On Field's *"Saving Truth from Paradox"*

P.D.Welch

(I) In current semantic theories
(a) there is a lack of a viable conditional $\supset$
(b) there may be occurrences of failure of the T-scheme:
(that for some sentence(s) A we do not have '$T(\ulcorner A \urcorner) \leftrightarrow A$')
(c) there may be failure of *the intersubstitutivity* of $T(\ulcorner A \urcorner)$ for A where the latter is a subformulae of some B.
So Field:
• (II) introduces a *binary operator* $\longrightarrow$ to function as a form of generalised conditional;
• (III) provides for set models (one or more) semantics that remedies (a)-(c) in a non-classical logic;
• (IV) provides an analysis of the 'defectiveness' of *e.g.* the liar sentence through a hierarchy of *determinateness relations*.

$\mathcal{M}^+$ Expansions: *G-solutions*

- $\mathcal{M}$ an $\mathcal{L} = \mathcal{L}_{\mathcal{M}}$-structure : expanded to $\mathcal{M}^+$ in $\mathcal{L}^+ = \mathcal{L}^+_M$ (containing additionally T, $\longrightarrow$)
- $\mathcal{L}^+$ evaluated in a 3-valued $\{0, \frac{1}{2}, 1\}$ logic, (or in a De Morgan function algebra V_c).

• He takes issue with the idea that we can *define* 'real truth' by using notions of designated semantic values obtained by a variety of methods (Kripkean minimal fixed points, strong Kleene, supervaluational ... or *via* revision theory. In particular Tarski like constructions using *set-sized* models for which we can give mathematical or inductive definitions of 'designated truth value' cannot deliver for us a theory of 'real truth.'

• But in order to give some description of the theory he is aiming for, he has given a number of G-solutions or G-models, in particular a 'principal' one deriving the consistency of the naive theory of truth (The T-scheme), the Intersubstitutivity Principle, and with $\longrightarrow$.

• We thus have G-models $\mathcal{M}^+$ over *e.g.* ground models $\mathcal{M}$:

$$\langle \mathbb{N}, +, \times, 0, S \rangle, \quad \langle V_\alpha, \in \rangle, \ldots$$

Real Validity

> *"it might be better to adopt the view that what is validated by a given version of the formal semantics [i.e. a G-model] outruns "real validity": that the genuine logical validities are some* effectively generable subset of those inferences *that preserve value 1"*

(emphasis now mine). He continues:

> *"... there would doubtless be some arbitrariness in which effectively generable subset to choose, but that is perfectly acceptable unless one wants to put high (and I think unreasonable) demands on the significance of the distinction between those inferences that are valid and those that are not."*

The complexity of the principal model: a concern

• The principal model over $\mathbb{N}$ is supposed to deliver a *first order* theory of truth with $\longrightarrow$;

• however it requires (as a piece of applied mathematics) a stronger subsystem of second order number theory (Π^1_3-CA) than any other piece of 'ordinary' mathematics

• Well beyond the reach of any current proof-theoretical ordinal analysis

• Can there be a simpler 'consistency proof'? A simpler 'principal model'?

Internal Structure of the Principal Model

- Recall that we have a determinateness operator $D(A) \equiv A \wedge \neg(A \rightarrow \neg A)$
- $D^{n+1}(A) \equiv D(D^n A))$; $D^{\omega}(A) \equiv \forall n \forall y(y = \ulcorner D^n(A) \urcorner \rightarrow T(y))$.
- Comes with 'determinate liars': $Q^{\alpha} \leftrightarrow \neg D^{\alpha}(\ulcorner Q^{\alpha} \urcorner)$.
- How far can these hierarchies go?

Taking $\mathcal{M}$ as $\mathbb{N}$

- To go beyond recursive ordinals let *sentences of* $\mathcal{L}^+$ *themselves* stand in for ordinal notations:

Definition

$\rho(A) \simeq$ least ρ such that semantic value of ρ is constant from ρ onwards.

We abbreviate $A \prec B$ for $P_\prec(\ulcorner A \urcorner, \ulcorner B \urcorner) = 1$ *etc.*

• If $\|A\| = 1$ (or 0) say, then $\{B : B \prec A\} = \{B : \|P_\prec(\ulcorner A \urcorner, \ulcorner B \urcorner)\| = 1\}$ is a prewellordering of order type some ordinal $\xi < \Delta_0$.

• We let Field$(\prec)$ denote the set of sentences stabilizing on 0 or 1. Moreover:

Lemma

There is a formula $P_\prec(v_0, v_1)$ in $\mathcal{L}^+$ so that for any sentences $A, B \in \mathcal{L}^+$, we have $\|P_\prec(\ulcorner A \urcorner, \ulcorner B \urcorner)\| = 1$ iff $\rho(A)\downarrow, \rho(B)\downarrow$ and $\rho(A) < \rho(B)$;

$$= 0 \text{ iff } \rho(A)\downarrow, \rho(B)\downarrow \text{ and } \rho(A) \geq \rho(B);$$
$$= \tfrac{1}{2} \text{ otherwise.}$$

Lemma

For any $\xi < \Delta_0$ there is a sentence $A = A_\xi$ in Field$(\prec)$ with the order type of $\{B | B \prec A\}$ equalling ξ.

• We may define for *any* sentence C

$$D^C(A) \equiv \forall B \prec C \forall y(y = \ulcorner D^B(A) \urcorner \to T(y)).$$

• For $C \in \text{Field}(\prec)$ this defines a bivalent determinateness hierarchy of length $\rho(C)$.

• However it is not a bivalent matter as to whether a general C is or is not in $\text{Field}(\prec)$. (In other words $\text{Field}(\prec)$ is not a crisp subclass of $\mathbb{N}$.) However if $C \in \text{Field}(\prec)$ then it can be shown that it is a bivalent matter whether a general B is $\prec$-below C or not.

• Consequently the expression

$$\text{“}\langle D^B(v_0) | B \prec C \quad \textit{forms a determinateness hierarchy} \rangle\text{”}$$

is not in the classical part of the language $\mathcal{L}^+$ to which the Law of Excluded Middle holds.

• Thus the internally defined determinateness hierarchy over $\mathbb{N}$ breaks down, not fuzzily, but precisely, at Δ_0. There is no internally definable maximal hierarchy.

Axiomatising $F =_{\mathrm{df}} \{A : \|A\| = 1\}$

• (Martin) We have an *open game representation* in $\mathcal{L}$ of the least Strong Kleene fixed point over $\mathbb{N}$ as an *open game*.

• A game for F can be formulated but is an $\exists\forall\exists$ game.

• One can have an open game representation of F over $\mathbb{N}$ in a language with a generalised quantifier $\mathcal{L}^{+}(Q)$. where

• ' $Qx\varphi x$ ' iff for 'path-many' x $\varphi(x)$; that is

$$Qx\varphi x \Leftrightarrow \exists A \in \mathrm{Field}(\prec)(\{n \in \mathbb{N} | \varphi(n)\} \supseteq \{B \mid B \preceq A\}).$$

• Speculatively this suggests a possible axiomatisation of a theory of truth (with a $\longrightarrow$) *together with* determinateness satisfying the laws or properties Field has already given for T and D.