Justification Logic

Sergei Artemov (Graduate Center CUNY)

Workshop on ***Proof, Justification and Learning,***
Nancy, May 26, 2008

Knowledge as Justified True Belief

Can be traced to Plato, was widely accepted until 1963 when a paper by **Edmund Gettier** (Analysis 23 (1963):121-123) provoked widespread attempts to revise or replace it.

Related Developments

- **Brouwer**: mathematical truth = provability.
- **Skolem**: quantifiers = ghosts of functions.
- **Kolmogorov**: problem solutions (proofs) have an abstract structure, hence 'logic of solutions' and intuitionistic logic.
- **BHK**-semantics: informal 'proof tables' for intuitionistic logic.
- **Gödel**: modal logic of provability, the first (incomplete) sketch of the Logic of Proofs

Related Developments

- **Curry-Howard**:
  Combinatory terms ≈ Hilbert-style proofs,
  λ-terms ≈ natural derivations (no iterations yet).
- **Kleene** realizability semantics:
  constructive witnesses = computational programs.
- **Boolos, Solovay**: a complete modal logic of formal provability **GL**; implicit representation of proofs via '∃.'
- **S.A.** - Logic of Proofs **LP** consistent with Gödel's design.

Gettier Example: Case I

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.

In This Talk

We introduce a formal theory of justification, a.k.a. **Justification Logic,** based on classical epistemic logic augmented by justification assertions ***t:F***

t *is a justification for* ***F*** .

This theory grew from the Logic of Proofs **LP** (1995) in which main structural theorems were found. It became an epistemic subject after its long-anticipated Kripke-style semantics was suggested by Fitting on 2003.

We apply Justification Logic for formalizing Gettier examples.

Significance for Epistemology

Formal logical methods do not determine philosophical positions, but rather provide a tool for analyzing assumptions and making correct conclusions. We hope that Justification Logic does just that.

Justification Logic provides **a framework** capable of formalizing a significant portion of epistemic reasoning; epistemologists may find it useful, e.g., in the way they now use formal logic in their studies.

Comparisons with previous work

Relations with previous work in CS/AI, e.g., the Logic of Knowledge (Hintikka, Halpern, et al.):

Justification Logic extends the Logic of Knowledge:

1. It adds a long-anticipated formal notion of justification which makes logic more expressible;

2. It provides new evidence-based semantics for knowledge and belief;

3. It supports the basic Hintikka-style systems but on a new evidence-based foundations. This helps to bridge the gap between epistemology and the modal logic of knowledge/belief.

Initial design decisions

1. Propositional and quantifier-free systems first.
2. Classical Boolean logic on the background first.
3. Theory of partial (not factive) justification first.
4. One agent first.
5. Multiple systems, not one silver bullet.
6. Justification Logic is an open system with all meaningful extensions welcome.

Notational convention (single agent)

Notational convention: **K*F*** stands for

agent knows that F .

Justification assertions have a format ***t:F*** that reads

t is accepted by agent as a justification of F.

There is also a 'potential' reading of epistemic assertions in which **K*F*** stands for

agent can know F

and ***t:F*** reads

t is a justification of F.

Epistemic logics (single knower)

$$\mathbf{K} = \textit{classical logic} \quad + \quad \mathbf{K}(F \rightarrow G) \rightarrow (\mathbf{K}F \rightarrow \mathbf{K}G) \quad +$$

$$+ \quad \textit{Necessitation Rule:} \quad \vdash F \Rightarrow \vdash \mathbf{K}F$$

$$\mathbf{T} = \mathbf{K} \quad + \quad \mathbf{K}F \rightarrow F$$

$$\mathbf{K4} = \mathbf{K} \quad + \quad \mathbf{K}F \rightarrow \mathbf{KK}F$$

$$\mathbf{S4} = \mathbf{K4} \quad + \quad \mathbf{K}F \rightarrow F$$

$$\mathbf{K45} = \mathbf{K4} \quad + \quad \neg\mathbf{K}F \rightarrow \mathbf{K}(\neg\mathbf{K}F)$$

$$\mathbf{S5} = \mathbf{S4} \quad + \quad \neg\mathbf{K}F \rightarrow \mathbf{K}(\neg\mathbf{K}F)$$

Justification Logic

Justification Logic is a family of logic systems which axiomatize justification (***t:F***) and knowledge (**K*F***).

At the beginning, we build our systems on the simplest base: classical Boolean logic, thus leaving more elaborate logical models, for further studies. Justifications provide a sufficiently serious challenge even in the Boolean base. The paradigmatic examples which we will consider can be handled with the Boolean Logic.

Justification Logic

*Besides, the core of Epistemic Logic consists of modal systems with a classical Boolean base (***K***,* **T***,* **K4***,* **S4***,* **K45***,* **KD45***,* **S5***, etc.). We provide each of them with a corresponding Justification Logic companion based on Boolean logic.*

Preliminary Assumptions

- *Justifications are **abstract objects** which have **structure**. We introduce a set of basic operations on justifications and establish their connection to epistemic modal logic.*
- *The usual **potential executability** assumptions: atomic justifications are feasible in time and space for an agent to inspect and accept; basic operations on justifications are feasible; agent does not lose or forget justifications; agent applies the laws of classical logic and accepts their conclusions; etc.*
- *We consider both: **partial justifications** and **factive justifications.***

Learning from the Logic of Proofs

The Logic of Proofs **LP** (**Gödel**, 1938; **S.A.**, 1995) contains the principles:

$$t{:}F \rightarrow F\ ,$$

$$[y{:}(P \rightarrow Q) \wedge x{:}P] \rightarrow (y \cdot x){:}Q\ ,$$

$$t{:}F \rightarrow (t+s){:}F\ ,$$

and

if $\vdash F$*, then* $s{:}F$ *for some* s ,

along with **Positive Introspection** (cf. below).

A quest for new principles

Goldman's reliabilism:

a subject's belief is justified (t:F) only if (→) the truth of a belief (F) has caused the subject to have that belief (in the appropriate way),

Formal representation: ***t:F →F*** ,

and for a justified true belief to count as knowledge, the subject must also be able to 'correctly reconstruct' (mentally) that causal chain.

Formal representation: **there should also be a special justification *c* for *t:F →F,* i.e., *c:(t:F →F)*** .

A quest for new principles

Lehrer-Paxson's undefeasibility condition:

knowledge is undefeated justified true belief - which is to say that a justified true belief counts as knowledge if and only if it is also the case that there is no further truth which, had the subject known it, would have defeated her present justification for the belief.

Formal representation: **if *t:F*, then for any other piece of evidence *s*, a joint evidence *t+s* is still an evidence for *F*** :

$$t{:}F \rightarrow (t+s){:}F \, .$$

A quest for new principles

Dretske's conclusive reasons, Nozick's truth-tracking

A reason must exist for a belief that would not be true if the belief itself were false.

Formal representation: $\neg F \rightarrow \neg t{:}F$, which is logically equivalent to $t{:}F \rightarrow F$.

If, for example, I believe that there is a chair in front of me, the reason for believing that it is there would not exist if the belief were false (that is, if the chair were not there).

Formal representation: $t{:}F \rightarrow (\neg F \rightarrow \neg t{:}F)$, which is logically equivalent to $t{:}F \rightarrow F$.

Basic Principles: Applicability

Application operation *takes justifications* **s** *and* **t** *and produces a justification* **s·t** *such that*

if **s**:(**F**→**G**) *and* **t**:**F**, *then* (**s·t**):**G.**

Symbolically

$$\mathbf{s{:}(F \to G) \wedge t{:}F \to (s \cdot t){:}G.}$$

This is a basic property of justifications, assumed in the Logic of Proofs **LP**, combinatory logic, λ-calculi, BHK-semantics, realizability, etc.

Basic Principles: Applicability

The corresponding modal epistemic principle

$$\mathbf{K}(F \rightarrow G) \wedge \mathbf{K}F \rightarrow \mathbf{K}G ,$$

is widely accepted in formal epistemology.

However, this principle smuggles the *logical omniscience* defect into modal epistemic logic because the latter does not have the capacity to measure how hard it is to attain knowledge.

Justification Logic naturally escapes logical omniscience by keeping track of the size of evidence terms.

Basic Principles: **Monotonicity**

(cf. Lehrer-Paxson Principle)

If ***s:F****, then whatever evidence* ***t*** *occurs, the combined evidence* ***s+t*** *remains a justification for* ***F****.*

Operation '**+**' takes justifications ***s*** and ***t*** and produces ***s+t***, which is a justification for everything justified by ***s*** or ***t***.

$$\boldsymbol{s{:}F \rightarrow (s+t){:}F} \quad \text{and} \quad \boldsymbol{s{:}F \rightarrow (t+s){:}F}\ .$$

This is not meant to be an update, but rather a consistency of evidence condition.

Operation '**+**' first appeared in the Logic of Proofs **LP**, where it denoted a sum (concatenation) of proofs.

Basic Principles: Logical Awareness

Logical axioms are justified ex officio.

An agent accepts logical axioms (including the ones concerning justifications) as justified at any depth.

The natural way of formalizing this principle is postulating that for each evidence constant ***c*** *and for each axiom* ***A,*** ***c:A*** *is again an axiom.*

First appeared in the Logic of Proofs **LP**.

Cf. the Necessitation Rule in modal epistemic logic:

$$\vdash F \quad \Rightarrow \quad \vdash \mathbf{K}F.$$

Additional Principles: **Adequacy**

A justification *t:F* is factive, i.e., sufficient for an agent to conclude that *F* is true.

Adequacy yields the **Reflexivity Axiom**

$$t{:}F \to F$$

similar to the epistemic axiom

$$\mathbf{K}F \to F\ ,$$

which is widely accepted as a basic property of knowledge (Plato, Wittgenstein, Hintikka, Nozick, etc.).

Reflexivity of justification assertions also first appeared in the Logic of Proofs **LP** as a principal feature of mathematical proofs.

Additional Principles: **Adequacy**

Note that according to the Logical Awareness principle the agent should also have a justification of the reflexivity axiom

$$c{:}(t{:}F \rightarrow F)\ ,$$

a justification of this new axiom

$$b{:}c{:}(t{:}F \rightarrow F)\ ,$$

yet another justification,

$$a{:}b{:}c{:}(t{:}F \rightarrow F)\ ,$$

etc.

Additional Principles: Introspection

One of the fundamental principles of knowledge is identifying

knowing

and

knowing that one knows.

In the formal modal setting, this corresponds to

$$\mathbf{K}F \to \mathbf{KK}F\ .$$

This principle has an adequate explicit counterpart

justified

yields

verifiably justified,

hence justifications are assumed to be verifiable.

Additional Principles: **Introspection**

The mere fact that the agent accepts ***t*** as a sufficient evidence of ***F*** serves as a sufficient evidence that ***t:F***. Often, such 'meta-evidence' has a physical form, e.g.,

- a referee report certifying that ***t:F***,
- a computer verification output on ***t:F***,
- a formal proof that ***t*** is a proof of ***F***, etc.

Positive Introspection assumes that given ***t***, an agent produces a justification **!*t*** of ***t:F***

$$t{:}F \rightarrow\, !t{:}(t{:}F)\ .$$

Negative Introspection $\neg t{:}F \rightarrow\, ?t{:}(\neg t{:}F)$.

Summary of Justification Principles

Basic

- **Applicability,**
- **Monotonicity,**
- **Logical Awareness,**

Additional

- **Adequacy,**
- **Positive Introspection,**
- **Negative Introspection.**

We should not expect a one-size-fits-all universal logic here; a variety of customized systems should do the job.

Justification terms (polynomials)

Built from variables ***x, y, z , . . .*** and constants ***a, b, c, . . .*** by means of operations.

- *application* ‘ · ’
- *sum* ‘ **+** ’

Constants denote atomic justifications which the system no longer analyzes.

Variables denote unspecified justifications.

Justification terms (polynomials)

More elaborate models could also use additional operations on justifications, e.g.,

- *verifier* ‘ **!** ’
- *negative verifier* ‘ **?** ’

Logic of (Partial) Justifications **J**

- **Classical propositional logic,**
- **Application** $s{:}(F \to G) \to (t{:}F \to (s{\cdot}t){:}G)$,
- **Sum** $s{:}F \to (s+t){:}F$, $s{:}F \to (t+s){:}F$,
- **Axiom Internalization** *for each axiom* ***A*** *and a constant* **c**, ***c:A*** *is again an axiom.*

Logic of Adequate Justifications **AJ**

AJ = J + Reflexivity $t{:}F \rightarrow F$

Adding Positive Introspection

J4 = J +

+ Positive Introspection $t{:}F \rightarrow !t{:}(t{:}F)$

AJ4 = AJ +

+ Positive Introspection $t{:}F \rightarrow !t{:}(t{:}F)$

Adding Negative Introspection

J45 = J4 +
+ Negative Introspection $\neg t{:}F \to ?t{:}(\neg t{:}F)$

AJ45 = AJ4 +
+ Negative Introspection $\neg t{:}F \to ?t{:}(\neg t{:}F)$

Checklist of basic justification systems

Justified belief systems	Justified knowledge systems
J	**AJ**
J4	**AJ4 (=LP)**
J45	**AJ45**

Forgetful projection

A meaningful way of getting from a justification assertion to one of knowledge:

$$\boldsymbol{s{:}F} \quad \Rightarrow \quad \mathbf{K}\boldsymbol{F}\,.$$

(***s*** *justifies believing in* ***F***) (***F*** *is known)*

Examples (***P***,***Q*** are atomic propositions):

$$\boldsymbol{t{:}P\rightarrow P} \quad \Rightarrow \quad \mathbf{K}\boldsymbol{P\rightarrow P},$$

$$\boldsymbol{t{:}P\rightarrow !t{:}(t{:}P)} \quad \Rightarrow \quad \mathbf{K}\boldsymbol{P}\rightarrow\mathbf{KK}\boldsymbol{P},$$

$$\boldsymbol{s{:}(P\rightarrow Q)\rightarrow(t{:}P\rightarrow(t\cdot s){:}Q)} \quad \Rightarrow \quad \mathbf{K}\boldsymbol{(P\rightarrow Q)\rightarrow(}\mathbf{K}\boldsymbol{P\rightarrow}\mathbf{K}\boldsymbol{Q)}.$$

Forgetful projection

Forgetful projection sometimes forgets too much:

a triviality $\boldsymbol{x{:}P \to x{:}P}$ ⇨ $\mathbf{K}P \to \mathbf{K}P$,

a meaningful principle $\boldsymbol{x{:}P \to (x+y){:}P}$ ⇨ $\mathbf{K}P \to \mathbf{K}P$,

a non-valid formula $\boldsymbol{x{:}P \to y{:}P}$ ⇨ $\mathbf{K}P \to \mathbf{K}P$.

However, it always maps valid formulas of Justification Logic (JL) to valid formulas of Epistemic Logic (EL).

The converse also holds:

any valid formula of EL is a forgetful projection of some valid formula of JL .

This follows from the Correspondence Theorem (below).

Consolidated Correspondence Theorem

Let ➪ be a forgetful projection, then

J ➪ K,

J4 ➪ K4,

J45 ➪ K45,

and

AJ ➪ T,

AJ4 ➪ S4,

AJ45 ➪ S5.

The first three cases concern logics of belief, and the last three - the logics of knowledge.

From Knowledge to Justifications?

The core of the Correspondence Theorem is the Realization Theorem. *There is an algorithm that recovers justification terms for all modal knowledge operators in valid principles of epistemic modal logics:*

S4 - S.A., 1995 (via cut elimination);
S5 - S.A., Kazakov, Shapiro, 1999;
K, T, K4 - Brezhnev, 2000;
S4 - Fitting, 2004 (semantical proof);
S4 - Kuznets, Brezhnev, 2005 (polynomial algorithm);
S5, K45 - Rubtsova, 2006 (with ‘ **?** ’, by Fitting’s method);
S4 - Fitting, 2006 (another algorithm).

JTB as a syntactic transformation

The JTB definition of knowledge also defines a translation from the language of justifications to the modal language of knowledge:

$$\boldsymbol{F \wedge t{:}F} \quad \rightarrow \quad \mathbf{K}\boldsymbol{F}\,.$$

*(**F** is true and **t** justifies believing in **F**)* → *(**F** is known)*

Like the forgetful projection, JTB loses information:

a triviality $\boldsymbol{P\wedge x{:}P \rightarrow P\wedge x{:}P}$ → $\mathbf{K}\boldsymbol{P}\rightarrow\mathbf{K}\boldsymbol{P}$,

a deep principle $\boldsymbol{P\wedge x{:}P \rightarrow P\wedge (x+y){:}P}$ → $\mathbf{K}\boldsymbol{P}\rightarrow\mathbf{K}\boldsymbol{P}$,

a non-valid formula $\boldsymbol{P\wedge x{:}P \rightarrow P\wedge y{:}P}$ → $\mathbf{K}\boldsymbol{P}\rightarrow\mathbf{K}\boldsymbol{P}$.

However, JTB transformation maps logics of adequate justifications exactly to the modal logics of knowledge!

Correspondence Theorem and JTB

Corollary of the Correspondence Theorem:

$$\mathbf{AJ} \rightarrow \mathbf{T},$$
$$\mathbf{AJ4} \rightarrow \mathbf{S4},$$
$$\mathbf{AJ45} \rightarrow \mathbf{S5}.$$

In the core of this matter there is an algorithm that for any valid formula of a modal logic of knowledge finds a valid formula of the corresponding logic of adequate justifications such that

$\mathbf{K}F$ is being decoded as $F \wedge t{:}F$

for an appropriate justification t.

Correspondence Theorem and Gettier

For basic justification systems, JTB projection does not distinguish between partial and adequate justifications:

J → **T** and **AJ** → **T**,

J4 → **S4** and **AJ4** → **S4**.

When the justifications in JTB definition are only partial, then the resulting knowledge-like modal operator obeys the same modal principles as does knowledge and so may appear to be knowledge.

Properties of Justification Systems

Usual logical properties:

- ***Deduction Theorem***,
- ***closure under substitutions***.

Properties unique to Justification Logic:

- ***Internalization***
  *if $\vdash$**F**, then $\vdash$**p:F** for some evidence term **p** (every established fact is justified).*
- ***Realization Theorem***.

Examples of derivations in **K** *and* **J**

Derivation in K

$A \to A \vee B$

$\mathrm{K}(A \to A \vee B)$

$\mathrm{K}A \to \mathrm{K}(A \vee B)$

$B \to A \vee B$

$\mathrm{K}(B \to A \vee B)$

$\mathrm{K}B \to \mathrm{K}(A \vee B)$

$(\mathrm{K}A \vee \mathrm{K}B) \to \mathrm{K}(A \vee B)$

Derivation in J

$A \to A \vee B$

$a{:}(A \to A \vee B)$

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

$B \to A \vee B$

$b{:}(B \to A \vee B)$

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

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

One more example (needed below)

Derivation in K

$A \to (B \to (A \wedge B))$

$\mathbf{K}(A \to (B \to (A \wedge B)))$

$\mathbf{K}A \to \mathbf{K}(B \to (A \wedge B))$

$\mathbf{K}A \to (\mathbf{K}B \to \mathbf{K}(A \wedge B))$

$\mathbf{K}A \wedge \mathbf{K}B \to \mathbf{K}(A \wedge B)$

Derivation in J

$A \to (B \to (A \wedge B))$

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

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

$x{:}A \to (y{:}B \to ((c \cdot x) \cdot y){:}(A \wedge B))$

$x{:}A \wedge y{:}B \to ((c \cdot x) \cdot y){:}(A \wedge B)$

Epistemic Semantics

Fitting model = *Kripke model* + ***evidence function*** $\mathcal{E}(t,F)$, *which specifies whether a justification* **t** *is acceptable for a formula* **F** *at a world* **w**; *format* $w \in \mathcal{E}(t,F)$.

w⊨t:F *iff*

- ***F*** *holds at all worlds accessible from* ***w*** *(traditional Kripke requirement);*
- ***t*** *is acceptable evidence for* ***F*** *in* ***w***, *i.e.,* $w \in \mathcal{E}(t,F)$.

≈ Halpern-Moses awareness function + justifications

Epistemic Semantics

In Fitting models (as in Kripke) the accessibility relation

- for **J** is arbitrary,
- for **AJ** is reflexive,
- for **J4** is transitive,
- for **AJ4** is reflexive and transitive,
- for **AJ45** is reflexive, transitive, and symmetric.

Fitting Completeness Theorem (+ Rubtsova, Pacuit, S.A.):
Each of these logics is sound and complete with respect to the corresponding class of Fitting models.

Formalizing Gettier Examples, Case I

Syntax and its intended interpretation:

- ***JJ*** = *Jones gets the job,*
- ***SJ*** = *Smith gets the job,*
- ***JC*** = *Jones has 10 coins in his pocket,*
- ***SC*** = *Smith has 10 coins in his pocket,*
- ***x*** = *whatever evidence Smith had about* ***JJ***,
- ***y*** = *whatever evidence Smith had about* ***JC***.

The system of choice for formalization is **J**: as we saw, it has all the logical tools that Gettier hints at.

Formalizing Gettier Examples, Case I

Explicitly made non-logical assumptions:

1. ***x*:*JJ*** *(**x** is a justification of 'Jones gets the job')*
2. ***y*:*JC*** *(**y** is a justification of 'Jones has 10 coins in his pocket')*
3. **¬*JJ*** *(Jones does not get the job)*
4. ***SJ*** *(Smith gets the job)*
5. ***SC*** *(Smith has 10 coins in his pocket)*

Strictly speaking, these assumptions are not sufficient to derive Gettier's conclusion ***Smith is justified in believing*** '***the man who will get the job has 10 coins in his pocket***.'

Formalizing Gettier Examples, Case I

In our setting, the sentence ‘the man who will get the job has 10 coins in his pocket’ is represented by the formula

$$(JJ \rightarrow JC) \wedge (SJ \rightarrow SC).$$

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

$$t{:}[(JJ \rightarrow JC) \wedge (SJ \rightarrow SC)]\ ,$$

is derivable from the assumptions $x{:}JJ,\ y{:}JC,\ \neg JJ,\ SJ,\ SC$.

By the Deduction Theorem for **J**, it suffices to find a Fitting model where at a certain node all the assumptions hold, and $t{:}[(JJ \rightarrow JC) \wedge (SJ \rightarrow SC)]$ does not hold.

Formalizing Gettier Examples, Case I

Indeed,

$$\boldsymbol{x{:}JJ,\ y{:}JC,\ \neg JJ,\ SJ,\ SC} \vdash \boldsymbol{t{:}[(JJ \rightarrow JC) \wedge (SJ \rightarrow SC)]}$$

iff

$$\vdash \boldsymbol{x{:}JJ \wedge y{:}JC \wedge \neg JJ \wedge SJ \wedge SC \rightarrow t{:}[(JJ \rightarrow JC) \wedge (SJ \rightarrow SC)]}\ .$$

By the Completeness Theorem for **J**, to refute the latter, it suffices to find a Fitting model where at a certain node all

$$\boldsymbol{x{:}JJ,\ y{:}JC,\ \neg JJ,\ SJ,\ SC}$$

hold, but

$$\boldsymbol{t{:}[(JJ \rightarrow JC) \wedge (SJ \rightarrow SC)]}$$

does not hold for any justification term ***t***.

Formalizing Gettier Examples, Case I

A countermodel in J

W={1,2}, R={(1,2)}, *ℰ is total, i.e.,* ***i***∈*ℰ*(***t***,*F*) *for each* ***F***,***t***,*i.*

'beliefs' **2 •** ***JJ***, ***JC***, ***SJ***, **¬*SC***

↑

'real world' **1 •** ***SJ***, ***SC***, ***JC***, **¬*JJ***

At **1,** all Gettier's assumptions hold, including ***x:JJ*** and ***y:JC*** (check this out!), but formulas

$$t:[(JJ\rightarrow JC)\wedge(SJ\rightarrow SC)]$$

are false at **1** for all ***t***'s, since ***(JJ→JC)∧(SJ→SC)*** is false at **2** which is accessible from **1**.

Formalizing Gettier Examples, Case I

Making default assumptions explicit

Apparently, Gettier meant to use additional default assumptions of a non-logical character in his reasoning in Case 1, e.g., that Smith **had a justified belief that 'Jones and Smith cannot both have the job,'**

$$z{:}(JJ \rightarrow \neg SJ).$$

Adding this assumption makes the reasoning complete. Note that assuming just $JJ \rightarrow \neg SJ$ is not sufficient! This fact should be justified to Smith.

Formalizing Gettier Examples, Case I

Augmented syntax:

- ***JJ*** = *Jones gets the job,*
- ***SJ*** = *Smith gets the job,*
- ***JC*** = *Jones has 10 coins in his pocket,*
- ***SC*** = *Smith has 10 coins in his pocket,*
- ***x*** = *whatever evidence Smith had about* ***JJ,***
- ***y*** = *whatever evidence Smith had about* ***JC,***
- ***z*** = *whatever evidence Smith had about* ***JJ → ¬SJ***.

Formalizing Gettier Examples, Case I

Augmented set of non-logical assumptions:

1. ***x:JJ*** *(**x** is a justification of 'Jones gets the job')*
2. ***y:JC*** *(**y** is a justification of 'Jones has 10 coins in his pocket')*
3. ***¬JJ*** *(Jones does not get the job)*
4. ***SJ*** *(Smith gets the job)*
5. ***SC*** *(Smith has 10 coins in his pocket)*
6. ***z:(JJ→¬SJ)*** *(**z** is a justification of 'Jones and Smith cannot both have the job')*

Formalizing Gettier Examples, Case I

Augmented set of assumptions is consistent in J.

Its natural **Fitting model** is:

'beliefs' **2 • *JJ, JC, SC*, ¬*SJ***

↑

'real world' **1 • *SJ, SC, JC,* ¬*JJ***

Evidence function justifies axioms by constants

$\mathcal{E}(c,A) = \mathcal{E}(x,JJ) = \mathcal{E}(y,JC) = \mathcal{E}(z,JJ \rightarrow \neg SJ) = \{1,2\}$.

In the whole model $1,2 \vDash x{:}JJ,\quad y{:}JC,\quad z{:}(JJ \rightarrow \neg SJ)$.

Formalizing Gettier Examples, Case I

Derivation from the augmented set of assumptions

7. ***(z·x): ¬SJ***, *from 1,6, by* ***Application***

8. ***p:(¬SJ→(SJ→SC))***, ***Internalization*** *of a tautology*

9. ***(z·x):¬SJ→(p·(z·x)):(SJ→SC)***, *by* ***Application***

10. ***(p·(z·x)):(SJ→SC)***, *from 7,9, by* ***Modus Ponens***

11. ***c:(JC→(JJ→JC))***, *by* ***Axiom Internalization***

12. ***y:JC→(c·y):(JJ→JC)***, *by* ***Application***

13. ***(c·y):(JJ→JC)***, *from 2,12, by* ***Modus Ponens***

14. ***t:[(JJ→JC) ∧ (SJ→SC)],*** *for some* ***t***, *from 10 and 13 (see 'Another example')*

Formalizing Gettier Examples, Case I

Hence, Gettier's conclusion:

Smith's belief that 'the man who will get the job has 10 coins in his pocket' was justified and true.

can be formalized and derived in **J** from the augmented set of Gettier assumptions.

Formalizing Gettier Examples, Case I

What principles have been used?

The above derivation relies on **classical logic**, as well as **Application** and **Axiom Internalization**.

Steps 8, 11, and 14 use the Axiom Internalization and specify certain constants as atomic justifications of the corresponding logical axioms.

Note that ***Monotonicity*** (***Sum***) has not been used.

Formalizing Gettier Examples, Case I

Let us look again at the natural **J**-model for Case I:

$$\text{'beliefs'}\ \ \mathbf{2} \bullet \textbf{\textit{JJ, JC, SC}}, \neg\textbf{\textit{SJ}}$$

$$\uparrow$$

$$\text{'real world'}\ \ \mathbf{1} \bullet \textbf{\textit{SJ, SC, JC}}, \neg\textbf{\textit{JJ}}$$

$$\mathcal{E}(x,JJ) = \mathcal{E}(y,JC) = \mathcal{E}(z,JJ\rightarrow\neg SJ) = \{1,2\}$$

This **J***-model is good for partial justifications only; it is* ***not an adequate justification model*** *(not* **AJ***-model), e.g.,* **1** *forces both* ***x:JJ*** *and* **¬*JJ***, which contradicts reflexivity.

Formalizing Gettier Examples, Case I

Moreover, Case I is apparently inconsistent in AJ.

Here is an obvious derivation of a contradiction:

1. ***x:JJ*** *(**x** is a justification of 'Jones will get the job')*
2. ***x:JJ→JJ*** *(**Reflexivity Axiom** of **AJ**)*
3. ***JJ*** *(by **Modus Ponens**)*
4. ***¬JJ*** *(another assumption)*

Streamlined Case I: no coins/pockets!

Smith has applied for a job, but has a justified belief that 'Jones will get the job.' Smith therefore (justifiably) concludes ... that 'either Jones or Smith will get the job.'

In fact, Jones does not get the job. Instead, Smith does. So Smith's belief that ' either Jones or Smith will get the job' was justified and true. But it does not appear to be knowledge.

Formalizing Gettier: Summary

- Gettier's Examples are formalizable in the most basic system of Justification Logic. Formalization has made visible using non-logical assumptions in Gettier's reasoning in Case I, which were not listed explicitly.
- Formal analysis has also demonstrated redundancies in Gettier example I: coins and pockets are irrelevant.
- Gettier Examples are **inconsistent in logics of factive justifications**, the ones that correspond to knowledge.

Whether old methods of modal logic work here as well?

Gettier example can be easily formalized in modal logic **K**, but an appropriate motivation of its relation to justifications goes through the Realization Theorem for **K** which claims that a derivation in **K** generates a justification reasoning.

Hence MORE Justification Logic!

Job-seeking epistemologist example

*Smith has a strong piece of evidence (**b**) that he will obtain a **F**aculty position (proposition **F**) based on the fact that his **b**ook on epistemology is universally admired. In addition, Smith has good reason (**v**) to believe that his earlier Silicon **V**alley experience alone is also quite sufficient to win this job. In fact, the hiring committee could not care less about epistemology and Smith gets the job based on his Silicon Valley experience. So Smith's belief that **F** based on `**b**' is a case of Justified True Belief, but not knowledge.*

Job-seeking epistemologist example

*Smith has a strong piece of evidence (**b**) that he will obtain a **F**aculty position (proposition **F**) based on the fact that his **b**ook on epistemology is universally admired. In addition, Smith has good reason (**v**) to believe that his earlier Silicon **V**alley experience alone is also quite sufficient to win this job. In fact, the hiring committee could not care less about epistemology and Smith gets the job based on his Silicon Valley experience. So Smith's belief that **F** based on `**b**' is a case of Justified True Belief, but not knowledge.*

Naturally formalized in JL : ***{b:F, v:F, v:F→F}***. Any proof of F here requires v, truth tracking by proving!

Job-seeking epistemologist example

*Smith has a strong piece of evidence (**b**) that he will obtain a **F**aculty position (proposition **F**) based on the fact that his **b**ook on epistemology is universally admired. In addition, Smith has good reason (**v**) to believe that his earlier Silicon **V**alley experience alone is also quite sufficient to win this job. In fact, the hiring committee could not care less about epistemology and Smith gets the job based on his Silicon Valley experience. So Smith's belief that **F** based on `**b**' is a case of Justified True Belief, but not knowledge.*

Modal logic does not capture the idea: $\{\mathbf{B}F, \mathbf{B}F, \mathbf{B}F \rightarrow F\}$

Other Kinds of Knowledge: Empirical, Perceptual, A Priori, etc.

It remains to be seen to what extent Justification Logic can be useful for analysis of *empirical, perceptual, and a priori* types of knowledge. From the perspective of Justification Logic, such knowledge may be considered as justified by

constants = atomic justifications,

ready to be incorporated into reasoning with other justifications according to standards of **Applicability, Monotonicity, Logical Awareness,** and **Adequacy**.

Applications (so far)

- A complete axiomatization of mathematical proofs by means of the Logic of Proofs **LP** (=**AJ4**). This answers a long-standing question discussed by Kolmogorov and Gödel in 1930s.
- **New foundations** for Hintikka epistemic modal logic. According to the Correspondence Theorem, '***F is known***' can be read as

  '***there is a sufficient justification of F***.'
- Non-Kripkian 'existential' semantics for modal logics.

Applications (so far)

- A new approach to the **Logical Omniscience Problem**; justification terms show how hard it is to obtain knowledge from initial assumptions.
- A new approach to common knowledge in AI: **justified common knowledge** provides a more efficient alternative here.
- Applications are anticipated in the areas where epistemic modal logic is used, e.g., Game Theory and Economics, Decision Theory, etc.

Future work

Major **foundational problems** here are

- *structure of realizations,*
- *multi-agent justifications,*
- *belief revision,*
- *justifications in non-monotonic reasoning,*
- *more applications.*