Revisioning Revision Theory - *Philip Welch, Amsterdam March 15, 2013.*

Die Wahrheit liegt weder in der unendlichen Annährung an einer objektiv Gegebenes noch in der Mitte, sondern rundherum wie ein Sack, der mit jeder neuen Meinung, die man hineinstopft, seine Form ändert, aber immer fester wird.

R.Musil

(Truth lies neither in an endless approximation to an objective given, nor in the centre, but rather all around like a sack, that changes its shape but becomes firmer and firmer with every new opinion that is stuffed into it.)

Revision Theories

- - of sentences - Herzberger;
- - of truth, and of "circular definitions" - Gupta; Belnap;
- - of definability -
  "Arithmetical Quasi-Inductive Definitions" - Burgess;
- - of a conditional operator $\longrightarrow$ - Field;
- - as a transfinite computational model -
  "Infinite Time Turing Machines" of Hamkins and Kidder.

Revision Mechanisms

$Q_0 = H_0 = D_0^{\Phi} = F_0 = P_{e,0} = \emptyset;$

$H_{\alpha+1} =_{df} \{\ulcorner\varphi\urcorner \mid \varphi \in \mathcal{L}_T, \langle \mathbb{N}, +, \times, \ldots, H_\alpha \rangle \models \varphi \ \};$

$D_{\alpha+1}^{\Phi} =_{df} \{n \mid \langle \mathbb{N}, +, \times, \ldots, D_\alpha^{\Phi} \rangle \models \Phi[\bar{n}]\}$ some fixed $\Phi \in \mathcal{L}_T$;

$F_{\alpha+1} =_{df} \{\ulcorner A \longrightarrow B \urcorner \mid |A| \leq |B|$ in the least sk-fpoint over F_α in $\mathcal{L}_{T,\rightarrow}\}$;

$P_{e,\alpha+1} =_{df} \{n \mid$
Cell C_n has a '1' after the e'th TM has performed the α'th comput'l step$\}$;

Then for each of these:

$Q_\lambda = Liminf_{\alpha \rightarrow \lambda} Q_\alpha = \{u \mid \exists \alpha < \lambda \forall \beta < \lambda (\alpha < \beta \rightarrow u \in Q_\beta)\}.$

Let P_U be the universal ITTM.

Let P_U be the universal ITTM.

Theorem

There is a single "stability ordinal" ζ so that for any $Q \in \{H, D^{\Phi}, F, P_U\}$

$$Q_\zeta = Q_{On}.$$

Moreover

$$H_\zeta \equiv_1 F_\zeta \equiv_1 P_{U,\zeta}$$

(where "$\equiv_1$" denotes recursive isomorphism.)

- Let T^2_ζ denote the Σ_2-$Th(L_\zeta)$. Then:

Theorem

(a) (Burgess) $D^{\Phi}_\zeta \leq_1 H_\zeta \equiv_1 T^2_\zeta$.
(b) For each of the other Q above:

$$Q \equiv_1 T^2_\zeta.$$

Uniform recoverability

Theorem

$\langle F_\alpha \mid \alpha < \beta \rangle$ is uniformly definable from F_β for any $\beta \leq \zeta$.

Similarly for the H-sets.

Interpretation: F_β contains within itself the whole history of the revision up to that stage β. When we talk about sets "internal to Field's model" we mean the sum total of those recursive in such stages (rather those up to the next repeat point Σ).

Field's Determinateness Operators

• Field introduces a 'determinateness operator' to show that the defectiveness of the liar sentence can be expressed:

$D(\ulcorner A \urcorner) =_{df} A \wedge \top \rightarrow A$

Then $D(L_0) \Leftrightarrow \neg T(L_0) \wedge \top \rightarrow L_0$
which is continually of semantic value 0.

• He then defines a strengthened liar: $L_1 \Leftrightarrow \neg D(T(L_1))$.
This flips values in a sequence $1, 1, 0, 0, 1, 1, 0, 0, \ldots$.

• But then: $DD(L_1)$ is continually 0.

• But this can be strengthened again: $L_2 \Leftrightarrow \neg DD(T(L_2))$.

Then by induction, define Strengthened Determinateness Operators:

$$\begin{array}{rcl} D^{k+1}(A) & \Leftrightarrow & D(D^k(A)) \\ D^{\omega}(A) & \Leftrightarrow & \forall n < \omega \forall y (y = D^n(\ulcorner A \urcorner) \rightarrow T(y)) \\ & \vdots & \end{array}$$

But also define Strengthened Liars:

$$\begin{array}{lcl} L_{k+1} & \Leftrightarrow & \neg D^{k+1}(T(L_{k+1})); \\ L_{\omega} & \Leftrightarrow & \neg D^{\omega}(T(L_{\omega})) \\ & \vdots & \end{array}$$

• He then discusses possibly extending these hierarchies along internally definable paths.

Definition

Let $\rho(A) \simeq \mu\alpha(\forall\beta > \alpha |A|_\alpha = |A|_\alpha = \|A\|)$.

Set $A \preceq B$ iff $\rho(A) \leq \rho(B)$.

Definition

Let $\rho(A) \simeq \mu\alpha(\forall\beta > \alpha |A|_\alpha = |A|_\alpha = \|A\|)$.

Set $A \preceq B$ iff $\rho(A) \leq \rho(B)$.

• There are formulae $P_{\preceq}, P_{\prec}$ in $\mathcal{L}_{\to,T}$ so that $A \preceq B$ iff $\|P(\ulcorner A \urcorner, \ulcorner B \urcorner)\| = 1$ (Internal Reflection). • For any $\alpha < \zeta$ there is a sentence A_α that

stabilizes precisely at α.

• The import of the latter is that we may use *sentences as notations* for ordinals $< \zeta$, but that there are no *stable notations* beyond that.

Definition

Define for *any* sentence $C \in \mathcal{L}_T$:

$$D^C(A) \equiv \forall B[P(\ulcorner A \urcorner, \ulcorner B \urcorner) \to (\forall y(y = \ulcorner D^B(A) \urcorner \to \mathrm{Tr}(y)))].$$

$$L_C :\equiv \neg D^C(T(L_C)).$$

Ineffable Liars

Theorem

There are sentences $C \in \mathcal{L}_T$ so that for any internally definable determinateness predicate D^B (i.e. *with* $\rho(B)\downarrow$) $\|D^B(Q_C)\| = \frac{1}{2}$.
Thus the defectiveness of L_C is not measured by any such determinateness predicate definable within the $\mathcal{L}_T$ language. They are "ineffable" liars.

• Thus such a L_C is a liar sentence that has diagonalized past all the internally definable determinateness predicates of the model.
• In Summary: within his model, (i) we have given a clear delineation of the notion of *internal path* and delimited their possible lengths; (ii) defined liar sentences which are immune to any notion of determinateness internal to the model.

What's wrong with the revision theory of truth?

What is wrong with the revision theory of truth?

What is wrong with the revision theory of truth?

Overheard complaints:

(C1): "The stable truths of the *RTT* have no good axiomatisation". Or perhaps, more specifically, the Herzberger stability set lacks this.

(C2): " No hope for a *KF*-like theory S such that

$$(\mathbb{N}, H) \models S \Leftrightarrow S \text{ is a stability set for some starting hypothesis}''$$

(C3): " The limit rule(s) of the RTT (or Herzberger revision theory) are poorly justified and have little to do with truth."

(C4): "The stable truths are ungrounded".

Revisioning revision theory

Ad (C1):
Question: To what extent does the revision theory give rise to attractive axiomatic truth theories formulated in the object language?
It is possible to make some remarks here.

1: (Cantini) $\mathfrak{M}_\zeta = \langle \mathbb{N}, H_\zeta \rangle \models VF$.

2 Observation:

Theorem (Leigh-W)

$$\mathfrak{M}_\alpha = \langle \mathbb{N}, H_\alpha \rangle \models VF \Leftrightarrow {}^{\omega}\omega \cap L_\alpha \models \Delta^1_3\text{-}CA_0$$

Question: Can anything extra be said truth-theoretically for the same levels of the Fieldian hierarchy? Or weaker levels, such as models of Δ^1_2*-*CA_0*?*

...continued

Question: Which H_α harbour possible interesting strong Kleene fixed points?

• For the same α with ${}^{\omega}\omega \cap L_\alpha \models \Delta^1_3\text{-}CA_0$ we have that H_α contains a *maximal strong Kleene* fixed point.

• For any sentence σ, if there is a maximal sKleene-fixedpoint I with $\sigma \in I \subseteq H_\alpha$ for some α then there is such an $I \subseteq H_\lambda$; and hence $\sigma \in H_\lambda$.

(Here $\lambda = \sup\{\tau \mid \tau$ *is the halting time of an ITTM* P_e *on integer input* $\}$.)

(C2) No *KF* like theory

(C2): *" No hope for a KF-like theory S such that*

$$(\mathbb{N}, H) \models S \Leftrightarrow S \textit{ is a stability set for some starting hypothesis}"$$

This hits home: there is no recursively axiomatisable first order theory T so that

$$(\mathbb{N}, H_\zeta) \models T$$

without there being unboundedly many $\alpha < \zeta$ with

$$(\mathbb{N}, H_\alpha) \models T.$$

What about: (C3): *" The limit rule(s) of the RTT (or Herzberger revision theory) are poorly justified and have little to do with truth."?*

Justifications for the Limit Rule(s):

Anti-justification for the Limit Rule(s):

[1] N. Belnap, *"Gupta's Rule of Revision Theory of Truth "*, JPL, **11**, No. 1, 1982, 103-116.

Anti-justification for the Limit Rule(s):

Belnap:

> *"The purpose of my suggestion is to make maximally arbitrary those features of the rule of revision that require arbitrariness at all. Only then will we be sure that what survives this arbitrariness really fully depends on the model with which we began with" (p106)*[1]

[1] N. Belnap, *"Gupta's Rule of Revision Theory of Truth "*, JPL, **11**, No. 1, 1982, 103-116.

Anti-justification for the Limit Rule(s):

Belnap:

> *"The purpose of my suggestion is to make maximally arbitrary those features of the rule of revision that require arbitrariness at all. Only then will we be sure that what survives this arbitrariness really fully depends on the model with which we began with" (p106)*[1]
>
> *" The more arbitrary, the more capricious, the more unreasonable, and the less patterned the permitted [limit rules] the more interesting is what remains invariant over all [them]." (p107)*

[1] N. Belnap, *"Gupta's Rule of Revision Theory of Truth "*, JPL, **11**, No. 1, 1982, 103-116.

C4 "The (Herzbergerian) Stable truths are ungrounded"

This could turn on what one means by groundedness.
• The arithmetically quasi-inductively defined predicates *AQI* form a *Spector class*.

• A *generalised quantifier* is any non-empty $Q \subsetneq \mathcal{P}(\mathbb{N})$ which is monotone, *i.e.* $X \in Q \wedge Y \supseteq X \rightarrow Y \in Q$. (Thus $\exists = \{X \subseteq \mathbb{N} \mid X \neq \emptyset\}$.)
We often write $QxR(x)$ for $R \in Q$.

• The *dual quantifier* $\breve{Q}$ is given by:

$$\breve{Q}xR(x) \Leftrightarrow \neg Qx\neg R(x)$$

. Then $\forall = \breve{\exists} = \{\mathbb{N}\}$.

• The relations inductive over $(\mathbb{N}, +, \times, 0, Succ)$ in *FOL* are precisely the Π^1_1-relations.

Theorem (Harrington)

For any Spector class Γ *there is a quantifier* Q_Γ *so that the relations inductive in* FOL_{Q_Γ} *are precisely those of* Γ.

• We thus may regard the H_ζ set as after all grounded by a monotone inductive definition, albeit not by one in *FOL* but in a logic enhanced by a non-standard quantifier.

• The relations inductive over $(\mathbb{N}, +, \times, 0, Succ)$ in *FOL* are precisely the Π^1_1-relations.

Theorem (Harrington)

For any Spector class Γ *there is a quantifier* Q_Γ *so that the relations inductive in* FOL_{Q_Γ} *are precisely those of* Γ.

• We thus may regard the H_ζ set as after all grounded by a monotone inductive definition, albeit not by one in *FOL* but in a logic enhanced by a non-standard quantifier.

$$
\begin{aligned}
&Q_H(X) \Leftrightarrow \\
&X \supseteq \{(A, B, 0) \mid B \in Field(\preceq) \wedge A \preceq B\} \cup \\
&\qquad\qquad \cup \{(A, B, 1) \mid B \in Field(\preceq) \wedge (A \notin Field(\preceq) \vee B \prec A)\}
\end{aligned}
$$

Game representations

A further remark:

Theorem (Aczel)

With $Q = Q_\Gamma$ as above, the relations $R(\vec{y})$ over $\mathbb{N}$ can be represented as:

$$R(\vec{y}) \Leftrightarrow Qx_0\breve{Q}x_1Qx_2\breve{Q}x_3Qx_4\ldots\exists nR_0(\langle x_0,\ldots,x_n\rangle,\vec{y}).$$

• We thus have an *open game-theoretic* representation of relations in this class. In particular the set of Herzbergerian stable truths can be so represented.

Theorem

$\ulcorner\sigma\urcorner \in H_\zeta \Leftrightarrow$

Player I has a winning strategy in an 'open' game of the form above in FOL_Q.

Closing remarks

- We thus may regard more complex "truth" sets as arising by processes familiar to arithmetic and Π^1_1 definability, open game theoretic semantics *etc.*, but now in a *stronger* logic.

- This begs the question about studying truth theories over, say arithmetic, in logics stronger than first order classical logic rather than weaker.