Journal of Applied Non-Classical Logics
Vol. 00, No. 00, Month 201X, 1–16

RESEARCH ARTICLE

A new framework for justification logic

A. Giordani[a*]

[a] *Department of Phylosophy, Catholic University of Milan, Milan, Italy*;
(*Received 00 Month 201X; final version received 00 Month 201X*)

The logic of justification provides an in-depth analysis of the epistemic states of an agent. This paper aims at solving some of the problems to which the common interpretation of the operators of justification logic is subject, by providing a framework in which a crucial distinction between potential and explicit justifiers is exploited. The paper is subdivided into three sections. The first section offers an introduction to a basic system **LJ** of justification logic and to the problems concerning its interpretation. In the second section, three new systems of justification logic are introduced and characterized with respect to an appropriate semantics. The final section shows why the highlighted problems do not afflict the new systems and how it is possible to interpret **LJ** in the new framework.

Keywords: justification logic; epistemic logic; explicit knowledge; Fitting's semantics

1. Introduction

The logic of justification is a well-developed discipline where a number of logical systems, which allow us to provide a more in-depth analysis of the epistemic states of an agent, are studied. The present paper aims at solving some of the problems to which the common interpretation of the operators of justification logic is subject by providing a new framework in which a crucial distinction between potential and explicit justifiers is introduced.

The paper is subdivided into three sections. In the present section, a short introduction to a basic system **LJ** of justification logic is offered and the principal problems concerning its interpretation are outlined. In the central section, three new systems of justification logic are introduced and characterized with respect to an appropriate modal semantics. Finally, in the last section, it is shown why the highlighted problems do not afflict the new systems and how it is possible to interpret **LJ** in this new framework.

1.1 *Justification logic*

The basic idea underlying the logic of justification is to analyze propositions like $\langle\varphi$ is known$\rangle$ as $\langle\varphi$ is known in virtue of justifier $t\rangle$, thus making explicit the epistemic ground, or *justifier*, in virtue of which a proposition is *known* to be true.[1] The internal structure

*Email: alessandro.giordani@unicatt.it

[1]See (Artemov, 2001, 2008; Artemov & Nogina, 2004, 2005) and (Fitting, 2008, 2009) for a comprehensive introduction to the topic and a presentation of the connections between operators on justifiers and operators on proofs. In (Artemov & Nogina, 2005) the related correspondence between such operators and functions on codes of proof in metamathematics is also proposed.

of a justifier is then further analyzed. In particular, composite justifiers are introduced, which are constructed by applying different operators to basic building blocks. Typically, the following operators are taken into account.

(1) the $\cdot$ operator, which takes two justifiers t and s as inputs and returns a justifier $t \cdot s$ that justifies all the propositions that follow by applying *modus ponens* to propositions justified by t and s.
(2) the $+$ operator, which takes two justifiers t and s as inputs and returns a justifier $t + s$ that justifies all the propositions that are either justified by t or justified by s.
(3) the ! operator, which takes one justifier t as input and returns a justifier $!t$ that justifies all the propositions of the form $t : \varphi$, provided that φ is justified by t.

Hence, the language $\mathcal{L}(\mathbf{LJ})$ of basic systems of justification logic includes a set $Tm(\mathcal{L}(\mathbf{LJ}))$ of terms and a set $Fm(\mathcal{L}(\mathbf{LJ}))$ of formulas defined by the following rules:

- $t := c_i \mid x_i \mid (t \cdot s) \mid (t + s) \mid !t$
- $\varphi := p_i \mid \neg\varphi \mid \varphi \wedge \psi \mid t : \varphi$

where $\{c_i\}_{i\in\mathbb{N}}$ is a set of justification constants, $\{x_i\}_{i\in\mathbb{N}}$ is a set of justification variables, and $\{p_i\}_{i\in\mathbb{N}}$ is a set of propositional variables.

1.2 *A basic system of justification logic*

The standard axiomatic system for a basic logic of factive justification **LJ** is as follows:

LJ0: a sufficient set of classically valid formulas
LJ1: $t : (\varphi \to \psi) \to (s : \varphi \to t \cdot s : \psi)$
LJ2: $t : \varphi \vee s : \varphi \to t{+}s : \varphi$
LJ3: $t : \varphi \to !t : t : \varphi$
LJ4: $t : \varphi \to \varphi$
LJR: $c_i : \varphi$, where φ is an axiom in **LJ0-LJ4** and c_i is justification constant.[1]
MP: if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

Remark 1. It is assumed that explicit justification is factive, i.e., sufficient for knowledge. Actually, $t : \varphi$ states that t is one of the explicit justifiers in virtue of which φ is known to be true. In order to capture the notion of explicit justifier that is not sufficient for knowledge, axiom $t : \varphi \to \varphi$ has to be dropped.

A Kripke-style semantic framework for several systems of justification logic was introduced by Fitting (2005), exploiting ideas proposed by Fagin & al. (1995). In Fitting's semantics for **LJ**, a frame is a tuple $\langle W, \mathcal{R}, \mathcal{E}\rangle$, where W is a non-empty set of states, $\mathcal{R}$ is a reflexive and transitive relation on W, and $\mathcal{E}$ is a selection function from states and justifiers to sets of formulas. Intuitively, $\mathcal{R}$ is the relation that connects a possible state w with all the states that are consistent with the epistemic perspective that characterizes the epistemic agent at w, while $\mathcal{E}$ is the function that selects the set of formulas that can be justified by a certain justifier at a certain state. Hence, $\varphi \in \mathcal{E}(w, t)$ states that, at w, t is a justifier that can serve as possible evidence for φ. It is worth noting that possible evidence can be both *objectively defective*, since $\varphi \in \mathcal{E}(w, t)$ does not imply that φ is true at w, and *subjectively defective*, since $\varphi \in \mathcal{E}(w, t)$ does not imply that φ

[1] **LJ** is considered here as the basic system of standard factive justification logic, even though a simpler system can be obtained by dropping axiom **LJ3**: see (Fitting, 2008), where **LJ** is situated in a wider context. It is also to notice that the reference to a constant specification is avoided by allowing that any justification constant justifies any axiom: see (Fitting, 2009).

is implicitly known at w. The fundamental conditions on $\mathcal{E}$, which ensure a complete characterization of the system, are as follows:

- $\varphi \rightarrow \psi \in \mathcal{E}(w,t)$ and $\varphi \in \mathcal{E}(w,s) \Rightarrow \psi \in \mathcal{E}(w,t$ $cdots)$
- $\mathcal{E}(w,t) \cup \mathcal{E}(w,s) = \mathcal{E}(w,t+s)$
- $\varphi \in \mathcal{E}(w,t) \Rightarrow t:\varphi \in \mathcal{E}(w,!t)$
- $\mathcal{R}(w,v) \Rightarrow \mathcal{E}(w,t) \subseteq \mathcal{E}(v,t)$

The truth condition of $t:\varphi$ is then

$$M,w \models t:\varphi \Leftrightarrow \forall v(\mathcal{R}(w,v) \Rightarrow M,v \models \varphi) \text{ and } \varphi \in \mathcal{E}(w,t)$$

In this way, $M,w \models t:\varphi$ implies

(i) that φ is true, since $\mathcal{R}$ is reflexive
(ii) that φ is justified, in particular by t, since $\varphi \in \mathcal{E}(w,t)$
(iii) that φ is implicitly known, since $\forall v(\mathcal{R}(w,v) \Rightarrow M,v \models \varphi)$

Hence, $t:\varphi$ is interpreted as φ is explicitly known, where t is one of the explicit justifiers in virtue of which φ is known.

1.3 *Problems with justification logic*

In Fitting's semantics, t justifies the knowledge of φ when t is admissible for φ and φ is implicitly known, where being admissible might be intended as being a possible evidence for φ. It is known that this conception gives rise to some interesting problems.[1] In particular, suppose that an agent implicitly knows φ and that t and s are admissible for φ. Then

(1) it is necessary for the agent to explicitly know φ in virtue of t.
(2) it is necessary for the agent to be justified in knowing φ in virtue of both t and s.
(3) it is necessary for the agent to explicitly know that she explicitly knows φ in virtue of t.

These consequences appear to be problematic because we acknowledge an intuitive distinction between being an *admissible* piece of evidence for a proposition and being an *admitted* piece of evidence for it. Hence, we typically acknowledge a distinction between a *potential* sense of being a justifier for φ, according to which t justifies φ independently of the existence of some agent that explicitly thinks of t as such a justifier, and an *explicit* sense of being a justifier for φ, according to which t actually justifies φ for a certain agent.

In particular, let us mention two significant cases where the introduction of this distinction is crucial for modeling important problems in the analysis of knowledge.[2]

Case 1: *Aristotle's treatment of Meno's dilemma.*

In the first chapter of the first book of *Posterior Analytics*, Aristotle provides his solution to Meno's dilemma[3]. The version of Meno's dilemma explicitly discussed by Aristotle is

[1]See (Baltag & al., 2014) for a critical analysis focusing on the tension between the fact that the knowledge provided by a justifier, in the case in which the justifier actually provides knowledge, is explicit and the fact that agents cannot be in a position to acknowledge a justifier as such.
[2]I'm grateful to an anonymous referee for having suggested to discuss the following cases, which are indeed paradigmatic for the application of the distinction I'm dealing with.
[3]A delightful presentation of both the history of the dilemma and Aristotle's solution is provided in (Fine, 2014).

an argument aimed at showing that a man is unable to inferentially come to know what he do not already know. To be sure, so the argument, either you already know what you are learning, and so you can't actually coming to know it, or you don't know what you are learning, and so you can't come to know it either, as you would be unable to acknowledge it. The argument is based on the general premise to the effect that either you actually know what you are learning or you absolutely don't know what you are learning. According to Aristotle, this premise is unjustified, since you might be in a different and intermediate position. Indeed, he assumes a distinction between what is known universally and what is known *simpliciter*, and to illustrate the distinction he considers the case of a learner who knows that every triangle has interior angles whose sum is equal to the sum of two right angles, but ignores that a certain figure inscribed in a semicircle is a triangle. Then, what Aristotle assumes is that, while the learner does not know *simpliciter* that this figure has interior angles whose sum is equal to the sum of two right angles, since he does not know that it is a triangle, still he does know universally that this figure has interior angles whose sum is equal to the sum of two right angles, since this figure is actually a triangle and he knows that every triangle has such interior angles. Hence, what Aristotle assumes is that nothing prevents you from being in a position in which (i) you know universally what you are learning and still (ii) you do not know *simpliciter* it, so that you can actually learn it, since nothing prevents one's knowing already in a sense, and not knowing in a different sense, what one is learning. In conclusion, in such cases, the learner is considered as having potential knowledge and learning is considered as turning this potential knowledge into explicit knowledge.

Case 2: *Prawitz's treatment of Dummett's dilemma.*

In a number of places[1], Dummett advances that a constructivist can accept a timeless notion of truth, provided that he accepts (i) to equate truth with provability and (ii) to interpret provability *neither* as existence of a proof living in a realm of proofs which is independent of our knowledge *nor* as existence of a proof that is actually constructed, or that will eventually be constructed. As finding a way between the last two constraints has proved to be particularly difficult, Dummett's indication turns into a dilemma for the constructivist, for either you accept that truth is relative to time or you accept that truth is grounded on entities that are independent of our knowledge. The solution proposed by Prawitz is to distinguish between actual and potential existence of a proof and to identify abstract existence of a proof with potential existence:[2]

> We can then say that the correctness of an assertion requires the actual existence of a proof, while the truth of the asserted proposition requires only (and is identical with) the potential existence of a proof of the proposition (Prawitz, 1998, p. 48).

This move allows us to obtain an intelligible account of truth from a constructivist point of view. In particular, it allows us to maintain both that proofs are in principle knowable, since they are defined as something whose existence is actualized by epistemic acts of an agent, and that it is possible for a proposition to be indeterminate as to the truth value, since no claim is made as to the abstract existence of a proof for any proposition or its negation.

The preceding cases are about proofs, which are prototypes of justifiers, and highlight the limitations of a framework in which the distinction between potential and explicit proofs is neglected. In particular, the standard semantic framework of justification logic

[1] See, for instance, (Dummett, 2000, p.12) and (Dummett, 1998).
[2] See (Prawitz, 1994, 1998) for a full account. See also (Martin-Löf, 1987, 1998) for developments of a very similar conception, and (Dean & Kurokawa, 2010) for a general discussion.

is unable to model them, both because it does not allow for two kinds of justifiers and because, once the distinction is granted, the interpretation of $\varphi \in \mathcal{E}(w,t)$ turns out to be tricky. To be sure, assuming that $\varphi \in \mathcal{E}(w,t)$ states that t is a *potential* justifier for φ seems to be the better choice, but it is at odds with the fact that $\varphi \in \mathcal{E}(w,t)$ implies that φ is *explicitly* known, and not only potentially known. Alternatively, assuming that $\varphi \in \mathcal{E}(w,t)$ states that t is an *explicit* justifier for φ seems to be at odds with the foregoing conditions on $\mathcal{E}$, since it seems too strong to assume that any justifier which is admissible for a proposition is always explicitly admitted by the agent for that proposition. Finally, in the last two cases, an implicit or explicit claim is made to the effect that all the potentially existing proofs are accessible, so that we are encouraged to find a tool for describing the potential dynamics of the epistemic agent.

In general, three approaches are available in order to face the aforementioned problems:

(1) sticking to the implicit part of the system, while changing its explicit part;
(2) sticking to the explicit part of the system, while changing its implicit part;
(3) changing both the implicit and the explicit part of the system.

Approach 1 was recently pursued in (Baltag & al., 2014) and some interesting results are emerged. Approach 2 is the one I am going to pursue here, while the last approach is left for future work. Note that I don't want to imply that the solution proposed here is more insightful that the one proposed in (Baltag & al., 2014), but only to propose a solution within a semantical framework where the distinction between the *potential and objective* side of the justification and the *explicit and subjective* side of the justification is mirrored in the distinction between what is interpreted just in terms of the accessibility relations on the set of epistemic states and what is interpreted both in terms of the accessibility relations on the set of epistemic states and in terms of the explicit selection function.

2. Logic of potential justifiers and justification

Let us introduce the following language.

Definition 1. *Language $\mathcal{L}$*

The language $\mathcal{L}$ of basic systems of justification logic includes a set $Tm(\mathcal{L})$ of terms and a set $Fm(\mathcal{L})$ of formulas defined by the following rules:

- $t := c_i \mid x_i \mid (t \cdot s) \mid (t + s) \mid\, !t$
- $\varphi := p_i \mid \neg\varphi \mid \varphi \wedge \psi \mid [\mathbf{F}]\varphi \mid [t]\varphi \mid \mathbf{E}(t,\varphi)$

where $\{c_i\}_{i\in\mathbb{N}}$ is a set of justification constants, $\{x_i\}_{i\in\mathbb{N}}$ is a set of justification variables, and $\{p_i\}_{i\in\mathbb{N}}$ is a set of propositional variables. The other logical connectives are defined in the usual way.

In this framework, an agent is able to increase her knowledge by performing inferential steps,[1] thought of as constructions of composite justifiers from available ones. The intended interpretation of the modal and epistemic propositions is then as follows.

(1) $[\mathbf{F}]\varphi$ states that φ is actually true and it will be always true after any number of inferential steps. Hence, $[\mathbf{F}]$ encodes a notion of stability through epistemic dynamics. This operator is interpreted by introducing a relation R on the set of epistemic states,

[1]Thus, the point of view developed by Duc (1997) on the epistemic dynamics is adopted here. See also (van Benthem & Velazquez-Quesada, 2010) for a development of this idea along slightly different lines.

in such a way that w is R-related to v if and only if v is the result of performing a finite number of inferential steps. Hence, R is assumed to be reflexive and transitive.

(2) $[t]\varphi$ states that φ can be objectively justified by t, namely that t is an *admissible justifier* for φ, and this relation is assumed to be wholly objective, and so independent of any agent admitting t as a justifier for φ. Any operator $[t]$ is interpreted by introducing a function S which returns the set of epistemic states that are consistent with what can be justified by t at a world w. Thus, $v \in S(w,t)$ if and only if v makes true every proposition that can be justified by t in w.

(3) $\mathbf{E}(t,\varphi)$ states that φ is subjectively justified by t, namely that t is an *admitted justifier* for φ, i.e. that there exists an agent which admits t as a justifier for φ. Any operator $\mathbf{E}(t,\cdot)$ is interpreted by introducing a selection function E which returns the set of formulas that are justified, for the agent, by t at a world w. Thus, $\varphi \in E(w,t)$ if and only if the agent acknowledges φ as justified by t.

Remark 2. Since objective justification is intended as potential justification, it is assumed that, if t justifies φ, then t is also a justifier for any proposition that is logically implied by φ. Hence, objective justification, in contrast to subjective justification, is closed with respect to logical consequence.

Let us summarize our terminology in the following table.

objective sense: $[t]\varphi$	subjective sense: $\mathbf{E}(t,\varphi)$
t objectively justifies φ	t subjectively justifies φ
t is an admissible justifier for φ	t is an admitted justifier for φ
t is a potential justifier for φ	t is an explicit justifier for φ

Remark 3. The fact that t is an admissible justifier for φ implies that t is a potential justifier for φ in two different senses: t is potential both because it can be acknowledged by the agent as a justifier for φ (potential basis of subjective justification) and because it can be constituted by a set of premises from which φ follows (potential basis of objective justification).

Remark 4. The fact that t is an admissible justifier for φ is not sufficient for the truth of φ, since objective justification is typically not assumed to be a sufficient condition for truth. The fact that t is an admitted justifier for φ is not sufficient for concluding that φ is indeed justified by t, since subjective justification is not assumed here to be a sufficient condition for objective justification.

Remark 5. The fact that the agent has an objective justifier for φ is stated by a proposition like $[t]\varphi \wedge \mathbf{E}(t,\varphi)$. The difference between $\mathbf{E}(t,\varphi)$ and $[t]\varphi \wedge \mathbf{E}(t,\varphi)$ is then that $\mathbf{E}(t,\varphi)$ states that the agent just *admits* that φ is justified by t, while $[t]\varphi \wedge \mathbf{E}(t,\varphi)$ states that the agent *acknowledges*, or correctly admits, that φ is justified by t, thus adding the information that the subjective assumption is indeed right.

2.1 *Systems*

In the light of what we said above, let us consider the following three systems of justification logic.

System LJ0
Group 1. Propositional axioms and rule

P: φ, provided φ is propositionally valid
MP: if $\vdash \varphi$ and $\vdash \varphi \rightarrow \psi$, then $\vdash \psi$

Group 2. Modal axioms and rule

F1: $[\mathbf{F}](\varphi \to \psi) \to ([\mathbf{F}]\varphi \to [\mathbf{F}]\psi)$	**J1**: $[t](\varphi \to \psi) \to ([t]\varphi \to [t]\psi)$
F2: $[\mathbf{F}]\varphi \to \varphi$	**J2**: $[t]\varphi \vee [s]\varphi \to [t+s]\varphi$
F3: $[\mathbf{F}]\varphi \to [\mathbf{F}][\mathbf{F}]\varphi$	**J3**: $[t+s]\varphi \to [t \cdot s]\varphi$
F4: $[t]\varphi \to [\mathbf{F}][t]\varphi$	**J4**: $[t]\varphi \to [!t][t]\varphi$
F5: $\neg[t]\varphi \to [\mathbf{F}]\neg[t]\varphi$	**J5**: $[t]\varphi \to \varphi$

Group 2 introduces the axioms characterizing $[\mathbf{F}]$ and $[t]$, for any t. Axioms **F1-F3** and **RF** characterize $[\mathbf{F}]$ as a standard modal operator corresponding to a reflexive and transitive relation on the set of states, while axioms **F4** and **F5** state that all the justification relations are stable with respect to $[\mathbf{F}]$. Axiom **J1** and **RJ** characterize $[t]$ as a standard modal operator corresponding to a general relation on the set of states. **J2** states that $t+s$ is just the sum of justifiers t and s, so that a proposition is justified by $t+s$ when it is justified by t or by s. **J3** states that $t \cdot s$ is at least as strong as $t+s$, in accordance with the assumption that $t \cdot s$ is the closure $t+s$ with respect to the application of modus ponens. **J4** states that $!t$ justifies $[t]\varphi$ precisely when t justifies φ. Finally, **J5** is a sufficiency assumption on justifiers: justifiers are required to provide knowledge, and so to ensure the truth of the justified proposition.

LJ0 is our starting system. It is very elementary concerning the description of the agent, since no axioms on subjective justification is introduced. The following two systems are developed in order to improve that description.

System LJI (*ideal system*)**: LJ0** plus the axioms of the following group

Group 2. Ideal epistemic axioms

IE1: $\mathbf{E}(t, \varphi \to \psi) \to (\mathbf{E}(s, \varphi) \to \mathbf{E}(t \cdot s, \psi))$
IE2: $\mathbf{E}(t, \varphi) \vee \mathbf{E}(s, \varphi) \to \mathbf{E}(t+s, \varphi)$
IE3: $\mathbf{E}(t, \varphi) \to \mathbf{E}(!t, [t]\varphi)$
IE4: $\mathbf{E}(t, \varphi) \to \mathbf{E}(t, \top)$

System LJ1 (*non-ideal system*)**: LJ0** plus the axioms of the following group

Group 3.2. non-ideal epistemic axioms

E1: $\mathbf{E}(t, \varphi \to \psi) \to (\mathbf{E}(s, \varphi) \to \neg[\mathbf{F}]\neg\mathbf{E}(t \cdot s, \psi))$
E2: $\mathbf{E}(t, \varphi) \vee \mathbf{E}(s, \varphi) \to \neg[\mathbf{F}]\neg\mathbf{E}(t+s, \varphi)$
E3: $\mathbf{E}(t, \varphi) \to \neg[\mathbf{F}]\neg\mathbf{E}(!t, [t]\varphi)$
E4: $\mathbf{E}(t, \varphi) \to \mathbf{E}(t, \top)$

The epistemic axioms provide a characterization of the abilities of the agent. Axioms in group 3.1 correspond to the standard principles assumed in justification logic. Still, in the present framework, they can only be introduced as a set of ideal assumptions on the agent. To be sure, **IE1-IE3** state that, if the agent has admitted that certain justifiers justify certain propositions, then she *immediately admits* that an appropriate composite justifier justifies some other proposition, e.g., if the agent has admitted that t justifies φ, then she immediately admits that $t+s$ justifies φ. Axioms in group 3.2 are the corresponding non-ideal version. Thus, **IE1-IE3** state that, if the agent has admitted that certain justifiers justify certain propositions, then she only is *in a position to admit* that an appropriate composite justifier justifies some other proposition, since she can achieve the appropriate justifier by performing some inferential steps. As a consequence, a significant distinction between what the agent actually admits and what she's able to achieve can now be drawn. Finally, axiom **IE4** = **E4** is common to both groups and states that, if t is admitted as a justifier for some proposition, then it is admitted as a justifier for the true proposition $\top$. This axiom is just a means for deriving some interesting results concerning the construction of justifiers, since $\mathbf{E}(t, \top)$ can be construed as stating that the agent has t at her disposal.

2.2 *Semantics*

Definition 2. *Basic frame for $\mathcal{L}$.*

A basic frame for $\mathcal{L}$ is a tuple $\langle W, R, S, E \rangle$, where

W is a set of epistemic states
$R \subseteq W \times W$ is a reflexive and transitive relation on W
$S : W \times Tm(\mathcal{L}) \to \wp(W)$, is a selection function for terms in $\mathcal{L}$
$E : W \times Tm(\mathcal{L}) \to \wp(Fm(\mathcal{L}))$, is a subjective selection function for terms in $\mathcal{L}$

In addition S has to satisfy the following conditions.

S1: $S(w, t+s) \subseteq S(w, t) \cap S(w, s)$
S2: $S(w, t \cdot s) \subseteq S(w, t+s)$
S3: $v \in S(w, !t) \Rightarrow S(v, t) \subseteq S(w, t)$
S4: $w \in S(w, t)$
SR: $R(w, v) \Rightarrow S(w, t) = S(v, t)$

Definition 3. *Explicit frame for $\mathcal{L}$.*

An explicit frame for $\mathcal{L}$ is a basic frame for $\mathcal{L}$ where E satisfies:

E1: $\varphi, \varphi \to \psi \in E(w, t) \Rightarrow \psi \in E(v, t)$, for some v, such that $R(w, v)$
E2: $E(w, t) \cup E(w, s) \subseteq E(w, t+s)$, for some v, such that $R(w, v)$
E3: $\varphi \in E(w, t) \Rightarrow [t]\varphi \in E(v, !t)$, for some v, such that $R(w, v)$
E4: $E(w, t) \neq \varnothing \Rightarrow \top \in E(v, t)$

Definition 4. *Ideal explicit frame for $\mathcal{L}$.*

An ideal explicit frame for $\mathcal{L}$ is a basic frame for $\mathcal{L}$ where E satisfies:

IE1: $\varphi, \varphi \to \psi \in E(w, t) \Rightarrow \psi \in E(v, t)$
IE2: $E(w, t) \cup E(w, s) \subseteq E(w, t+s)$, for some v
IE3: $\varphi \in E(w, t) \Rightarrow [t]\varphi \in E(v, !t)$
IE4: $E(w, t) \neq \varnothing \Rightarrow \top \in E(v, t)$

Definition 5. *Model for $\mathcal{L}$.*

A model for $\mathcal{L}$ is a tuple $M = \langle W, R, S, E, V \rangle$, where

$\langle W, R, S, E \rangle$ is a frame for $\mathcal{L}$
$V : \mathsf{P} \to \wp(W)$, is a valuation function for propositional variables in $\mathcal{L}$

As usual, a valuation function for propositional variables is introduced as a function that assigns to any propositional variable a set of epistemic states, to be considered either as a true proposition or as the set of worlds where the proposition denoted by the variable is true.

Definition 6. *Truth at a world in a model for $\mathcal{L}$.*

The notion of truth of a formula is recursively defined as follows:

$M, w \models p \Leftrightarrow w \in V(p)$
$M, w \models \neg\varphi \Leftrightarrow M, w \not\models \varphi$

$M, w \models \varphi \wedge \psi \Leftrightarrow M, w \models \varphi$ *and* $M, w \models \psi$
$M, w \models [\mathbf{F}]\varphi \Leftrightarrow M, w \models \varphi$, for all v such that $R(w, v)$
$M, w \models [t]\varphi \Leftrightarrow M, w \models \varphi$, for all v such that $v \in S(w, t)$
$M, w \models \mathbf{E}(t, \varphi) \Leftrightarrow \varphi \in E(w, t)$

2.3 *Characterization*

Let us now show that the previous systems can be completely characterized by the class of basic frames, the class of explicit frames, and the class of ideal explicit frames. It is not difficult to show that the axioms in groups 1 and 2 are valid with respect to the class of all basic frames and that the rules preserve validity. Similarly, it is not difficult to show that the explicit epistemic axioms and the ideal explicit epistemic axioms are valid with respect to the class of all explicit frames and the class of all ideal explicit frames. Hence, let us focus on completeness.[1]

Theorem 7. ***LJ0*** *is sound and strongly complete with respect to the class of all basic frames for* $\mathcal{L}$.

Completeness: the proof is based on a canonicity argument.

Proof. Let $w/[\mathbf{F}] = \{\varphi \,|\, [\mathbf{F}]\varphi \in w\}$ and $w/[t] = \{\varphi \,|\, [t]\varphi \in w\}$ for all $t \in Tm(\mathcal{L})$. Then, the canonical model is the tuple $M = \langle W, R, S, E, V \rangle$, where

- W is the set of maximally **LJ0**-consistent sets of formulas
- R is such that $R(w, v) \Leftrightarrow w/[\mathbf{F}] \subseteq v$
- S is such that $v \in S(w, t) \Leftrightarrow w/[t] \subseteq v$
- E is such that $E(w, t) = \{\varphi \,|\, \mathbf{E}(t, \varphi) \in w\}$

Corollary 8. $v \in S(w, t) \cap S(w, s) \Leftrightarrow w/[t] \cup w/[s] \subseteq v$.

Straightforward:
$v \in S(w, t) \cap S(w, s) \Leftrightarrow v \in S(w, t)$ and $v \in S(w, s)$
$v \in S(w, t) \cap S(w, s) \Leftrightarrow w/[t] \subseteq v$ and $w/[s] \subseteq v$
$v \in S(w, t) \cap S(w, s) \Leftrightarrow w/[t] \cup w/[s] \subseteq v$

Lemma 9. *M is a model for $\mathcal{L}$.*

Step 1. $R \subseteq W \times W$ is a reflexive and transitive relation on W.
This is straightforward, by the definition of R and axioms **F2** and **F3**.

Step 2. $S : W \times Tm(\mathcal{L}) \to \wp(W)$: by definition of S.
The conditions on S are satisfied.

$S1 : S(w, t+s) \subseteq S(w, t) \cap S(w, s)$. Since $w \in W$, $[t]\varphi \vee [s]\varphi \in w \Rightarrow [t+s]\varphi \in w$, by **J2**. Therefore $w/[t] \cup w/[s] \subseteq w/[t+s]$, and so $w/[t+s] \subseteq v \Rightarrow w/[t] \cup w/[s] \subseteq v$. Hence, $S(w, t+s) \subseteq S(w, t) \cap S(w, s)$, by corollary 1.

$S2 : S(w, t \cdot s) \subseteq S(w, t+s)$. Since $w \in W$, $[t+s]\varphi \in w \Rightarrow [t \cdot s]\varphi \in w$, by **J3**. Therefore $w/[t+s] \subseteq w/[t \cdot s]$, and so $w/[t \cdot s] \subseteq v \Rightarrow w/[t+s]$. Hence, $S(w, t \cdot s) \subseteq S(w, t+s)$, by the definition of S.

[1] See (Blackburn & al., 2001, chap. 4) for an extensive introduction to modal completeness proof, and in particular completeness by canonicity. In what follows I will omit the standard parts and definitions and focus on the new parts of the proofs only.

$S3 : v \in S(w, !t) \Rightarrow S(v, t) \subseteq S(w, t)$. It suffices to prove that, if $w/[!t] \subseteq v$, then $w/[t] \subseteq v/[t]$. Suppose $w/[!t] \subseteq v$ and $\varphi \in w/[t]$. Then $[t]\varphi \in w$, so that $[!t][t]\varphi \in w$, by **J4** and $w \in W$. Therefore $[t]\varphi \in v$, and so $w/[t] \subseteq v/[t]$.

$S4 : w \in S(w, t)$. It is an immediate consequence of **J5**.

$SR : R(w, v) \Rightarrow S(w, t) = S(v, t)$. It suffices to prove that, if $w/[\mathbf{F}] \subseteq v$, then $w/[t] = v/[t]$. Suppose $w/[\mathbf{F}] \subseteq v$ and $\varphi \in w/[t]$. Then $[t]\varphi \in w$, so that $[\mathbf{F}][t]\varphi \in w$, by **F4** and $w \in W$. Thus, $[t]\varphi \in w/[\mathbf{F}]$, and so $\varphi \in w/[t]$. Suppose now $\varphi \notin w/[t]$. Then $[t]\varphi \notin w$, so that $[\mathbf{F}]\neg[t]\varphi \in w$, by **F5**, and $w \in W$. Thus, $\neg[t]\varphi \in w/[\mathbf{F}]$, and so $\varphi \notin v/[t]$.

Step 3. $E : W \times Tm(\mathcal{L}) \to \wp(Fm(\mathcal{L}))$: by definition of E.

Lemma 10. *(Truth Lemma)* $M, w \models \varphi \Leftrightarrow \varphi \in w$.

The only interesting cases are the epistemic and the ontic ones. however, in this cases the proof is well known since both $[\mathbf{F}]$ and $[t]$, $t \in Tm(\mathcal{L})$, are standard modalities, in view of **F1**, **RF** and **J1**, **RJ**.

Theorem 1 follows from lemma 1 and lemma 2 in the standard way.

Theorem 11. ***LJ1*** *and* ***LJI*** *are sound and strongly complete with respect to the class of all explicit frames and ideal explicit frames for* $\mathcal{L}$.

Soundness is left as an exercise. Completeness is proved by virtue of a canonicity argument.

Proof. The canonical model is as in theorem 1. The new part of the proof consists in showing that conditions *E1-E5* and *IE1-IE5* on E are satisfied. The axioms **E1**-**E5** and **IE1**-**IE5** are just what we need in order to ensure that. Let us check, by way of example, the less straightforward condition *E1*.

$E1 : \varphi, \varphi \to \psi \in E(w, t) \Rightarrow \psi \in E(v, t)$, for some v, such that $R(w, v)$. Suppose $\varphi, \varphi \to \psi \in E(w, t)$. It has to be shown that there is a v, such that $R(w, v)$ and $\psi \in E(v, t)$, i.e., such that $w/[\mathbf{F}] \subseteq v$ and $\mathbf{E}(t, \varphi) \in v$. As it is known, in order to prove that, it suffices to show that $w/[\mathbf{F}] \cup \{\mathbf{E}(t, \varphi)\}$ is a **LJ1**-consistent set. Hence, suppose $w/[\mathbf{F}] \cup \{E(t, \varphi)\}$ is not **LJ1**-consistent, so that $w/[\mathbf{F}] \vdash \neg\mathbf{E}(t, \varphi)$. Then $w \vdash [\mathbf{F}]\neg\mathbf{E}(t, \varphi)$. Still, by assumption, $\mathbf{E}(t, \varphi), \mathbf{E}(t, \varphi \to \psi) \in w$, and so $\neg[\mathbf{F}]\neg\mathbf{E}(t, \varphi) \in w$, by axiom **E1**. Thus $w \vdash \neg[\mathbf{F}]\neg\mathbf{E}(t, \varphi)$, and so w is not consistent, in contradiction with $w \in W$. In conclusion, $w/[\mathbf{F}] \cup \{\mathbf{E}(t, \varphi)\}$ is **LJ1**-consistent.

3. Solutions of the problems

We are now in a position to solve the problems initially presented and to assess the principles of standard justification logic. Let us introduce the following definition.

Definition 12. *(*actual justification*):* $t : \varphi := [t]\varphi \wedge \mathbf{E}(t, \varphi)$.

Hence, φ is actually known in virtue of justifier t precisely when t is a justifier for φ and the agent explicitly acknowledges that t is a justifier for φ. In accordance with this definition, the condition of being justified in assuming that a certain proposition is true is the conjunction of an objective condition, according to which t is indeed a justifier for the proposition, and a subjective condition, according to which the agent assumes t as a justifier for that proposition.

3.1 *Principles on justification: solutions*

Let us suppose that t and s are admissible justifiers for a true proposition φ and that knowing that φ is defined as being actually justified in assuming the truth of a true proposition, which is: $\varphi \wedge [t]\varphi \wedge \mathbf{E}(t, \varphi)$, for some justifier t.

1. it is *not* necessary for the agent to explicitly know φ in virtue of t.

It is possible for the agent to be in the situation described by: $\varphi \wedge [t]\varphi \wedge \neg\mathbf{E}(t, \varphi)$. Consider a model $M = \langle W, R, S, E, V\rangle$ and a world w such that $w \in S(w, t) = V(\varphi)$ and $E(w, t) = \varnothing$. In this case, while t is an admissible justifier for φ, it is not an admitted justifier for φ, and so the first problem is solved.

2. it is *not* necessary for the agent to be justified in knowing φ in virtue of t and s.

It is possible for the agent to be in the situation described by: $\varphi \wedge [t]\varphi \wedge \mathbf{E}(t, \varphi) \wedge \neg\mathbf{E}(s, \varphi)$. Consider a model $M = \langle W, R, S, E, V\rangle$ and a world w such that $w \in S(w, t) = S(w, s) = V(\varphi)$, $\varphi \in E(w, t)$ and $E(w, s) = \varnothing$. In this case, while t is both an admissible and an admitted justifier for φ, s is not admitted, and so the second problem is solved.

3. it is *not* necessary for the agent to explicitly know that she explicitly knows φ in virtue of t.

It is possible for the agent to be in the situation described by: $\varphi \wedge [t]\varphi \wedge \mathbf{E}(t, \varphi) \wedge \neg\mathbf{E}(!t, [t]\varphi)$. Consider a model $M = \langle W, R, S, E, V\rangle$ and a world w such that $w \in S(w, t) = S(w, !t) = V(\varphi)$, $\varphi \in E(w, t)$ and $E(w, !t) = \varnothing$. In this case, it is possible for the agent to explicitly know φ without explicitly knowing that she knows φ, and so this final problem is also solved.

Case 1: *Aristotle's solution.*

As we have seen, Aristotle's solution aims at showing that, even if one already knows the conclusion in a way, one can learn it, by coming to know it in a different way.[1] Aristotle is confronted with this argument:

(P1) Either the learner knows the conclusion or he does not.
(P2) If he knows the conclusion, he can't learn it.
(P3) If he doesn't know the conclusion, he can't learn it.
(C) Therefore, he can't learn the conclusion.

In providing his solution, Aristotle distinguishes two ways of interpreting **(P2)** and **(P3)**:

(P2.1) If he knows the conclusion *simpliciter*, he can't learn it.
(P2.2) If he knows the conclusion universally, he can't learn it.
(P3.1) If he doesn't know the conclusion universally, he can't learn it.
(P3.2) If he doesn't know the conclusion *simpliciter*, he can't learn it.

He proceeds then in accepting **(P2.1)**, **(P3.1)** and rejecting **(P2.2)**, **(P3.2)** as unjustified. In the particular example he considers, he accepts that someone who knows universally that this figure has angles whose sum is equal to the sum of two right angles, by knowing that every triangle has such angles, can indeed learn what he does not know *simpliciter*, which is that this figure has angles whose sum is equal to the sum of two right angles. In our framework, Aristotle's solution can be represented by assuming the following model of knowledge. Let a be an entity of type T, say a triangle. Let P be a universal property of triangles, say having angles whose sum is equal to the sum of two

[1] See (Fine, 2014, pp. 207–209) for a detailed exposition.

right angles. Let the learner know how to handle universal propositions, so that he has an a priori justifier $c : \forall x(T(x) \rightarrow P(x)) \rightarrow (T(a) \rightarrow P(a))$. Then, the learner

(i) knows *simpliciter* that $P(a)$ iff, for some t and s,

UK. $t : \forall x(T(x) \rightarrow P(x))$
AK. $s : T(a) \wedge \mathbf{E}(c \cdot t \cdot s, P(a))$

(ii) knows universally that $P(a)$ iff, for some t and s,

UK. $t : \forall x(T(x) \rightarrow P(x))$
PK. $[s]T(a) \wedge \neg[\mathbf{F}]\neg\mathbf{E}(c \cdot t \cdot s, P(a))$

Note that in **(i)** what is required is both that the learner has *actual* knowledge of $T(a)$ and that he has constructed the proof $c \cdot t \cdot s$ for $P(a)$, while in **(ii)** it is only required that the learner has *potential* knowledge of $T(a)$ and that he *can* construct $c \cdot t \cdot s$. Hence, knowledge *simpliciter* is grounded in the construction of a proof for the consequent of a known universal implication, given that the antecedent is known, while potential knowledge is grounded in the existence of such a proof and the power the agent has to construct it. Thus, knowing *simpliciter* that $P(a)$ excludes learning that $P(a)$, as learning implies absence of actual knowledge, but knowing universally is consistent with learning, since in order to come to know that $P(a)$ the learner has to actualize his power to construct an admissible proof.

As we can see, both the distinction between actual and potential justification and the introduction of the dynamical operator $[\mathbf{F}]$ are essential for modeling Aristotle's standpoint.

Case 2: *Prawitz's solution.*

Prawitz's solution can be represented as follows. According to a constructivist conception, the distinction between actual knowledge and potential knowledge corresponds to the distinction between having constructed a proof and having access to an abstractly existent proof, and this distinction is mirrored by the distinction between $t : \varphi$ and $[t]\varphi$. In addition, the thesis that every truth is knowable can be analyzed as a conjunction of two more specific principles: (i) the principle that every true proposition has a proof in virtue of which it is true and (ii) the principle that every such proof is accessible to the epistemic agent.[1] In our framework, these principles can be captured by assuming the following axioms.

PC (*Correspondence*): $\varphi \rightarrow [t]\varphi$, for some t
PK (*Knowability*): $[t]\varphi \rightarrow \neg[\mathbf{F}]\neg\mathbf{E}(t, \varphi)$, for all t

In particular, **PK** states that any admissible proof for φ is such that the agent can admit it as a proof of φ by following a certain epistemic path. Let us now define *objective provability* in terms of existence of a proof and *subjective provability* in terms of possibility of achieving such a proof. Then, given **PK**, the following form of equivalence between objective and subjective provability becomes derivable.

Corollary: $[t]\varphi \leftrightarrow \neg[\mathbf{F}]\neg(t : \varphi)$ is derivable in **LJ0**+**P**K.

(i): from right to left

$\neg[\mathbf{F}]\neg(t : \varphi) \vdash_{\mathbf{LJ0}} \neg[\mathbf{F}]\neg[t]\varphi$, by **RF** and **F1**

[1]These principles are put forward in (Martin-Löf, 1998) for satisfying two requirements for the notion of truth originally introduced by Dummett: (C) If a statement is true, there must be something in virtue of which it is true; (K) If a statement is true, it must be in principle possible to know that it is true.

$\neg[\mathbf{F}]\neg(t : \varphi) \vdash_{\mathbf{LJ0}} [t]\varphi$, by **F5**

(ii) from left to right
$[t]\varphi, \neg[\mathbf{F}]\neg\mathbf{E}(t, \varphi) \vdash_{\mathbf{LJ0}} \neg[\mathbf{F}]\neg(t : \varphi)$, by **RF**, **F1** and **F4**
$[t]\varphi \vdash_{\mathbf{LJ0+PK}} \neg[\mathbf{F}]\neg(t : \varphi)$, by **PK**

Hence, if every objective proof can be achieved by the epistemic agent, all that is objectively provable is also subjectively provable. Again, the distinction between actual and potential justification and the introduction of the dynamical operator $[\mathbf{F}]$ are essential for articulating in an appropriate way the thesis according to which every truth is knowable, and so Prawitz's standpoint.[1]

3.2 *Principles on justification: the ideal case*

Suppose we are working in the ideal logic **LJI**. Then, the following principles are derivable concerning actual justification.

JI.1: $t : (\varphi \to \psi) \to (s : \varphi \to t \cdot s : \psi)$

$\vdash_{\mathbf{LJ1}} [t](\varphi \to \psi) \to ([s]\varphi \to [t \cdot s]\psi)$
$\vdash_{\mathbf{LJ1}} \mathbf{E}(t, \varphi \to \psi) \to (\mathbf{E}(s, \varphi) \to \mathbf{E}(t \cdot s, \psi))$
$\vdash_{\mathbf{LJ1}} t : (\varphi \to \psi) \to (s : \varphi \to t \cdot s : \psi)$

JI.5: $t : \varphi \vee s : \varphi \to t + s : \varphi$

$\vdash_{\mathbf{LJ1}} [t]\varphi \vee [s]\varphi \to [t + s]\varphi$
$\vdash_{\mathbf{LJ1}} \mathbf{E}(t, \varphi) \vee \mathbf{E}(s, \varphi) \to \mathbf{E}(t + s, \varphi)$
$\vdash_{\mathbf{LJ1}} t : \varphi \vee s : \varphi \to t + s : \varphi$

JI.3: $t : \varphi \to !t : [t]\varphi$

$\vdash_{\mathbf{LJ1}} [t]\varphi \to [!t][t]\varphi$
$\vdash_{\mathbf{LJ1}} \mathbf{E}(t, \varphi) \to \mathbf{E}(!t, [t]\varphi)$
$\vdash_{\mathbf{LJ1}} t : \varphi \to !t : [t]\varphi$

JI.4: $t : \varphi \to \varphi$

$\vdash_{\mathbf{LJ1}} [t]\varphi \to \varphi$
$\vdash_{\mathbf{LJ1}} t : \varphi \to \varphi$

Thus, **LJI** is almost powerful enough to capture all the principles of the standard justification logic. To be sure, the only axiom that is not derivable in its original version is **LJ3**, since what we can derive is just the slightly different version **J1.3**. In fact, to see that **LJ3** is not derivable in the present system it is sufficient to consider the model $M = \langle W, R, S, E, V\rangle$, where

- $W = \{w, v\}$
- R is such that $R(w, w)$ and $R(v, v)$
- S is such that $S(w, t) = S(v, t) = \{v\}$, for all $t \in Tm(\mathcal{L})$

[1]It is also worth noting that the analysis of the knowability thesis in terms of both **PC** and **PK** is significant for a further discussion of the solution to the paradox of knowability proposed in (Dean & Kurokawa, 2010). I plan to discuss all that in a future work.

- E is such that $E(w,t) = Fm(\mathcal{L})$ and $E(v,t) =$, for all $t \in Tm(\mathcal{L})$
- V is (an otherwise generic modal valuation) such that $V(p) = W$

is a model of $\mathcal{L}$ based on an ideal explicit frame, and so it makes true all the axioms of **LJI**.

Lemma 13. *The following formulas are not true in M:*

(i) $[t]p \wedge \mathbf{E}(t,p) \rightarrow [!t]\mathbf{E}(t,p)$
(ii) $[t]p \wedge \mathbf{E}(t,p) \rightarrow [!t]([t]p \wedge \mathbf{E}(t,p))$

Proof. Since $S(w,t) \subseteq V(p)$, $M,w \models [t]p$. Since $E(w,t) = Fm(\mathcal{L})$, $p \in E(w,t)$, and so $M,w \models \mathbf{E}(t,p)$. Since $E(v,t) =$, $M,v \not\models \mathbf{E}(t,p)$, and so $M,v \not\models [!t]\mathbf{E}(t,p)$ and $M,v \not\models [t]\varphi \wedge \mathbf{E}(t,p)$, since $S(w,t) = v$. □

As a consequence:

$\not\vdash_{\mathbf{LJI}} [t]\varphi \wedge \mathbf{E}(t,\varphi) \rightarrow [!t]\mathbf{E}(t,\varphi)$
$\not\vdash_{\mathbf{LJI}} [t]\varphi \wedge \mathbf{E}(t,\varphi) \rightarrow [!t]([t]\varphi \wedge \mathbf{E}(t,\varphi))$

Still, $[t]\varphi \wedge \mathbf{E}(t,\varphi) \rightarrow [!t]([t]\varphi \wedge \mathbf{E}(t,\varphi))$ is precisely **LJ3**: $t : \varphi \rightarrow !t : t\varphi$, and so **LJ3** is not a theorem of **LJI**. This conclusion immediately raises the question of how **LJ3** can be warranted. **LJ3** is considered intuitive as a principle of a logic of justification on the assumption that $!t$ is precisely what counts as an objective possible piece of evidence for the proposition that t justifies φ. Hence, a first principle we might want to introduce is actually $[t]\varphi \rightarrow [!t][t]\varphi$, which is a pretty intuitive proposition[1]. Furthermore, on the assumption that the agent is able to have an immediate access to $!t$, a second principle we might want to introduce is $t : \varphi \rightarrow !t : [t]\varphi$, which is still an intuitive proposition, stating that $!t$ is soundly held as a justifier for $[t]\varphi$, provided t is soundly held as a justifier for φ. However, what is difficult to see is why $!t$, which is an objective justifier for $[t]\varphi$, should be held as an objective justifier for the subjective state $\mathbf{E}(t,\varphi)$. In conclusion, $[t]\varphi \rightarrow [!t][t]\varphi$ appears to be the better interpretation in the present system of a version of **LJ3** concerning potential justifiers, while $t : \varphi \rightarrow !t : [t]\varphi$ appears to be the better interpretation in the present system of a version of **LJ3** concerning actual justifiers.

3.3 *Principles on justification: the non-ideal case*

Suppose now we are working in the non-ideal logic **LJ1**. Then, the following principles concerning actual justification are derivable.

J1.1: $t : (\varphi \rightarrow \psi) \rightarrow (s : \varphi \rightarrow \neg[\mathbf{F}]\neg t \cdot s : \psi)$

$\vdash_{\mathbf{LJ0}} [t](\varphi \rightarrow \psi) \rightarrow ([s]\varphi \rightarrow [t \cdot s]\psi)$
$\vdash_{\mathbf{LJ0}} [t](\varphi \rightarrow \psi) \rightarrow ([s]\varphi \rightarrow [\mathbf{F}][t \cdot s]\psi)$
$\vdash_{\mathbf{LJ0}} \mathbf{E}(t, \varphi \rightarrow \psi) \rightarrow (\mathbf{E}(s,\varphi) \rightarrow \neg[\mathbf{F}]\neg\mathbf{E}(t \cdot s, \psi))$
$\vdash_{\mathbf{LJ0}} [\mathbf{F}][t \cdot s]\psi \wedge \neg[\mathbf{F}]\neg\mathbf{E}(t \cdot s, \psi) \rightarrow \neg[\mathbf{F}]\neg([t \cdot s]\varphi \wedge \mathbf{E}(t \cdot s, \psi))$
$\vdash_{\mathbf{LJ0}} t : (\varphi \rightarrow \psi) \rightarrow (s : \varphi \rightarrow \neg[\mathbf{F}]\neg t \cdot s : \psi)$

Hence, if t is an actual justifier for $(\varphi \rightarrow \psi)$ and s is an actual justifier for φ, then it is possible for the agent to achieve an actual justifier for ψ, this actual justifier being precisely $t \cdot s$.

[1]It corresponds to the well-known metamathematical principle $Proof_T(n, \ulcorner\varphi\urcorner) \rightarrow Proof_T(f(n), Proof_T(\ulcorner\varphi\urcorner))$, which states that, if n is the code of a proof in T of φ, then $f(n)$ is the code of a proof in T of $Proof_T(n, \ulcorner\varphi\urcorner)$, where f is an appropriate primitive recursive function and $\ulcorner\varphi\urcorner$ is the code of φ. As we can see, no subjective element is involved is this principle.

J1.2: $t : \varphi \vee s : \varphi \rightarrow \neg[\mathbf{F}]\neg t + s : \varphi$

$\vdash_{\mathbf{LJ0}} [t]\varphi \vee [s]\varphi \rightarrow [t+s]\varphi$
$\vdash_{\mathbf{LJ0}} [t]\varphi \vee [s]\varphi \rightarrow [\mathbf{F}][t+s]\varphi$
$\vdash_{\mathbf{LJ0}} \mathbf{E}(t,\varphi) \vee \mathbf{E}(s,\varphi) \rightarrow \neg[\mathbf{F}]\neg\mathbf{E}(t+s,\varphi)$
$\vdash_{\mathbf{LJ0}} [\mathbf{F}][t+s]\varphi \wedge \neg[\mathbf{F}]\neg\mathbf{E}(t+s,\varphi) \rightarrow \neg[\mathbf{F}]\neg([t+s]\varphi \wedge \mathbf{E}(t+s,\varphi))$
$\vdash_{\mathbf{LJ0}} t : \varphi \vee s : \varphi \rightarrow \neg[\mathbf{F}]\neg t + s : \varphi$

Hence, if t and s are actual justifiers, then it is possible for the agent to achieve an actual justifier for all the propositions that are justified either by t or by s, this actual justifier being precisely $t + s$.

J1.3: $t : \varphi \rightarrow \neg[\mathbf{F}]\neg !t : [t]\varphi$

$\vdash_{\mathbf{LJ0}} [t]\varphi \rightarrow [!t][t]\varphi$
$\vdash_{\mathbf{LJ0}} [t]\varphi \rightarrow [\mathbf{F}][!t][t]\varphi$
$\vdash_{\mathbf{LJ0}} \mathbf{E}(t,\varphi) \rightarrow \neg[\mathbf{F}]\neg\mathbf{E}(!t,[t]\varphi)$
$\vdash_{\mathbf{LJ0}} [\mathbf{F}][!t][t]\varphi \wedge \neg[\mathbf{F}]\neg\mathbf{E}(!t,[t]\varphi) \rightarrow \neg[\mathbf{F}]\neg([!t][t]\varphi \wedge \mathbf{E}(!t,[t]\varphi))$
$\vdash_{\mathbf{LJ0}} t : \varphi \rightarrow \neg[\mathbf{F}]\neg !t : [t]\varphi$

Hence, if t is an actual justifier for φ, then it is possible for the agent to achieve an actual justifier for the proposition that t is an objective justifier for φ, this actual justifier being precisely $!t$.

J1.4: $t : \varphi \rightarrow \varphi$
Similar to **JI.4**.

As we can see, within **LJ1** the epistemic limits of a standard agent are taken into consideration. In particular, it is required that the construction of new composite justifiers from the current ones take time, so that the final justifier is the outcome of some course of thought, in the spirit of Duc (1997). As a consequence, while the standard principles of justification logic are no longer derivable, the principles concerning actual justifiers are completely intuitive, since no idealization on the epistemic power of the agent is involved.

A final observation. In some recent studies on the logic of justification, see (Dean & Kurokawa, 2010) and (Baltag & al., 2014), a predicate of existence is introduced that enables us to state that a certain justifier has actually been constructed by an agent, or that a certain justifier is currently accessed by the agent. In our context, such a predicate can be introduced by definition, in view of axioms **E4** and **IE4**.

Definition 14. *(existence of a justifier):* $\mathrm{E!}(t) := \mathbf{E}(t, \top)$.

It is not difficult to see that the following theorems are derivable:

non-ideal case	ideal case
$\vdash_{\mathbf{LJ0}} \mathrm{E!}(t) \wedge \mathrm{E!}(s) \rightarrow \neg[\mathbf{F}]\neg\mathrm{E!}(t \cdot s)$	$\vdash_{\mathbf{LJ1}} \mathrm{E!}(t) \wedge \mathrm{E!}(s) \rightarrow \mathrm{E!}(t \cdot s)$
$\vdash_{\mathbf{LJ0}} \mathrm{E!}(t) \wedge \mathrm{E!}(s) \rightarrow \neg[\mathbf{F}]\neg\mathrm{E!}(t + s)$	$\vdash_{\mathbf{LJ1}} \mathrm{E!}(t) \wedge \mathrm{E!}(s) \rightarrow \mathrm{E!}(t + s)$
$\vdash_{\mathbf{LJ0}} \mathrm{E!}(t) \wedge \mathrm{E!}(s) \rightarrow \neg[\mathbf{F}]\neg\mathrm{E!}(t!)$	$\vdash_{\mathbf{LJ1}} \mathrm{E!}(t) \wedge \mathrm{E!}(s) \rightarrow \mathrm{E!}(!t)$

As expected, in the non-ideal case, the intuitive epistemic dynamics of an agent is better represented. Indeed, all composite justifiers are then thought of as possible results of epistemic actions, as witnessed by the prefix $\neg[\mathbf{F}]\neg$. Hence, the ideal case can be viewed as the case that arises when the epistemic steps are compressed and all that can be achieved by performing inferences has in fact been achieved.

References

Artemov, S. (2001). Explicit provability and constructive semantics. *The Bulletin for Symbolic Logic*, *7*: 1–36.

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

Artemov. S. and Nogina, E. (2004). Logic of knowledge with justifications from the provability perspective. *Technical Report*, CUNY 2004.

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

Baltag, A., Renne, B. and Smets, S. (2014). The logic of justified belief, explicit knowledge, and conclusive evidence. *Annals of Pure and Applied Logic*, *165*: 49–81.

Blackburn, B., de Rijke, M., and Venema, Y. (2001). *Modal Logic*. Cambridge: Cambridge University Press.

Dean, W. and Kurokawa, H. (2010). From the Knowability Paradox to the existence of proofs. *Synthese*, *176*: 177–225.

Duc, H. N. (1997). Reasoning about rational, but not logically omniscient, agents. *Journal of Logic and Computation*, *7*: 633–648.

Dummett, M. (1998). Truth from the constructive standpoint. *Theoria*, *64*, 123–138.

Dummett, M. (2000). *Elements of Intuitionism*. Second Edition. Oxford: Clarendon Press.

Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (1995). *Reasoning about Knowledge*. MA: The MIT Press.

Fine, G. (2014). *The Possibility of Inquiry: Meno's Paradox from Socrates to Sextus*. Oxford: Oxford University Press.

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

Fitting, M. (2008). Justification logics, logics of knowledge, and conservativity. *Annals of Mathematics and Artificial Intelligence*, *53*: 153-167.

Fitting, M. (2009) Reasoning with justifications. In Makinson, D., Malinowski, J. and Wansing, H. (Eds.), *Towards Mathematical Philosophy*, Trends in Logic 28, Springer, 2009. pp. 107–123.

Martin-Löf, P. (1987). Truth of a Proposition, Evidence of a Judgement, Validity of a Proof. *Synthese*, *73*: 407–420.

Martin-Löf, P. (1998). Truth and knowability: On the principles C and K of Michael Dummett. In G. Dales & G. Oliveri (Eds.), *Truth in mathematics* (pp. 105–114). New York: Oxford University Press.

Prawitz, D. (1994). Meaning Theory and Anti-Realism. In G. Oliveri & B. McGuinness (Eds.), *The Philosophy of Michael Dummett* (pp. 79–92). Dordrecht: Kluwer.

Prawitz, D. (1998). Truth and Objectivity from a verificationist point of view. In G. Dales & G. Oliveri (Eds.), *Truth in mathematics* (pp. 41–51). New York: Oxford University Press.

van Benthem, J. and Velazquez-Quesada, F. (2010). The dynamics of awareness. *Synthese*, *177*: 5–27.