8. An unary probability is as follows:

$$P(B) = P(B/\top).$$

Remark 2.9. A shortcoming of Definition 2.8 is that the underlying language must be able to express a tautology, which is not the case for some languages ($\{\wedge\}$, for example).

It is easy to see that

Lemma 2.10. Definition 2.8 meets the provisions of Definition 2.2.

Definition 2.5 could be equivalently reformulated in the traditional "unary probability" way now. It is worth noting that it lacks the usual problem connected with division by zero.

Definition 2.11. (An "unary-probability-style" formulation of Definition 2.5) Following Definition 2.8, a binary probability of B given $A_1, \ldots, A_k$, $k \neq 0$, is determined via a *joint* truth table of $A_1, \ldots, A_k, B$ as follows, where n is the number of its rows simultaneously containing T, for each $A_1, \ldots, A_k$, and m, $m \leq n$, is the number of its rows simultaneously containing T, for each $A_1, \ldots, A_k$ and B:

- ⋄ $P(B/\bigwedge_{i=1}^{k} A_i) = 1$, if $n = P(\bigwedge_{i=1}^{k} A_i) = 0$;
- ⋄ $P(B/\bigwedge_{i=1}^{k} A_i) = \frac{m}{n} = \frac{P((\bigwedge_{i=1}^{k} A_i) \wedge B)}{P(\bigwedge_{i=1}^{k} A_i)}$, otherwise.

The readers can easily determine that $P(p) = P(p/\top) = \frac{1}{2}$, $P(\bot) = P(\bot/\top) = 0$, and $P(p \to p/\top) = 1$.

A probabilistic entailment relation between Γ and B (denoted by $\Gamma \approx B$) as well as $\approx$-validity is defined as follows.

Definition 2.12. $\Gamma \approx B$ iff $\frac{1}{2} < P(B/\bigwedge_{i=1}^{k} A_i) \leq 1$.[3] In particular, $\approx B$ iff $P(B/\top) = 1$.

The readers can easily determine that $P(p) = P(p/\top) = \frac{1}{2}$, $P(\bot) = P(\bot/\top) = 0$, and $P(p \to p/\top) = 1$. that $p, q \approx p \to q$, $\approx p \to p$, and $\not\approx p \to q$. Notice that $p \not\approx q$ and $p \approx q \vee r$.

Before investigating the properties of $\approx$, let me consider some arguments contra the criterion $\frac{1}{2} < P(B/\bigwedge_{i=1}^{k} A_i) \leq 1$. According to Douven, (I slightly unify the original notation; the *italic* is not mine):

> Formally, the intuition that if E is to qualify as evidence for H, E should make H probable, or very probable, would come down to imposing the requirement

[3] The notation $P(B/\bigwedge_{i=1}^{k} A_i) \in (\frac{1}{2}, 1)$ is employed, too.

> that $P(H/E)$ be above some specified threshold value $\mathbf{t}$, which one might take to be .5 or even .9 or still higher (though it would be wrong to require that $\mathbf{t} = 1$, as surely I do not pretheoretically consider E to be evidence for H only if E makes H *certain*) (Douven, 2011: 487–488).

In a later paper, the thesis that the target criterion is wrong, gets the *consensus gentium* flavor (I slightly unify the original notation; the *italic* is not mine; the **boldface** is mine):[4]

> Some have said that the Bayesian notion of evidence fails to completely capture my intuitive notion of evidence. What I mean when I say that A is evidence for B is — according to these authors — not just that A makes B *more* probable, but also that A makes B *highly* probable. Formally, A is evidence in this strengthened sense iff (*i*) $P(B/A) > P(B)$ and (*ii*) $Pr(B/A) > \theta$, for some value θ close, but unequal, to 1. (Different authors hold different views about what the threshold value should be; but **all agree** [...] that $0,5 \leq \theta < 1$) (Douven, 2014: 264).

As seen from the quotes above, the criterion $\frac{1}{2} < P(B/\bigwedge_{i=1}^{k} A_i) \leq 1$ in Definition 2.12 contains two "abnormalities": $\frac{1}{2} < P(B/\bigwedge_{i=1}^{k} A_i)$ rather than $\frac{1}{2} \leq P(B/\bigwedge_{i=1}^{k} A_i)$ and $P(B/\bigwedge_{i=1}^{k} A_i) \leq 1$ rather than $P(B/\bigwedge_{i=1}^{k} A_i) < 1$. With regard to the latter, it is in line with my approach to generalize **CPL** in terms of its entailment relation *as nontrivially as possible* (see Remark 2.6 above).

Hence, I purportedly consider an event which makes another event *certain* as a kind of evidence. On the other hand, probability theory allows for events whose probability measures equal to 1: hence, classical valid formulae turn out to be natural analogs of such events. The former "abnormality" has a purely formal justification: if $\frac{1}{2} \leq P(B/\bigwedge_{i=1}^{k} A_i)$ were the criterion, then it would be the case that $p \approx q$, due to $P(p/q) = \frac{1}{2}$. And this is an entailment one would certainly try to avoid: any event follows from any other event.[5]

Let me investigate the properties of $\approx$. As purportedly intended, $\approx$ is a generalization of $\models$:

Lemma 2.13. If $\Gamma \models B$, then $\Gamma \approx B$. In particular, if $\models B$, then $\approx B$.

[4] A detailed analysis of this passage is in subsection 3.1 below.

[5] I notice that this entailment does not hold for the two probabilistic entailment relations that Douven analyzes thoroughly: one could easily assign the respective probabilities to different rows in a truth table for the target entailment. In my approach — let me stress it again — the probabilities of all the rows are equal. For example, in the case of 3 variables, a probability of each row is $\frac{1}{8}$.

Proof. By the definition of $\models$, if $\Gamma \models B$, then $P(B/\bigwedge_{i=1}^{k} A_i) = 1$. Hence, $\Gamma \approx B$. In order to prove that if $\models B$, then $\approx B$, one employs the fact that by the definition of $\models$, $A \models B$ iff $\models B$, given $P(A) = 1$. □

Corollary 2.14. There are Γ, B such that $\Gamma \approx B$ and $\Gamma \not\models B$.
Proof. $p \vee q \approx q$ and $p \vee q \not\models q$.[6] □

Here some basic properties of $\approx$ are explored. A more detailed study that classifies $\approx$ on the other probabilistic entailments is in Chapter 3 below. It is crucial to remember the notation $r(A, B)$: in a joint truth table for A and B, r is the number of rows containing T simultaneously for A and B.

Lemma 2.15. $\approx$ is reflexive, contractive, permutative, and neither symmetrical nor transitive nor monotonic.

Proof. $A \approx A$ follows from $P(A/A) = 1$.
$A, A \approx C$ iff $A \approx C$ follows from $r(A) = r(A, A)$.
$A, B \approx C$ iff $B, A \approx C$ follows from $r(A, B) = r(B, A)$.
It is not symmetrical due to $p \approx \top$ and $\top \not\approx p$.[7]
It is not transitive due to both $p \approx p \vee q$ and $p \vee q \approx q$, but $p \not\approx q$.[8]
It is not monotonic due to $p \vee q \approx p$, but $\neg p, p \vee q \not\approx p$.[9] □

It is only a weak form of inconsistency called (CNC) and considered on page 225 below that holds for $\approx$. In this aspect, $\approx$ behaves the same as $\models$.

The nontriviality of $\approx$ comes from Milne's argument (Milne, 2000: 311), too:

> As is well known, the following two principles are incompatible:
>
> 1. if h entails e then e confirms h, at least when h is not logically false and e is not logically true;
> 2. if e confirms h and h entails h' then e confirms h', at least when h' is not logically true.
>
> Since $e\&e'$ entails both e and e', it follows from (1) and (2) that any logically contingent statement confirms any other with which it is logically compatible.

$p \wedge q \models p$, but $p \not\approx p \wedge q$, due to $P(p \wedge q/p) = \frac{1}{2}$. Hence, (1) fails. To show the failure of (2), let me notice that $h \models h'$ implies $h \approx h'$, by Lemma 2.13. In order to derive $e \approx h'$ from $e \approx h$ and $h \approx h'$, one needs $\approx$ to be transitive, which is not the case by Lemma 2.15.

[6] $P(q/p \vee q) = \frac{2}{3}$.
[7] $P(\top/p) = 1$ and $P(p/\top) = \frac{1}{2}$.
[8] $P(p \vee q) = \frac{3}{4}$, $P((p \vee q)/p) = 1$, and $P(q) = \frac{1}{2}$, $P(q/(p \vee q)) = \frac{2}{3}$, but $P(q/p) = \frac{1}{2}$.
[9] $P(p) = \frac{1}{2}$, $P(p/p \vee q) = \frac{2}{3}$, but $P(p/(\neg p \wedge (p \vee q))) = 0$.

3. CLASSIFYING $\approx$

The purpose of this chapter is to deepen the present investigation of $\approx$ by two means. On the one hand, I am to find out more properties than the Gentzenian ones from Lemma 2.15 and, on the other hand, to classify the place of $\approx$ among the nonmonotonic and/or nontransitive logics. I believe that such semantic analysis would allow me to come up with another paper devoted to an adequate syntactic axiomatization of $\approx$. For this purpose, I employ Douven's analysis of a *similar* probabilistic entailment (Douven, 2014) as well as the *shopper's guide* by Hlobil to *choosing your nonmonotonic logic* (Hlobil, 2018) and Cobreros et al.'s entailments for tolerant reasoning (Cobreros & Egré & Ripley, 2021).

3.1. THE DOUVEN PROPERTIES

In (Douven, 2014), Douven provides a detailed analysis of two notions of evidential support, the *B*ayesian $A \multimap_B C$ and its *S*trengthen case $A \multimap_S C$. Note that both $\multimap_B$ and $\multimap_S$ do not belong to the object-language, i.e., neither $A \multimap_B C$, nor $A \multimap_S C$ are *conditionals*.[10] The former is short for $P(C|A) > P(C)$, where (I unify Douven's notation) "$C \multimap_B A$ means that A is evidence in the Bayesian sense for C $\langle \dots \rangle$ P designates a specific (but unspecified) person's degrees-of-belief function, to which all sentences containing the symbol $\multimap_B$ are taken to implicitly refer" (Douven, 2014: 263). The latter is short for (i) $P(C|A) > P(C)$ and (ii) $0,5 \leq P(C|A) < 1$, where (ii) is the commonly accepted interval to a threshold value such that A makes C *highly* probable, not just *more* probable.[11] It is in this sense that $A \multimap_S C$ strengthens $A \multimap_B C$. For the reasons of this paper, I will not discuss $A \multimap_B C$ and henceforth, $\multimap$ means $\multimap_S$ only.

The main result of Douven's paper is that out of 33 principles (see Table 1 in ibid.: 265), the below 11 ones hold for $A \multimap C$, where $\vdash$ is classical derivability relation and ANT, CNC, M2, M3, MOD, RCE, RCEA, RCEC, REF, WAND, XOR are their Douven labels.[12]

- ◇ (ANT) Whenever $A \multimap B$, then $A \multimap (A \wedge B)$;
- ◇ (CNC) Whenever $A \nvdash \bot$, $A \multimap B$ and $A \multimap \overline{B}$, then $\bot$;
- ◇ (M2) Whenever $A \multimap (B \wedge C)$ and $A \multimap (B \vee C)$, then $A \multimap B$ or $A \multimap C$;

[10] For conditionals, see, for example, (Flaminio, Godo, Hosni, 2020).

[11] I discuss this passage preliminary when I clarify Definition 2.12 above.

[12] Douven deciphers some of them (for example, he attributes M2–M3 to Milne, 2000) and also mentions their alternative labels. I do not replace his $\overline{A}$-notation with my $\neg A$-notation.

- ⋄ (M3) Whenever $\vdash \overline{B \wedge C}$, $A \rightarrowtriangle B$ and $A \rightarrowtriangle C$, then $A \rightarrowtriangle (B \vee C)$;
- ⋄ (MOD) Whenever $\overline{A} \rightarrowtriangle A$, then $B \rightarrowtriangle A$;
- ⋄ (RCE) Whenever $\vdash A \supset B$, then $A \rightarrowtriangle B$;
- ⋄ (RCEA) Whenever $\vdash A \equiv B$, then $(A \rightarrowtriangle C) \equiv (B \rightarrowtriangle C)$;
- ⋄ (RCEC) Whenever $\vdash A \equiv B$, then $(C \rightarrowtriangle A) \equiv (C \rightarrowtriangle B)$;
- ⋄ (REF) $A \rightarrowtriangle A$;
- ⋄ (WAND) Whenever $A \rightarrowtriangle B$ and $A \wedge \overline{C} \rightarrowtriangle \bot$, then $A \rightarrowtriangle (B \wedge C)$;
- ⋄ (XOR) Whenever $\vdash \overline{A \wedge B}$, $A \rightarrowtriangle C$ and $B \rightarrowtriangle C$, then $(A \vee B) \rightarrowtriangle C$.

Notice that Douven's proofs are not automatically applicable to $\mid\!\approx$. In fact, it turns out to be possible in the case of (XOR) only. On the other hand, I highlight every time I employ Douven's proofs.

To prove (ANT) one needs a special case of axiom 5 in Definition 2.4: $P(A \wedge B)/A = P(A/B \wedge A) \cdot P(B/A)$ and the fact that $B \wedge A \models A$. The latter implies $B \wedge A \mid\!\approx A$, by Lemma 2.13, with $P(A/B \wedge A) = 1$.

To prove (CNC) one needs axiom 4 in Definition 2.4 that guarantees the unsatisfiability of the "whenever" clause of (CNC).

A stronger version of (M2) which I label $(M2_{str})$ is valid for $\mid\!\approx$:

$(M2_{str})$ whenever $A \mid\!\approx B \wedge C$, then $A \mid\!\approx B$ and $A \mid\!\approx C$. W. l. g., I assume that $P(A) \neq P(\bot)$: otherwise, $A \mid\!\approx B$ and $A \mid\!\approx C$ hold via Lemma 2.13.

1. $A \mid\!\approx B \wedge C$ — given
2. $P(B \wedge C/A) \in (\frac{1}{2}, 1]$ — from 1 by Definition 2.12
3. $\frac{P((B \wedge C) \wedge A)}{P(A)} \in (\frac{1}{2}, 1]$ — from 2 by Definition 2.11
4. $\frac{P((B \wedge C) \wedge A)}{P(A)} \leq \frac{P(A \wedge C)}{P(A)}$ — by truth-table calculations[13]
5. $\frac{P(A \wedge C)}{P(A)} \in (\frac{1}{2}, 1]$ — from 3 and 4 by math
6. $A \mid\!\approx C$ — from 5 by Definition 2.11
7. $\frac{P((B \wedge C) \wedge A)}{P(A)} \leq \frac{P(A \wedge B)}{P(A)}$ — by the above truth-table calculations
8. $\frac{P(A \wedge B)}{P(A)} \in (\frac{1}{2}, 1]$ — from 3 and 7 by math
9. $A \mid\!\approx B$ — from 8 by Definition 2.11

(M3) is $\mid\!\approx$-invalid. To show its invalidity, let me follow Douven and apply the probability law $P(B/A) + P(C/A) = P(B \wedge C/A) + P(B \vee C/A)$[14] because $A \mid\!\approx B$, $A \mid\!\approx C$, and $\vdash \overline{B \wedge C}$. Hence, $P(B \vee C/A) > 1$ which is absurd.

[13] In a joint truth-table for A, B, C, if one takes into account only the rows, where A is true, then the number of rows where $B \wedge (C \wedge A)$ is true does not exceed the number of rows where $A \wedge C$ is true.

[14] Note that I do not argue Douven's proof to be erroneous.

However, a modified version of (M3) which I label ($M3_{mdf}$) is $\approx$-valid:
($M3_{mdf}$) $\vdash \overline{B \wedge C}$, $A \rightarrowtriangle B$, then $A \rightarrowtriangle (B \vee C)$.[15]

A proof of ($M3_{mdf}$) is essentially the Douvenian one of (M3) (Douven, 2014: 271–272), who cites Milne, in turn (Minle, 2000: 316). W. l. g., I assume that $P(A) \neq P(\bot)$: otherwise, $A \rightarrowtriangle (B \vee C)$ holds via Lemma 2.13.

(MOD) is proven via $P(A/\overline{A}) = 0$ (Douven, 2014: 273) unless $P(A) \neq P(\top)$. In that case, however, $P(B/T) = 1$.

(RCE) is proven via $\models A \supset B \Leftrightarrow A \models B$ and Lemma 2.13.

(RCEA) and (RCEC) are proven via Lemma 2.13.

(REF) is proven via Lemma 2.15.

To prove (WAND) let me notice that $A \wedge \overline{C} \approx \bot \Leftrightarrow P(A \wedge \overline{C}) = P(\bot)$. There are three cases: (1) $P(A) = P(\bot)$, (2)$P(\neg C) = P(\bot)$, and (3) $P(A \wedge \overline{C}) = P(\bot)$, whilst neither (1), nor (2).

(1) implies $\bot \approx B \wedge C$ that holds, by Lemma 2.13. (2) implies $P(C) = P(\top)$. Hence, by **CPL**, (WAND) reduces to the trivially valid formulation: whenever $A \approx B$ and $\bot \approx \bot$, then $A \approx B$. At last, (3) implies $P(A) = P(C)$. Hence, by **CPL**, (WAND) reduces to the following formulation: whenever $A \approx B$ and $A \wedge \overline{A} \approx \bot$, then $A \approx (B \wedge A)$. By Lemma 2.13 and axiom 7 from Definition 2.4, it then reduces to the above-proven (ANT): whenever $A \approx B$, then $A \approx (A \wedge B)$.

(XOR) is provable by Douven (Douven, 2014: 276).

Now let me consider the two principles that Douven highlights: they, and only they, are both $\rightarrowtriangle_B$-valid and $\rightarrowtriangle_S$-invalid.

- ◇ (Contraposition) Whenever $A \rightarrowtriangle \overline{B}$, then $B \rightarrowtriangle \overline{A}$;
- ◇ (M1) Whenever $A \rightarrowtriangle B$ and $A \rightarrowtriangle C$, then $A \rightarrowtriangle (B \wedge C)$ or $A \rightarrowtriangle (B \vee C)$.

(Contraposition) fails for $\approx$, too: $\top \approx \overline{(p \wedge q)}$, but $p \wedge q \not\approx \neg\top$.[16]

However, (M1) is valid for $\approx$. In fact, its stronger version, which is labeled ($M1_{str}$) is already valid:
($M1_{str}$) whenever $A \approx C$, then $A \approx B \vee C$.[17]

1. $A \approx C$ — given
2. $A \not\approx B \vee C$ — assuming for the sake of contradiction
3. $P(A) \neq 0$ — from 2 by Definition 2.12

[15] Another variant of ($M3_{str}$) is $\vdash \overline{B \wedge C}$, $A \rightarrowtriangle C$, then $A \rightarrowtriangle (B \vee C)$. Its proof is analogous to the one below.

[16] $P(\overline{(p \wedge q)}/\top) = \frac{3}{4}$ and $P(\overline{\top}/p \wedge q) = 0$.

[17] Another variant contains $A \approx C$, and its proof is analogous to the one below.

4. $P(C) \neq 1$ — from 2 by Definition 2.12
5. $P(C/A) \in (\frac{1}{2}, 1]$ — from 1 by Definition 2.12
6. $P(B \vee C/A) < \frac{1}{2}$ — from 2 by Definition 2.12
7. $P(\overline{(\overline{B} \wedge \overline{C})}/A) < \frac{1}{2}$ — from 6 by **CPL**
8. $1 - P(\overline{B} \wedge \overline{C}/A) < \frac{1}{2}$ — from 3 and 7 by axiom 4 in Definition 2.4
9. $\frac{1}{2} < P(\overline{B} \wedge \overline{C}/A)$ — from 8 by math
10. $\frac{1}{2} < P(\overline{B}/\overline{C} \wedge A) \cdot P(\overline{C}/A)$ — from 9 by axiom 5 in Definition 2.4
11. $\frac{1}{2} < \big(1 - P(B/\overline{C} \wedge A)\big) \cdot \big(1 - P(C/A)\big)$ — from 3, 4, 10 by axiom 4 in Definition 2.4
12. $P(B/\overline{C} \wedge A) - P(B/\overline{C} \wedge A) \cdot P(C/A) < \frac{1}{2} - P(C/A)$ — from 11 by math
13. $x - xa < \frac{1}{2} - a$ — from 12 by obvious substitutions[18]
14. $x(1 - a) < \frac{1}{2} - a$ — from 13 by math
15. $x < \frac{\frac{1}{2} - a}{1 - a}$ — from 14 by math
16. $P(B/\overline{C} \wedge A) < 0$ — from 15 and 5
17. $P(B/\overline{C} \wedge A) \in [0, 1]$ — by Lemma 2.7

Last but not least, to check $\approx$-validity of the 20 remaining Douvenian principles, which are both $\twoheadrightarrow_B$-invalid and $\twoheadrightarrow_S$-invalid, deserves a separate paper.

3.2. THE HLOBILIAN SHOPPER'S GUIDE TO NONMONOTONIC LOGICS

In (Hlobil, 2018: 3), Hlobil *presents an exhaustive menu of nonmonotonic logics*:

> You cannot get a nonmonotonic logic without having to give up some principles that many find desirable. The good news is that you get a choice regarding which principles you want to give up. [...] I will go through some of these choices. The result will be an exhaustive (but not exclusive) classification of nonmonotonic logics into seventeen types.

It might be helpful to refer to that paper for details, especially to the tree of choices in Figure 1 (Hlobil, 2018: 5).[19] For the reasons of the current study, it would make sense to list Hlobil's choices and specify $\approx$ on them (for

[18] Note that $P(B/\overline{C} \wedge A)$ is an unknown in the inequality and hence denoted by x whilst by 5, $P(C/A)$ is a parameter and hence denoted by a.

[19] The full tree is not included here, but a direct reference to Hlobil's paper on the web has been provided in References below for convenience.

unifying reasons, I change the original notation and, if any, add [Douven's labels] from (Douven, 2014):[20]

- ⋄ (CO) If $A \in \Gamma$, then $\Gamma \approx A$;
- ⋄ (RE) [REF] $A \approx A$;[21]
- ⋄ (Mixed-Cut) If $\Gamma \approx A$ and $\Delta, A \approx B$, then $\Gamma, \Delta \approx B$;
- ⋄ (DDT) If $\Gamma \approx A \to B$, then $\Gamma, A \approx B$;
- ⋄ (CT) [CT] If $\Gamma \approx A$ and $\Gamma, A \approx B$, then $\Gamma \approx B$;
- ⋄ (PEM) $\Gamma \approx A \vee \neg A$;
- ⋄ (CM) [Cmon] If $\Gamma \approx A$ and $\Gamma \approx B$, then $\Gamma, B \approx A$;
- ⋄ (PF) [SDA] $\Gamma, A \vee B \approx C$ iff $\Gamma, A \approx C$ and $\Gamma, B \approx C$;
- ⋄ (DI) If $\Gamma, A \vee (B \wedge C) \approx D$, then $\Gamma, (A \vee B) \wedge (A \vee C) \approx D$;
- ⋄ (FU) $\Gamma, A \wedge B \approx C$ iff $\Gamma, A, B \approx C$.

Lemma 3.1. CO, RE, PEM, FU, and DI are the only $\approx$-valid principles from the above list.

Proof. For the invalidity of CT, CM, PF, and FU see (ibid.: 270, 268, 275–276, 266–267, respectively). The from-left-to-right part of DDT fails if $r \approx p \to q$ and $r, p \not\approx q$.[22] Mixed-Cut fails if $r, p \approx p \vee q$ and $r, p \vee q \approx q$, but $r, p \not\approx q$.[23] The $\approx$-validity of CO follows both from the fact that if $A \in \Gamma$, then $\Gamma \models A$, and Lemma 2.13. The $\approx$-validity of RE follows from Lemma 2.15. The $\approx$-validity of PEM follows from $\Gamma \models A \vee \neg A$ and Lemma 2.13. The $\approx$-validity of FU follows from the truth-table fact that its rows, where $A \wedge B$ is true, are the same, where both A and B are true. At last, the $\approx$-validity of DI follows from $P(A \vee (B \wedge C)) = P((A \vee B) \wedge (A \vee C))$. □

Lemma 3.1 indicates that $\approx$ is in 2 out of the 17 Hlobil types specified in the quote from the beginning of this subsection. In the tree in Figure 1 (Hlobil, 2018: 5),[24] the 3 branches having the node *rej–CO* are discarded for the reason that *CO* is $\approx$-valid. The 6 branches having the node *rej–PEM* are discarded for the reason that *PEM* is $\approx$-valid. Each branch having the end-node *rej–FU* or the end-node *rej–DI* is discarded for the reason that *FU* is $\approx$-valid or *DI* is $\approx$-valid, respectively, which leaves 4 branches in total. For the reason of the $\approx$-invalidity of *CM*,[25] the 2 branches having

[20] As in the case of Douven, I refer to the Hlobil paper for decoding the labels below.

[21] Note that in Douven's [REF], A is plausible.

[22] $P(p \to q/r) = \frac{3}{4}$ and $P(q/r, p) = \frac{1}{2}$.

[23] $P(p \vee q/r, p) = 1$, $P(q/r, p \vee q) = \frac{2}{3}$, and $P(q/r, p) = \frac{1}{2}$.

[24] See footnote 19 above.

[25] Note that Hlobil's choice branchings are not mutually exclusive: rejecting a principle does not inherently mean the rejection of its "counterpart".

the end-node *rej–PF* on the third level [C3] are discarded. As a result, only two types remain: the branch having the nodes *rej–MO*, *rej–Mixed-Cut*, *rej–CT*, *rej–CM*, *rej–PF* and the branch having the nodes *rej–MO*, *rej–Mixed-Cut*, *rej–DDT*, *rej–CM*, *rej–PF*. They are obviously reducible to the unique $\approx$-friendly type: the branch having the nodes *rej–MO*, *rej–Mixed-Cut*, *rej–CT*, *rej–DDT*, *rej–CM*, *rej–PF*.[26]

So, what is $\approx$ even if the Hlobil classification answers this question apagogically only? I answer this question by employing the four nonmonotonic logics that Hlobil mentions explicitly. The $\approx$-validity of *CO* implies it is not a relevance-like logic like **R** (Anderson & Belnap, 1975) or Hlobil's **NM-LR**.[27] The $\approx$-invalidity of *CT* implies it is not a cumulative logic like **KLM** (Kraus & Lehmann & Magidor, 1990). At last, the $\approx$-invalidity of *DI* implies it is not like Hlobil's **NM-G3cp**.[28] With regard to the *motivating choices* that Hlobil discusses (Hlobil, 2018: 6–7), $\approx$ prioritizes *staying supraclassical* over *rejecting as few structural principles as possible.* On the other hand, $\approx$ does not make a choice between rejecting principles regarding the behavior of connectives and rejecting structural principles: $\approx$ fails *PF* and *DDT* and hence $\to$ and $\vee$ do not behave properly (but it is not the case for $\wedge$ because *FU* is $\approx$-valid) as well as $\approx$ anticipatedly fails *Mixed-Cut* to avoid the supraclassical trivialization discussed at the end of Section 2. As a result, $\approx$ is not suitable for *inferentialism-friendly* nonmonotonic logics such as **NM-G3cp** and **NM-LR** which are Hlobil's favorite kind of nonmonotonic logic.[29]

3.3. COBREROS ET AL.'S PRAGMATIC-TO-TOLERANT ENTAILMENT FOR TOLERANT REASONING

Furthermore, it would be beneficial for this research to turn to Cobreros et al.'s *nonmonotonic and/or nontransitive* approaches to tolerant reasoning

[26] According to Hlobil (Hlobil, 2018: 5), "Figure 1 should be read as follows: Every nonmonotonic logic must reject all the principles that occur on at least one complete branch of the tree. [...] Of course, a logic can always reject more principles than what the tree in Figure 1 requires. Hence, a logic can belong to several of my seventeen types".

[27] The latter is a nonmonotonic variant of Bimbo's distribution-free relevance logic **LR** (Bimbo, 2015).

[28] **NM-G3cp** is a nonmonotonic variant of Troelstra and Schwichtenberg's classical sequent calculus **G3cp** in which the structural rule of *weakening* is absorbed (Troelstra, Schwichtenberg, 2000). Hlobil also employs the name **G3cp-NM**.

[29] He proposes sequent-style axiomatizations for two of them (Hlobil, 2018: 11, 13) which preserve Makinson's result that nonmonotonic entailment relation is not closed under substitution (Makinson, 2003).

(Cobreros & Egré & Ripley, 2021: 682) which (appropriate for the topic of this paper) are aimed at disproving the following thesis (the *italic* is original):

> According to one influential view of the sorites paradox, the tolerance principle — the constraint whereby if someone is tall, for example, then someone whose height is imperceptibly shorter is tall too — is an *unsound* rule of reasoning (see Williamson, 1994).

To that end, the authors propose three specific consequence relations, with one of them (called *pragmatic-to-tolerant* consequence and denoted by $\models^{prt}$ as well as its updated version called *Pragmatic-to-tolerant* consequence and denoted by $\models^{Prt}$) being reflexive, contractive, nonmonotonic, and nontransitive, i. e., it is closer to $\approx$ than the other two. Despite the fact that analyzing Cobreros et al.'s approaches exceeds the scope of this paper and the fact that formulating the tolerance principle needs a first-order language, it may well be prospective for further applications.

It is not the case that all classically valid modes of reasoning are $\models^{Prt}$-valid as in the case of $\approx$.[30] The reason to update $\models^{prt}$ to $\models^{Prt}$ is a flaw of the former in that it separates the premisses and their conjunction, i. e., $p, \neg p \models^{prt} q$, but $p \wedge \neg p \not\models^{prt} q$.[31] With regard to $\approx$, it is simple to confirm that the target separation is not the case for $\approx$: $A, \neg A \approx B$ iff $A \wedge \neg A \approx B$, due to truth-table calculations. I highlight two *Prt*-features. The first one is $A, \neg A \not\models^{Prt} B$ which is $\approx$-valid, on the other hand, because the explosion is $\models$-valid. The second *Prt*-feature — the standard $\wedge$-elimination entailments $A \wedge B \models^{Prt} A$ and $A \wedge B \models^{Prt} B$ — holds for $\approx$, too.

4. RELATED WORK

In this section, the main focus lies with related papers that share the approach of the present study in that all the 2^n truth table distributions for a formula containing n distinct propositional variables are equiprobable. Hence, a classic approach by Carnap, for example, is beyond the scope of this section (Carnap, 1962). Section 4.1 considers the approach by Bocharov and Markin (Bocharov & Markin, 2008) who avoid employing the principle of *reverse deduction* which is considered together with two approaches that apply it in Section 4.2. Section 4.2.1 considers the approach by Voishvillo and Degtyarev (Voishvillo & Degtyarev, 2001) whilst Section 4.2.2 considers

[30] Note that another entailment proposed there under the name of *s*trict-to-*t*olerant entailment validates all classical modes of reasoning.

[31] It is not a $\models^{prt}$-specific feature. For example, see Weir's nontransitive trivalent logic of *n*eo-*c*lassical entailment NC_3 (Weir, 2013). And this feature troubles him in no way at all.

Ivlev's method (Ivlev, 2008; 2015). This Section ends with a summary table that contains a comparative analysis of the classical and the four probabilistic entailments of this paper. While exposing the approaches in question, I unify the original notations.

4.1. AN APPROACH THAT IS NOT BASED ON THE PRINCIPLE OF REVERSE DEDUCTION

Definition 4.1. (Bocharov & Markin, 2008: 450–451) Let $n, n > 0$, be the total number of rows in a truth table for A and let $m, m \leq n, m \geq 0$, be the number of rows in this truth table, where A is true. Then an unary probability P^* of A is determined as follows:

$$P(A) = \frac{m}{n}.$$

It is clear that

Lemma 4.2. Definition 4.1 meets the provisions of Definition 2.2.

Definition 4.3. (ibid.: 451) A binary probability of B given $A_1, \ldots, A_k$, $k \neq 0$ (denoted by $P^*(B/\bigwedge_{i=1}^{k} A_i)$), is determined via a *joint* truth table of $A_1, \ldots, A_k, B$ as follows, where $P(\bigwedge_{i=1}^{k} A_i) \neq 0$:

$$P(B/\bigwedge_{i=1}^{k} A_i) = \frac{P((\bigwedge_{i=1}^{k} A_i) \wedge B)}{P(\bigwedge_{i=1}^{k} A_i)}.$$

It is clear that

Lemma 4.4. Definition 4.3 meets the provisions of Definition 2.4.

A probabilistic entailment relation between Γ and B (denoted by $\Gamma \approx^* B$) is defined as follows.

Definition 4.5. (ibid.: 452) $\Gamma \approx^* B$ iff $P(B) < P(B/\bigwedge_{i=1}^{k} A_i)$.

Let me move on to the investigation of the properties of $\approx^*$. In contrast to $\approx$ (Lemma 2.13), $\approx^*$ is not a generalization of $\models$:

Lemma 4.6. There are Γ, B such that $\Gamma \approx^* B$ and $\Gamma \not\models B$. On the other hand, there are Γ, B such that $\Gamma \not\approx^* B$ and $\Gamma \models B$. In particular, there is no A such that $\approx^* A$.

Proof. A proof of Corollary 2.13 contains an example that proves the first part of this Lemma. $p \models \top$ and $p \not\approx^* \top$ prove its second part. At last, $\not\approx^* B$, for any B, follows from the nonemptiness of Γ in Definition 4.5. $\square$

Lemma 4.7. $\approx^*$ is symmetric, contractive, and permutative, i. e., if $A \approx^* B$, then $B \approx^* A$; if $A, A \approx^* B$, then $A \approx^* B$; if $A, C \approx^* B$, then $C, A \approx^* B$, respectively.

$\approx^*$ is not reflexive, monotonic, and transitive, i. e. $A \not\approx^* A$; if $A \approx^* B$, then $A, C \not\approx^* B$; if $A \approx^* B$ and $B \approx^* C$, then $A \not\approx^* C$, respectively.

Proof. The first part of this Lemma derives straight from Definition 4.5. As regards the second part of this Lemma, proof of Lemma 2.15 contains the respective examples that prove the lack of both monotonicity and transitivity. The lack of reflexivity follows from $\top \not\approx^* \top$ or $\bot \not\approx^* \bot$. □

Despite $\approx^*$ not being reflexive, I highlight the following easy-provable (see also Definition 2.1 above)

Lemma 4.8. $A \approx^* A$ iff A is plausible.

Lemma 4.9. $\approx^*$ is inconsistent, i. e., it is not the case that $A \approx^* B$ and $A \approx^* \neg B$.

Proof. (On contrary)

1. $A \approx^* B$ and $A \approx^* \neg B$ – given
2. $A \approx^* B$ – from 1
3. $A \approx^* \neg B$ – from 1
4. $P(B) < \frac{P(A \wedge B)}{P(A)}$ – from 2 by Definition 4.5
5. $P(\neg B) < \frac{P(A \wedge \neg B)}{P(A)}$ – from 3 by Definition 4.5
6. $1 - P(B) < \frac{P(A \wedge \neg B)}{P(A)}$ – from 5 by Remark 2.3
7. $P(B) * P(A) < P(A \wedge B)$ – from 4
8. $P(A) - P(B) * P(A) < P(A \wedge \neg B)$ – from 6
9. $P(A) - P(A \wedge \neg B) < P(B) * P(A)$ – from 8
10. $P(A) - P(A \wedge \neg B) < P(A \wedge B)$ – from 7, 9
11. $P(A) < P(A \wedge B) + P(A \wedge \neg B)$ – from 10
12. $P(A) < P(A)$ – from 11 by axiom 3 in Definition 2.2
13. it is not the case that $A \approx^* B$ and $A \approx^* \neg B$ – from 12 □

Due to Definition 4.5, A in Lemma 4.9 is readily generalized to Γ. This is avoided for the sake of simplicity.

Another key thing worth noting is the form of inconsistency in Lemma 4.9 which is stronger than the $\approx$-one discussed on pages 223, 225 above.

4.2. TWO APPROACHES BASED ON THE PRINCIPLE OF REVERSE DEDUCTION

Next, two mutually related approaches to plausible entailment that are based on the so-called *principle of reverse deduction* will be analyzed: see quotes on pages 233 and 233 below.

4.2.1. Voishvillo and Degtyarev's Approach. Voishvillo and Degtyarev put it as follows (the quote is changed cosmetically):[32]

> It is essential to pay attention to the fact that if $B \models A$(A deductively follows from B), then $A \approx_{VD} B$. The opposite is not true, though. This way of establishing inductive entailment between A and B on the basis of deductive entailment between B and A is said to be the principle of *reverse deduction.* Additionally, for the relation of deductive entailment that is under consideration here, one excludes paradoxical cases of the relation [...], i.e., the cases when A is a negation of some logical law of the system under consideration or when B is some logical law [...] (Voishvillo & Degtyarev, 2001: 389).

On the previous page, they propose the following definition of plausible entailment, which they call *inductive entailment*:

Definition 4.10. (ibid.: 388) $A \approx_{VD} B$ iff $B \not\models A$ and $P(B) < P(B/A)$, where A, B are plausible.

Regretfully, there is a contradiction between the fact that "if $B \models A$ (A deductively follows from B), then $A \approx_{VD} B$" mentioned in the quote on page 233 and Definition 4.10. Due to the former, $p \approx_{VD} p$, due to $p \models p$ whilst due to the latter, $p \not\approx_{VD} p$.

Even if one considers Definition 4.10 rather than the fact under question to be the proper source that explicates their account on plausible entailment, then one cannot consider it satisfactory still. Definition 4.10 implies $\approx_{VD}$ to be irreflexive, i.e., $A \not\approx_{VD} A$, for any A. This property of plausible entailment is very unlikely to have some philosophical background, let alone that Voishvillo and Degtyarev never mention it explicitly. Notice also that according to Lemma 4.8, $\approx^*$ is not reflexive rather than irreflexive, i.e., $A \approx^* A$, for any plausible A.

4.2.2. Ivlev's Approach. Ivlev's approach is slightly different from the one by Voishvillo and Degtyarev (Ivlev, 2015: 94); the quote is changed cosmetically:

> *Reverse deduction* is as follows. One needs to justify a sentence A. One establishes that each sentence $B_1, B_2, \ldots, B_n (n \geq 1)$ follows from A or, equivalently,

[32] All the translations below belong to me.

a conjunction of these sentences follows from it. Additionally, A is not contradictory, whilst $B_1, B_2, \ldots, B_n$ are not valid. One concludes that the sentences $B_1, B_2, \ldots, B_n$ support the sentence A, i. e.

$$\frac{A \models B_1 \wedge B_2 \wedge \ldots \wedge B_n, \not\models \neg A, \quad \not\models B_1, \not\models B_2, \ldots, \not\models B_n}{B_1, B_2, \ldots, B_n \approx_* A} \tag{4.1}$$

In other words, Ivlev accepts the restriction for A, B to be plausible.

Thus defined, $\approx_*$ needs an auxiliary condition nevertheless in order to avoid undesired plausible entailments. For example, $p \approx_* p \wedge q$, due to $p \wedge q \models p$. However, the former entailment states that p supports its conjunction with an arbitrary sentence, which one could hardly accept. Moreover, if one generalizes the previous example by conjuncting p and a conjunction of n arbitrary sentences, then $p \approx_* p \wedge (q_1 \wedge \ldots \wedge q_n) \ldots)$, due to $p \wedge (q_1 \wedge \ldots \wedge q_n) \ldots) \models p$, etc. And this example states the absurdity that p supports any conjunction consisting of it and n arbitrary sentences.

To this end, one additionally imposes the condition which one calls *positive relevancy* and which is nothing but the right-side condition of Definition 4.5: $P(A) < P(A/B)$.[33] As a result, one obtains $\approx_*$ to be *quite the same* as $\approx^*$, where two differences need to be highlighted.

$\approx_*$ might be determined in the same way as $\approx^*$ and it can be determined differently by employing the machinery of crossing out formulae in a truth table while calculating conditional probability (both approaches in question determine an unconditional probability in the same way).

I take an example of calculations from (Ivlev, 2015: 95) and apply the needed changes.

$p \vee q$	$p \wedge q$
~~F~~	~~F~~
T	F
T	F
T	T

Table 1. Truth table with crossing out for $p \vee q$, $p \wedge q$

$p \vee q$ supports $p \wedge q$, due to the fact $p \wedge q$ classically implies $p \vee q$. The conditional probability of $p \wedge q$ is determined as follows. One establishes a probability of the proposition in question given the truth of the proposition $p \vee q$, i. e., one establishes the degree of support of the initial proposition by the proposition $p \vee q$. One builds up joint truth tables for these propositions: see Table 1.

One crosses out those rows where the proposition $p \vee q$ is false, i. e., one is presupposed to have received the information about the truth of $p \vee q$: see Table 1.

The probability of the sentence $p \wedge q/p \vee q = \frac{1}{3}$. Notation: $P(p \wedge q/p \vee q)$. (It reads: the probability of $p \wedge q$ given $p \vee q$.)

[33] I repeat the mantra on the possibility of generalizing A to Γ.

The second difference between $\approx_*$ and $\approx^*$ is that the former is reflexive.

Lemma 4.11. $\approx_*$ is reflexive, contractive, permutative, symmetrical and neither transitive nor monotonic.

Proof. $A \approx A$ follows from $P(A/A) = 1$. The other properties are proven in Lemma 4.7. □

Lemma 4.12. $\approx^*$ is inconsistent, i. e., it is not the case that $A \approx^* B$ and $A \approx^* \neg B$.

Proof. It is analogous to the one in Lemma 4.9. □

With regard to $\approx^*$ and $\approx_*$, their equality is established with the following

Lemma 4.13. $A \approx_* B$ iff $A \approx^* B$.

Proof. It is obvious in the case from left to right. The case from right to left holds because $A \not\approx^* B$, if A or B are not plausible: (1) if A is $\bot$, then $P(A) = 0$; (2) if A is $\top$, then $P(B) = P(B/\top)$; (3) if B is $\bot$, then $P(B) = P(A/B) = 0$; (4) if B is $\top$, then $P(B) = P(A/B) = 1$. □

To summarize, a comparative analysis of the four probabilistic entailments discussed in the paper: $\approx$, $\approx^*$, $\approx_{VD}$, and $\approx_*$ in sections 2, 4.1, 4.2.1 and 4.2.2, respectively, is presented in Table 2, below.

The following properties hold for each probabilistic entailment in question: permutation, contraction, and the lack of both transitivity and monotonicity.

	INCONS	CONPRIM	SUP	DEF	REF	SYM
$\models$	–	inapp	–	inapp	+	–
$\approx$	–	+	+	$\frac{1}{2} < P(B/A) \leq 1$	+	–
$\approx^*$	+	–	–	$P(B) < P(B/A)^\blacklozenge$	–	+
$\approx_{VD}$	co	nt	ra	di	ct	ion
$\approx_*$	+	+	–	$P(B) < P(B/A)^\lozenge$	+	+

Table 2. A comparison of $\models$, $\approx$, $\approx^*$, $\approx_{VD}$, $\approx_*$

CONS, CONPRIM, SUP, DEF, REF, SYM mean a strong form of inconsistency, primitiveness of a conditional probability, supraclassicality, definition of an entailment, reflexivity, symmetricity, respectively, whilst + and – mean the fact an entailment holds or does not hold this property; at last, "inapp" means inapplicable.[34] The entries of the $\approx_{VD}$-row are filled with "contradiction" (see page 233). The conditions $\blacklozenge$ and $\lozenge$ mean A is not

[34]An entailment holds the property of the primitiveness of a conditional probability iff it does not have the property of the primitiveness of an unconditional probability. The analogous equivalence is true with respect to the strong vs. the weak forms of inconsistency.

$\perp$ and A, B are plausible, accordingly. As before, the definitions of these entailments are for the particular case, due to simplicity reasons.

5. CONCLUSION

In the paper, a nontrivial plausible probabilistic entailment relation is proposed. Its original feature is a combination of the primitiveness of a conditional probability, which one calculates with the method of truth tables for **CPL**, and supraclassicality. Moreover, a comparison with some closely related probabilistic entailments is provided, along with a position on certain nomenclatures in related literature. The main topic for future research on the surface is to set up proof-theoretic axiomatizations of each consistent entailment discussed in the paper as well as to continue checking the $\approx$-validity of the other Douven properties to whom Section 3.1 is devoted.

REFERENCES

Anderson, A. R., and N. D Belnap. 1975. *Entailment: The Logic of Relevance and Necessity*. Princeton: Princeton University Press.

Bimbo, K. 2015. *Proof Theory: Sequent Calculi and Related Formalisms*. Boca Raton: CRC Press.

Bocharov, V. A., and V. I. Markin. 2008. *Vvedeniye v logiku [Introduction to Logic]* [in Russian]. Moskva [Moscow]: ID Forum / INFRA-M.

Carnap, R. 1962. *Logical Foundations of Probability*. 2nd ed. Chicago: University of Chicago Press.

Cobreros, P., P. Egré, and D. Ripley. 2021. "Tolerant Reasoning: Nontransitive or Nonmonotonic?" *Synthese* 199 (3): 681–705.

Douven, I. 2011. "Further Results on the Intransitivity of Evidential Support." *Review of Symbolic Logic*, no. 4, 487–497.

——— . 2014. "Reasoning About Evidence." *Journal of Applied Logic*, no. 12, 263–278.

Hlobil, U. 2018. "Choosing Your Nonmonotonic Logic: A Shopper's Guide." PhilPapers. Accessed Oct. 21, 2023. `https://philpapers.org/go.pl?aid=HLOCYN`.

Horn, K. van. 2017. "From Propositional Logic to Plausible Reasoning: A Uniqueness Theorem." *International Journal of Approximate Reasoning*, no. 88, 309–332.

Ivlev, Yu. V. 2008. *Logika. Uchebnik [Logic. Textbook]* [in Russian]. Moskva [Moscow]: Prospekt.

——— . 2015. *Logika dlya yuristov [Logic for Law Students]* [in Russian]. Moskva [Moscow]: Prospekt.

Keynes, J. M. 1921. *Treatise on Probability*. London: Macmillan & Co.

Kolmogorov, A. 1956. *Foundations of the Theory of Probability*. New York: Chelsea.

Kraus, S., D. Lehmann, and M. Magidor. 1990. "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics." *Artificial Intelligence*, no. 44, 167–207.

Leblanc, H. 1983. "Alternatives to Standard First-Order Semantics." In *Handbook of Philosophical Logic*, ed. by D. M. Gabbay and F. Guenthner, 1:189–274. Dordrecht: Reidel.

Lorenz, D., B. Kooi, and J. Sack. 2019. "Logic and Probability." Ed. by E. N. Zalta. The Stanford Encyclopedia of Philosophy. Accessed June 7, 2023. https://plato.stanford.edu/archives/sum2019/entries/logic-probability/.

Makinson, D. 2003. "Bridges Between Classical and Nonmonotonic Logic." Research Gate. Accessed June 7, 2023. https://www.researchgate.net/publication/30967223_Bridges_between_Classical_and_Nonmonotonic_Logic.

Mendelson, E. 1997. *Introduction to Mathematical Logic*. New York: Chapman & Hall.

Milne, P. 2000. "Is There a Logic of Confirmation Transfer?" *Erkenntnis*, no. 53, 309–335.

Popper, K. 1955. "Two Autonomous Axiom Systems for the Calculus of Probabilities." *British Journal of the Philosophy of Science*, no. 6, 51–57.

von Wright, G. H. 1969. "Wittgenstein's Views on Probability." *Revue Internationale de Philosophie* 23:259–279.

Voyshvillo, Ye. K., and M. G. Degtyarev. 2001. *Logika. Uchebnik dlya studentov [Logic. Textbook for Students]* [in Russian]. Moskva [Moscow]: Vlados-press.

Weir, A. 2013. "A Robust Non-Transitive Logic." *Topoi* 34 (1): 99–107.

Williamson, T. 1994. *Vagueness*. London: Routledge.

Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus*. London: Kegan Paul / Trench / Trubner & Co.

Shangin V. O. [Шангин В. О.] A Supraclassical Probabilistic Entailment Relation [Суперклассическое вероятностное отношение следования] // Философия. Журнал Высшей школы экономики. — 2023. — Т. 7, № 4. — P. 215–238.

Василий Шангин
к. филос. н., доцент
Московский государственный университет имени М. В. Ломоносова (Москва)
Институт логики, когнитологии и развития личности (Москва);
ORCID: 0000–0002–7380–3213

Суперклассическое вероятностное отношение следования

Получено: 07.06.2023. **Рецензировано:** 20.09.2023. **Принято:** 30.09.2023.

Аннотация: В статье задается оригинальное суперклассическое нетривиальное правдоподобное отношение следования $\approx$, основанное на колмогоровской теории вероятностей.

Его важной чертой является примитивность условной вероятности, которая вычисляется с помощью метода таблиц истинности для классической логики высказываний. Мы изучаем свойства заданного отношения. В частности, мы показываем, что, будучи суперклассическим, т.е. все классически общезначимые формулы и выводимости имеют место для $\approx$, но обратное утверждение неверно, оно нетривиально и для него имеет место такой же вариант свойства непротиворечивости, что и для классического следования $\models$. мы определяем место предложенного следования в некоторых классификациях, найденных в соответствующей литературе. В частности, мы используем дювеноский анализ некоторых вероятностных отношений следования, содержащий десятки свойств, которые являются важными для любого вероятностного отношения следования, а также хлобиловский *руководитель покупателя при выборе своего немонотонного отношения следования*, благодаря немонотонности $\approx$, и предложенные Кобреросом, Эгром, Рипли и ван Руем отношения следования для толерантных рассуждений. Наконец, мы делаем сравнительный анализ классического, предложенного и некоторых тесно связанных со вторым отношений следования: то, что предложено В. А .Бочаровым, В. И. Маркиным, то, что предложено Е. К. Войшвилло, М. Г. Дегтяревым, а также то, что предложено Ю. В. Ивлевым, где два последних отношения основаны на так называемом принципе *обратной дедукции*, который является интуитивно приемлемым способом, который позволяет связать классическое и вероятностные отношения следования.

Ключевые слова: классическая логика, вероятностная логика, Байес, свидетельство, следование, обратная дедукция.

DOI: 10.17323/2587–8719–2023–4–215–238.