Explicit and Implicit Belief in an Hyperintensional Framework

Jiří Raclavský
(raclavsky@phil.muni.cz)

Department of Philosophy, Masaryk University, Brno (Gödel’s hometown)

Jednota filozofická, KF FF UP Olomouc, 24/2/2016
Abstract

- analysis of belief attitudes is one of central tasks in epistemology, logic and philosophy of language.
- according to intensional (possible world) semantics, agent’s believing is a relation towards a proposition $P$; $P$ is modelled as possible world proposition.
- in consequence of this analysis, an agent so-called implicitly believes (the case of implicit belief) all logical consequences of $P$, it is thus logically omniscient.
- contrary to it, hyperintensional semantics treats an agent as a logical idiot who is not capable of any such inference.
- according to hyperintensional semantics, the agent has an attitude explicitly towards $P$ (the case of explicit belief), whereas $P$ is modelled in a certain hyperintensional way.
- many argued that both approaches are too extreme (one is too benevolent, the latter is too restrictive) and that we need a middle approach.
- in this lecture, I offer a proposal which proposes an explicit belief modelled a specific way, and also an implicit belief which stems from the explicit approach.
The story is based on ...

**Figure:** Raclavský, Jiří, Kuchynka, Petr, Pezlar, Ivo (2015): *Transparentní intenzionální logika jako characteristic universalis a calculus ratiocinator*. ISBN 978-80-210-7973-1, Brno: Masarykova univerzita (Munipress).
1. Frege’s diagnosis and the cure of (SI) failure

2. TIL4.0 as a tool of logical analysis

3. Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes

4. Implicit belief and TIL4.0

5. Conclusions

6. References
Content

1. Frege’s diagnosis and the cure of (SI) failure
2. TIL4.0 as a tool of logical analysis
3. Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes
4. Implicit belief and TIL4.0
5. Conclusions
6. References
General aim – controlling transmission of knowledge

- the aim of logic and also formal epistemology is a conscious supervision of the process of *deriving knowledge* from the set of already given knowledge
- knowledge is transmitted thorough *inferences*, which can be understood simply as *arguments*
- every linguistically articulated argument is of form

\[ S_1, S_2, \ldots, S_n \therefore S \]

where \( S_i \) are sentences

- *logical analysis* explicates the meanings of \( S_1, S_2, \ldots, S_n, S \) as

\[ C_1, C_2, \ldots, C_n, C \]

where \( C_i \) are (extralinguistic) objects of a certain logical system
Principle of substitutivity of identicals (SI)

- Leibniz’s *Principle of substitutivity of identicals (SI)*

\[
\begin{align*}
\ldots E_1 \ldots \\
E_1 = E_2 \\
\ldots E_2 \ldots
\end{align*}
\]

if the identity statement \(E_1 = E_2\) holds, the conclusion \(\ldots E_2 \ldots\) is logically equivalent to the major premise \(\ldots E_1 \ldots\).

- note that we formulated here (SI) as a rule operating on expressions.
- many theoreticians uncritically assume that (SI) is isomorphic to (SI*) which operates on meanings of those expressions.
Frege’s finding: (SI) fails in intensional contexts

- Frege (1892) first noticed that an application of (SI) fails in cases which deploy belief sentences

\[
\begin{align*}
&\text{“}X\text{ believes that } \ldots E_1 \ldots\text{”} \\
\Rightarrow &\text{“}E_1 = E_2\text{”} \\
\Rightarrow &\text{“}X\text{ believes that } \ldots E_2 \ldots\text{”}
\end{align*}
\]

- Frege’s diagnosis of failure: expressions “\(E_1\)” and “\(E_2\)” occur
  a. within “\(E_1 = E_2\)”\(\) in extensional (gerade, transparent) context, while
  b. within the major premise and conclusion: in intensional (ungerade, oblique) context
Frege’s cure: (SI*)

- Frege’s instructive step was a move from (SI) to (SI*)
- Frege (Church, Tichý, i.e. (neo-)fregeans) suggested to differentiate
  a. meaning (Sinn, concept) of an expression
  b. denotatum (Bedeutung, denotation) of an expression
- example: the meaning of “morning star” is MORNING STAR, whereas its denotatum is the planet Venus
- Frege’s logical analysis of the argument about belief:

  $X$ believes that ... the meaning of (“$E_1$”) ...  
  the denotatum of (“$E_1$”) = the denotatum of (“$E_2$”) 
  $X$ believes that ... the meaning of (“$E_2$”) ...

- Frege’s cure of failure: when applying (SI*), it is not allowed to exchange the meaning of (“$E_i$”) and the denotatum of (“$E_i$”) because the meaning of (“$E_i$”) $\neq$ the denotatum of (“$E_i$”)
Content

1. Frege’s diagnosis and the cure of (SI) failure

2. TIL4.0 as a tool of logical analysis

3. Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes

4. Implicit belief and TIL4.0

5. Conclusions

6. References
Transparent Intensional Logic (TIL) 4.0

- *Transparent Intensional Logic* (TIL) was founded by Tichý in the very beginning of 1970s, i.e. in the time of much more famous system by Montague (1974)
- TIL is intensional in the sense of accepting *possible worlds* ($W$), which is the parameter of logical modality; it is a “possible world semantics” (PWS)
- *Intensions* are functions from possible worlds and moments of time to extensions; examples:
  - *Individual concepts-objects* (their extensions are individuals)
  - *Propositions* are intensions, the extensions for which (in various $W$’s and $T$’s) are truth-values $\top$, $\bot$ or nothing (partiality gap)

- see TIL4.0 in Raclavský, Kuchyňka, Pezlar (2015)
### Semantic scheme

- **neofregean semantic scheme in TIL4.0:**

<table>
<thead>
<tr>
<th>expression</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>meaning</strong></td>
<td>hyperintension C</td>
</tr>
<tr>
<td><strong>denotatum</strong></td>
<td>intension / non-intension</td>
</tr>
<tr>
<td>referent in $W$ at $T$</td>
<td>the value of the intension in $W$ at $T$ / non-intension</td>
</tr>
</tbody>
</table>

- whether the denotatum of an expression $E$ is an intension or not, depends on whether $E$'s referent intuitively varies along $W$s and $T$s

<table>
<thead>
<tr>
<th>the denotatum of expression:</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Yes’</td>
</tr>
<tr>
<td>“Fido is a dog”</td>
</tr>
<tr>
<td>“$1 + 2 = 3$”</td>
</tr>
<tr>
<td>is:</td>
</tr>
<tr>
<td>a truth value (viz. $\top$)</td>
</tr>
<tr>
<td>a non-constant proposition</td>
</tr>
<tr>
<td>a constant proposition</td>
</tr>
</tbody>
</table>

- hyperintensions are inevitable for explication of meanings – as has been argued recently by many writers (hyperintensions are sufficiently fine-grained, etc.)
Constructions as explicata of hyperintensions

- *constructions* – i.e. our explicata of hyperintensions – are structured, abstract *procedures* of algorithmic nature, which *construct objects*

- every object is constructed by infinitely many (*ν*-)*congruent* constructions

- for example, the analytically true proposition *Verum* is constructed by the following congruent *propositional constructions*
  - \( \lambda w \lambda t [^0 1^0 + ^0 2^0 = ^0 3^0] \)
  - \( \lambda w \lambda t [^0 1^0 = ^0 1^0] \)
  - \( \lambda w \lambda t[^0 \forall \lambda x^0 \forall \lambda y^0 \forall \lambda z^0 \forall \lambda n[[x^n 0 + y^n 0 = z^n]^0 \rightarrow [n^0 < ^0 3^0]]] \) (i.e. Fermat’s Last Theorem)
  - ...
Main semantic relations and identity statements

<table>
<thead>
<tr>
<th>expressions</th>
<th>synonymy</th>
<th>equivalence</th>
<th>coreference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_1$</td>
<td>$E_2$</td>
<td>$E_1$</td>
<td>$E_2$</td>
</tr>
<tr>
<td>$C$</td>
<td>$C_1$</td>
<td>$C_2$</td>
<td>$C_1$</td>
</tr>
<tr>
<td>$D$</td>
<td>$D$</td>
<td>$D_1$</td>
<td>$D_2$</td>
</tr>
<tr>
<td>$R$</td>
<td>$R$</td>
<td>$R$</td>
<td>$R$</td>
</tr>
</tbody>
</table>

- for expressions not denoting intensions, their coreference in $W$ and $T$ coincides with their co-denotation-equivalence

- due to this rich ‘vertical’ analysis by TIL, the identity statement

  \[ X = Y \]

  has three main readings

- where $X$ and $Y$ are constructions (expressed by “$X$” and “$Y$”, respectively) and $\equiv$ is type-theoretically appropriate
  
  a. $\lambda w \lambda t [^0 X \equiv^0 Y]$ (synonymy)
  b. $\lambda w \lambda t [X \equiv Y]$ (equivalence)
  c. $\lambda w \lambda t [X_{wt} \equiv Y_{wt}]$ (coreference)
Logical analysis and attitude logic in TIL4.0

Content

1. Frege’s diagnosis and the cure of (S1) failure

2. TIL4.0 as a tool of logical analysis

3. Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes

4. Implicit belief and TIL4.0

5. Conclusions

6. References
Explicit vs implicit belief

Levesque (1984) suggested to construe the analysis of belief sentences i. by means of PWS and ii. hyperintensional logic as presenting i. implicit and ii. explicit belief

- **in the case of implicit belief**, an agent’s attitude is understood as related exclusively towards a proposition $P$ constructed by the construction $C$

  $\rightarrow$ hence, inferences to attitudes towards any logical consequence of $C$ are not blocked

- i.e. an agent is considered to have unlimited inferential/epistemic abilities (omniscience)

- **in the case of explicit belief**, an agent’s attitude is understood as related exclusively towards a propositional construction $C$

  $\rightarrow$ hence, inferences to attitudes towards any logical consequence of $C$ are blocked
Golden rule of attitude logic (GRAL)

- general reason of invalidity of substitution during the application of $(SI^*)$, and thus also $(SI)$, is that the object of attitude $O$ was changed.

Golden rule of attitude logic (GRAL)

- When substituting during the application of $(SI^*)$ or $(SI)$ it is not allowed to change the original object $O$ of an attitude $R$ of an agent $A$.

- Intuitive justification of GRAL: when $A$ has an attitude $R$ towards $O$, in that moment, she has an attitude exclusively towards $O$.

- Note that an advantage of $(SI^*)$ in contrast to $(SI)$ is that the breach of GRAL is immediately apparent.
Wrong logical analysis of belief sentences

- conventional PWS of Montague (1974), but also Hintikka (1962), analyzed belief sentences in a *wrong way*
- belief attitudes are explicated as *relations towards propositions*
- for example, “Fermat believes that $\forall x \forall y \forall z \forall n((x^n + y^n = z^n) \rightarrow (n < 3))$.”:

$$\lambda w \lambda t[0]{\text{Believes}}_{wt}^\pi 0{\text{Fermat}}$$

$$[\lambda w \lambda t[0]{\forall x}^0{\forall y}^0{\forall z}^0{\forall n}^0[([x^n + y^n] = z^n) \rightarrow [n^0 < 0^3]]]]$$

- the construction expressed by Fermat’s Last Theorem is *not the object* of belief here, but a *proxy for an object of belief*, which is a certain proposition $P$ (namely Verum) to determine $P$, any other construction congruent with $C$ can be utilized

- *incorrectness* consists in that, on natural reading of the sentence, Fermat has an attitude towards a structured procedure, not towards an unstructured proposition (state of affairs) $P$
Implicit belief and PWS-propositions

- linguistic formulation of the following argument is often used to show that on implicit belief, an agent has unrealistic inferential/epistemic abilities

\[
\lambda w \lambda t[^0 \text{Believes}^\pi_w t][^0 A C]
\]

\[
\lambda w \lambda t[^0 C = C']
\]

\[
\lambda w \lambda t[^0 \text{Believes}^\pi_w t[^0 A C']]
\]

where \( C \) is e.g. \( \lambda w \lambda t[^0 1^0 + 0^2] = 0^3 \)

and \( C' \) is e.g. \( \lambda w \lambda t[^0 \forall \lambda x^0 \forall \lambda y^0 \forall \lambda z^0 \forall \lambda n[^0 [x^n + y^n = z^n] \rightarrow [n < 0^3]]] \)

1. however, on this reading of the linguistic argument, it is valid
2. nevertheless, on such analysis, an agent in fact does not have a belief; the analysis is quite wrong – it offers an attitude towards state-of-affairs
3. inferential abilities here demonstrated pertain to the speaker – the impression of unlimited inferential abilities is in fact based on a confusion with explicit belief (on such reading, the argument would be invalid because of GRAL)
Right logical analysis of belief sentences

- Tichý (1988:222) suggested a convincing analysis of belief sentences
  - belief attitudes are explicated as relations towards propositional constructions
- for example:

\[ \lambda w \lambda t [ ]^0 \text{Believes}_{wt}^k \text{Fermat} \]

\[ 0 [ \lambda w \lambda t [ ]^0 \forall \lambda x [ ]^0 \forall \lambda y [ ]^0 \forall \lambda z [ ]^0 \forall \lambda n [ ]^0 [ x^n + y^n = z^n ]^0 \rightarrow [ n^0 < 0^3 ] ] ] ] \]

- construction of the form \(^0C\) (so-called trivialization of \(C\)) constructs immediately \(C\)
- according to this analysis, an agent has an attitude exclusively towards the propositional construction in the scope of \(^0\)
Belief attitudes and an invalid form of (SI*)

- the following form of (SI*) is *invalid*, and thus rejected by us

\[
\lambda w \lambda t [^0 \text{Believes}_w t ^0 A ^0 C] \\
\lambda w \lambda t [C ^0 = C'] \\
\lambda w \lambda t [^0 \text{Believes}_w t ^0 A ^0 C'] \quad \text{(invalid scheme)}
\]

- though \( C \) and \( C' \) are congruent, we are *not allowed*, when conforming to GRAL, to change the object of attitude from \( C \) to \( C' \)

\( \rightarrow \) blocking thus inferences of attitudes towards consequences of \( C \)

- our cure is thus an analogue of Frege’s cure
  - the congruence of \( C \) with \( C' \) is based on identity of denotata, not meanings
  - the object of an attitude *is the meaning of the subordinate sentence*, i.e. \( C \)
  - the object of an attitude *is not a denotatum of the subordinate sentence*
substitution realized by (SI*) would consist in replacing a subconstruction \( C \) in

\[
\lambda w \lambda t [^0 \text{Believes}^k_{wt} A^0 C]
\]

technical implementation of Fregean cure within TIL4.0:

it is not possible to replace \( C \) by \( C' \) in the major premise, because \( C \) is not a proper subconstruction of the major premise

cf. the definition of subconstruction (cf. Raclavský et al. 2015)
however, there exist valid forms of (SI*) on such construal, cf. e.g.

\[
\lambda w \lambda t [^0 \text{Believes}_{wt}^0 A^0C] \\
\lambda w \lambda t [^0C^0 = D] \\
\lambda w \lambda t [^0 \text{Believes}_{wt}^k 0 A D]
\]

(SI*) is justified by substitution – the consequence is \(\nu\)-congruent with

\[
\lambda w \lambda t [^0 \text{Subst}^k 0 D^00C^0 [\lambda w \lambda t [^0 \text{Believes}_{wt}^k 0 A 0C]]]
\]

in the major premise, we substitute whole construction \(0C\) by \(D\); unlike \(C\), \(0C\) is a subconstruction of the major premise

if the minor premise is true, \(0C\) and \(D\) are \(\nu\)-congruent, i.e. they both \(\nu\)-construct \(C\); \(\rightarrow\) GRAL is preserved
Content

1. Frege’s diagnosis and the cure of (SI) failure

2. TIL4.0 as a tool of logical analysis

3. Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes

4. Implicit belief and TIL4.0

5. Conclusions

6. References
above we have seen the TIL4.0-approach to belief: belief is explicated as explicit belief

does it mean that all consideration about implicit belief are illegitimate?
- No!, the following inferences can be intuitively valid:

\[
\begin{align*}
\text{"Fermat believes } 1 + 2 &= 3." \\
\implies \\
\text{"Fermat believed that } \forall x \forall y \forall z \forall n ((x^n + y^n = z^n) \implies (n < 3))."
\end{align*}
\]

TIL4.0 is capable treat such inferences as valid
- the question is: How exactly?
before offering a proposal, consider this: could Fermat derive Fermat’s Last Theorem from the mathematics known to him, e.g. from $1 + 2 = 3$?

- intuitively, it depends on whether Fermat dealt with a sufficiently strong derivation system

*derivation system (DS)* is a couple $\langle Cs, R \rangle$, where $Cs$ is a class of constructions and $R$ a class of derivation rules

- examples of known DSs: Russell’s concept of propositional function, marxism, . . . , Peano arithmetic, Presburger arithmetic, . . .
- for details see (Raclavský, Kuchyňka 2011; Raclavský, Kuchyňka, Pezlar 2015); for *derivation rules* see the same sources
Implicit belief and TIL4.0: hidden premise

- when attributing to $A$ a belief attitude $R$ towards $C$ and then also towards $C'$, we assume that $A$ deals with a $DS$, in which it is possible for $A$ to derive $C'$ from $C$
- valid arguments about an implicit belief thus contains the corresponding hidden premise

$$
\lambda w \lambda t^k 0 \text{Believes}_{wt} 0 A \{ 0 C \}
\lambda w \lambda t^k 0 \text{IsCapableToDeriveFrom}_{wt}^S A \{ 0 C' \}
\lambda w \lambda t^k 0 \text{Believes}_{wt} 0 A \{ 0 C' \}
$$

- discussion (violation of GRAL?, implicit/explicit/new belief sets, ...)

Jiří Raclavský (raclavsky@phil.muni.cz) (Department of Philosophy, Masaryk University, Brno (Gödel's hometown), Jednotna Lozocká, KFF FF UP Olomouc, 24/2/2016)
Content

1 Frege’s diagnosis and the cure of (Sl) failure

2 TIL4.0 as a tool of logical analysis

3 Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes

4 Implicit belief and TIL4.0

5 Conclusions

6 References
**Brief conclusion**

<table>
<thead>
<tr>
<th>belief:</th>
<th>is <em>analyzed</em> as attitude towards:</th>
<th><em>inferences</em> to attitudes towards objects distinct from the given objects:</th>
</tr>
</thead>
<tbody>
<tr>
<td>explicit</td>
<td>propositional constructions</td>
<td>are blocked</td>
</tr>
<tr>
<td>implicit</td>
<td>propositional constructions</td>
<td>are not blocked: $DS_A$-derivability</td>
</tr>
</tbody>
</table>
Content

1. Frege’s diagnosis and the cure of (SI) failure
2. TIL4.0 as a tool of logical analysis
3. Logical analysis and attitude logic in TIL4.0
   - Golden rule of attitude logic
   - Logical analysis of belief attitudes – explicit belief
   - Logic of belief attitudes
4. Implicit belief and TIL4.0
5. Conclusions
6. References
References


