



Article

Probabilistic Justification Logic

Joseph Lurie 

Department of Philosophy, University of Connecticut, Storrs, CT 06269, USA; joseph.lurie@uconn.edu;
Tel.: +1-201-527-7252

Received: 30 November 2017; Accepted: 13 February 2018; Published: 16 February 2018

Abstract: Justification logics are constructive analogues of modal logics. They are often used as epistemic logics, particularly as models of evidentialist justification. However, in this role, justification (and modal) logics are defective insofar as they represent justification with a necessity-like operator, whereas actual evidentialist justification is usually probabilistic. This paper first examines and rejects extant candidates for solving this problem: Milnikel’s Logic of Uncertain Justifications, Ghari’s Hájek–Pavelka–Style Justification Logics and a version of probabilistic justification logic developed by Kokkinis et al. It then proposes a new solution to the problem in the form of a justification logic that incorporates the essential features of both a fuzzy logic and a probabilistic logic.

Keywords: logic; justification; epistemology; evidentialism; justification logic; fuzzy logic; probabilistic logic; epistemic logic

1. Introduction

1.1. Justification Logic

Justification logics are constructive analogues of modal logics. Syntactically, justification logics are built from the following constituents:

- Propositional constants: p, q , etc.
- Boolean propositional connectives: \neg, \supset , etc.
- Proof constants: a, b, c , etc.
- Proof variables: x, y , etc.
- Operators ranging over proof polynomials: $!, \cdot$ and $+$.
- A special proof operator, denoted with a colon ($:$), which takes a proof polynomial as its left input, a proposition as its right input and outputs a proposition.

Additionally, the symbols t, s , etc., shall be used as schematic letters representing proof polynomials, and φ, ψ , etc., shall be used as schematic letters representing propositional formulae. Furthermore, we adopt the convention that formulae of justification logics will be written with the $:$ operator taking the narrowest possible scope and the \supset operator taking the widest possible scope, so as to minimize the need for parentheses.

The phrase “proof polynomials”, as used in the above specification, denotes the class of grammatical terms that can be formed using the proof constants and variables: formally, this can be defined as follows:

Definition 1 (Proof polynomials). *The set of proof polynomials (PP) is defined by the following:*

- All proof variables x, y , etc., are members of PP.
- All proof constants a, b, c , etc., are members of PP.
- If $s, t \in PP$, then $s \cdot t \in PP$, $s + t \in PP$, and $!s \in PP$.

- Nothing else is a member of PP.

The first justification logic was the system LP,¹ which is an analogue of the well-known modal system S4 of Lewis and Langford [2]. LP was developed by Artemov [3] for the purpose of providing a constructive theory of provability for intuitionistic logic. The canonical presentation of LP is an axiomatic system, as follows:

Definition 2. The system LP is the closure of the axiom schemas PC, K^j, T^j and 4^j under the inference rules modus ponens and simple axiom justification.² These axiom schemas and inference rules are in turn defined as follows:

PC Any set of axiom schemas whose closure under modus ponens is sound and complete with respect to classical propositional logic

K^j $\vdash t : (\varphi \supset \psi) \supset (s : \varphi \supset (t \cdot s) : \psi)$. This axiom is more commonly referred to as the axiom of application.

T^j $\vdash t : \varphi \supset \varphi$. This is called the axiom of reflection.

4^j $\vdash t : \varphi \supset !t : \varphi$. This is sometimes called the axiom of positive introspection, or the proof checker axiom.

Modus ponens If $\vdash \psi \supset \varphi$ and $\vdash \psi$, then $\vdash \varphi$.

Simple axiom justification If φ is an instance of one of the axiom schemas included in a particular system, then there is a proof constant c such that $\vdash c : \varphi$.

A simple model theory for LP has been provided by Mkrtychev [4]. There is also an alternative model theory, given by Fitting [5], which is essentially an embedding of the Mkrtychev model theory into the Kripke frame semantics that is typically used to model modal logics. Many authors prefer the Fitting model theory because it more clearly emphasizes the relationship between justification and modal logics. However, for the purposes of this paper, it is more convenient to use the simpler Mkrtychev theory, which can be defined as follows:

Definition 3. A Mkrtychev model of the system LP consists of an evidence relation \mathcal{E} for that system (as defined below) plus a valuation function v from the set of sentences to $\{0, 1\}$ that satisfies the following:

- $v(\neg\varphi) = 1$ iff $v(\varphi) = 0$.
- $v(\psi \supset \varphi) = 1$ iff $v(\psi) = 0$ or $v(\varphi) = 1$.
- $v(t : \varphi) = 1$ iff $\mathcal{E}(t, \varphi)$.
- If there is any t such that $\mathcal{E}(t, \varphi)$, then $v(\varphi) = 1$.

Definition 4. An evidence relation for the system LP is a relation \mathcal{E} between the set of proof polynomials and the set of sentences, which satisfies the following:

- If φ is an instance of any axiom schema in the axiomatic presentation of LP given above, then there is a proof constant c such that $\mathcal{E}(c, \varphi)$.
- For all proof polynomials s, t and sentences φ, ψ , if $\mathcal{E}(s, \varphi \supset \psi)$ and $\mathcal{E}(t, \varphi)$, then $\mathcal{E}(s \cdot t, \psi)$.
- If $\mathcal{E}(t, \varphi)$, then $\mathcal{E}(!t, t : \varphi)$ for all proof polynomials t and sentences φ .

Definition 5 (Consequence for Mkrtychev models). $\Gamma \vDash_{LP} \varphi$ iff for every Mkrtychev model of LP such that $v(\gamma) = 1$ for every $\gamma \in \Gamma$, it is also the case that $v(\varphi) = 1$.

As mentioned above, LP is the justification analogue of S4. Justification analogues of other normal modal logics are well known; Artemov [6] provides a good overview of how to construct these systems.

¹ The abbreviation LP, in this context, stands for “logic of proofs”. It is important that Artemov’s system LP not be confused with the LP of Priest [1], which abbreviates “logic of paradox”.

² The sum schemas, $\vdash s : \varphi \supset (s + t) : \varphi$ and $\vdash t : \varphi \supset (s + t) : \varphi$, may also be conservatively included in LP, but will not be required for any of the applications considered in this paper; they are included in the model theory of Section 3 for the convenience of the reader who wishes to use them, but may be omitted there without harm.

There are also justification analogues of normal justification logics that contain axioms not found in LP; these are created by extending the justification logic syntax with additional operators on proof polynomials, which are then used to formulate justification analogues of the required modal axioms. Fitting [7] gives details of several justification logics of this sort. Furthermore, some research has been conducted into justification versions of non-normal and non-classical modal logics; see Lurie [8] for details on those systems. The primary result of this paper will be presented as a modification of LP, but nothing will crucially depend on the choice of that system. The same changes that are used to transform LP into an analogue of any other desired modal logic can be applied directly to the system presented in Section 3 below.

1.2. Epistemology

The most prominent application that has been proposed for justification logic is an epistemic application, as a model of the concepts of knowledge and/or justification.³ The major motivation underlying this usage is that many accounts of these concepts (especially the popular “evidentialist” accounts of epistemic justification) are fundamentally constructivist and thus are more perspicuously modeled by justification logics than by non-constructive modal logics. In addition, Artemov [6,9] argues that justification logics provide expressive resources that can allow for better formal resolution of epistemic puzzles such as the fake barn examples of Goldman [10] and Kripke [11].

However, if the purpose of employing justification logic is to create a perspicuous model of real epistemic situations, particularly as understood under an evidentialist theory of justification, then there is a sense in which all the justification systems mentioned in the previous section miss the mark. Evidence rarely guarantees the truth of a proposition, but rather provides a certain probability that the proposition is true. One is generally considered justified in believing a proposition if the probability given by the available evidence exceeds a certain epistemic threshold. This understanding of evidence is widely accepted, but the question of how this threshold is determined has spawned a major debate in epistemology, with at least three major competing views: contextualism (the view of DeRose [12]), which claims that the threshold is set by the conversational context of an utterance of “I know that p”; subject-sensitive invariantism (the view of Stanley [13]), which claims that conversational contexts do not affect the epistemic threshold, but that the threshold does vary depending on the real-world situation of an epistemic agent; and pure invariantism (the view of Williamson [14]), which fixes a single epistemic threshold for all agents in all epistemic circumstances. In addition to the major three, there are other epistemic theories that seek to accomplish the same goal using a different sort of framework, such as the relativist theory of MacFarlane [15]. In this paper, I do not advocate any of these views, but rather I seek to devise a logical model of the proposition-evidence relation that can be applied to any of them.⁴

1.3. Probability Theory and Fuzzy Logic

The mathematical tools required to properly model this non-deterministic notion of evidence are probability theory and fuzzy logic. Probability theory is a mathematical model of uncertainty. Fuzzy logic is a mathematical model of vagueness. Giangiacomo Gerla, the fuzzy logician who has done the most work investigating probabilistic logics, emphasizes the distinction between these topics in [19] using the example sentence, “The rose on the table is red”. In a case where you can see the rose on the table and its color is something intermediate between red and pink, the truth of the sentence is vague but not uncertain. In a case where you are shown two roses, one determinately red and one determinately not-red, and then one of the roses is placed on a table out of your sight, the truth of the

³ Indeed, the name “justification logic” was chosen for this class of formal systems precisely because of the epistemic application.

⁴ There is also a significant class of theories addressing the same general epistemic question which cannot be modeled by the justification logic presented in this paper. Notable examples include the contrastivist theory of Schaffer [16] and the older relevant alternatives theories of Dretske [17,18] and Goldman [10].

sentence is uncertain but not vague. However, when epistemic operators are added to the language, this sharp distinction begins to collapse. If I have a small degree of uncertainty concerning the truth of proposition p , this creates vagueness in the truth of the proposition that I know that p . If the proposition p were vague as well as uncertain, this would presumably give an even weaker degree of truth to the knowledge claim, but it is unclear how those two sources of vagueness should interact. For the purposes of this paper, let us assume that the only vagueness present is that which is caused by uncertainty in the propositions to which epistemic operators are applied. The logic developed under this assumption could be used directly by those who advocate a reductive account of vagueness, or it could perhaps be modified to account for the separate phenomena of vagueness and uncertainty. I shall take no stand here as to which application is to be preferred, but the latter must be postponed to another paper.

In Section 3, I will set forth a justification logic based on this assumption; a probabilistic fuzzy logic similar to Gerla's. However, first, let us examine competing proposals for logically modeling the sorts of epistemic applications discussed in Section 1.2.

2. Previous Justification Logic Approaches to the Vagueness of Epistemic Justification

2.1. Milnikel's Logic of Uncertain Justifications

Definition 6. *The Logic of Uncertain Justifications, J^U , is formed by augmenting the axiomatic system of classical propositional logic with a family of additional binary operators $:_r$, where r is a rational number in the interval $(0, 1]$. All of these operators shall take proof polynomials as left inputs and formulae as right inputs and shall be governed by the following axioms and inference rules:*

Application $\vdash s :_p (\varphi \supset \psi) \supset (t :_q \varphi \supset (s \cdot t) :_{pq} \psi)$

Monotonicity₁ $\vdash s :_r \varphi \supset (s + t) :_r \varphi$

Monotonicity₂ $\vdash t :_r \varphi \supset (s + t) :_r \varphi$

Confidence weakening *If $p \leq q$, then $\vdash t :_q \varphi \supset t :_p \varphi$*

Iterated axiom justification *If φ is a substitution instance of any schema listed above, or of any axiom schema of the chosen axiomatization of classical propositional logic, then we may select an arbitrary sequence of proof constants c_1, \dots, c_n and infer $\vdash c_n :_1 c_{n-1} :_1 \dots c_1 :_1 \varphi$*

The intended interpretation of a formula $t :_r \varphi$ in J^U is something along the lines of "Evidence t provides justification for believing φ with at least confidence level r ". With this interpretation, J^U has the resources to at least express the sort of epistemic uncertainty that was set forth in the introduction to this paper. Indeed, J^U appears to be the simplest possible logic possessing those expressive resources. Additionally, it has the advantage of being fully expressible as an axiomatic system; Milnikel also provides a model theory and the associated strong completeness proof. By contrast, most forms of fuzzy logic (including the probabilistic fuzzy logic that will be the basis of my probabilistic justification logic in Section 3) are only axiomatizable with weak completeness (or some intermediate completeness property; compare Pavelka [20]). Having weak but not strong completeness means that although an axiomatization can be used for certain metatheoretic tasks, it does not contain all of the information that is found in the model theory; in particular, axiomatizations of fuzzy logics do not have any feature corresponding to the "fuzziness" of the truth values, but rather differ from axiomatic classical logic only in omitting those theorems that are invalidated by fuzzy counterexamples. In an Artemov-style justification logic, the axiomatic system is taken to be the canonical presentation of the logic, and alternative presentations such as model theories or sequent calculi are derived from the axiomatic system. For example, the definition of the Mkrtychev model theory in Section 1.1 includes a clause that directly refers to the axiomatic system. This methodology requires strong completeness, as we cannot take the axiomatic system to be semantically fundamental when it lacks core semantic information.

The advantages of Milnikel's approach are offset by a major weakness: J^U does not provide a genuine model of the phenomenon of epistemic uncertainty. The knowledge⁵ operator of natural language is semantically vague, and the most important source of this vagueness is the probabilistic uncertainty of the truth of the embedded proposition. A true model of knowledge thus requires both a fuzzy logic and a probabilistic system. J^U is neither. It does not provide any formula that expresses the vague knowledge operator, but rather an infinite collection of deterministic uncertain-justification formulae. Furthermore, the only specifications made in J^U as to which proof polynomials are to justify which propositions at which level are that axioms must be justified to degree one by proof constants, and that justification must be transmissible by modus ponens. Nothing more is specified; in particular, nothing is specified to ensure that the justifications obey the laws of probability. Even such extreme violations such as a single proof constant c satisfying both $c :_{0.7} A$ and $c :_{0.7} \neg A$ are not ruled out by J^U .

2.2. Kokkinis' Probabilistic Justification Logic⁶

Kokkinis et al. [21,22] develop a class of genuinely probabilistic justification logics using methods analogous to Milnikel's. Instead of Milnikel's "uncertain justification" operator, Kokkinis' logics employ two separate operators: a probability operator $P_{>r}$ and a standard justification operator $:$. The semantics is straightforward: the justification operator is handled in the style of the Mkrtychev evidence function, while the probability operator is handled by incorporating an algebraic probability theory.

The outstanding feature of the Kokkinis logic is that it is a simple justification logic that also contains a probability operator. As a logic of probabilistic justification, however, it is decidedly lacking. For one thing, there is no syntax in the language that can be identified as a good representation of a particular case of probabilistic justification. In presenting a formalization of the lottery paradox in [22], Kokkinis et al. are forced to translate extra-logically by stipulating that for every proof polynomial t , there exists another proof polynomial $pb(t)$ such that $t : (P_{>0.99} \varphi) \supset pb(t) : \varphi$. Admittedly, that case involves belief rather than justification, but there does not seem to be any adequate approach to probabilistic justification proper aside from the same strategy of defining a new operation on proof polynomials; the P operators are not justifications, and the $:$ operator does not distinguish between probabilistic and non-probabilistic cases.

The Kokkinis logic also suffers from the converse problem that it cannot adequately represent certain justification. A typical epistemic agent possesses some knowledge that is derived from completely certain sources (e.g., mathematical proof) and other knowledge that is subject to probabilistic uncertainty. It is therefore important that an epistemic logic be able to represent both types of knowledge. The fundamental logical property of certain knowledge is factivity. The $:$ operator of Kokkinis logic is not factive. Normally, this would be a minor issue; factivity can be provided to a justification logic by adding T^J to the axiomatic system, making the Fitting model reflexive, and so forth. However, the proposed approach to probabilistic justification requires that the $:$ operator not be factive, as certain instances of $:$ are defined as representing merely probabilistic justification. Perhaps the problem might be solved by defining certain justification as the special case of probabilistic justification where the probability is one, and then reworking the account of probabilistic justification so that the principal operator is $P_{\geq r}$ rather than $:$. This would require $P_{\geq 1}$ to be factive; as specified in Kokkinis et al. [21,22], it is not, but that feature can probably be changed without difficulty.

⁵ Or epistemic justification; I phrase the point in terms of knowledge because that is the notion that is generally employed in natural language, but what really matters here is the justification aspect of the "justified true belief" analysis of knowledge, and not either the truth or belief aspects.

⁶ I owe thanks to an anonymous reviewer for pointing out to me the works that are discussed in this section and in Section 2.3.

2.3. Ghari's Hájek-Pavelka-Style Justification Logics

Ghari [23] presents a class of fuzzy justification logics whose semantics is substantially similar to that of the probabilistic justification logic that I present below. Ghari's logics are all justification versions of well-known non-probabilistic fuzzy logics such as Łukasiewicz continuum-valued logic and Pavelka logic. These logics do have an important epistemic application, which Ghari demonstrates with the following example:

Suppose that you are invited to the fourth birthday party of your nephew Mark. When you meet Mark, based on your observation, you are justifying to believe that 'Mark is a child'. One second after, your first observation in the birthday party is still an evidence to believe that he is a child, and one second after that, you believe that he is still a child for the same evidence, and so on. Hence, you believe that Mark is a child for the same evidence after any number of seconds have elapsed. However, after an appropriate number of seconds have elapsed, e.g. when Mark is aged thirty-five, your first observation in the birthday party is not an evidence to believe that he is a child. ([23], p. 771)

In this example, your evidence (and thus knowledge) that Mark is a child is completely certain at the time when it is acquired (Mark's fourth birthday party). Moreover, the quality of the evidence itself does not vary over time; discounting cases where you forget information or Mark dies prematurely, you are absolutely certain of Mark's age at the initial observation, 13 years after the initial observation (when Mark is 17), and indeed 21 years after the initial observation (when Mark is 35). What does change is your certainty regarding the proposition that Mark is a child: you are certain of this proposition's truth at the first time reference, uncertain at the second, and certain of its falsity at the third. The judgments of certainty and uncertainty here are not genuinely epistemic phenomena at all; they are entirely due to the semantic vagueness of the word "child".

The form of uncertainty resulting from epistemic judgments involving vague predicates is an important topic of study. Indeed, it is an unavoidable topic, given that nearly all natural language predicates are semantically vague. However, as explained in Section 1.3, they are not the intended target of the present inquiry. In this paper, we are seeking to build a logical model of situations where the evidence itself is uncertain, and that task requires a probabilistic fuzzy logic rather than any of Ghari's non-probabilistic systems.

3. Probabilistic Justification Logic

Given our arguments for the inadequacy of all of the aforementioned logics, how can we give an adequate logical model of epistemic vagueness? As indicated above, we must develop a system of justification logic that functions as both a fuzzy logic and a probabilistic logic.

We begin setting forth our new logic, probabilistic LP (henceforth, pr-LP), as a variant of the Mkrttychev model theory for LP given in Section 1.1. Because this will be a fuzzy logic, the assignment function v will now map sentences to an interval of truth values $[0, 1]$, and the evidence function \mathcal{E} will likewise map ordered pairs of proof polynomials and sentences to $[0, 1]$. The model must also include an additional component, a constant $\kappa \in [0, 1]$, representing the epistemic threshold discussed in Section 1.2 of this paper. It seems felicitous to use this same threshold κ as the cut-off for proper assertability (which is known as designation in the multi-valued logic literature); given the assumption that there is no vagueness outside of epistemic contexts, the assertability criterion will only be needed for epistemic operators and their truth-functional compounds.

Definition 7. A Mkrttychev premodel of pr-LP is a triple $\mathcal{M} = \langle v, \mathcal{E}, \kappa \rangle$, where v is a function from the set of sentences to $[0, 1]$, \mathcal{E} is a binary function from the sets of proof polynomials and sentences to $[0, 1]$, $\kappa \in [0, 1]$, and the following conditions are satisfied:

- $v(\perp) = 0$.

- $v(\top) = 1$.
- $v(\neg\varphi) = 1 - v(\varphi)$.
- $v(\varphi \vee \psi) \geq \max(v(\varphi), v(\psi))$.
- The connective \supset can be defined from \neg and \vee in the usual manner and evaluated per the above.⁷
- $v(\varphi \wedge \psi) \geq v(\varphi) + v(\psi) - 1$.
- The conventional relationship between probabilities of conjunctions and disjunctions must also hold: $v(\varphi \vee \psi) = v(\varphi) + v(\psi) - v(\varphi \wedge \psi)$.⁸
- $v(t : \varphi) = \mathcal{E}(t, \varphi)$.
- If $\mathcal{E}(t, \varphi) = 1$ for any proof polynomial t , then for the corresponding φ , $v(\varphi) = 1$.
- $\mathcal{E}(s \cdot t, \psi) \geq \mathcal{E}(s, \varphi \supset \psi) + \mathcal{E}(t, \varphi) - 1$.
- If $\mathcal{E}(t, \varphi) \geq \kappa$, then $\mathcal{E}(!t, t : \varphi) \geq \kappa$ ⁹
- If $\mathcal{E}(s, \varphi) = x$ and $\mathcal{E}(t, \psi) = y$, where φ and ψ have no common subformula, then $\mathcal{E}(s + t, \varphi) \geq x$, $\mathcal{E}(s + t, \psi) \geq y$, and $\mathcal{E}(s + t, \varphi \wedge \psi) \geq x + y - 1$.
- If φ is an instance of one of the axioms of the axiomatic presentation of LP given in Section 1.1, except for axiom T^j , then there must exist some proof constant c such that $\mathcal{E}(c, \varphi) = 1$.
- If ψ is a sentence of the form $t : \varphi \supset \varphi$ (that is, an axiom T^j instance) and $\mathcal{E}(t, \varphi) = 1$, then there must exist some proof constant c such that $\mathcal{E}(c, \psi) = 1$.

Definition 8. A model of pr-LP is a Mkrtychev premodel $\langle v, \mathcal{E}, \kappa \rangle$ that satisfies the following additional properties:

Classicality For any atomic proposition p , $v(p) = 0$ or $v(p) = 1$.

Projective consistency For every proof polynomial t , there is an evidence function \mathcal{F} such that the triple $\langle \lambda x. \mathcal{E}(t, x), \mathcal{F}, \kappa \rangle$ is also a Mkrtychev premodel of pr-LP.

Definition 9 (Satisfaction). A model of pr-LP, $\mathcal{M} = \langle v, \mathcal{E}, \kappa \rangle$, satisfies a formula φ (written $\mathcal{M} \models \varphi$) iff $v(\varphi) \geq \kappa$.

Definition 10 (Consequence). $\Gamma \models \varphi$ iff every model of pr-LP that satisfies all of the formulae within Γ also satisfies φ .

⁷ Reed Solomon suggested to me that instead of using the material conditional, we might understand $\varphi \rightarrow \psi$ as a conditional probability $P(\psi|\varphi)$. Computationally, this would give the valuation:

$$v(\varphi \rightarrow \psi) = \begin{cases} \pm 1, & \text{if } v(\varphi) \leq 0; \\ \frac{v(\varphi \wedge \psi)}{v(\varphi)}, & \text{otherwise.} \end{cases}$$

where the ± 1 option represents an interpretational choice between having false antecedent cases interpreted as true-by-default (as in the material conditional) or as undefined/error conditions (as in the definition of conditional probability). This valuation represents an interesting alternative to the material conditional, and it might, in principle, be a more perspicuous translation of some natural-language conditionals. Unfortunately, the proposal has one fatal flaw: it is ultimately grounded in an abuse of notation. Conditionals, like all logical connectives, can be embedded in other sentential contexts. To perform an analogous embedding with conditional probability, however, would require us to treat $P(A|B)$ as if it were merely a substitution instance of $P(A)$, which it is not; notations like $P(A \vee (B|C))$ or $P((A|B)|C)$ are meaningless. If, instead of conditional probability, we take the specified valuation function as fundamental, then nested conditionals will at least be grammatical, but the semantic clauses will not provide any interpretation of such a sentence beyond the general condition $v(\varphi) \in [0, 1]$. This follows from the triviality result proved by Lewis [24].

⁸ Note that the three restrictions given for conjunction and disjunction are not redundant. For example, algebraically combining the \wedge inequality with the relating equation results only in a statement that $v(\varphi \vee \psi) \leq 1$, which tells us nothing useful about the behavior of \vee .

⁹ In a previous draft of this paper, I used the condition “If $\mathcal{E}(t, \varphi) = x$, then $\mathcal{E}(!t, t : \varphi) = x$ ” instead of the present form. This has the benefit that it provides a more concrete valuation for the $!$ operator. However, in most cases there is no particular philosophical justification for such a restriction, and I have been convinced by a reviewer’s suggestion that mere simplification is not enough of a motive to prefer the concrete valuation over the more general form. If one prefers the concrete form, making this change will not have any significant effect on the resulting logic.

It is also helpful to supplement the language with an additional sentential operator K representing knowledge, with semantics given by

$$v(K\varphi) = \begin{cases} 1, & \text{if there is a proof polynomial } t \text{ such that } \mathcal{E}(t, \varphi) \geq \kappa; \\ 0, & \text{otherwise.} \end{cases}$$

This additional operator becomes unnecessary if we allow quantification over proof polynomials, as the quantified formula $\exists x(x : \varphi)$ would have truth value $\geq \kappa$ (and thus be assertible) iff $K\varphi$ would have truth value one. Quantification is usually omitted in justification logics for historic reasons,¹⁰ LP in particular was created to serve as a constructive account of provability for intuitionistic logic, and allowing quantification would risk allowing non-constructive demonstrations of provability. However, in the present context, I see no strong reason to prefer the primitive K operator over quantification. Nor do I see any strong reason to prefer quantification over a primitive K operator. Given some particular epistemic theory, one of these methods will likely be a more perspicuous model than the other, but for now, there is no need to make a decision on this matter.

The semantic clauses for the truth-functional connectives are based on the account of probabilistic logic given in Adams [26]; Gerla builds his probabilistic fuzzy logics using an algebraic semantics, though he notes in [19] an equivalence between one of his systems and a system similar to that given here. The given truth clauses will seem strange to a reader who is familiar with probability theory from a high school or undergraduate math course; in such settings, for example, there is a theorem giving a precise probability for a conjunction by $P(A \wedge B) = P(A) \times P(B)$.¹¹ However, the truth of that theorem depends crucially on the assumption that the conjuncts are probabilistically independent. This assumption holds for the examples normally discussed in math classes: coin tosses, dice rolls, unbiased statistical samples from a sufficiently large population, etc. However, the assumption of independence fails in epistemic contexts. For example, if you go to a reputable zoo, you can conclude that there is a very strong probability that the animal in the zebra cage is a zebra, and likewise that the animal in the tapir cage is a tapir. These probabilities are clearly not independent; if the “zebra” should turn out to be a disguised mule, this would greatly increase the probability that the “tapir” is actually a pig. In this example, the cause of the dependence is quite obvious; a major part of our evidence that the animals are what they are claimed to be is that the zoo is reputable, and displaying fake animals would undermine this reputability. In more general cases, there may not be such blatant dependence, but the possibility of it must always be accounted for. Probabilistic dependence may either increase or decrease the probability of a conjunction (or disjunction, etc.) as compared with the independent case; this is why most of the probability values are given as lower bounds rather than as exact equations. It can be shown by simple algebraic computation that the values that conjunctions and disjunctions take in the independent case fall within the range of possible values that is given for the general case presented here. It is also routine to demonstrate that in the cases where all subformulae receive the values zero and one, the probabilistic conjunction, disjunction, and negation are equivalent to the standard Boolean connectives.

The preferred general theory of probability in the mathematics literature is known as Kolmogorov probability, after its development in [28]. The probabilistic logic of pr-LP does in fact encode a version of Kolmogorov probability, as we will now show:

¹⁰ Quantified justification logic is investigated in Fitting [25]. Even after the publication of that paper, almost all work in justification logic has been conducted in purely propositional systems. This constitutes a striking divergence from the majority of other fields of logic, where first-order systems are standard.

¹¹ This theorem is the motivation underlying the multiplication of indices in the application schema of J^U . Milnikel [27] addresses the challenge that independence may fail by suggesting that the product of indices be replaced with the minimum. This solution coheres with the general probability theory presented here, but the logic J^U still is not genuinely probabilistic for the reasons discussed above.

Theorem 1. *Every Mkrtychev premodel of pr-LP is also a model of (finitely additive) Kolmogorov probability.*

Proof. Let the atomic sentences of pr-LP serve as events of a probability theory, with \perp as the impossible event and \top as the certain event. We define the complement of an event as the negation of the corresponding sentence, the intersection of events as the conjunction of sentences, and the union of events as the disjunction of sentences. This gives us a field of events with the required structure for at least a finitely additive formulation of probability theory. Note, however, that the atomic sentences are not all elementary events in Kolmogorov's sense, as some of them have non-empty intersections. We then define the probability of an event as the semantic value given to the corresponding sentence by the v function. The union of all events is the disjunction of all atomic propositions, including \top , which has value 1. Given that $v(\varphi \vee \psi) \geq \max(v(\varphi), v(\psi))$, the value of this universal disjunction is also 1, as specified by one of the Kolmogorov axioms. The semantic rules for pr-LP also include a version of the addition rule, $v(\varphi \vee \psi) = v(\varphi) + v(\psi) - v(\varphi \wedge \psi)$, which is a well-known theorem of Kolmogorov probability, and which suffices to entail the remaining Kolmogorov axiom, that the probability of a union of disjoint events is the sum of the individual probabilities. \square

Thus, we have it that the atomic sentences of a premodel encode a probability theory. What about the evidential sentences $t : \varphi$? After all, these sentences are the aspect of the logic that is intended to be probabilistic. That problem is solved by the projective consistency criterion of the full pr-LP model. Projective consistency mandates that the interpretation of each proof polynomial, as used in the $:$ operation, must itself be a Mkrtychev premodel, and thus also a probability model. Thus, we have it, in contrast to J^U , that pr-LP rules out justification instances that contradict the laws of probability, such as $v(c : A) = 0.7$ and $v(c : \neg A) = 0.7$, as there can be no Mkrtychev premodel whose valuation assigns the value 0.7 to both A and $\neg A$, given that $v(\neg A)$ is required to be $1 - v(A)$.

Philosophically, the interpretation of this projective consistency criterion is similar to that of the modular semantics of justification logic proposed by Artemov [29]. The underlying idea of modular semantics is that the semantic interpretation of a proof polynomial is the set of propositions for which it serves as evidence. By interpreting these propositions as sets of worlds in a Fitting model,¹² we have it the interpretation of a proof polynomial in a modular semantics is an epistemic situation: the set of worlds which are compatible with the given evidence. Because pr-LP is developed exclusively in Mkrtychev semantics, we do not have this sort of world-talk, but what we do have is that the interpretation of a proof polynomial is a Mkrtychev premodel. As with any propositional model theory, this Mkrtychev premodel can be viewed as a set of propositions that are interpreted as true; in this case, with the truth value being interpretable as the degree to which the evidence represented by the proof polynomial justifies the proposition.

The classicality condition on pr-LP models is philosophically licensed by the assumption declared at the end of Section 1.3, that there is no genuine non-epistemic vagueness. One might assume that we can dispense with this assumption simply by eliminating the value constraint, but there are two major problems with that proposal. First, pr-LP does not have the resources to account for how the non-epistemic vagueness of an atomic sentence affects the semantic value of a knowledge claim about that atom. Intuitively, knowledge should be weaker given the same evidence in a vague case than in a non-vague case, but perhaps we could accept pr-LP as it stands by pushing back against this intuition. The more fundamental problem is that non-epistemic vagueness may not be probabilistic in nature. All of the arguments made above for the use of probabilistic logic proceed under the assumption that the vagueness being analyzed derives from epistemic uncertainty. Indeed, most logicians who apply fuzzy logics to model non-epistemic vagueness find the most perspicuous model of the phenomenon to be a non-probabilistic fuzzy logic; for example, see Goguen [30] or Hájek [31]. The fuzzy justification

¹² Treating propositions as sets of worlds is ubiquitous in philosophical interpretation of modal logic, so this is an uncontroversial move.

logics of Ghari [23] provide a good model for epistemic reasoning involving such non-epistemic vagueness; unfortunately, there is no way to unify the approach to that phenomenon with the approach presented here to epistemic vagueness.

The assumption that non-epistemic propositions which are not subject to epistemic vagueness behave classically also justifies the decision to base pr-LP's axiom justification principle on the axioms of normal LP rather than an axiomatization of pr-LP itself. All of the axioms of LP are true for analytic reasons—either as pure logic or as conceptual analyses of epistemic concepts—and as such do not depend on evidence that is subject to probabilistic uncertainty. This design decision also makes the model theory of pr-LP much easier to employ, as its underlying axiomatic system is a well-studied normal justification logic rather than an obscure probabilistic logic. Unfortunately, it makes it more difficult to devise an axiomatization of pr-LP itself; at present, no such axiomatization is known.

The restrictions placed on the \mathcal{E} function are fairly straightforward. The factivity clause and its corresponding justification case are restricted to apply only to value 1 evidence for the obvious reason that if we have only uncertain evidence (however strong) for the truth of a proposition, then it might actually turn out to be false. The clause governing proof polynomials of the form $s \cdot t$ has a conjunction-like formula because these proof polynomials represent the result of modus ponens arguments, and such arguments provide no information about the conclusion unless both premises are true.

As noted in Section 1.1, the $+$ operator and its semantic clauses may be omitted, and are only included for the convenience of the reader who may wish to employ it. The function of the $+$ operator is to monotonically concatenate two pieces of evidence, without explicitly drawing any new inferences. Thus, the first two requirements of the semantic clause, which state that $s + t$ constitutes evidence for everything evinced by s and t , with strength not lower than the strength of the original evidence. This only makes sense in cases where φ and ψ are unrelated formulae, as otherwise one could apply summation to a case where $v(b : A) = 0.7$ and $v(c : \neg A) = 0.7$ to get $v((b + c) : A) = 0.7$ and $v((b + c) : \neg A) = 0.7$, violating projective consistency. The conjunction part of the semantic clause for $+$ is purely a notational convenience. In the absence of this assumption, we have that $((c \cdot (s + t)) \cdot (s + t)) : (\varphi \wedge \psi)$, where $s : \varphi$, $t : \psi$, and c is the proof constant corresponding to the propositional axiom $\varphi \supset (\psi \supset (\varphi \wedge \psi))$. To be completely pedantic, this only holds if the chosen axiomatization of propositional logic includes that tautology as an axiom; if not, c must be replaced by some complex proof polynomial. If this justification system were intended to be applied as a logical metatheory (as was the case in the original development of LP by Artemov [3]), it would be necessary to track such polynomials, but in the epistemic application, it is pointless. Hence the proposed simplification. The value assigned to the conjunction as justified by the sum coheres with the result of the repeated application principle in the longer form, given that all of the constituents except $s + t$ itself are axiomatic, and their evidence functions thus output value one.

The $!$ operator, whose semantics is given by the clause that if $\mathcal{E}(t, \varphi) \geq \kappa$, then $\mathcal{E}(!t, t : \varphi) \geq \kappa$ (or the footnoted alternative), corresponds in the epistemic interpretation to the KK principle. The KK principle lies at the heart of one of the major divides in epistemology: practically all epistemic internalists are committed to it as a fundamental law of epistemology, whereas most externalists reject it. We can certainly model an epistemology that lacks the KK principle using the methods advocated in this paper. However, it is not quite as simple as merely removing the $!$ semantic clause from the model theory and the 4^\dagger clause from the underlying LP axiomatization. As is noted in Artemov [6] and other justification logic survey texts, it is also necessary to strengthen the axiom justification rules to their iterated forms. In the model theory, instead of requiring a single proof constant c such that $\mathcal{E}(c, \varphi) = 1$, these clauses now require an infinite sequence of proof constants c_i such that $\mathcal{E}(c_0, \varphi) = 1$ and $\mathcal{E}(c_i, c_{i-1} : \dots : c_0 : \varphi) = 1$ for all $i > 0$. Note that this change must be made for both the clause pertaining to generic axiom instances and the special clause for T^\dagger instances.

Another technical consideration relating to the behavior of pr-LP as a justification logic is the feasibility of proving projection and realization theorems. These results, which relate each justification

logic to a corresponding modal logic, are considered to be essential to the intended function of the entire justification logic research program. Before we can address whether these results are provable, we must decide how they ought to be stated. One option is to formulate a modal extension of the probabilistic propositional logic that is employed in the Mkrtychev premodels. However, this seems to be the wrong track to take. The intended interpretation of the “forgetful projection” operator is precisely to discard the distinctions between the sources of justification represented by the various proof polynomials, and simply to output a modal formula that indicates, for every formula that is justified within the justification logic, that it is justified. Given this interpretation, it ought to be that part of the information that is “forgotten” in the pr-LP case is whether the particular justification that is given for a formula is certain or uncertain. Because the intended modal logic is to ignore the distinction between certain and uncertain justification, it need only be a normal modal logic rather than a probabilistic modal logic. Which modal logic is it? One might be tempted to assume S4, but this is incorrect. In pr-LP, the T^j axiom and its axiom justification result are restricted to only hold for certain justifications; they have countermodels involving uncertain justification. The only formulae that are required to possess certain justifications are logical truths: the set of classical consequences of the propositional, K^j , and J^j axioms. Moreover, instances of $\Box\varphi \supset \varphi$ where φ is a logical truth are theorems of the T-free system K4, simply by virtue of these formulae being material conditionals with logically true consequents. Thus, it appears that K4 is the suitable logic to be related to pr-LP via projection theorem.¹³ Indeed, the actual proof of projection of pr-LP into K4 is a routine induction on the semantic clauses of the pr-LP model theory.

We have it that, with the right formulation of the theorem, projection is established easily. However, that is true for most justification logics; realization is generally the more difficult of the two results. There are two routes that one might take. Considering that pr-LP is formulated in an entirely model-theoretic manner, one might attempt to adapt the model-theoretic realization proof of Fitting [5]. Perhaps this procedure will work, but in previous research (e.g., Su [32]), it has been found that details of Fitting’s proof fail when applied to non-classical justification logics, and so, I am not certain whether the proof is feasible in the probabilistic setting. A more promising route is to adapt the realization proof that is used in Artemov [33] and Brezhnev [34], which proceeds by induction on a cut-free sequent formulation of the modal logic. Because the sequent calculus that is doing the essential work is that of the modal logic K4 rather than the justification logic pr-LP, it should not matter that we do not presently have a sequent formulation of the justification logic. The construction of proof polynomials to validate a realization of each modal sequent step can be accomplished using the pr-LP model theory in the absence of a sequent calculus.

In a typical epistemic application of pr-LP, the proof constants are used to represent particular pieces of evidence.¹⁴ The corresponding outputs of the \mathcal{E} function are populated with whatever information that evidence grants about various features of the world (as represented by probabilities of the truth of propositions; an appropriate representation given that evidence is probabilistic in nature). The values (v) of atomic propositions are filled in by the actual facts about how the world really is.

¹³ We can also formulate another probabilistic justification logic that forgetfully projects to K4: the system pr-J4. This logic will have the same set of theorems as pr-LP, and thus it is fitting that they have the same forgetful projection. The difference between the two logics is that, in the case of a particular model that contains a certain justification of a formula that is not logically true, pr-J4 will permit that the formula be false on that model, whereas pr-LP forbids this.

¹⁴ Some authors (e.g., Antonakos [35]) approach the semantics of justification logic in such a way that proof constants can only be interpreted as justifications of logical axioms; if justifications of any other information are wanted, proof variables must be used. This usage, however, is not good practice. Treating constants and variables in such a manner makes it difficult to add quantification to the language without engendering confusion. It also does not cohere with the use of constants and variables in the majority of logical and mathematical practice, where constants are used to denote any object that is explicitly specified and variables are reserved for objects that are unknown, or whose identity is genuinely variable in the sense of not being fixed over all situations. Individual pieces of evidence are constants according to this standard usage, and so ought to be represented by proof constants in justification logic. If one desires to make a formal separation between justifications of axioms and extra-logical epistemic justifications, it is better to make this separation by partitioning the set of proof constants rather than by involving proof variables.

The value of κ is filled in by the appropriate (physical or linguistic) context. The result is a fairly simple and usable logical system with a K operator that models the physical and epistemic world about as accurately and perspicuously as is possible.¹⁵

Although the formal system is committed to the notion that the semantic values involved are probabilities, it makes no commitment as to exactly what these probabilities represent. The answer to this question seems to depend on the ontological foundation of one's chosen epistemology. For an externalist, the probabilities in question will be Bayesian conditional probabilities, as this is the only coherent account of the objective probability of an event depending on a condition. For an internalist, the probabilities might also be Bayesian, or they might be something more abstract, such as the subjective credences of a rational agent who is attending to specific pieces of evidence.

Acknowledgments: This research has been funded solely by its author. Thanks are due to the faculty of the University of Connecticut (particularly to Reed Solomon and David Ripley) and to various anonymous reviewers for helpful comments.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Priest, G. The Logic of Paradox. *J. Philos. Log.* **1979**, *8*, 219–241.
2. Lewis, C.I.; Langford, C.H. *Symbolic Logic*; Appleton-Century-Crofts: New York, NY, USA, 1932.
3. Artemov, S. Logic of proofs. *Ann. Pure Appl. Log.* **1994**, *67*, 29–59.
4. Mkrtychev, A. Models for the logic of proofs. In *Logical Foundations of Computer Science; Lecture Notes in Computer Science*; Adian, S., Nerode, A., Eds.; Springer: Heidelberg, Germany, 1997; Volume 1234, pp. 266–275.
5. Fitting, M. The logic of proofs, semantically. *Ann. Pure Appl. Log.* **2005**, *132*, 1–25.
6. Artemov, S. The Logic of Justification. *Rev. Symb. Log.* **2008**, *1*, 477–513.
7. Fitting, M. Modal Logics, Justification Logics, and Realization. *Ann. Pure Appl. Log.* **2016**, *167*, 615–648.
8. Lurie, J. New Directions in Justification Logic. Ph.D. Thesis, University of Connecticut, Storrs, CT, USA, forthcoming.
9. Artemov, S. Why do we need justification logic? In *Games, Norms and Reasons: Logic at the Crossroads*; Synthese Library; van Benthem, J., Gupta, A., Pacuit, E., Eds.; Springer: Dordrecht, The Netherlands; New York, NY, USA, 2011; Volume 353.
10. Goldman, A. Discrimination and perceptual knowledge. *J. Philos.* **1976**, *73*, 771–791.
11. Kripke, S.A. Nozick on knowledge. In *Philosophical Troubles*; Oxford University Press: Oxford, UK, 2011; pp. 162–224.
12. DeRose, K. Contextualism and Knowledge Attributions. *Philos. Phenomenol. Res.* **1992**, *52*, 913–929.
13. Stanley, J. *Knowledge and Practical Interests*; Oxford University Press: Oxford, UK, 2005.
14. Williamson, T. Contextualism, Subject-Sensitive Invariantism and Knowledge of Knowledge. *Philos. Q.* **2005**, *55*, 213–235.
15. MacFarlane, J. *Assessment Sensitivity: Relative Truth and Its Applications*; Oxford University Press: Oxford, UK, 2014.
16. Schaffer, J. From Contextualism to Contrastivism. *Philos. Stud.* **2004**, *119*, 73–103.
17. Dretske, F. Epistemic Operators. *J. Philos.* **1970**, *67*, 1007–1023.

¹⁵ An anonymous reviewer suggested that I support this claim with a concrete example. This suggestion turns out to be surprisingly unhelpful. Simple toy examples will show that the model coheres with some intuitive principles of uncertain reasoning, for example, that reasoning from two certain premises preserves certainty, whereas reasoning from two uncertain premises magnifies uncertainty. However, I have not been able to devise a more complex example that yields interesting conclusions. I did search the literature on Bayesian epistemology in the hopes of finding usage examples of probabilistic epistemic reasoning that I could adapt and, to my surprise, found no such examples in the literature. The closest thing to a useful concrete example would be an epistemic Dutch book argument, but this is just as contrived as any of the simple toy examples, and really does not show anything other than that an agent's total body of evidence, when collected together by something like the justification logic + operator, must still conform to the laws of probability on pain of incoherence. This criterion is satisfied by pr-LP, given that the projective consistency requirement holds for all proof polynomials and not merely for individual proof constants.

18. Dretske, F. *Knowledge and the Flow of Information*; MIT Press: Cambridge, MA, USA, 1981.
19. Gerla, G. *Fuzzy Logic: Mathematical Tools for Approximate Reasoning*; Trends in Logic; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001; Volume 11.
20. Pavelka, J. On Fuzzy Logic I, II, and III. *Zeitschrift für Math. Logik und Grundlagen der Mathematik* **1979**, *25*, 45–52, 119–134, 447–464.
21. Kokkinis, I.; Maksimović, P.; Ognjanović, Z.; Studer, T. First Steps Towards Probabilistic Justification Logic. *Log. J. IGPL* **2015**, *23*, 662–687.
22. Kokkinis, I.; Ognjanović, Z.; Studer, T. Probabilistic Justification Logic. In *Proceedings of Logical Foundations of Computer Science (LFCS'16)*; Lecture Notes in Computer Science; Artemov, S., Nerode, A., Eds.; Springer: Cham, Switzerland, 2016; Volume 9537, pp. 174–186.
23. Ghari, M. Pavelka-Style Fuzzy Justification Logics. *Log. J. IGPL* **2016**, *24*, 743–773.
24. Lewis, D. Probabilities of Conditionals and Conditional Probabilities. *Philos. Rev.* **1976**, *85*, 297–315.
25. Fitting, M. A quantified logic of evidence. *Ann. Pure Appl. Log.* **2008**, *152*, 67–83.
26. Adams, E.W. *A Primer of Probability Logic*; CSLI Publications: Stanford, CA, USA, 1998.
27. Milnikel, R.S. The logic of uncertain justifications. *Ann. Pure Appl. Log.* **2014**, *165*, 305–315.
28. Kolmogorov, A.N. *Grundbegriffe der Wahrscheinlichkeitsrechnung*; Springer: Berlin, Germany, 1933.
29. Artemov, S. The Ontology of Justifications in the Logical Setting. *Stud. Logica* **2012**, *100*, 17–30.
30. Goguen, J.A. The Logic of Inexact Concepts. *Synthese* **1969**, *19*, 325–373.
31. Hájek, P. On Vagueness, Truth Values, and Fuzzy Logics. *Stud. Logica* **2009**, *91*, 367–382.
32. Su, C.P. Paraconsistent Justification Logic: A Starting Point. In *Advances in Modal Logic (AiML)*; Gore, R., Kooi, B., Kurucz, A., Eds.; College Publications: London, UK, 2014; Volume 10.
33. Artemov, S. Explicit provability and constructive semantics. *Bull. Symb. Log.* **2001**, *7*, 1–36.
34. Brezhnev, V. *On Explicit Counterparts of Modal Logics*; CFIS 2000-06; Cornell University: Ithaca, NY, USA, 2000.
35. Antonakos, E. Explicit Generic Common Knowledge. In *Logical Foundations of Computer Science*; Lecture Notes in Computer Science; Artemov, S.N., Nerode, A., Eds.; Springer: Heidelberg, Germany, 2013; Volume 7734, pp. 16–28.



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).