Philosophy Colloquium
CUNY Graduate Center

The Logic of Justification

Sergei Artemov
The City University of New York

New York, September 3, 2008

Formal methods

do not directly solve philosophical problems, but rather provide a tool for analyzing assumptions and making sure that we draw correct conclusions.

Probability Theory,
Boolean Logic,
Proof Theory,
etc.

Our hope is that **Justification Logic** will do just that.

Mainstream Epistemology:

Starting point: tripartite approach to knowledge (usually attributed to Plato)

$$\textit{Knowledge} \quad \sim \quad \textit{Justified True Belief.}$$

In the wake of papers by Russell, Gettier, and others: questioned, criticized, revised; now is generally regarded as a necessary condition for knowledge.

Logic of Knowledge: the model-theoretic approach (von Wright, Hintikka, ...) has dominated modal logic and formal epistemology since the 1960s.

$$\Box F \quad \sim \quad F \text{ holds at all possible epistemic situations.}$$

Logic of Knowledge: the model-theoretic approach (von Wright, Hintikka, ...) has dominated modal logic and formal epistemology since the 1960s.

$$\Box F \quad \sim \quad F \textit{ holds at all possible epistemic situations.}$$

Basic principles:

Axioms and rules of classical propositional logic +

$\Box(F \to G) \to (\Box F \to \Box G)$ *Epistemic Closure*

$\Box F \to F$ *Factivity*

$\Box F \to \Box\Box F$ *Positive Introspection*

$\neg\Box F \to \Box(\neg\Box F)$ *Negative Introspection*

Necessitation Rule: $\dfrac{\vdash F}{\vdash \Box F}$.

Logic of Knowledge: the model-theoretic approach (von Wright, Hintikka, ...) has dominated modal logic and formal epistemology since the 1960s.

$$\Box F \quad \sim \quad F \text{ holds at all possible epistemic situations.}$$

Easy, visual, useful in many cases, but misses the mark considerably: *What if F holds at all possible worlds, e.g., a mathematical truth, say $P \neq NP$, but the agent is simply not aware of the fact due to lack of evidence, proof, justification, etc.?*

Logic of Knowledge: the model-theoretic approach (von Wright, Hintikka, ...) has dominated modal logic and formal epistemology since the 1960s.

$\Box F$ $\sim$ F *holds at all possible epistemic situations.*

Easy, visual, useful in many cases, but misses the mark considerably: *What if F holds at all possible worlds, e.g., a mathematical truth, say $P \neq NP$, but the agent is simply not aware of the fact due to lack of evidence, proof, justification, etc.?*

Speaking informally: modal logic offers a limited formalization

Knowledge $\sim$ *True Belief.*

There were no justifications in the modal logic of knowledge, hence a principal gap between mainstream and formal epistemology.

Obvious defect: Logical Omniscience

A basic principle of modal logic (of knowledge, belief, etc.):

$$\Box(F \to G) \to (\Box F \to \Box G).$$

At each world, the agent is supposed to “know” all logical consequences of his/her assumptions.

Obvious defect: Logical Omniscience

A basic principle of modal logic (of knowledge, belief, etc.):

$$\Box(F \rightarrow G) \rightarrow (\Box F \rightarrow \Box G).$$

At each world, the agent is supposed to "know" all logical consequences of his/her assumptions.

"Each agent who knows the rules of Chess should know whether there is a winning strategy for White."

Obvious defect: Logical Omniscience

A basic principle of modal logic (of knowledge, belief, etc.):

$$\Box(F \to G) \to (\Box F \to \Box G).$$

At each world, the agent is supposed to "know" all logical consequences of his/her assumptions.

"Each agent who knows the rules of Chess should know whether there is a winning strategy for White."

"Suppose one knows a product of two (very large) primes. In what sense does he/she know each of the primes, given that factorization may take billions of years of computation?"

Adding justifications into the language

$$t{:}F$$

t is a justification of F for a given agent

t is accepted by agent as a justification of F

t is a sufficient resource for F

F satisfies conditions t

etc.

Basic Justification Logic J, the language

Justification polynomials are terms built from *variables* $x, y, z, \ldots$ and *constants* $a, b, c, \ldots$ by means of operations: '$\cdot$' and '$+$'

x

a

$a{\cdot}x + b{\cdot}y$

$z{\cdot}(a{\cdot}x + b{\cdot}y)$, etc.

Formulas: usual, with addition of new constructions '$t{:}F$'

$c{:}(A \wedge B \rightarrow A)$

$x{:}A \rightarrow (c{\cdot}x){:}B$

$x{:}A \vee y{:}B \rightarrow (a{\cdot}x + b{\cdot}y){:}(A \vee B)$, etc.

Basic Justification Logic J

- The standard axioms and rules of classical propositional logic,
- $t{:}(F \rightarrow G) \rightarrow (s{:}F \rightarrow (t \cdot s){:}G)$ *Application*
- $s{:}F \rightarrow (s+t){:}F, \quad t{:}F \rightarrow (s+t){:}F$ *Sum*

Basic Justification Logic J

- The standard axioms and rules of classical propositional logic,
- $t{:}(F \rightarrow G) \;\rightarrow\; (s{:}F \rightarrow (t\cdot s){:}G)$ *Application*
- $s{:}F \rightarrow (s+t){:}F, \quad t{:}F \rightarrow (s+t){:}F$ *Sum*

Reflects basic reasoning about justifications.

Justifications are not assumed to be factive.

No logical truths are assumed *a priori* as justified for the agent.

Basic Justification Logic J

- The standard axioms and rules of classical propositional logic,
- $t{:}(F \rightarrow G) \rightarrow (s{:}F \rightarrow (t \cdot s){:}G)$ *Application*
- $s{:}F \rightarrow (s+t){:}F, \quad t{:}F \rightarrow (s+t){:}F$ *Sum*

Reflects basic reasoning about justifications.

Justifications are not assumed to be factive.

No logical truths are assumed *a priori* as justified for the agent.

Good for conditional statements:

if x is a justification for A, then t(x) is a justification for B

Basic Justification Logic J

- The standard axioms and rules of classical propositional logic,
- $t{:}(F \rightarrow G) \rightarrow (s{:}F \rightarrow (t \cdot s){:}G)$ *Application*
- $s{:}F \rightarrow (s+t){:}F, \quad t{:}F \rightarrow (s+t){:}F$ *Sum*

Reflects basic reasoning about justifications.

Justifications are not assumed to be factive.

No logical truths are assumed *a priori* as justified for the agent.

Good for conditional statements:

if x is a justification for A, then t(x) is a justification for B

Old Epistemic Modal language: $\Box A \rightarrow \Box B$

New Justification Logic language: $x{:}A \rightarrow t(x){:}B$

Introducing some a priori justified knowledge

Reasoning with justifications treats some logical truths as *a priori* justified. Consider a logical axiom:

$$A \wedge B \rightarrow A$$

To assume it justified, use a constant

$$c{:}(A \wedge B \rightarrow A)$$

This new axiom may also be assumed justified

$$d{:}c{:}(A \wedge B \rightarrow A), \text{etc.}$$

Constant Specifications range from empty (Cartesian skeptic) to the total (all axioms are justified to any depth) at our will.

Internalization: 'F is derived' yields '$t{:}F$ is derived' for some t.

Examples of reasoning in J

$A \wedge B \rightarrow A$ - logical axiom

$a{:}(A \wedge B \rightarrow A)$ - constant specification

$a{:}(A \wedge B \rightarrow A) \rightarrow (x{:}(A \wedge B) \rightarrow (a \cdot x){:}A)$ - Application Axiom of J

$x{:}(A \wedge B) \rightarrow (a \cdot x){:}A$ - by *Modus Ponens*

If x *is a justification for* $A \wedge B$ *then* $a \cdot x$ *is a justification for* A*, provided* a *is a proof (justification) for the logical axiom* $A \wedge B \rightarrow A$.

Examples of reasoning in J

$a{:}(A \rightarrow A \vee B)$ - constant specification
$x{:}A \rightarrow (a\cdot x){:}(A \vee B)$ - by Application and *Modus Ponens*

$b{:}(B \rightarrow A \vee B)$ - constant specification
$y{:}B \rightarrow (b\cdot y){:}(A \vee B)$ - by Application and *Modus Ponens*

$(a\cdot x){:}(A \vee B) \rightarrow (a\cdot x + b\cdot y){:}(A \vee B)$ - by Sum
$(b\cdot y){:}(A \vee B) \rightarrow (a\cdot x + b\cdot y){:}(A \vee B)$ - by Sum

$x{:}A \vee y{:}B \rightarrow (a\cdot x + b\cdot y){:}(A \vee B).$

Sum '$+$' is used here to reconcile distinct justifications for the same formula $(a\cdot x){:}(A \vee B)$ and $(b\cdot y){:}(A \vee B)$.

Epistemic models for J (Fitting-style)

Kripke model + **possible evidence function** $\mathcal{E}(t, F)$:

t is a possible evidence for F at world u.

Principal definition $t{:}F$ holds at u iff

1. $v \Vdash F$ whenever uRv (the usual Kripke condition for $\Box F$);
2. t is a possible evidence for F at u.

Soundness and Completeness take place.

Justification Logic J is capable of formalizing paradigmatic epistemic examples involving justifications: Gettier, Kripke's red barn example, Russell's prime minister example, etc.

Red Barn Example (Goldman – Kripke)

Suppose I am driving through a neighborhood in which, unbeknownst to me, papier-mâché barns are scattered, and I see that the object in front of me is a barn. Because I have barn-before-me percepts, I believe that the object in front of me is a barn. Our intuitions suggest that I fail to know barn. But now suppose that the neighborhood has no fake red barns, and I also notice that the object in front of me is red. Then I come to know that the object in front of me is a barn. Therefore, being a red barn, which I know, entails there being a barn, which I do not.

Formalization of *RBE* in epistemic modal logic

B - 'the object which I see is a barn'

R - 'the object which I see is red'

$\Box$ is my belief modality.

1. $\Box B$
2. $\Box(B \wedge R)$

Case (2) is knowledge, whereas (1) is not knowledge, by the problem's description. On the other hand, (1) logically follows from (2) in any epistemic modal logic:

$(B \wedge R) \rightarrow B$, logical axiom

$\Box[(B \wedge R) \rightarrow B]$, by Necessitation

$\Box(B \wedge R) \rightarrow \Box B$, by the Normality axiom.

This is a paradox, which is faithfully reproduced in modal logic.

Justification Logic provides a clean resolution of this paradox

Let us use the language of explicit justifications here.

Assumptions:

1. $u{:}B$.
2. $v{:}(B \wedge R)$

Reasoning:

3. $(B \wedge R) \rightarrow B$, logical axiom
4. $a{:}[(B \wedge R) \rightarrow B]$, Constant Specification
5. $v{:}(B \wedge R) \rightarrow (a \cdot v){:}B$, by Application.

Justification Logic provides a clean resolution of this paradox

Let us use the language of explicit justifications here.

Assumptions:

1. $u{:}B$.
2. $v{:}(B \wedge R)$

Reasoning:

3. $(B \wedge R) \rightarrow B$, logical axiom
4. $a{:}[(B \wedge R) \rightarrow B]$, Constant Specification
5. $v{:}(B \wedge R) \rightarrow (a \cdot v){:}B$, by Application.

The paradox disappears! Instead of deriving (1) from (2), we have derived $(a \cdot v){:}B$, but not $u{:}B$, i.e., I know B for reason $a \cdot v$, NOT for reason u.

(1) remains a case of belief rather then knowledge.

Gettier example

Smith has applied for a job, but has a justified belief that ‘Jones will get the job.’ He also has a justified belief that ‘Jones has 10 coins in his pocket.’ Smith therefore (justifiably) concludes ... that ‘the man who will get the job has 10 coins in his pocket.’

In fact, Jones does not get the job. Instead, Smith does. However, as it happens, Smith also has 10 coins in his pocket. So his belief that ‘the man who will get the job has 10 coins in his pocket’ was justified and true. But it does not appear to be knowledge.

Goal: **to formalize Gettier’s reasoning faithfully, to verify it, to perform the assumption and redundancies analysis.**

Formalizing the data

JJ = *Jones gets the job*

JS = *Smith gets the job*

CJ = *Jones has 10 coins in his pocket*

CS = *Smith has 10 coins in his pocket*

x = *whatever evidence Smith had about* JJ

y = *whatever evidence Smith had about* CJ

Explicitly made assumptions:

1. $x{:}JJ$ (x is a justification of 'Jones gets the job')
2. $y{:}CJ$ (y is a justification of 'Jones has 10 coins in his pocket')
3. $\neg JJ$ (Jones does not get the job)
4. JS (Smith gets the job)
5. CS (Smith has 10 coins in his pocket)

Justification Logic methods show that these assumptions are not sufficient to derive Gettiers conclusion:
Smith is justified in believing that 'the man who will get the job has 10 coins in his pocket.'

In this setting, the sentence 'the man who will get the job has 10 coins in his pocket' can be represented by the formula

$$(JJ \to CJ) \wedge (JS \to CS).$$

No justified knowledge assertion for this formula, i.e.,

$$t{:}[(JJ \to CJ) \wedge (JS \to CS)],$$

is derivable from the assumptions $x{:}JJ$, $y{:}CJ$, $\neg JJ$, JS, CS.

Countermodel for Gettier's claim

$W = \{1,2\}$, $R = \{(1,2)\}$, $\mathcal{E}$ is total.

$$\begin{array}{rcl} \text{`possible belief world' } 2 & \bullet & JJ,\ CJ,\ JS,\ \neg CS \\ & \uparrow & \\ \text{`real world' } 1 & \bullet & \neg JJ,\ CJ,\ JS,\ CS \end{array}$$

All assumptions hold at 1,2. At 2 both men have jobs and Smith does not have coins.

$$2 \not\Vdash (JS \rightarrow CS), \quad 2 \not\Vdash (JJ \rightarrow CJ) \wedge (JS \rightarrow CS),$$

hence

$$1 \not\Vdash t{:}[(JJ \rightarrow CJ) \wedge (JS \rightarrow CS)], \quad \text{for each } t.$$

Augmented set of assumptions

Is it now easy to spot a missing assumption?

Jones and Smith cannot both have this job

which is, of course, a default here. This is NOW ENOUGH too. What we need is *Smith is justified in believing 'Jones and Smith cannot both have this job'*.

6. $z:(JJ \to \neg JS)$ (z is a justification of 'Jones and Smith cannot both have the job')

Derivation of Gettier's claim

7. $(z\cdot x){:}(\neg JS)$, from 1,6, by Application
8. $p{:}[\neg JS \rightarrow (JS \rightarrow CS)]$, Internalization of a tautology
9. $(z\cdot x){:}(\neg JS) \rightarrow (p\cdot(z\cdot x)){:}(JS \rightarrow CS)$, by Application
10. $(p\cdot(z\cdot x)){:}(JS \rightarrow CS)$, from 7,9, by Modus Ponens
11. $c{:}[CJ \rightarrow (JJ \rightarrow CJ)]$, by Internalization
12 $y{:}CJ \rightarrow (c\cdot y){:}(JJ \rightarrow CJ)$, by Application
13. $(c\cdot y){:}(JJ \rightarrow CJ)$, from 2,12, by Modus Ponens
14. $t{:}[(JJ \rightarrow CJ) \wedge (JS \rightarrow CS)]$, for an appropriate t, from 10 and 13.

Metatheory of the Gettier example

Missing assumption analysis has just been performed.

Actually, we can also eliminate redundancies: no coins/pockets are needed...

What does Justification Logic bring to the logic of knowledge?

1. It adds a long-anticipated mathematical notion of justification, making the logic more expressive. We now have the capacity to reason about justifications, simple and compound. We can compare different pieces of evidence pertaining to the same fact. We can measure the complexity of justifications, thus connecting the logic of knowledge to a rich complexity theory, etc.

What does Justification Logic bring to the logic of knowledge?

2. Justification logic furnishes a new, evidence-based foundation for the logic of knowledge, according to which ‘F is known’ is interpreted as ‘F has an adequate justification.’

What does Justification Logic bring to the logic of knowledge?

3. Justification logic provides a novel, evidence-based mechanism of truth tracking which can be a valuable tool for extracting robust justifications from a larger body of justifications which are not necessarily reliable.

What does Justification Logic bring to the logic of knowledge?

4. Applications to well-known problems in epistemology: Gettier, Kripke, etc. (S.A.); the Knowability Paradox and the Knower Paradox (Dean & Kurokawa), Logical Omniscience Problem (S.A. & Kuznets), etc.

Reading

S.Artemov. Explicit provability and constructive semantics. Bulletin of Symbolic Logic, 7(1):136, 2001.

S.Artemov. Justification Logic. Technical Report TR-2007019, CUNY Ph.D. Program in Computer Science, 2007. To appear in the Review of Symbolic Logic, 2008.

S.Artemov & R.Kuznets. Logical omniscience via proof complexity. Springer Lecture Notes in Computer Science, v. 4207, pp. 135 - 149, 2006.

W.Dean & H.Kurokawa. The Knower Paradox and the Quantified Logic of Proofs. Austrian Ludwig Wittgenstein Soc., v. 31, 2008.

M.Fitting. The logic of proofs, semantically. Annals of Pure and Applied Logic, 132(1):125, 2005.