A Logic for Reasoning about Justified Uncertain Beliefs

Tuan-Fang Fan
Department of CSIE,
National Penghu University of Science and Technology, Penghu 880, Taiwan
dffan@npu.edu.tw

Churn-Jung Liau
Institute of Information Science,
Academia Sinica, Taipei 115, Taiwan
liaucj@iis.sinica.edu.tw

Abstract

Justification logic originated from the study of the logic of proofs. However, in a more general setting, it may be regarded as a kind of explicit epistemic logic. In such logic, the reasons why a fact is believed are explicitly represented as justification terms. Traditionally, the modeling of uncertain beliefs is crucially important for epistemic reasoning. While graded modal logics interpreted with possibility theory semantics have been successfully applied to the representation and reasoning of uncertain beliefs, they cannot keep track of the reasons why an agent believes a fact. The objective of this paper is to extend the graded modal logics with explicit justifications. We introduce a possibilistic justification logic, present its syntax and semantics, and investigate its meta-properties, such as soundness, completeness, and realizability.

1 Introduction

Since the seminal work by Hintikka [1962], modal logic has been a standard approach for reasoning about knowledge and belief of intelligent agents [Fagin *et al.*, 1996]. In the epistemic/doxastic reading of modal logics[1], the formula $\Box\varphi$ is interpreted as "φ is believable" or "φ is knowable". However, without explicit justifications, the reasons why φ is believed or known are not represented in the logic. By contrast, justification logics (JL) supply the missing component by adding justification terms to epistemic formulas [Artemov and Nogina, 2005; Artemov, 2008; Artemov and Fitting, 2012; Fitting, 2005]. The first member of the JL family is the logic of proofs (LP) proposed in [Artemov, 2001]. Although the original purpose of LP is to formalize the Brouwer-Heyting-Kolmogorov semantics for intuitionistic logic and establish the completeness of intuitionistic logic with respect to this semantics, in a more general setting, LP may be regarded as a device that makes reasoning about knowledge explicit and keeps track of the justifications [Artemov, 2008; Artemov and Nogina, 2005].

[1]For the purpose of the paper, the difference between belief and knowledge is not important. Hence, hereafter, we use epistemic reasoning to denote reasoning about any kind of informational attitude for an agent.

One key issue in epistemic reasoning is the modeling of uncertain beliefs [Halpern, 2003]. While there exist many different uncertainty representation formalisms, the possibility theory-based approach [Zadeh, 1978] is the most appropriate one to be integrated into modal logics, because the necessity measure bears a striking similarity to the epistemic modality. In particular, graded modalities with possibility theory semantics have been successfully applied to the representation and reasoning of uncertain beliefs [Dubois *et al.*, 2013; Esteva *et al.*, 1997; Liau and Lin, 1992; 1996]. However, as in the case of classical epistemic logic, graded modal logics cannot keep track of the reasons of uncertain beliefs. Hence, the objective of the paper is to extend the graded modal logics with explicit justifications. We introduce a possibilistic justification logic (PJL) which is regarded as the explicit version of the quantitative modal logic (QML) proposed in [Liau and Lin, 1992]. We present the syntax and semantics of PJL and investigate its meta-properties, such as soundness, completeness, and realizability.

The remainder of the paper is organized as follows. In Section 2, we review the basic ideas of justification logic and quantitative modal logic. In Section 3, we introduce possibilistic justification logic and investigate its meta-properties. In Section 4, we discuss application of PJL to inconsistency-tolerant reasoning and compare PJL with related logical systems. In Section 5, we present the conclusions and indicate future research directions.

2 Preliminaries

2.1 Modal logic and justification logic

We start with the syntax and semantics of classical modal logic [Blackburn *et al.*, 2001]. Let PV denote the set of propositional variables. Then, the well-formed formulas (wffs) of the propositional modal logic are defined as follows:

$$\varphi ::= p \mid \bot \mid \varphi \supset \varphi \mid \Box\varphi,$$

where $p \in PV$, $\bot$ is the logical constant representing *falsum*, $\supset$ is the material implication, and $\Box$ is the epistemic modality. Other logical connectives such as $\top, \neg, \wedge, \vee, \equiv$ and modality $\Diamond$ are defined as abbreviations as usual. We use $\mathcal{L}^{\Box}$ and $\mathcal{L}$ to denote the propositional modal language and the underlying propositional language respectively.

The wffs of modal logic are semantically interpreted in a Kripke model (possible world model), which is defined as a triple $\mathfrak{M} = \langle W, R, V \rangle$, where W is a set of possible worlds, $R \subseteq W \times W$ is a binary *accessibility relation* on W, and $V : PV \to 2^W$ assigns to each propositional symbol in PV a subset of W. Given a model $\mathfrak{M}$, we can define a forcing relation $\Vdash_{\mathfrak{M}}$ between W and $\mathcal{L}^{\Box}$ by the semantic rules (we usually omit the subscript $\mathfrak{M}$ for simplicity). In addition to the standard rules for classical connectives, we have $w \Vdash \Box\varphi$ iff for any u such that $(w, u) \in R$, $u \Vdash \varphi$. For a given Kripke model $\mathfrak{M} = \langle W, R, V \rangle$ and a wff φ, we can define the truth set of φ with respect to $\mathfrak{M}$ by

$$|\varphi|_{\mathfrak{M}} = \{w \mid w \in W, w \Vdash \varphi\}.$$

We usually drop the subscript and simply write $|\varphi|$ for the truth set of φ when the model $\mathfrak{M}$ is clear from the context. Let Σ be a set of wffs. Then, we use $w \Vdash_{\mathfrak{M}} \Sigma$ to denote that $w \Vdash_{\mathfrak{M}} \psi$ for all $\psi \in \Sigma$. Furthermore, a wff φ is a logical consequence of Σ, denoted by $\Sigma \models \varphi$, if, for any $\mathfrak{M}$ and w, $w \Vdash_{\mathfrak{M}} \Sigma$ implies $w \Vdash_{\mathfrak{M}} \varphi$. A wff φ is valid, denoted by $\models \varphi$, if it is a logical consequence of the empty set.

In JL, in addition to the category of formulas, there is a second category of justifications, which are formal terms built up from constants and variables using various operation symbols. Constants represent justifications for commonly accepted truths—-typically axioms, whereas variables denote unspecified justifications. Different JL's allow different operation symbols. However, most of them contain application and sum. Specifically, the justification terms and wffs of the basic JL are formalized as follows:

$$t ::= a \mid x \mid t \cdot t \mid t + t,$$

$$\varphi ::= p \mid \bot \mid \varphi \supset \varphi \mid t : \varphi,$$

where $p \in PV$, a is a justification constant, and x is a justification variable. We use $\mathcal{L}_J$ to denote the basic JL language. The set of variables appearing in a term t is denoted by $\mathcal{V}(t)$ and for a set S of terms, we define $\mathcal{V}(S)$ as $\bigcup_{t \in S} \mathcal{V}(t)$. A term is *ground* if $\mathcal{V}(t) = \emptyset$.

JL furnishes an evidence-based foundation for the logic of knowledge with assertions of format $t : \varphi$, which denotes "t is a justification of φ", or more strictly, "t is accepted as a justification of φ" [Artemov, 2008]. Semantically, the formula $t : \varphi$ can be regarded as that t is an admissible evidence for φ and based on the evidence, φ is believed. Thus, the model of JL is the Kripke model of modal logics enriched with an additional evidence component [Artemov, 2008; Fitting, 2005]. This kind of model, called Kripke-Fitting model, is formally defined as a quadruple $\mathfrak{M} = \langle W, R, E, V \rangle$, where $\langle W, R, V \rangle$ is a Kripke model and E is an *admissible evidence function* such that $E(t, \varphi) \subseteq W$ for any justification term t and wff φ. Intuitively, $E(t, \varphi)$ specifies the set of possible worlds in which t is regarded as admissible evidence for φ. In this paper, we consider the minimal justification logic J. For J, it is required that E must satisfy the closure condition with respect to the application and sum operations:

1. Application: $E(s, \varphi \supset \psi) \cap E(t, \varphi) \subseteq E(s \cdot t, \psi)$.
2. Sum: $E(s, \varphi) \cup E(t, \varphi) \subseteq E(s + t, \varphi)$.

The first condition states that an admissible evidence for $\varphi \supset \psi$, which can be regarded as a function that transforms a justification of φ to a justification of ψ, can be applied to an admissible evidence for φ to obtain an admissible evidence for ψ. The second condition guarantees that adding a piece of new evidence does not defeat the original evidence. That is, $s + t$ is still an admissible evidence for φ whenever either s or t is an admissible evidence for φ. Note that conflicting evidence is allowed in JL. For example, if s is an admissible evidence for φ and t is an admissible evidence for $\neg\varphi$, then by the second condition above, $s + t$ is an admissible evidence for both φ and $\neg\varphi$. In this case, it means that s and t are in conflict with each other, and we do not exclude such evidence. The forcing relation $\Vdash_{\mathfrak{M}}$ between W and the justification formula $t : \varphi$ then follows the rule:

- $w \Vdash t : \varphi$ iff $w \in E(t, \varphi)$ and for any u such that $(w, u) \in R$, $u \Vdash \varphi$.

2.2 Possibilistic reasoning and modal logic

Possibility theory is developed from fuzzy set theory by Zadeh [1978]. Given a universe W, a *possibility distribution* on W is a function $\pi : W \to [0, 1]$. For a given π, we can define possibility and necessity measures $\Pi, N : 2^W \to [0, 1]$ respectively as $\Pi(X) = \sup_{w \in X} \pi(w)$ and $N(X) = 1 - \Pi(\overline{X})$, where $\overline{X}$ is the complement of X with respect to W. Dubois and Prade [1988; 1991] propose the possibilistic logic (PL) based on possibility theory. The wffs of PL based on a propositional language $\mathcal{L}$ are of the forms $(\varphi\ N\alpha)$ or $(\varphi\ \Pi\alpha)$, where $\varphi \in \mathcal{L}$ and $\alpha \in (0, 1]$. Informally, $(\varphi\ N\alpha)$ (resp. $(\varphi\ \Pi\alpha)$) means that the necessity (resp. possibility) of φ is at least α. Although PL is useful in reasoning about an agent's uncertain beliefs, it is not suitable for introspective agents, i.e., the agents reasoning about the beliefs of itself. However, by the analogy between necessity (resp. possibility) measure and the modal operator $\Box$ (resp. $\Diamond$) indicated in [Dubois and Prade, 1988], it is quite easy to extend PL to a graded modal logic that can model uncertain epistemic reasoning [Dubois *et al.*, 2013; Esteva *et al.*, 1997; Liau and Lin, 1992; 1996]. In this paper, we employ the formalism of quantitative modal logic (QML) proposed in [Liau and Lin, 1992; 1996].

To represent nested necessity and possibility measures, QML uses a less cumbersome notation that is compatible with modal operators. In fact, QML can be viewed as a logic with multiple modal operators. Let $[\alpha]$ and $[\alpha]^+$ denote modal operators for any rational $\alpha \in [0, 1]$. Then, the formation rules of QML wffs are as follows:

$$\varphi ::= p \mid \bot \mid \varphi \supset \varphi \mid [\alpha]\varphi \mid [\alpha]^+\varphi,$$

where $p \in PV$. We also abbreviate $\neg[1 - \alpha]\neg\varphi$ and $\neg[1 - \alpha]^+\neg\varphi$ as $\langle\alpha\rangle^+\varphi$ and $\langle\alpha\rangle\varphi$ respectively. The intuitive interpretation of $[\alpha]\varphi$ (resp. $[\alpha]^+\varphi$) is that an agent believes φ with certainty at least (more than) α. We use $\mathcal{L}^{\Pi}$ to denote the QML language.

A QML model is a triple $\mathfrak{M} = \langle W, R, V \rangle$, where W and V are defined as those for modal logic, and $R : W \times W \to [0, 1]$ is a fuzzy relation on W. For each $w \in W$, a possibility distribution π_w can be defined as $\pi_w(u) = R(w, u)$ for all $u \in W$. Let N_w denote the necessity measure corresponding to π_w for each $w \in W$. Then, the forcing relation between W and $\mathcal{L}^{\Pi}$ includes the following two rules:

- $w \Vdash [\alpha]\varphi$ iff $N_w(|\varphi|) \geq \alpha$;
- $w \Vdash [\alpha]^+\varphi$ iff $N_w(|\varphi|) > \alpha$.

3 Possibilistic Justification Logic

As JL provides an evidence-based foundation to epistemic logic, we can also combine justification terms with graded modalities in QML to represent uncertain beliefs justified by a piece of evidence. The situations often occur when the evidence is not strong enough to fully support the belief. The resultant logic is called possibilistic justification logic(PJL). The formation rules of PJL wffs are as follows:

$$\varphi ::= p \mid \bot \mid \varphi \supset \varphi \mid t :_\alpha \varphi \mid t :_\alpha^+ \varphi,$$

where $p \in PV$, $\alpha \in [0,1]$ is a rational number, and t is a justification term formed by the same rules as those in JL. We use $\mathcal{L}_J^\Pi$ to denote the PJL language. A rational number α appearing in a wff is called a *grade* or a *degree* and we use $\mathcal{G}(\varphi)$ to denote the set of grades appearing in a wff φ. Also, for a set Σ of wffs, we define $\mathcal{G}(\Sigma)$ as $\bigcup_{\varphi\in\Sigma} \mathcal{G}(\varphi)$. Moreover, for any subset $F \subseteq [0,1]$, we use $\mathcal{L}_J^\Pi(F)$ to denote the fragment of $\mathcal{L}_J^\Pi$ in which only grades in F occur.

As the Kripke-Fitting model is an extension of the Kripke model, a PJL model is defined as a quadruple $\mathfrak{M} = \langle W, R, E, V\rangle$, where $\langle W, R, V\rangle$ is a QML model and the admissible evidence function E is the same as in the Kripke-Fitting model. The forcing relation $\Vdash_\mathfrak{M}$ between W and the uncertain justification formulas $t :_\alpha \varphi$ and $t :_\alpha^+ \varphi$ are defined as follows:

- $w \Vdash t :_\alpha \varphi$ iff $w \in E(t, \varphi)$ and $N_w(|\varphi|) \geq \alpha$
- $w \Vdash t :_\alpha^+ \varphi$ iff $w \in E(t, \varphi)$ and $N_w(|\varphi|) > \alpha$

The intuitive interpretation of $t :_\alpha \varphi$ is that, according to the evidence t, φ is believed with certainty at least α, and $t :_\alpha^+ \varphi$ can be interpreted analogously. The validity of wffs and the notion of logical consequence are defined in the same way as in modal logic. We use $\models_{J^\Pi}$ to denote the validity and logical consequence in PJL.

The validity in PJL can be characterized by the following Hilbert-style axiomatization J^Π.

1. Axiom schemata:
   (a) The standard set of axioms for classical propositional logic
   (b) $s :_\alpha (\varphi \supset \psi) \supset (t :_\alpha \varphi \supset s \cdot t :_\alpha \psi)$
   (c) $s :_\alpha^+ (\varphi \supset \psi) \supset (t :_\alpha^+ \varphi \supset s \cdot t :_\alpha^+ \psi)$
   (d) $s :_\alpha \varphi \supset s + t :_\alpha \varphi$ and $t :_\alpha \varphi \supset s + t :_\alpha \varphi$
   (e) $s :_\alpha^+ \varphi \supset s + t :_\alpha^+ \varphi$ and $t :_\alpha^+ \varphi \supset s + t :_\alpha^+ \varphi$
   (f) $(s :_0 \varphi \wedge t :_\alpha \varphi) \supset s :_\alpha \varphi$
   $(s :_0 \varphi \wedge t :_\alpha^+ \varphi) \supset s :_\alpha^+ \varphi$
   (g) $t :_\alpha \varphi \supset t :_\beta^+ \varphi$, if $\alpha > \beta$
   (h) $t :_\alpha^+ \varphi \supset t :_\alpha \varphi$
   (i) $\neg t :_1^+ \varphi$
2. Rules of inference:
   (a) Modus Ponens: from φ and $\varphi \supset \psi$, infer ψ;
   (b) Axiom Internalization Rule[2]: from any instance φ of the axiom schemata above and any sequence of justification constants $c_1, c_2, \ldots, c_n$, infer $c_n :_1 c_{n-1} :_1 \cdots c_1 :_1 \varphi$.

Let Σ be a subset of PJL wffs. Then, an axiomatic *proof* (or *derivation*) from Σ is a finite sequence of wffs, each of which is an instance of an axiom schema, an element of Σ, or follows from earlier items by one of the rules of inference; and a wff φ is *derivable* from Σ, denoted by $\Sigma \vdash_{J^\Pi} \varphi$, if φ is the last line of a proof from Σ. Moreover, φ is an axiomatic *theorem* of the system J^Π, denoted by $\vdash_{J^\Pi} \varphi$, if it is derivable from the empty set. A set Σ is *inconsistent* if $\Sigma \vdash_{J^\Pi} \bot$, otherwise, Σ is consistent.

The axiom internalization rule behaves like the necessitation rule in modal logic or QML. For the soundness of the rule, we impose a further restriction on the admissible evidence function:

- for any instance φ of the axiom schemata above and any sequence of justification constants $c_1, c_2, \ldots, c_n (n \geq 1)$, $E(c_n, c_{n-1} :_1 \cdots c_1 :_1 \varphi) = W$.

Note that the premises of the axiom internalization rule are restricted to instances of axioms instead of any theorems. The main reason of such restriction is that in JL or PJL, justification terms can track the derivation of a theorem. Because justification constants represent atomic derivations which cannot be analyzed further, an axiom that holds without resorting to any deduction process can be simply justified by any justification constant. However, a theorem that is not an axiom must be derived through a complex deduction process whose structure is reflected by justification terms of the theorem. Thus, justification constants are not necessarily applicable to any theorem. Despite the restriction on the axiom internalization rule, internalization of any proof is possible via the following result.

Theorem 1 (Internalization) *The system J^Π enjoys internalization. That is, for any wff $\varphi_1, \ldots, \varphi_n, \psi \in \mathcal{L}_J^\Pi$, if*

$$\{\varphi_1, \ldots, \varphi_n\} \vdash_{J^\Pi} \psi$$

then for any terms $t_1, \ldots, t_n$, rational numbers $\alpha_1 \ldots, \alpha_n \in [0,1]$, and $I \subseteq \{1, 2, \ldots, n\}$, there exists a term t such that $\mathcal{V}(t) \subseteq \mathcal{V}(\{t_1, \ldots, t_n\})$ and

$$\{t_i :_{\alpha_i} \varphi_i \mid i \in I\} \cup \{t_i :_{\alpha_i}^+ \varphi_i \mid i \notin I\} \vdash_{J^\Pi} t :_\alpha \psi$$

where $\alpha = \min_{1\leq i\leq n} \alpha_i$. Furthermore, if $\min_{i\in I} \alpha_i > \alpha$, then the conclusion of the derivation above can be strengthened to $t :_\alpha^+ \psi$.

Proof. The theorem can be easily proved by induction on the length of the derivation $\{\varphi_1, \ldots, \varphi_n\} \vdash_{J^\Pi} \psi$. ■

A special case of the internalization theorem shows that any proof of an axiomatic theorem can be internalized.

Corollary 1 *For any wff $\varphi \in \mathcal{L}_J^\Pi$, if $\vdash_{J^\Pi} \varphi$ then there exists a ground term t such that $\vdash_{J^\Pi} t :_1 \varphi$.*

[2]The rule corresponds to using total constant specification [Artemov, 2008]. For brevity, we do not introduce the notion of constant specification in this paper.

The system J^{Π} is sound and complete with respect to PJL models. Nevertheless, because of the failure of the compactness theorem, the strong completeness of the system is restricted. For example, let $\Sigma = \{c :_{1-\frac{1}{n}} \varphi \mid n > 1\}$. Then, $c :_1 \varphi$ is a logical consequence of Σ. However, because any finite subset of $\Sigma \cup \{\neg c :_1 \varphi\}$ is obviously satisfiable and any deduction must be finitary, $c :_1 \varphi$ is not derivable from Σ. Let us consider a set of wffs $\Sigma \cup \{\varphi\} \subseteq \mathcal{L}_J^{\Pi}$ such that $\mathcal{G}(\Sigma \cup \{\varphi\})$ is finite and let $F = \{\alpha, 1-\alpha \mid \alpha \in \mathcal{G}(\Sigma \cup \{\varphi\})\} \cup \{0,1\}$. Then, we can define the *world-alternative* functions $AL, AL^+ : 2^{\mathcal{L}_J^{\Pi}(F)} \times F \to 2^{\mathcal{L}_J^{\Pi}(F)}$ as follows:

$$AL(\Gamma,\alpha) = \{\varphi \mid \exists\alpha', t(\alpha' \geq \alpha, t :_{\alpha'} \varphi \in \Gamma \text{ or } t :_{\alpha'}^{+} \varphi \in \Gamma)\}$$

$$\begin{aligned} AL^+(\Gamma,\alpha) &= \{\varphi \mid \exists\alpha', t(\alpha' > \alpha, t :_{\alpha'} \varphi \in \Gamma)\} \\ &\cup \{\varphi \mid \exists\alpha', t(\alpha' \geq \alpha, t :_{\alpha'}^{+} \varphi \in \Gamma)\}. \end{aligned}$$

Informally, $AL(\Gamma,\alpha)$ (resp. $AL^+(\Gamma,\alpha)$) contains all wffs φ such that $N(|\varphi|) \geq \alpha$ (resp. $N(|\varphi|) > \alpha$) is implied by formulas in Γ. To establish completeness, we use the standard canonical model construction. A model $\mathfrak{M} = \langle W, R, E, V\rangle$ is called a canonical model for $\mathcal{L}_J^{\Pi}(F)$ if

- W is the set of all maximal consistent subsets of $\mathcal{L}_J^{\Pi}(F)$. We use Γ, Δ, etc. to denote elements of W.
- R satisfies the following two conditions for all $\Gamma, \Gamma' \in W$ and $\alpha \in F$:
  - if $\alpha > 0$, then $AL(\Gamma,\alpha) \subseteq \Gamma'$ iff $R(\Gamma,\Gamma') > 1-\alpha$;
  - $AL^+(\Gamma,\alpha) \subseteq \Gamma'$ iff $R(\Gamma,\Gamma') \geq 1-\alpha$.
- $E(t,\varphi) = \{\Gamma \in W \mid \exists\alpha(t :_{\alpha} \varphi \in \Gamma \text{ or } t :_{\alpha}^{+} \varphi \in \Gamma)\}$.
- $V(p) = \{\Gamma \in W \mid p \in \Gamma\}$.

The following lemma shows that canonical models indeed exist.

Lemma 1 *There exists R satisfying the two conditions for the accessibility relation of a canonical model.*

Proof. The result relies on that F is finite. Let us assume that the elements of F are enumerated decreasingly as $1 = \alpha_1 > \alpha_2 > \cdots > \alpha_k = 0$. Then, for any $\Gamma \in W$, we have $AL^+(\Gamma,\alpha_1) = \emptyset$ and for any $1 \leq i < k$,

$$AL^+(\Gamma,\alpha_i) \subseteq AL(\Gamma,\alpha_i) \subseteq AL^+(\Gamma,\alpha_{i+1}).$$

Hence, for any $\Gamma, \Gamma' \in W$, $R(\Gamma,\Gamma')$ can be defined in the following way:

1. if $AL^+(\Gamma,\alpha_i) \subseteq \Gamma' \not\supseteq AL(\Gamma,\alpha_i)$ for some $1 \leq i < k$, then $R(\Gamma,\Gamma') = 1-\alpha_i$
2. if $AL(\Gamma,\alpha_i) \subseteq \Gamma' \not\supseteq AL^+(\Gamma,\alpha_{i+1})$ for some $1 \leq i < k$, then $R(\Gamma,\Gamma') = 1-\alpha$ for some $\alpha \in (\alpha_i, \alpha_{i+1})$
3. if $AL^+(\Gamma,\alpha_k) \subseteq \Gamma'$, then $R(\Gamma,\Gamma') = 1$.

Then, it is easily verified that R satisfies the requirements. ■

The key fact for canonical models is the following Truth Lemma.

Lemma 2 *Let $\mathfrak{M} = \langle W, R, E, V\rangle$ be a canonical model for $\mathcal{L}_J^{\Pi}(F)$. Then, for any $\varphi \in \mathcal{L}_J^{\Pi}(F)$ and $\Gamma \in W$, $\Gamma \Vdash_{\mathfrak{M}} \varphi$ iff $\varphi \in \Gamma$.*

Proof. The lemma can be proved by induction on the complexity of φ. The atomic cases are covered by the definition of V and the cases for Boolean connectives and logical constants are standard. The only interesting cases are $\varphi = t :_{\alpha} \psi$ or $\varphi = t :_{\alpha}^{+} \psi$. We consider the former case and the latter one is proved analogously.

($\Leftarrow$): let us consider two cases:

1. If $t :_{\alpha} \psi \in \Gamma$ for some $\alpha > 0$, then by definition $\Gamma \in E(t,\psi)$ and $\psi \in AL(\Gamma,\alpha)$. Hence, $\psi \in \Gamma'$ for any Γ' such that $R(\Gamma,\Gamma') > 1-\alpha$ by the first condition of R. By the induction hypothesis, this means that $\Pi_{\Gamma}(|\neg\psi|) \leq 1-\alpha$ and $N_{\Gamma}(|\psi|) \geq \alpha$, which leads to $\Gamma \Vdash t :_{\alpha} \psi$.
2. If $t :_0 \psi \in \Gamma$ (i.e. $\alpha = 0$), then $\Gamma \in E(t,\psi)$ is sufficient to show that $\Gamma \Vdash t :_0 \psi$ because $N_{\Gamma}(|\psi|) \geq 0$ holds trivially.

($\Rightarrow$): there are also two cases:

1. If $t :_0 \psi \notin \Gamma$, then by axioms (g) and (h), $\Gamma \notin E(t,\psi)$. Hence, $\Gamma \nVdash t :_0 \psi$.
2. If $t :_{\alpha} \psi \notin \Gamma$ for some $\alpha > 0$, then we must consider whether $t :_0 \psi \in \Gamma$. In the case that $t :_0 \psi \notin \Gamma$, we have seen that $\Gamma \notin E(t,\psi)$ and hence $\Gamma \nVdash t :_{\alpha} \psi$. Thus, let us assume that $t :_0 \psi \in \Gamma$. We first show that $AL(\Gamma,\alpha) \cup \{\neg\psi\}$ is consistent. If this is not the case, then by axioms (b), (c), (g), (h) and the modus ponens rule, we can find a term s such that $s :_{\alpha} \psi \in \Gamma$ and by axiom (f), $t :_{\alpha} \psi \in \Gamma$, which is contradictory with the assumption. Therefore, there exists a maximal consistent Γ' such that $\neg\psi \in \Gamma'$ and $R(\Gamma,\Gamma') > 1-\alpha$, which means that $\Pi_{\Gamma}(|\neg\psi|) > 1-\alpha$ and $N_{\Gamma}(|\psi|) < \alpha$ by the induction hypothesis. Consequently, $\Gamma \nVdash t :_{\alpha} \psi$. ■

To prove the completeness, let us assume that φ is not derivable from Σ. Then, $\Sigma \cup \{\neg\varphi\}$ is consistent and can be extended to a maximal consistent subset of $\mathcal{L}_J^{\Pi}(F)$. Hence, by the Truth Lemma, we can find a possible world in a canonical model that satisfies $\Sigma \cup \{\neg\varphi\}$. That is, φ is not a logical consequence of Σ. Because the soundness of the system can be proved in a standard way, we have the following theorem.

Theorem 2 (Soundness and Completeness) *For any $\Sigma \cup \{\varphi\} \subseteq \mathcal{L}_J^{\Pi}$ such that $\mathcal{G}(\Sigma \cup \{\varphi\})$ is finite, we have $\Sigma \vdash_{J^{\Pi}} \varphi$ iff $\Sigma \models_{J^{\Pi}} \varphi$.*

Obviously, the logic J^{Π} is a conservative extension of the logic J. In other words, we can embed the language $\mathcal{L}_J$ into $\mathcal{L}_J^{\Pi}$ by using a straightforward translation mapping $\tau : \mathcal{L}_J \to \mathcal{L}_J^{\Pi}$ that maps each propositional variable into itself, respects the Boolean connectives, and satisfies the following condition:

$$\tau(t : \varphi) = t :_1 \tau(\varphi).$$

Then, we have the following result.

Theorem 3 (Conservative extension) *For any wff $\varphi \in \mathcal{L}_J$, we have $\models_J \varphi$ iff $\models_{J^{\Pi}} \tau(\varphi)$.*

Proof. Because all axioms and inference rules of J are instances of axioms and inference rules of J^{Π} under the translation, $\models_J \varphi$ implies $\models_{J^{\Pi}} \tau(\varphi)$ by the soundness and completeness of both systems. On the other hand, if $\not\models_J \varphi$, then there exists a Kripke-Fitting model $\mathfrak{M} = \langle W, R, E, V \rangle$ and a possible world w such that $w \not\Vdash \varphi$. However, a Kripke-Fitting model can be regarded as a special case of PJL model and $w \Vdash \varphi$ in a Kripke-Fitting model iff $w \Vdash \tau(\varphi)$ in the corresponding PJL model. Therefore, $\models_{J^{\Pi}} \tau(\varphi)$ implies $\models_J \varphi$. ■

It is well-known that JL can be regarded as an explicit epistemic logic. Thus, the connection between these two logics is that the epistemic formula $\Box\varphi$ can be informally interpreted as "there exists some justification x such that $x : \varphi$". This observation leads to the notion of *forgetful projection* from $\mathcal{L}_J$ to $\mathcal{L}^{\Box}$ which replaces each occurrence of $t : \varphi$ by $\Box\varphi$. The forgetful projection always maps valid formulas of JL to valid formula of epistemic logic; and the converse also holds: any valid formulas of epistemic logic is a forgetful projection of some valid formula of JL [Artemov, 2008; Artemov and Fitting, 2012; Fitting, 2013]. This kind of correspondence theorem also generalizes to PJL and QML. A forgetful projection from $\mathcal{L}_J^{\Pi}$ to $\mathcal{L}^{\Pi}$ replaces each occurrence of $t :_{\alpha} \varphi$ with $[\alpha]\varphi$ for any $\alpha > 0$, each occurrence of $t :_0 \varphi$ with $\top$, and each occurrence of $t :_{\alpha}^{+} \varphi$ with $[\alpha]^{+}\varphi$ for any $\alpha \geq 0$. For each wff $\varphi \in \mathcal{L}_J^{\Pi}$, let φ^0 denote its forgetful projection and for any subset $\Sigma \subseteq \mathcal{L}_J^{\Pi}$, let $\Sigma^0 = \{\varphi^0 \mid \varphi \in \Sigma\}$. Then, we have the following result.

Theorem 4 (Correspondence theorem) *Let J^{Π} denote the set of all PJL theorems and let K_{+}^{Π} denote the set of all QML theorems in which the modality $[0]$ does not appear. Then, $(J^{\Pi})^0 = K_{+}^{\Pi}$.*

Proof. (Sketch) The easy part is to show that $(J^{\Pi})^0 \subseteq K_{+}^{\Pi}$ because the forgetful projection of the system J^{Π} is derivable from the basic QML system. The hard part is to prove a Realization Theorem which essentially claims that, for any theorem φ in K_{+}^{Π}, there is some way of replacing $[\alpha]$ and $[\alpha]^{+}$ symbols with justification terms $t :_{\alpha}$ and $t :_{\alpha}^{+}$ respectively to produce a theorem of J^{Π}. There are essentially two methods of establishing realization theorems: the constructive syntactic method due to Artemov [2001] that makes use of cut-free sequent systems for modal logics and the non-constructive semantic method due to Fitting [2005; 2013]. We will use the semantic method in our proof. In the proof, negative occurrences of modalities become justification variables and positive occurrences of modalities become justification terms that may involve those variables. The modality $[0]$ must be specially treated because $[0]\varphi$ is a QML theorem but $t :_0 \varphi$ is not. Note that each QML wff can be equivalently transformed into one wff without any occurrence of the $[0]$ modality by replacing each outermost occurrence of $[0]\varphi$ with $\top$. ■

4 Discussion and Related Work

4.1 Discussion

As QML is a modal formulation of PL, it inherits the advantage of PL to allow inconsistency-tolerant reasoning [Dubois *et al.*, 1994]. More specifically, let Σ denote a knowledge base (i.e. a set of wffs) in QML. Then, the inconsistency level of Σ, denoted by $Inc(\Sigma)$, is the maximum number α such that $\Sigma \vdash [\alpha]\bot$. A deduction of $[\alpha]\varphi$ is *nontrivial* only when $\alpha > Inc(\Sigma)$. In other words, the derivation of a QML formula $[\alpha]\varphi$ from Σ is *trivial* if $[\alpha]\bot$ is also derivable from Σ. In this way, QML allows a knowledge base to be inconsistent to some extent; and uncertain beliefs with necessity measures above the inconsistency level are still acceptable. However, instead of simply discarding a large part of information, the occurrence of inconsistency sometimes yields the requirement to update the knowledge base. By adding justification terms to the formulas, it is possible to track the origin of inconsistency which can be used in the update process. The following example illustrates this point.

Example 1 *Let us use φ_1, φ_2, φ_3, and ψ to denote "Tweety is a penguin", "Tweety is a bird", "Tweety lives in Zoologischer Garten Berlin", and "Tweety can fly" respectively; and let us assume that our knowledge base is $\Sigma = \{c :_1 \varphi_1, s :_1 (\varphi_1 \supset \varphi_2), d_1 :_{0.9} (\varphi_1 \supset \neg\psi), d_2 :_{0.8} (\varphi_2 \supset \psi), n :_{0.8} \varphi_3\}$ where c, s, d_1, d_2, n are justification constants denoting a reliable observation, a strict rule, a quite reliable default rule, a less reliable default rule and a partially reliable information source respectively. Then, by axioms (b), (g), and (h), we can derive $(d_1 \cdot c) \cdot (d_2 \cdot (s \cdot c)) :_{0.8} \bot$ (recalling that $\neg\psi$ is an abbreviation of $\psi \supset \bot$) which keeps track of sources of the inconsistency. By using inconsistency-tolerant reasoning, $(d_1 \cdot c) :_{0.9} \neg\psi$ is a nontrivial consequence of Σ, whereas $(d_2 \cdot (s \cdot c)) :_{0.8} \psi$ and $n :_{0.8} \varphi_3$ are both blocked because of the triviality. Although inconsistency-tolerant reasoning allows us to maintain highly certain information in the knowledge base, it is sometimes not good enough to simply discard less certain information. For example, it seems that $n :_{0.8} \varphi_3$ is unrelated to the inconsistent information in Σ, but the derivation of $n :_{0.8} \varphi_3$ is still blocked. In such case, the justification term $(d_1 \cdot c) \cdot (d_2 \cdot (s \cdot c))$ for $\bot$ will provide a clue to deal with the problem.*

By checking the term, we find that it is possible to modify the strength of the least reliable evidence d_2 to improve the consistency degree of the knowledge base. The most straightforward method is to remove $d_2 :_{0.8} (\varphi_2 \supset \psi)$ from Σ. Then, the inconsistency degree of Σ will be reduced to 0. However, the complete removal of $d_2 :_{0.8} (\varphi_2 \supset \psi)$ seems not very reasonable according to our commonsense. Hence, an alternative way is to change $d_2 :_{0.8} (\varphi_2 \supset \psi)$ to $d_2 :_{0.7} (\varphi_2 \supset \psi)$ because the defeating example Tweety may show that the strength of the default rule was set too high. Yet another approach is to keep the default rule unchanged but seek further evidence to increase the reliability of φ_3. This approach, despite not increasing the consistency of Σ, may suggest that some information independent of the inconsistent part of the knowledge base deserves further investigation. This usually occurs in the context of scientific inquiry. ■

In the Kripke-Fitting semantics of JL, the truth of a wff $t : \varphi$ in a possible world w depends on two conditions: (1) t is an admissible evidence for φ in w and (2) φ is believed. However, in the JL language, we cannot describe these two conditions separately. In particular, we cannot assert that t is an admissible evidence for φ by using JL formulas. By contrast, the wff $t :_0 \varphi$ precisely describes this fact because $w \Vdash t :_0 \varphi$ iff $w \in E(t, \varphi)$. Indeed, the formula is used in axiom (f) to formalize an implicit assumption in JL. That is, all admissible evidences for the same proposition are treated equally in the sense that if s and t are both admissible evidences for φ, then $s : \varphi$ is true iff $t : \varphi$ is true. An implication of the assumption in PJL is that multiple justifications of the same proposition will converge to the strongest justification among them. Indeed, by using axiom (f), we can derive that, if $s :_{\alpha} \varphi$ and $t :_{\beta} \varphi$ are true, then $s :_{\max(\alpha,\beta)} \varphi$ and $t :_{\max(\alpha,\beta)} \varphi$ are both true. Hence, if there exist multiple justifications for uncertain beliefs on φ, then the strongest justified uncertain belief will be chosen and all other justifications are regarded as supporting the uncertain belief to the same extent. The assumption is sometimes regarded as too restrictive. A possibility of lifting the restriction is to associate with each term t an accessibility relation R_t instead of using a single relation R in the PJL model. It may well be worth exploring the impact of the less restrictive semantics.

Mkrtychev semantics is a predecessor of Kripke-Fitting semantics [Mkrtychev, 1997]. Mkrtychev models essentially coincide with single-world Kripke-Fitting models. In such models, the truth value of $t : \varphi$ is determined only by the admissible evidence function E. Interestingly, the information about Kripke structure in Kripke-Fitting models can be completely encoded by the admissible evidence function and consequently, the system J is also complete with respect to Mkrtychev semantics [Artemov, 2008]. However, this is no longer true for PJL. For example, let us consider a set of PJL wffs $\Sigma = \{s :_{0.9} p, \neg s :^{+}_{0.9} p, t :_{0.7} q, \neg t :^{+}_{0.7} q\}$. Then, Σ is obviously satisfiable in PJL semantics. However, any models of Σ must contain at least two possible worlds to accommodate different necessity values on p and q.

4.2 Related works

In the last decades, the research on JL and its applications to formal epistemology, mathematics, computer science, and AI has been highly active [Artemov and Fitting, 2012]. However, it is only very recently that the integration of uncertainty reasoning with JL started to be an important topic in the discipline. Consequently, only a few works have been done along this direction. The most remarkable example is the logic $\mathbf{J}^U$ for uncertain justifications proposed in [Milnikel, 2014]. Syntactically, the wffs of $\mathbf{J}^U$ are defined by

$$\varphi ::= p \mid \bot \mid \varphi \supset \varphi \mid t :_{\alpha} \varphi.$$

However, its wff $t :_{\alpha} \varphi$ is interpreted quite differently. In a $\mathbf{J}^U$ model $\langle W, R, E, V\rangle$, the accessibility relation R is crisp but E is defined as a ternary function such that for any term t, possible world w, and wff φ, $E(w,t,\varphi) = [0,\beta)$ or $[0,\beta]$ for some rational number $\beta \in [0,1]$; and $w \Vdash t :_{\alpha} \varphi$ holds if $\alpha \in E(w,t,\varphi)$ and for any u such that $(w,u) \in R$, $u \Vdash \varphi$. Informally, this means that the same wff $t :_{\alpha} \varphi$ represents the uncertain justification of belief and the justification of uncertain belief in $\mathbf{J}^U$ and PJL respectively. A consequence of the semantic difference is that the forgetful projection of PJL is QML, whereas the forgetful projection of $\mathbf{J}^U$ should be classical modal logic[3].

[3]Although forgetful projection is not explicitly defined in [Milnikel, 2014], we think the claim is reasonable because by removing the admissible evidence function, a $\mathbf{J}^U$ model is simply a classical Kripke model.

The logic $\mathbf{J}^U$ is further extended to fuzzy justification logics in [Ghari, 2014b]. Instead of using classical two-valued logic, the proposed fuzzy justification logics BLJ and RPLJ are based on Hajek's fuzzy basic logic BL and rational Pavelka logic RPL respectively [Hájek, 1998], where BL has a classical syntax with conjunction ($\&$) and implication ($\rightarrow$) as primitive connectives and RPL is an extension of BL that allows rational truth constants $\overline{\alpha}$ for $\alpha \in [0,1]$. Hence, in BLJ and RPLJ, a justification formula also has the classical syntax. However, for a BLJ or RPLJ model $\langle W, R, E, V\rangle$, although the accessibility relation R is still crisp, V is now defined as a fuzzy valuation such that $V(w,p) \in [0,1]$ for any possible world w and propositional variable p, besides $E(w,t,\varphi) \in [0,1]$ for any term t, possible world w, and wff φ. Then, V is extended to all wffs according to the many-valued interpretation. In particular, $V(w, t : \varphi) = E(w,t,\varphi) \otimes \inf_{u:(w,u)\in R} V(u,p)$, where $\otimes$ is a t-norm that interprets the many-valued conjunction. Therefore, in BLJ and RPLJ, $t : \varphi$ can represent the uncertain justification of belief on vague propositions and their forgetful projections will be the respective fuzzy modal logics.

In addition to possibility theory-based representation, the combination of probabilistic reasoning and JL has been also proposed recently [Kokkinis *et al.*, 2014]. However, instead of considering probabilistic justification of belief or justification of probabilistic belief, the proposed logic PJ aims at reasoning about the probability of justification statements. Hence, the wffs of PJ are Boolean combinations of formulas of the form $P_{\geq\alpha}\varphi$ where $\varphi \in \mathcal{L}_J$ and interpreted by probability measures over the space of $\mathcal{L}_J$ models. By contrast, it is possible to develop a logic for reasoning about justified probabilistic beliefs using our approach by interpreting $t :_{\alpha} \varphi$ in a structure where the fuzzy accessibility relation is replaced by probability measures over possible worlds.

5 Conclusion

In this paper, we present an extension of JL to PJL in the same way that QML extends classical modal logic. As JL is regarded as an explicit modal logic, PJL can be regarded as an explicit version of QML. Because QML is a modal formulation of PL, it can represent and reason with uncertain beliefs and partial inconsistency. On the other hand, the justification terms in JL can track the derivation process of a deduction and record the structure of a proof. Thus, by combining the advantages of QML and JL, PJL can track the derivation of uncertain belief and locate the source of inconsistency if the belief is partially inconsistent.

Although we only present the basic system of PJL in this paper, it should be clear that PJL is a promising approach for integrating uncertainty formalisms into justification reasoning. As the starting point of a fruitful project, we can see several possible directions for further research including the extension of the basic system with axioms about beliefs like D, 4, 5; the decidability and complexity of the PJL; a constructive proof of the Realization Theorem based on sequent systems [Goetschi and Kuznets, 2012]; multi-agent PJL; probabilistic justification logic based on probability modalities [van der Hoek, 1997]; the applications of PJL to truth maintenance systems [de Kleer, 1993], formal argumentation [Ferretti *et al.*, 2014; Letia and Groza, 2012], and belief fusion [Ghari, 2014a].

Acknowledgments

This work was partially supported by the Ministry of Science and Technology of Taiwan under Grants 103-2410-H-346-007-MY2 and NSC 103-2221-E-001-032. We wish to thank the anonymous referees for their constructive suggestions.

References

[Artemov and Fitting, 2012] S. Artemov and M. Fitting. Justification logic. In E.N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2012 edition, 2012.

[Artemov and Nogina, 2005] S. Artemov and E. Nogina. Introducing justification into epistemic logic. *Journal of Logic and Computation*, 15(6):1059–1073, 2005.

[Artemov, 2001] S. Artemov. Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, 7:1–36, 2001.

[Artemov, 2008] S. Artemov. The logic of justification. *The Review of Symbolic Logic*, 1:477–513, 2008.

[Blackburn *et al.*, 2001] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, Cambridge, United Kingdom, 2001.

[de Kleer, 1993] J. de Kleer. A perspective on assumption-based truth maintenance. *Artificial Intelligence*, 59(1-2):63–67, 1993.

[Dubois and Prade, 1988] D. Dubois and H. Prade. An introduction to possibilistic and fuzzy logics. In P. Smets, A. Mamdani, D. Dubois, and H. Prade, editors, *Non-Standard Logics for Automated Reasoning*, pages 253–286. Academic Press, 1988.

[Dubois *et al.*, 1991] D. Dubois, J. Lang, and H. Prade. Fuzzy sets in approximate reasoning, part 2: logical approaches. *Fuzzy Sets and Systems*, 40:203–244, 1991.

[Dubois *et al.*, 1994] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In D.M. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol 3 : Nonmonotonic Reasoning and Uncertain Reasoning*, pages 439–513. Clarendon Press - Oxford, 1994.

[Dubois *et al.*, 2013] D. Dubois, E. Lorini, and H. Prade. Bipolar possibility theory as a basis for a logic of desires and beliefs. In W. Liu, V.S. Subrahmanian, and J. Wijsen, editors, *Scalable Uncertainty Management*, volume 8078 of *LNCS*, pages 204–218. Springer, 2013.

[Esteva *et al.*, 1997] F. Esteva, P. Garcia, L. Godo, and Ri. O. Rodríguez. A modal account of similarity-based reasoning. *International Journal of Approximate Reasoning*, 16(3-4):235–260, 1997.

[Fagin *et al.*, 1996] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1996.

[Ferretti *et al.*, 2014] E. Ferretti, M.L. Errecalde, A.J. Garcia, and G.R. Simari. A possibilistic defeasible logic programming approach to argumentation-based decision-making. *Journal of Experimental & Theoretical Artificial Intelligence*, 26(4):519–550, 2014.

[Fitting, 2005] M. Fitting. The logic of proofs, semantically. *Annals of Pure and Applied Logic*, 132(1):1–25, 2005.

[Fitting, 2013] M. Fitting. Realization using the model existence theorem. *Journal of Logic and Computation*, 2013.

[Ghari, 2014a] M. Ghari. Distributed knowledge justification logics. *Theory of Computing Systems*, 55(1):1–40, 2014.

[Ghari, 2014b] M. Ghari. Justification logics in a fuzzy setting. *preprint arXiv:1407.4647*, 2014.

[Goetschi and Kuznets, 2012] R. Goetschi and R. Kuznets. Realization for justification logics via nested sequents: Modularity through embedding. *Annals of Pure and Applied Logic*, 163(9):1271–1298, 2012.

[Hájek, 1998] P. Hájek. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publisher, 1998.

[Halpern, 2003] J.Y. Halpern. *Reasoning about Uncertainty*. MIT Press, 2003.

[Hintikka, 1962] J. Hintikka. *Knowledge and Belief*. Cornell University Press, 1962.

[Kokkinis *et al.*, 2014] I. Kokkinis, P. Maksimović, Z. Ognjanović, and T. Studer. First steps towards probabilistic justification logic, 2014. preprint.

[Letia and Groza, 2012] I. Letia and A. Groza. Arguing with justifications between collaborating agents. In P. McBurney, S. Parsons, and T. Rahwan, editors, *Argumentation in Multi-Agent Systems*, volume 7543 of *LNCS*, pages 102–116. 2012.

[Liau and Lin, 1992] C.J. Liau and I.P. Lin. Quantitative modal logic and possibilistic reasoning. In B. Neumann, editor, *Proceedings of the 10th ECAI*, pages 43–47. John Wiley & Sons. Ltd, 1992.

[Liau and Lin, 1996] C.J. Liau and I.P. Lin. Possibilistic reasoning - a mini-survey and uniform semantics. *Artificial Intelligence*, 88(1-2):163–193, 1996.

[Milnikel, 2014] R.S. Milnikel. The logic of uncertain justifications. *Annals of Pure and Applied Logic*, 165:305–315, 2014.

[Mkrtychev, 1997] A. Mkrtychev. Models for the logic of proofs. In S.I. Adian and A. Nerode, editors, *Proceedings of the 4th International Symposium on Logical Foundations of Computer Science (LFCS)*, volume 1234 of *LNCS*, pages 266–275. Springer, 1997.

[van der Hoek, 1997] W. van der Hoek. Some considerations on the logics $P_F D$: A logic combining modality and probability. *Journal of Applied Non-Classical Logics*, 7(3):287–307, 1997.

[Zadeh, 1978] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3–28, 1978.