First-order justification logic JT45

Felipe Salvatore[1]

[1]graduate student in philosophy
University of São Paulo

USP, April 02, 2015

Motivations

Justification logic: a brief introduction

First-order justification logic

First-order JT45

Discussion: Realization

Discussion: Interpolation

Motivations

- Introduce justification terms into epistemic first-order logic.
- Investigate the connection between justification logic and modal logic; specifically the role of the *Interpolation Theorem*.
  - The Interpolation Theorem fails for first-order S5 (FOS5) [4].
  - The Interpolation Theorem can be restored:
    - i) Restoration through the mechanism of hybrid logic [1].
    - ii) Restoration through propositional quantification [5].

Intuiitionistic logic

In the debate around foundations of mathematics one of the philosophical positions that arose was Brouwer's intuitionism.

Briefly, intuitionism says that the truth of a mathematical statement should be identified with the proof of that statement. Summarizing the core idea of this position in a slogan:

truth means provability

BHK semantics

The Brouwer–Heyting–Kolmogorov (BHK) semantics gives an informal meaning to the logical connectives $\bot, \wedge, \vee, \rightarrow, \neg$ in the following way:

- $\bot$ is a proposition which has no proof (a absurdity, e.g. $0 = 1$).
- A proof of $\varphi \wedge \psi$ consist of a proof of φ and a proof of ψ.
- A proof of $\varphi \vee \psi$ is given by exhibiting either a proof of φ or a proof of ψ.
- A proof of $\varphi \rightarrow \psi$ is a construction *f* transforming any proof *t* of φ into a proof $f(t)$ of ψ.
- A proof of $\neg\varphi$ is a construction which transforms any proof of φ into a proof of a contradiction.

Gödel 1933

In [7] Gödel introduced a new unary operator B to classical logic; $B\varphi$ should be read as 'φ is provable'. To describe the behavior of this operator Gödel constructed the following calculus (S4):

All tautologies

$B\varphi \to \varphi$

$B(\varphi \to \psi) \to (B\varphi \to B\psi)$

$B\varphi \to BB\varphi$

(*Modus Ponens*) $\vdash \varphi, \vdash \varphi \to \psi \Rightarrow \vdash \psi$

(*Internalization*) $\vdash \varphi \Rightarrow \vdash B\varphi$

Gödel 1933

Based on the intuitionistic notion of truth as provability, Gödel defined the following translation:

- $p^B = Bp$;
- $\bot^B = \bot$;
- $(\varphi \wedge \psi)^B = (\varphi^B \wedge \psi^B)$;
- $(\varphi \vee \psi)^B = (\varphi^B \vee \psi^B)$;
- $(\varphi \rightarrow \psi)^B = B(\varphi^B \rightarrow \psi^B)$.

It was shown that this translation 'makes sense', i.e., that the following theorem holds:

For every formula φ, Int $\vdash \varphi$ iff S4 $\vdash \varphi^B$.

Gödel 1933

In [7] Gödel pointed out that S4 does not correspond to the calculus of the predicate $\mathit{Prov}(x)$ – $\exists y \mathit{Proof}(y,x)$ – in **PA.** Simply because S4 proves the formula:

$$B(B(\bot) \to \bot)$$

If we translate this formula in the language of **PA**:

$$\mathit{Prov}(\ulcorner \mathit{Prov}(\ulcorner \bot \urcorner) \to \bot \urcorner)$$

Since the following sentences are equivalent in **PA**:

$$\mathit{Prov}(\ulcorner \bot \urcorner) \to \bot$$
$$\neg \mathit{Prov}(\ulcorner \bot \urcorner)$$
$$\mathit{Consist}(\mathbf{PA})$$

$\mathit{Prov}(\ulcorner \mathit{Prov}(\ulcorner \bot \urcorner) \to \bot \urcorner)$ means that the consistency of **PA** is internally provable in **PA**, which contradicts Gödel's Second Incompleteness Theorem.

Gödel 1938

In a lecture in 1938 [8] Gödel suggested a way to remedy this problem. Instead of using the implicit representation of proofs by the existential quantifier in the formula $\exists y Proof(y, x)$ one can use explicit variables for proofs (like t) in the formula $Proof(t, x)$. In these lines, Gödel proposed the following ternary operator

$$tB(\varphi, \psi)$$

which should be read as

't is a derivation of ψ from φ'

Gödel 1938

Using $tB(\varphi)$ as an abbreviation of $tB(\top, \varphi)$, Gödel formulate the following axiom system:

All tautologies

$tB(\varphi) \to \varphi$

$tB(\varphi, \psi) \to (sB(\psi, \theta) \to f(t, s)B(\varphi, \theta))$

$tB(\varphi) \to t'B(tB(\varphi))$

(*Modus Ponens*) $\vdash \varphi, \vdash \varphi \to \psi \Rightarrow \vdash \psi$;

(*Internalization*) $\vdash \varphi \Rightarrow \vdash tB(\varphi)$ (where t is an derivation of φ).

Logic of Proofs (LP)

Independently of Gödel's system presented in [8] (the lecture was published only in 1998), Artemov (in [2]) propose one new logic called Logic of Proofs (LP) which is axiomatized by the following system:

All tautologies

$t{:}\varphi \to \varphi$

$t{:}(\varphi \to \psi) \to (s{:}\varphi \to [t \cdot s]{:}\psi)$

$t{:}\varphi \to\, !t{:}t{:}\varphi$

$t{:}\varphi \to [t + s]{:}\varphi$

$s{:}\varphi \to [t + s]{:}\varphi$

(*Modus Ponens*) $\vdash \varphi, \vdash \varphi \to \psi \Rightarrow \vdash \psi$;

(*axiom necessitation*) $\vdash c{:}\varphi$, where φ is an axiom and c is a justification constant.

Logic of Proofs (LP)

LP can mirror derivations in S4. For example:

In S4:

1. $p \rightarrow (p \vee q)$ (tautology)
2. $\Box(p \rightarrow (p \vee q))$ (necessitation)
3. $\Box p \rightarrow \Box(p \vee q)$ (distribution)
4. $q \rightarrow (p \vee q)$(tautology)
5. $\Box(q \rightarrow (p \vee q))$ (necessitation)
6. $\Box q \rightarrow \Box(p \vee q)$ (distribution)
7. $(\Box p \vee \Box q) \rightarrow \Box(p \vee q)$ (classical reasoning)

Logic of Proofs (LP)

In LP:

1. $p \to (p \vee q)$ (tautology)
2. $c_1{:}(p \to (p \vee q))$ (axiom necessitation)
3. $c_1{:}(p \to (p \vee q)) \to (x{:}p \to [c_1 \cdot x]{:}(p \vee q))$
4. $x{:}p \to [c_1 \cdot x]{:}(p \vee q)$ (modus ponens)
5. $q \to (p \vee q)$(tautology)
6. $c_2{:}(q \to (p \vee q))$ (axiom necessitation)
7. $c_2{:}(q \to (p \vee q)) \to (y{:}q \to [c_2 \cdot y]{:}(p \vee q))$
8. $y{:}q \to [c_2 \cdot y]{:}(p \vee q)$ (modus ponens)
9. $[c_1 \cdot x]{:}(p \vee q) \to [[c_1 \cdot x] + [c_2 \cdot y]]{:}(p \vee q)$
10. $[c_2 \cdot y]{:}(p \vee q) \to [[c_1 \cdot x] + [c_2 \cdot y]]{:}(p \vee q)$
11. $(x{:}p \vee y{:}q) \to [[c_1 \cdot x] + [c_2 \cdot y]]{:}(p \vee q)$ (classical reasoning)

Logic of Proofs (LP)

If φ is a S4 formula, there is a mapping r (called a *realization*) from the occurrences of B's (or boxes) into terms. The result of this mapping on φ is denoted φ^r. For example:

$$((\Box p \vee \Box q) \rightarrow \Box(p \vee q))^r$$
$$=$$
$$(x{:}p \vee y{:}q) \rightarrow [[c_1 \cdot x] + [c_2 \cdot y]]{:}(p \vee q)$$

(Realization Theorem between S4 and LP) For every φ in the language of S4, there is a realization r such that

$$\text{S4} \vdash \varphi \text{ iff } \text{LP} \vdash \varphi^r$$

There is a way to define an interpretation $*$ of the LP formulas into the sentences of **PA** (for details see [2]). And with all this machinery Artemov was able to prove the following result:

(Provability Completeness of Intuitionistic Logic) For every φ, for every interpretation $*$, there is a realization r such that

$$\mathsf{Int} \vdash \varphi \text{ iff } \mathsf{S4} \vdash \varphi^B \text{ iff } \mathsf{LP} \vdash (\varphi^B)^r \text{ iff } \mathbf{PA} \vdash ((\varphi^B)^r)^*$$

LP is just one example of *Justification Logic*. Another example, that is interesting to us, is the one called JT45, it extends the language of LP with the unary justification operator ? and has the following additional axiom scheme:

$$\neg t{:}\varphi \to ?t{:}\neg t{:}\varphi$$

We can prove the realization theorem for this logic too!

(Realization Theorem between S5 and JT45) For every φ in the language of S5, there is a realization r such that

$$\mathrm{S5} \vdash \varphi \text{ iff } \mathrm{JT45} \vdash \varphi^r$$

Motivations

Justification logic: a brief introduction

First-order justification logic

First-order JT45

Discussion: Realization

Discussion: Interpolation

From propositional justification logic to first-order

Let $\varphi(x)$ be any tautology, and let t be the following derivation:

1. $\varphi(x)$
2. $\forall x \varphi(x)$ (generalization)
3. $\forall x \varphi(x) \to (Q(x) \to \forall x \varphi(x))$ (tautology)
4. $Q(x) \to \forall x \varphi(x)$ (Modus Ponens)

Although x is free in the formula $Q(x) \to \forall x \varphi(x)$, if c is a term we can not substitute c for x in t in order to obtain a derivation $t(c)$ of $Q(c) \to \forall x \varphi(x)$ (if we do that we ruin the derivation at 2.).

From propositional justification logic to first-order

Now, let s be the following derivation:

1. $\varphi(x)$
2. $\forall x \varphi(x)$ (generalization)
3. $\forall x \varphi(x) \rightarrow (Q(y) \rightarrow \forall x \varphi(x))$ (tautology)
4. $Q(y) \rightarrow \forall x \varphi(x)$ (Modus Ponens)

y is free in the formula $Q(y) \rightarrow \forall x \varphi(x)$ and moreover for every term c the result of substituting c for y in s, $s(c/y)$, is the derivation of $Q(c) \rightarrow \forall x \varphi(x)$.

From propositional justification logic to first-order

These examples show us that there are two different roles of variables in a derivation: a variable can be a *formal symbol* that can be subjected to generalization or a *place-holder* that can be substituted for.

In t, x is both a formal symbol and a place-holder. And in s, x is a formal symbol and y is a place-holder.

This consideration motivates the following definition:

x is free in the derivation t of the formula φ iff for every term c, $t(c/x)$ is the derivation of $\varphi(c/x)$.

From propositional justification logic to first-order

In propositional justification logic we write $t{:}\varphi$ to express that t is a derivation of φ. In order to represent the distinct roles of variables in the first-order justification logic, we are going to write formulas of the form:

$$t{:}Q(x) \rightarrow \forall x \varphi(x)$$

$$s{:}_{\{y\}} Q(y) \rightarrow \forall x \varphi(x)$$

The role of $\{y\}$ in $s{:}_{\{y\}} Q(y) \rightarrow \forall x \varphi(x)$ is to point out that y is free in the derivation s of $Q(y) \rightarrow \forall x \varphi(x)$.

Motivations

Justification logic: a brief introduction

First-order justification logic

First-order JT45

Discussion: Realization

Discussion: Interpolation

Language of first-order JT45

The basic definitions that we present here are taken from the technical report of Artemov and Yarvorskaya [3].

Justification Terms

$$t := p_i \mid c_i \mid (t_1 \cdot t_2) \mid (t_1 + t_2) \mid\, !t \mid\, ?t \mid gen_x(t)$$

Formulas

$$\varphi := Q(x_1, \ldots, x_n) \mid \bot \mid \varphi \to \psi \mid \forall x \varphi \mid t{:}_X \varphi$$

Language of first-order JT45

Where $X, Y, \ldots$ are variables for finite set of individual variables. We write Xy instead of $X \cup \{y\}$, in this case it is assumed that $y \notin X$. We use $t{:}\varphi$ as an abbreviation for $t{:}_{\emptyset}\varphi$. And we write L to denote the set of formulas.

We define the notion of free variables of φ, $fv(\varphi)$, by induction similarly as in the classical case, the new clause is

- If φ is $t{:}_X\psi$, then $fv(\varphi)$ is X.

First-order JT45: axiom system

First-order JT45 (FOJT45) is axiomatized by the following schemes and inference rules:

A1 classical axioms of first-order logic
A2 $t{:}_{Xy}\varphi \rightarrow t{:}_X\varphi$, provided y does not occur free in φ
A3 $t{:}_X\varphi \rightarrow t{:}_{Xy}\varphi$
B1 $t{:}_X\varphi \rightarrow \varphi$
B2 $s{:}_X(\varphi \rightarrow \psi) \rightarrow (t{:}_X\varphi \rightarrow [t \cdot s]{:}_X\psi)$
B3 $t{:}_X\varphi \rightarrow [t + s]{:}_X\varphi$, $s{:}_X\varphi \rightarrow [t + s]{:}_X\varphi$
B4 $t{:}_X\varphi \rightarrow !t{:}_X t{:}_X\varphi$
B5 $\neg t{:}_X\varphi \rightarrow ?t{:}_X\neg t{:}_X\varphi$
B6 $t{:}_X\varphi \rightarrow gen_x(t){:}_X\forall x\varphi$, provided $x \notin X$
R1 (*Modus Ponens*) $\vdash \varphi, \vdash \varphi \rightarrow \psi \Rightarrow \vdash \psi$
R2 (*generalization*) $\vdash \varphi \Rightarrow \vdash \forall x\varphi$
R3 (*axiom necessitation*) $\vdash c{:}\varphi$, where φ is an axiom and c is a justification constant.

Basic Results

Theorem

(*Internalization*) Let $p_0, \ldots, p_k$ be justification variables; $X_0, \ldots, X_k$ be finite sets of individual variables, and $X = X_0 \cup \cdots \cup X_k$. In these conditions, if $p_0{:}_{X_0}\varphi_0, \ldots, p_k{:}_{X_k}\varphi_k \vdash \psi$, then there is a justification term $t(p_0, \ldots, p_k)$ such that

$$p_0{:}_{X_0}\varphi_0, \ldots, p_k{:}_{X_k}\varphi_k \vdash t{:}_X\psi.$$

Basic Results

Proposition

(*Explicit counterpart of the Barcan Formula and its converse*)
For every formula $\varphi(x)$ and every justification term t, there are justification terms $CB(t)$ and $B(t)$ such that:

$$\vdash t{:}\forall x\varphi(x) \rightarrow \forall x CB(t){:}_{\{x\}}\varphi(x)$$

$$\vdash \forall x t{:}_{\{x\}}\varphi(x) \rightarrow B(t){:}\forall x\varphi(x)$$

- $t : \forall x\varphi(x) \rightarrow \forall x[c \cdot t]{:}_{\{x\}}\ \varphi(x)$
- $\forall x t{:}_{\{x\}}\varphi(x) \rightarrow [r \cdot ?[[c_3 \cdot [c_2 \cdot gen_x(c_1)]] \cdot ?t]]{:}\ \forall x\varphi(x)$

Basic Results

Proposition

(*Explicit counterpart of the Barcan Formula and its converse*)
For every formula $\varphi(x)$ and every justification term t, there are justification terms $CB(t)$ and $B(t)$ such that:

$$\vdash t{:}\forall x\varphi(x) \rightarrow \forall x CB(t){:}_{\{x\}}\varphi(x)$$

$$\vdash \forall x t{:}_{\{x\}}\varphi(x) \rightarrow B(t){:}\forall x\varphi(x)$$

- $t : \forall x\varphi(x) \rightarrow \forall x[c \cdot t]{:}_{\{x\}}\ \varphi(x)$
- $\forall x t{:}_{\{x\}}\varphi(x) \rightarrow [r \cdot ?[[c_3 \cdot [c_2 \cdot gen_x(c_1)]] \cdot ?t]]{:}\ \forall x\varphi(x)$

Basic Results

Proposition

(*Explicit counterpart of the Barcan Formula and its converse*)
For every formula $\varphi(x)$ and every justification term t, there are justification terms $CB(t)$ and $B(t)$ such that:

$$\vdash t{:}\forall x\varphi(x) \rightarrow \forall x CB(t){:}_{\{x\}}\varphi(x)$$

$$\vdash \forall x t{:}_{\{x\}}\varphi(x) \rightarrow B(t){:}\forall x\varphi(x)$$

- $t : \forall x\varphi(x) \rightarrow \forall x[c \cdot t]{:}_{\{x\}}\ \varphi(x)$
- $\forall x t{:}_{\{x\}}\varphi(x) \rightarrow [r \cdot ?[[c_3 \cdot [c_2 \cdot gen_x(c_1)]] \cdot ?t]]{:}\ \forall x\varphi(x)$

Semantics

A possible world semantics for first-order LP is presented in Fitting [6]. We have adopted the definitions of this paper for FOJT45 and with this definitions we were able to prove a completeness theorem for this logic.

But we leave the semantical part for a future talk...

Motivations

Justification logic: a brief introduction

First-order justification logic

First-order JT45

Discussion: Realization

Discussion: Interpolation

Realization

Let φ be a formula of FOS5. We define the realization of φ in the language of FOJT45, φ^r, as follows:

- If φ is atomic, then $\varphi^r = \varphi$.
- If $\varphi = \psi \rightarrow \theta$, then $\varphi^r = \psi^r \rightarrow \theta^r$
- If $\varphi = \forall x \psi$, then $\varphi^r = \forall x \psi^r$
- If $\varphi = \Box\psi(x_1, \ldots, x_n)$, then $\varphi^r = t{:}_{\{x_1,\ldots,x_n\}}\psi^r$

A realization is normal if all negative occurrences of $\Box$ are assigned justification variables. It can easily be checked that

For every φ, $fv(\varphi) = fv(\varphi^r)$

Realization

Let φ be a formula of FOJT45. The forgetful projection of φ, φ°, is defined as follows:

- If φ is atomic, then $\varphi^\circ = \varphi$.
- If $\varphi = \psi \to \theta$, then $\varphi^\circ = \psi^\circ \to \theta^\circ$
- If $\varphi = \forall x\psi$, then $\varphi^\circ = \forall x\psi^\circ$
- If $\varphi = t{:}_X\psi$, then $\varphi^\circ = \Box\forall\vec{y}\psi^\circ$ where $\vec{y} \in fv(\psi)\backslash X$.

As before, it can easily be checked that

$$\text{For every } \varphi,\ fv(\varphi) = fv(\varphi^\circ)$$

Realization

Proposition

For every justification formula φ,

$$\text{If FOJT45} \vdash \varphi, \text{ then FOS5} \vdash \varphi^{\circ}.$$

Theorem

(Realization) If FOS5 $\vdash \varphi$, then FOJT45 $\vdash \varphi^r$ for a normal realization r.

Motivations

Justification logic: a brief introduction

First-order justification logic

First-order JT45

Discussion: Realization

Discussion: Interpolation

Interpolation

- *The Interpolation Theorem* holds for FOS5 iff for every sentences φ and ψ if $\vdash \varphi \to \psi$, then there is a formula θ such that $\vdash \varphi \to \theta$, $\vdash \theta \to \psi$ and the non-logical symbols that occur in θ occur both in φ and ψ.
- *The Interpolation Theorem* holds for FOJT45 iff for sentences φ and ψ if $\vdash \varphi \to \psi$, then there is a formula θ such that $\vdash \varphi \to \theta$, $\vdash \theta \to \psi$ and the non-logical symbols and the justification terms that occur in θ occur both in φ and ψ.

Interpolation

Proposition

If the Realization Theorem holds between FOS5 and FOJT45, then the Interpolation Theorem fails for FOJT45.

Proof

Suppose that the Interpolation Theorem holds for FOJT45. By [4], let φ and ψ be sentences such that FOS5 $\vdash \varphi \to \psi$ and there is no interpolant between them.

By the Realization Theorem, there is a normal realization *r* such that

FOJT45 $\vdash \varphi^r \to \psi^r$

Interpolation

By hypothesis, there is a formula θ such that the non-logical symbols and the justification terms that occur in θ occur both in φ^r and ψ^r.

FOJT45 $\vdash \varphi^r \to \theta$
FOJT45 $\vdash \theta \to \psi^r$

By the forgetful projection:

FOS5 $\vdash (\varphi^r \to \theta)^\circ$
FOS5 $\vdash (\theta \to \psi^r)^\circ$

i.e.,

FOS5 $\vdash \varphi \to \theta^\circ$
FOS5 $\vdash \theta^\circ \to \psi$

Interpolation

Now, since there is no interpolant between φ and ψ, then there is no relation symbol occurring in θ°. Hence, θ° is a formula such that $\perp$ is the only atomic formula that occur in θ°. Thus, either θ° is valid or θ° is unsatisfiable.

If θ° is valid, then, since $\models \theta^\circ \to \psi$, ψ is valid. And so, $\varphi \to \psi$ has an interpolant, a contradiction.

If θ° is unsatisfiable, then, since $\models \varphi \to \theta^\circ$, φ is unsatisfiable. And so, $\varphi \to \psi$ has an interpolant, a contradiction.

Thank you
for your attention.

[1] Carlos Areces, Patrick Blackburn, and Maarten Marx, *Repairing the interpolation theorem in quantified modal logic*, Annals of Pure and Applied Logic **124** (2001), 287–299.

[2] Sergei Artemov, *Explicit provability and constructive semantics*, The Bulletin of Symbolic Logic **7** (2001), no. 1, 1–36.

[3] Sergei Artemov and Tatiana Yavorskaya (Sidon), *First-order logic of proofs*, Technical Report TR-20111005, CUNY Ph.D Program in Computer Science, May 2011.

[4] Kit Fine, *Failures of the interpolation lemma in quantified modal logic*, The Journal of Symbolic Logic **44** (1979), no. 2, 201–206.

[5] Melvin Fitting, *Interpolation for first-order S5*, The Journal of Symbolic Logic **67** (2002), no. 2, 621–634.

References II

[6] ______ , *Possible world semantics for first order LP*, Annals of Pure and Applied Logic **165** (2014), 225–240.

[7] Kurt Gödel, *An interpretation of the intuitionistic propositional calculus*, Kurt Gödel, Collected Works (Solomon Feferman et al, ed.), vol. 1, Oxford University Press, 1986, pp. 296–303.

[8] ______ , *Lecture at Zilsel's*, Kurt Gödel, Collected Works (Solomon Feferman et al, ed.), vol. 3, Oxford University Press, 1995, pp. 86–113.