

# A Logic of Explicit Knowledge

Melvin Fitting

Dept. Mathematics and Computer Science  
Lehman College (CUNY), 250 Bedford Park Boulevard West  
Bronx, NY 10468-1589  
e-mail: melvin.fitting@lehman.cuny.edu  
web page: comet.lehman.cuny.edu/fitting

## Abstract

A well-known problem with Hintikka-style logics of knowledge is that of logical omniscience. One knows too much. This breaks down into two subproblems: one knows all tautologies, and one's knowledge is closed under consequence. A way of addressing the second of these is to move from knowledge simpliciter, to knowledge for a reason. Then, as consequences become 'further away' from one's basic knowledge, reasons for them become more complex, thus providing a kind of resource measurement.

One kind of reason is a formal proof. Sergei Artemov has introduced a logic of explicit proofs, **LP**. I present a semantics for this, based on the idea that it is a logic of knowledge with explicit reasons. A number of fundamental facts about **LP** can be established using this semantics. But it is equally important to realize that it provides a natural logic of more general applicability than its original provenance, arithmetic provability.

## 1 Introduction

The study of the properties of knowledge with the tools of formal logic was begun by Hintikka, in [5], using the machinery of possible worlds. Essentially the idea is: one knows  $X$  in a certain state provided  $X$  is the case in all states that one cannot distinguish from that state. Then various interesting conditions on knowledge, such as positive and negative introspection, correspond to simple conditions on accessibility relations between states, such as transitivity and symmetry. The problem with this modal approach is that it inherits two standard principles of normal modal logics. One is the validity of  $K(X \supset Y) \supset (KX \supset KY)$  (where  $K$  is a modal operator with  $KZ$  customarily read " $Z$  is known"). The other is the soundness of the inference of  $KX$  from  $X$ . The first yields that one knows the consequences of what one knows, and the second, that one knows all tautologies, no matter how complex. Together these are usually referred to as the problems of *logical omniscience*. The standard way out is to say that what is being captured is not actual knowledge, but potential knowledge—loosely,  $KX$  should be read " $X$  is knowable," rather than " $X$  is known." In fact this seems quite acceptable, and Hintikka's approach formalizes the logic of potential knowledge reasonably well.

This still leaves us wanting a logic of *actual* knowledge. And that seems hard to come by. It may be that  $X$  is potentially knowable by me because it follows from things I already know, but I may not actually know  $X$  simply because I never thought about it. There does not seem to be much scope for the development of a formal logic here. Still, we can make some progress by considering, not a logic of knowledge *per se*, but a logic of *explicit* knowledge. We then would take as basic not  $KX$ , but  $t:X$ , which we read " $X$  is known for reason  $t$ ." Done right, this can turn one of the logical

omniscience problems into something more reasonable: I may know the consequences of what I know, but for more complex reasons. We have a kind of resource measurement on knowledge that seems potentially useful.

Then we should also have a calculus on reasons, allowing us to combine or reflect on reasons in a natural way. In fact, such a calculus already exists. Among the broad and somewhat amorphous family of things we might call reasons are mathematical proofs. Sergei Artemov has introduced a kind of propositional modal logic of explicit proofs, **LP**, [1, 2], and in this there are various operations introduced for what he calls proof polynomials. These operations make intuitive sense in a broader context as well.

What I will do in this paper is modify Hintikka’s semantics, by incorporating ideas from [7], that have the effect of allowing for explicit reasons for knowing. This semantics will accommodate the operations of Artemov in what I believe to be a natural way. I propose the result as a plausible candidate for a logic of explicit knowledge. Formally, it is equivalent to Artemov’s proof-theoretic approach, that is, a soundness and completeness theorem can be proved. But it is as a logic of knowledge that I would like you to consider this work, and not as a more narrow and technical mechanism.

## 2 Logic of Potential Knowledge

I present a quick summary of Hintikka’s approach, restricted to a single knower, so that it will be easier to see how the logic of explicit knowledge relates to it. This material is based on [5]. The language is propositional, say with formulas built up from propositional letters and  $\perp$  using  $\supset$ . In addition we have an operator  $K$  with the formation rule: if  $X$  is a formula, so is  $KX$ .

As usual with modal logics, a *frame* is a structure  $\langle \mathcal{G}, \mathcal{R} \rangle$  where  $\mathcal{G}$  is a non-empty set and  $\mathcal{R}$  is a binary relation on  $\mathcal{G}$ . We can think of  $\mathcal{G}$  as possible worlds, or states, and  $\mathcal{R}$  as relative accessibility, as usual. Informally, we may think of  $\Gamma \mathcal{R} \Delta$  as asserting: “if you (the knower) are actually in state  $\Gamma$ , you cannot tell it is not state  $\Delta$ .” For instance, if I am in New York City, and it is below freezing in Montreal, I may not be able to distinguish that state from one in which I am in New York City and it is above freezing in Montreal. To take a rather extreme example, suppose in state  $\Gamma$  I am a jellyfish, while in state  $\Delta$  I have evolved into Albert Einstein. In state  $\Gamma$  my capacity is too limited to distinguish between that state and  $\Delta$ , but in state  $\Delta$  I can certainly distinguish between actuality and state  $\Gamma$ , so under such extreme circumstances, distinguishability need not be symmetric.

A knowledge model is a structure  $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{V} \rangle$  where  $\langle \mathcal{G}, \mathcal{R} \rangle$  is a frame, and  $\mathcal{V}$  is a valuation, mapping propositional letters to sets of states. Then truth at worlds of a model is evaluated in the familiar Kripke-style way:

1.  $\mathcal{M}, \Gamma \Vdash P \iff \Gamma \in \mathcal{V}(P)$  for  $P$  a propositional letter;
2.  $\mathcal{M}, \Gamma \not\Vdash \perp$ ;
3.  $\mathcal{M}, \Gamma \Vdash X \supset Y \iff \mathcal{M}, \Gamma \not\Vdash X$  or  $\mathcal{M}, \Gamma \Vdash Y$ ;
4.  $\mathcal{M}, \Gamma \Vdash KX \iff \mathcal{M}, \Delta \Vdash X$  for all  $\Delta \in \mathcal{G}$  with  $\Gamma \mathcal{R} \Delta$ .

Item 4 above is the key one, of course. It says potential knowledge amounts to truth in all states I can’t distinguish from the actual one. Naturally one would want to assume indistinguishability is reflexive, which means  $KX \supset X$  is validated—one can’t know what isn’t so. Of course an attempt to model belief would drop this requirement. So-called *positive introspection* corresponds

to transitivity,  $KX \supset K K X$ —one has knowledge about ones own knowledge. Likewise *negative introspection* requires a Euclidean condition, or symmetry if positive introspection is present,  $\neg K X \supset K \neg K X$ —one has knowledge about ones own lack of knowledge.

Of course, as noted earlier, this semantics validates  $K(X \supset Y) \supset (K X \supset K Y)$ . Likewise validities are closed under the rule  $X/KX$ . These are the problems of *logical omniscience*, that makes this a semantics of *potential* rather than of *actual* knowledge.

### 3 Explicit Knowledge

Now we drop the operator  $K$  from the language, and introduce a family of explicit reasons instead—I'll use  $t$  as a typical one. Following [1, 2] I'll write  $t:X$  to indicate that  $t$  applies to  $X$ —read it as “ $X$  is known for reason  $t$ .” Formally, if  $t$  is a reason and  $X$  is a formula,  $t:X$  is a formula. Of course this means we must define a collection of reasons, but let us postpone doing so. To get things started, I'll just assume we have a family of variables,  $x, y, \dots$ , that stand for unspecified reasons, and other things will be introduced later.

The semantic idea is to add extra machinery to that of relational models, to capture the notion of *evidence* or *reason*. This is done by adapting machinery from [7]. We will consider structures  $\langle \mathcal{G}, \mathcal{R}, \mathcal{V}, \mathcal{E} \rangle$  where  $\langle \mathcal{G}, \mathcal{R}, \mathcal{V} \rangle$  is a knowledge model, essentially as in the previous section, and  $\mathcal{E}$  is an evidence function, mapping states and reasons to sets of formulas. The idea is,  $X \in \mathcal{E}(\Gamma, t)$  informally says that  $t$  is evidence for formula  $X$  in state  $\Gamma$ . Having evidence for a statement does not mean the statement is fact. An assertion that the moon is made of green cheese, coming from a three-year-old, is not evidence that the moon is made of green cheese. An assertion that there once was considerable water on Mars, coming from a space scientist, is evidence. We would be inclined to take it seriously, but it is not, of course, conclusive.

Various conditions will be imposed on the  $\mathcal{E}$  function later on, but for now it is simply an unrestricted mapping from states and reasons to formula sets. Given this, we define a simple version of a model with explicit reasons. Item (4) in the characterization of truth in knowledge models, Section 2, is replaced by a more complicated condition.

**Definition 3.1** A *basic model* is a structure  $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{E}, \mathcal{V} \rangle$  where  $\langle \mathcal{G}, \mathcal{R} \rangle$  is a reflexive frame,  $\mathcal{E}$  is an evidence function on  $\langle \mathcal{G}, \mathcal{R} \rangle$ , and  $\mathcal{V}$  is a mapping from propositional variables to subsets of  $\mathcal{G}$ . Then truth at states is characterized as follows. For each  $\Gamma \in \mathcal{G}$ :

1.  $\mathcal{M}, \Gamma \Vdash P \iff \Gamma \in \mathcal{V}(P)$  for  $P$  a propositional letter;
2.  $\mathcal{M}, \Gamma \not\Vdash \perp$ ;
3.  $\mathcal{M}, \Gamma \Vdash X \supset Y \iff \mathcal{M}, \Gamma \not\Vdash X$  or  $\mathcal{M}, \Gamma \Vdash Y$ ;
4.  $\mathcal{M}, \Gamma \Vdash (t:X)$  if and only if  $X \in \mathcal{E}(\Gamma, t)$  and, for every  $\Delta \in \mathcal{G}$  with  $\Gamma \mathcal{R} \Delta$ ,  $\mathcal{M}, \Delta \Vdash X$ .

We say  $X$  is *true at state*  $\Gamma$  if  $\mathcal{M}, \Gamma \Vdash X$ , and otherwise  $X$  is *false at*  $\Gamma$ . Likewise  $X$  is *valid* in a model if it is true at all states of the model.

Item 4 above is the only change from earlier. Informally it says that  $t:X$  is true at a state if  $X$  is knowable at that state, and  $t$  is evidence for  $X$ .

Here are a few examples, to show how this works.

**Example** Not surprisingly there are very simple models in which  $t:P \supset t:(P \wedge Q)$  are not valid. But this invalidity can come about for two quite different reasons. First, consider the model  $\mathcal{M}$

having a single state,  $\Gamma$ , accessible to itself, and with an evidence function such that  $\mathcal{E}(\Gamma, t)$  is the entire set of formulas. In this model,  $t$  serves as ‘universal’ evidence. Also, use a valuation such that  $\mathcal{V}(P) = \{\Gamma\}$  and  $\mathcal{V}(Q) = \emptyset$ . Then we have  $\mathcal{M}, \Gamma \Vdash t:P$  but  $\mathcal{M}, \Gamma \not\Vdash t:(P \wedge Q)$  because, even though  $t$  serves as universal evidence,  $P \wedge Q$  is not knowable at  $\Gamma$ , since  $P$  is true but  $Q$  is not.

Next consider the model  $\mathcal{N}$ , again having a single state  $\Gamma$  accessible to itself. This time take  $\mathcal{V}$  to be the mapping assigning  $\{\Gamma\}$  to every propositional letter. But also, set  $\mathcal{E}(\Gamma, t) = \{P\}$ . Then of course both  $P$  and  $P \wedge Q$  are knowable at  $\Gamma$ , but we have  $\mathcal{N}, \Gamma \Vdash t:P$  and  $\mathcal{N}, \Gamma \not\Vdash t:(P \wedge Q)$ , the latter because  $t$  does not serve as evidence for  $P \wedge Q$ .

Two important things can be learned from the example above. First, one might not know  $X$  for reason  $t$  either because  $X$  is not knowable, or because it is, but  $t$  is not a reason for  $X$ . Technically, the work can be done by the underlying knowledge model, or by the evidence function. In fact, the semantics presented in [7] amounts to single-world models, in which the evidence function carries the entire load.

The second item illustrated by the example above is that, while knowability is essentially a semantic notion, evidence is more syntactic. For example, the model  $\mathcal{N}$  also invalidates  $t:P \supset t:(P \wedge P)$ . At first glance this is surprising, since in any standard logic of knowledge  $KP \supset K(P \wedge P)$  is valid. But, just because  $t$  serves as evidence for  $P$ , it need not follow that it also serves as evidence for  $P \wedge P$ . They are syntactically different, and effort is needed to recognize that the later formula is a redundant version of the former. To take this to an extreme, consider the formula  $t:P \supset t:(P \wedge P \wedge \dots \wedge P)$ , where the consequent has a conjunction with as many conjuncts as there are elementary particles in the universe!

In the positive direction, we do have the validity of the following, for every  $t$  and  $X$ .

$$t:X \supset X \tag{1}$$

Very simply, if  $\mathcal{M}, \Gamma \Vdash t:X$  for some basic model  $\mathcal{M}$  then, in particular, we must have  $X$  true at every state accessible from  $\Gamma$ . Since the accessibility relation is required to be reflexive, we have  $\mathcal{M}, \Gamma \Vdash X$ .

## 4 Operations on Reasons

So far, evidence has no structure. Following [1, 2], there are interesting and natural operations to be considered. Here they are, along with the additional conditions that must be placed on basic models to accommodate them.

### 4.1 Multiplication

Suppose you are given  $s$  as evidence for something, then subsequently you are also given  $u$  as evidence for something else. You now might be able to combine these pieces of evidence to produce a reason for something that needs both to justify. It may happen that the order in which you are told these items is significant. It may even be that the use of evidence is like that of a resource—two applications are more costly than one. The following formal evidence-combining notion has been found to be useful in this context. We have a binary operation, denoted ‘ $\cdot$ ’ that combines  $s$  and  $t$  to produce  $s \cdot t$ . The corresponding requirement on evidence functions is the following.

$$\text{If } (X \supset Y) \in \mathcal{E}(\Gamma, s) \text{ and } X \in \mathcal{E}(\Gamma, t) \text{ then } Y \in \mathcal{E}(\Gamma, s \cdot t) \tag{2}$$

It is not assumed that  $\cdot$  is associative, or commutative, or idempotent, though any of these can be added as extra conditions. However, with the machinery for multiplication added, even with the minimal conditions allowed, we have the validity of

$$s:(X \supset Y) \supset (t:X \supset (s \cdot t):Y) \quad (3)$$

The corresponding formula for knowable,  $K(X \supset Y) \supset (KX \supset KY)$  essentially says the consequences of what is knowable are also knowable. Formula (3), on the other hand, says we know the consequences of what we know, but for more complex reasons.

## 4.2 Addition

There is another way of combining reasons that is also sometimes seen. If  $s$  and  $t$  are pieces of evidence, and we are speaking in a hurry, we might say that something is justified by one of  $s$  or  $t$ , without bothering to specify which. The formal version of this is an operation denoted ‘+’ that combines  $s$  and  $t$  to produce  $s + t$ , with the following requirements imposed on evidence functions.

$$\begin{aligned} \text{If } X \in \mathcal{E}(\Gamma, s) \text{ then } X \in \mathcal{E}(\Gamma, s + t) \\ \text{If } X \in \mathcal{E}(\Gamma, t) \text{ then } X \in \mathcal{E}(\Gamma, s + t) \end{aligned} \quad (4)$$

Just as with multiplication, there is no assumption of associativity, commutativity, idempotency, or distributivity with respect to multiplication. These can, of course, be added as extra conditions. Not surprisingly, we now have the validity of the following.

$$\begin{aligned} s:X \supset (s + t):X \\ t:X \supset (s + t):X \end{aligned} \quad (5)$$

## 4.3 Exponentiation

Don’t be alarmed by the terminology—it is by analogy with linear logic, with nothing deep intended. In everyday life, whatever that is, we are generally able to substantiate our reasons. For instance, if I say I know something because I found it in the Encyclopedia Britannica, you might ask, “so?” I then might explain why that is a reliable source of information—in effect I am giving a reason for my reason. Indeed, a purported piece of evidence is of no value without some justification for its being evidence. This idea leads to a formal operation denoted ‘!’ that maps reasons to reasons, turning reason  $t$  into  $!t$ , where the idea is, if  $t$  is a reason for  $X$ ,  $!t$  should be a reason for  $t:X$ . This time the requirements on our semantics are more complex. We first need a requirement on evidence functions.

$$\text{if } X \in \mathcal{E}(\Gamma, t) \text{ then } t:X \in \mathcal{E}(\Gamma, !t) \quad (6)$$

Next we impose an additional requirement on frames: the accessibility relation  $\mathcal{R}$  must be transitive. And finally, we need a monotonicity condition.

$$\text{if } \Gamma \mathcal{R} \Delta \text{ then } \mathcal{E}(\Gamma, t) \subseteq \mathcal{E}(\Delta, t) \quad (7)$$

These conditions combined entail the validity of the following.

$$t:X \supset !t:t:X \quad (8)$$

## 4.4 Constants

Finally, there are things we just know outright. It is not that we know them for no reason, but rather that the reason is one of the fundamental ones from which we begin our thinking. It is like playing a game, in a sense. If the game allows citation of the Encyclopedia Britannica, we may do so if it is useful. But if we question whether our eyes are correctly reporting the text to our brain, and want a reason for that, we are really playing a different game. We must start somewhere to get somewhere. In our context, unanalyzed reasons are represented by constants. The obvious condition is that a constant can only serve as a reason for a formula that is valid. But otherwise there is considerable flexibility. A constant might serve as a reason for more than one formula, or a formula might have more than one constant that serves as a reason for it. This leads to the following.

A *constant specification* is a mapping  $\mathcal{C}$  from constants to sets of valid formulas, where validity means true at all states of all the models meeting all the conditions imposed so far. If  $X \in \mathcal{C}(c)$ , where  $c$  is a constant, we think of  $c$  as a reason for  $X$ , a reason whose details are taken for granted.

## 5 Putting It All Together

Ideas have been introduced a few at a time. Now let's put them together.

A collection of terms informally called 'reasons' or 'evidence' is characterized as follows.

1. It is assumed there is an infinite family of constants, typically  $a, b, \dots$ , and an infinite family of variables, typically  $x, y, \dots$ . These are terms.
2. If  $t$  and  $u$  are terms, so are  $t \cdot u$ ,  $t + u$ , and  $!t$ .

Next, the class of formulas is characterized.

1. It is assumed there is an infinite family of propositional letters, typically  $P, Q, \dots$ . These are formulas.
2.  $\perp$  is a formula.
3. If  $X$  and  $Y$  are formulas so is  $X \supset Y$ .
4. If  $X$  is a formula and  $t$  is a term, then  $t:X$  is a formula.

This is the language we use from now on (and have been using informally before this). And now the notion of weak model, from [4], can be defined.

**Definition 5.1** A *weak model* is a structure  $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{E}, \mathcal{V} \rangle$ , meeting the conditions for a basic model, from Definition 3.1, and also the following additional conditions. The relation  $\mathcal{R}$  is not just reflexive, but also transitive. The evidence function  $\mathcal{E}$  also meets conditions (2), (4), and (6). The monotonicity condition, (7), holds. Truth at states is characterized exactly as in Definition 3.1. Finally, we say  $\mathcal{M}$  *meets constant specification*  $\mathcal{C}$  if  $\mathcal{C}(c) \subseteq \mathcal{E}(\Gamma, c)$  for each  $\Gamma \in \mathcal{G}$  and constant  $c$ .

I have been presenting this as a logic of explicit knowledge. Of course formal proofs are certainly among the things we would consider reasons. In [1, 2], Sergei Artemov introduced **LP**, which stands for *logic of proofs*. It was intended that terms, as defined above, should be thought of as representing proofs in first-order Peano arithmetic—indeed they are called *proof polynomials* in his treatment. Artemov gave an axiomatization for **LP**—here is a variant of it.

First, here are the axioms, which are given as schemata.

1. Some finite number of schemes to axiomatize classical propositional logic
2.  $s:(X \supset Y) \supset (t:X \supset (s \cdot t):Y)$
3.  $t:X \supset X$
4.  $s:X \supset (s + t):X$
5.  $t:X \supset (s + t):X$
6.  $t:X \supset !t:t:X$

There are two rules of inference. The first is the familiar *modus ponens*.

$$\frac{X, X \supset Y}{Y}$$

The other rule is a kind of explicit necessitation rule. Let  $\mathcal{C}$  be a constant specification that provides constants for all the axioms from the list above. We have already seen that these axioms are all valid in weak models, so this is actually a constant specification as characterized in Section 4.4. Let us call the following rule  $\mathcal{C}$  *necessitation*

$$\frac{}{c:A}$$

where  $A$  is an axiom from the list above, and  $A \in \mathcal{C}(c)$ . This final rule completes the presentation of the axiom system for **LP**.

In [2] a completeness theorem for **LP** relative to first-order Peano arithmetic is, in fact, established. Thus **LP** meets the conditions that prompted its creation—it provides a formalization of the notion of explicit provability in arithmetic.

A completeness theorem of a different kind was shown in [7], relative to an abstract semantics not directly connected with arithmetic. In [4] a completeness theorem was proved with respect to weak models. Here is a statement of it.

**Theorem 5.2** *Let  $\mathcal{C}$  be a constant specification providing constants for axioms of **LP**. Then  $X$  is valid in all weak models that meet constant specification  $\mathcal{C}$  if and only if  $X$  is provable in the **LP** axiom system, using  $\mathcal{C}$  necessitation as a rule.*

Given the discussion earlier, this means we can consider **LP** as a promising logic of explicit knowledge, as well as a logic about provability in arithmetic.

## 6 What We Know Outright

Part of the machinery in the weak model semantics is a constant specification. It assigns constants to formulas that, in some sense, we should just know. If the constant specification assigns constants to exactly the axioms of **LP**, there is a completeness theorem. But we can entertain other possibilities. A constant specification need not provide constants for anything. Or it could provide constants for only a small and inconsequential set of validities. It turns out that there are semantic consequences of the behavior of the constant specification, consequences of some interest.

Consider a weak model  $\mathcal{M} = \langle \mathcal{G}, \mathcal{R}, \mathcal{E}, \mathcal{V} \rangle$ . Suppose we say  $X$  is *knowable* at  $\Gamma \in \mathcal{G}$  provided  $\mathcal{M}, \Delta \Vdash X$  for every  $\Delta \in \mathcal{G}$  such that  $\Gamma \mathcal{R} \Delta$ . (There is no machinery in **LP** to symbolize this, but it is a significant and useful notion nonetheless.) Let us say  $\mathcal{M}$  is *fully explanatory* provided, whenever

$X$  is knowable at  $\Gamma$  (for any  $\Gamma \in \mathcal{G}$ ) then for some term  $t$ ,  $\mathcal{M}, \Gamma \Vdash t:X$ . Being fully explanatory means that everything knowable must have a reason. Being fully explanatory is a semantic condition—some weak models are fully explanatory, some are not. Let us say a *strong model* is a weak model that meets the fully explanatory condition. And let us say a constant specification  $\mathcal{C}$  *entails equivalence of semantics* provided every weak model meeting constant specification  $\mathcal{C}$  is a strong model.

The fully explanatory condition we have been discussing is a local notion—it imposes a condition that must be met state by state in a model. The next item is more global in nature. Suppose  $X$  is one of the **LP** axioms given above. A constant specification  $\mathcal{C}$  might be so weak that it does not assign a constant to  $X$ . And it can happen that this propagates, so that  $X$ , while valid, is valid for no reason at all, in the formal sense. This is not particularly desirable. Let us say a constant specification  $\mathcal{C}$  *entails internalization* provided that, for each formula  $X$  that is valid in all weak models meeting constant specification  $\mathcal{C}$ , there is a term  $t$  such that  $t:X$  is also valid in all weak models meeting  $\mathcal{C}$ . Notice that this is most definitely not a local condition.

The local and global conditions just introduced are intimately connected. It is shown in [4] that a constant specification  $\mathcal{C}$  entails equivalence of semantics if and only if it entails internalization. From now on I'll call a constant specification *full* if it entails equivalence of semantics, or equivalently if it entails internalization.

Artemov, in [2], showed that the axiomatically formulated **LP** meets a condition corresponding to internalization, provided we use a constant specification assigning a constant to each axiom. It follows from this, and Theorem 5.2, that we have a completeness theorem for **LP** with respect to strong models as well.

## 7 Realization Theorems

The conditions placed on accessibility relations in weak (or strong) models are those of **S4**. This suggests there should be a formal connection.

For a formula  $X$ , let  $X^\circ$  be the result of replacing every occurrence of  $t$  in  $X$  with  $\Box$ , for every term  $t$ . So, for instance,  $[s:P \supset !s:sP]^\circ$  is  $\Box P \supset \Box \Box P$ . It is easy to show that if  $X$  is valid in all weak models meeting constant specification  $\mathcal{C}$  (an arbitrary constant specification), then  $X^\circ$  is valid in **S4**. Remarkably, there is a converse to this—one was established proof theoretically by Artemov. In [3, 4] I gave a semantic proof of this converse result. Stated in terms of the present semantics, if a formula in the language of conventional monomodal logic is **S4** valid, there is some replacement of occurrences of  $\Box$  with terms that produces a formula valid in all weak models satisfying constant specification  $\mathcal{C}$ , but the constant specification must be of a particular form. I'll skip the details here.

Interestingly enough, there is also an embedding of **S4** into the logic of explicit reasons that does not need any conditions on constant specifications beyond being full, and that does not make use of the  $+$  operator, but the translation itself is of a more complex nature. I refer you to [4] for details.

## 8 Conclusion

Work on the logic of explicit reasons is an active area of research. Much has been done, from the point of view that it is a logic having to do with provability in Peano arithmetic, and almost all of this work has been proof-theoretic in nature. The primary source is [2]. A semantic investigation, using a different notion of model, can be found in [7], where decidability results are also established.



Complexity results can be found in [6]. A Gentzen system/tableau system is a useful tool for working with the logic, and one can be found in [2], along with a sketch of a constructive cut elimination proof. A semantic cut elimination argument is in [8] using the semantics of [7]—there is one using the semantics of this paper in [4]. One can drop some of the term operators and their corresponding semantic conditions, and get natural sublogics that correspond to **K**, **T**, or **K4** in the same way that the full system corresponds to **S4**. There has been some investigation in the direction of extending the logic towards **S5**, but a semantics for this is not yet well-developed.

In Section 6 a notion of knowability was discussed. It seems natural to introduce this formally into the system. That is, we should enlarge the language to allow  $\Box X$  as well as  $t:X$  to be part of the syntax. Something along these lines has been done in [9], giving the operator  $\Box$  behavior akin to the necessity operator of **GL**, the Gödel-Löb provability logic. I am currently working on expanding the semantics presented here to admit an **S4** operator in a natural way. This would allow for the realization results of Section 7 to be internalized, in a sense, remaining within a single more complex logic, instead of being between two different ones. This is work in progress.

Perhaps the most interesting question here is the vaguest. Various operations on explicit reasons have been introduced, but all these have their motivation in arithmetic. That is,  $\cdot$ ,  $+$ , and  $!$  all correspond to important facts about provability in Peano arithmetic. Are there any other natural operations on explicit reasons, where the notion of reason can come from a broader context than just arithmetic? Responses welcome.

## References

- [1] S. Artemov. Operational modal logic. Technical Report MSI 95-29, Cornell University, December 1995.
- [2] S. Artemov. Explicit provability and constructive semantics. *The Bulletin for Symbolic Logic*, 7(1):1–36, 2001. <http://www.cs.gc.cuny.edu/~sartemov/publications/BSL.ps>.
- [3] M. C. Fitting. A semantic proof of the realizability of modal logic in the logic of proofs. Technical Report TR-2003010, City University of New York, New York, Sept. 2003. <http://www.cs.gc.cuny.edu/tr/>.
- [4] M. C. Fitting. The logic of proofs, semantically. To appear *Annals of Pure and Applied Logic*. Available on web page, <http://comet.lehman.cuny.edu/fitting>, 2004.
- [5] J. Hintikka. *Knowledge and Belief*. Cornell University Press, 1962.
- [6] R. Kuznets. On the complexity of explicit modal logics. In *Computer Science Logic 2000*, volume 1862 of *Lecture Notes in Computer Science*, pages 371–383. Springer, 2000.
- [7] A. Mkrtychev. Models for the logic of proofs. In S. I. Adian and A. Nerode, editors, *Logical Foundations of Computer Science*, number 1234 in *Lecture Notes in Computer Science*, pages 266–275. Springer, 1997.
- [8] B. Renee. Completeness of a tableau system for the logic of proofs. Technical Report TR-2004001, City University of New York, 2004.
- [9] R. E. Yavorsky. Provability logics with quantifiers on proofs. *Annals of Pure and Applied Logic*, 113:373–387, 2001.