Ultimate truth *vis à vis* stable truth

P.D. Welch
School of Mathematics, University of Bristol, England.

September 14th, 2006

Abstract. We show that the set of *ultimately true* sentences in Hartry Field's Revenge-immune solution model to the semantic paradoxes is recursively isomorphic to the set of *stably true* sentences obtained in Hans Herzberger's revision sequence starting from the null hypothesis. We further remark that this shows that a substantial subsystem of second order number theory is needed to establish the semantic values of sentences in Field's relative consistency proof of his theory over the ground model of the standard natural numbers: Δ^1_3-CA_0 (second order number theory with a $\mathbf{\Delta^1_3}$-Comprehension Axiom scheme) is insufficient.

We briefly consider his claim to have produced a "revenge-immune" solution to the semantic paradoxes by introducing this conditional. We remark that the notion of a "determinately true" operator can be introduced in other settings.

Keywords: theories of truth, fixed points, revision sequence, constructible sets, admissibility.

1. Introduction.

In (Field, 2003) Field constructs a system of a theory of truth over ground models $\mathcal{M}$, with an additional conditional operator $\rightarrow$ in the language. We shall here use for illustrative purposes $\mathcal{M} = \mathbb{N}$, the standard model of arithmetic, with $\mathcal{L}$ the appropriate arithmetical language using the connectives $\wedge, \vee, \neg$ and the two quantifiers $\exists, \forall$; $\mathcal{L}^+$ will be, just as for him, this language augmented with the $\rightarrow$ connective, to be considered as an operator, and Tr as a predicate.)

The construction in (Field, 2003) Section 2 is in ordinal stages, where each successor ordinal stage performs an entire construction of the "next strong Kleene fixed point" *à la* Kripke. Thus the transition of $\alpha \rightarrowtail \alpha + 1$ involves the usual Kripkean monotonic construction over an assignment of values amongst $\{0, 1, \frac{1}{2}\}$ to the "conditionals" of the form $|A \rightarrow B|_\alpha$. (We are suppressing the secondary ordinal subscript of this latter intermediate construction, which Field calls the "mini-stages" - and also for the most part suppresses.) The latter assignments are determined by semantic values given at strictly previous (full) stages. We repeat his Clause 8 to illustrate:

$$|A \rightarrow B|_\alpha = \begin{cases} 1 & : \text{ iff } (\exists\beta < \alpha)(\forall\gamma \in [\beta, \alpha)(|A|_\gamma \leq |B|_\gamma), \\ 0 & : \text{ iff } (\exists\beta < \alpha)(\forall\gamma \in [\beta, \alpha)(|A|_\gamma > |B|_\gamma), \\ \frac{1}{2} & : \text{ otherwise.} \end{cases}$$

Although the motivations are entirely different (and this cannot be emphasised enough perhaps) the system here has striking *structural* similarities with the revision theory of truth using Herzberger style revision sequences (*cf.*(Herzberger, 1982b),(Herzberger, 1982a)). For a revision sequence, knowing all truth values at some stage $\alpha+1$, we can read off which truth value was assigned to B at stage α from Tr(B). In Field's system we have the same effect: the "1 or non-1" semantic value of B at the previous stage α is imported for the next stage $\alpha+1$ *via* $|\top \to B|_{\alpha+1}$: knowing these semantic values means we know what happened before. There are similar, but differing, considerations for limit stages. For Herzberger, at limit stages the extension of Tr is determined through similar looking rules to the first two clauses in the above definition, however there is no value of $\frac{1}{2}$, but in this third case an (arbitrary?) assignment of 0. In a single Herzberger sequence one can look at the set of *stable sentences*, that is, those that from some point on have a set truth-value. (We emphasise that we are *not* considering *categorical truth sets* in the style of Gupta and Belnap (Gupta and Belnap, 1993), where a global averaging is obtained by considering all possible starting hypotheses for the Herzberger revision process.)

The values $|A \to B|_\alpha$ are continuous at limits (*cf.* his Continuity Lemma for Conditionals). It then makes sense to define $||A||$ as the ultimate value of $|A|_\alpha$:

$$||A|| = \begin{cases} 1 & : \text{ iff } (\exists\beta)(\forall\gamma \geq \beta(|A|_\gamma = 1), \\ 0 & : \text{ iff } (\exists\beta)(\forall\gamma \geq \beta(|A|_\gamma = 0), \\ \frac{1}{2} & : \text{ otherwise} \end{cases}$$

In Section 4 we discuss his claim that his system, having successfully introduced a conditional operator, gives rise to a "revenge-immune" solution to the semantic paradoxes. We do not believe that the proposed system is any more immune to strengthened liar paradoxes than some already in existence.

Field asks if there are *acceptable points*, that is ordinals Δ so that for any A we have that $||A|| = |A|_\Delta$. Section 3 of (Field, 2003) is devoted to proving that there are such. For the paper it is sufficient for him to demonstrate the existence of such points but he adds: "I believe a more informative proof should be possible, which would show among other things that acceptable fixed points ... occur before Ω." Here Ω is the next initial ordinal of cardinality greater than that of the underlying model $\mathcal{M}$ that is being discussed - so here for us this is ω_1. In fact his proof is simple and direct, but again adds that "the price of its simplicity is that the proof is less informative than one might like about the way the values of sentences change as the level increases towards an acceptable point." The purpose of Section 2 of this note is to give a further proof (well, in fact 3 of them) for the existence of acceptable points (which are all below Ω). Probably none of them are "informative" in the latter sense he might like, and we are not sure that there could be anything very illuminating *in general*, meaning for the "general" sentence A, given the complexity of the notions involved. He shows that the least acceptable point is greater than a particular recursive ordinal λ_0 that is used in his iterations of a "determinately true" operator D. Theorem 2 below shows that the process of calculating semantic values is *necessarily* complicated, and thus "how" a general sentence achieves its ultimate value involves looking at the theory of quasi-inductive definitions, or equivalently at the levels of the Gödel constructible hierarchy up to L_ζ - the first Σ_2-extendible level. We show (Theorem 3) that the least acceptable point is in fact exactly this ζ, and so far beyond the first non-recursive ordinal ω_{1ck}. Remark (ii) below also indicates the strength of the proposed system: any system of analysis in which we can establish these eventual semantic values is highly impredicative.

DEFINITION 1.1. *(i) The set of* ultimate truths $D =_{df} \{\ulcorner A \urcorner : \|A\| = 1\}$.
(ii) The set of (Herzberger) stable truths
$H =_{df} \{\ulcorner A \urcorner :$ *A is stably true in Herzberger's revision sequence starting from the null hypothesis*$\}$.

In the above by "the null hypothesis" we mean the hypothesis that all sentences are false, (or equivalently for our purposes this could be taken as "all sentences are true": any recursive starting distribution of truth values would serve equally well). For the notion and properties of Herzberger's revision sequence the reader may consult that author's papers already cited, or the account in (McGee, 1991).

THEOREM 1.1. *D is recursively isomorphic to the set H.*

This set H has been noted to be equivalent to other sets of integers, defined independently. We shall comment on this further in Section 3, but briefly, if Q is the complete arithmetic quasi-inductive set, if $\widetilde{O}$ is the Σ_2-truth set of L_ζ (where ζ was named above), and if S is the set of (codes of) Turing programs that have eventually some constant value on their output tape, when allowed to run transfinitely (in the formalism of (Hamkins and Lewis, 2000)) then we have:

Fact 1 *The following are recursively isomorphic:*
(i) S ; (ii) $\widetilde{O}$; (iii) Q ; (iv) H.

Our third method of demonstrating the existence of acceptable points, is in fact directed to adding a fifth set D to this list. The following theorem is what we shall actually prove:

THEOREM 1.2. *Let D be the set of ultimate truths in Field's system. Then D is recursively isomorphic to $\widetilde{O}$.*

The statement of the abstract, Theorem 1, then follows from the Fact quoted above. We note the following further consequences:

Remark (i) Using the above Fact 1 and Theorem 1.2 the set of ultimately true sentences over arithmetic is a *complete quasi-inductive set,*; hence Field's logic $\models_{LCC}$ is not axiomatisable.

Remark (ii) The "conservativeness proof" of (Field, 2003) Sect.6 (there performed in ZFC) can be effected in a sufficiently strong fragment of second order number theory, but not in KP (Kripke-Platek set theory) nor $\mathbf{\Delta^1_3}$-CA_0. (For the latter theory see (Simpson, 1999);[1] for notions concerning KP and admissible ordinals see (Barwise, 1975).)

Similar results will hold for countable ground models $\mathcal{M}$ that are *acceptable* (this time in the sense of (Moschovakis, 1974)), that is, the least (Fieldian) acceptable point for $\mathcal{M}$ will be the least Σ_2-extendible ordinal relative to $\mathcal{M}$, and hence will be less than $|\mathcal{M}|^+$, the least cardinal greater than that of $\mathcal{M}$. We shall then have (in the appropriate language, and in the appropriate sense) that $D_{\mathcal{M}}$ is equivalent to $H_{\mathcal{M}}$. The proof of Theorem 1.2 (in the hard direction showing that $\widetilde{O}$ is (1-1) reducible to D) is essentially a demonstration that Field's system carries hidden within it a construction of the Gödel hierarchy of constructible sets up to the level ζ. We can show that at stage α in his construction we have a uniform way of recovering certain Σ_2-truth sets of the levels L_α of this hierarchy, whenever the latter is a model of a sufficient amount of set theory. For other levels we seemingly do not have a uniform way of doing this: although Field's clauses are uniform for all limit ordinals, the recovery of the truth sets for the L_α can be obtained from $\{A : |A|_\alpha = 1\}$, but perhaps not always *uniformly* so. [2] Happily below ζ there are many α where L_α models this theory, including ζ itself; this allows us then to exploit the uniformity at ζ, in order to find the (1-1) function decoding $\widetilde{O}$ from D.

We make a brief remark on our result concerning *recursive isomorphism* (essentially a "computable translation" - this will be explicitly defined below). One might hope for some more perspicaceous, or 'natural' translation of Herzberger stable truths directly into Field's 'ultimate truths', and hopefully *vice versa*. We think this unlikely. We do not really see one in either direction: this is to do with the fact that the process of revision according to revision sequences, has more in common with the supervaluation approach than the Kleene strong jump that Field employs in his 'mini-stages': direct translations seem to be ruled out by the incomparability of the methods. One should also remember however, that *all we ever do* with theories of truth employing a system of gödel numbering over a model such as the standard model of arithmetic, is to build a truth set using that particular coding scheme. There are infinitely many choices of (sensible) coding scheme, all of which only yield identical truth sets *up to recursive isomorphism*. However, even bearing this in mind, there seems a difficulty in finding a meaning preserving translation or mapping, between the stable truths and the ultimate truths. [3]

In the penultimate section we discuss a little the nature of the "revenge immunity" claim, and the import of the methods used in Field's consistency proof. We claim that our argument shows that an extraordinary amount of analysis - in the sense of a large fragment of second order number theory - is needed for Field's construction to work - well beyond anything that mathematicians working in analysis, (*i.e.* the structure of the real continuum rather than simply the natural numbers). In the final section we just remark that "determinately true" operators could have been introduced in a Herzbergerian revision theoretic setting.

2. The existence of acceptable points

First demonstration: The construction of (Field, 2003) Section 2 is performed within ZFC using the ground model $\mathcal{M}$. It is in fact a very absolute construction, and absolute for transitive models of ZFC^-, that is ZFC without the power set axiom (which is not used). Hence if H is a transitive set, $\alpha, \mathcal{M} \in H$ and $H \models ZFC^-$ then $H \models$ "$|A|_\alpha = j$" iff $|A|_\alpha = j$. Suppose, without loss of generality, that $\mathcal{M}$ is countable. By the Collection Axiom we may assume that there is some ordinal μ_0 so that for all $A \in \mathcal{L}^+$, if $|A|_\gamma = j$, for all sufficiently large γ, then there is an $\delta_A < \mu_0$ so that $\forall\beta > \delta_A |A|_\beta = |A|_{\delta_A} = j$. For the other A, *ie* those with $\|A\| = \frac{1}{2}$, which alternate value unboundedly in the ordinals, we can again appeal to Collection to find, for any ordinal $\nu \geq \mu_0,$ some cardinal $\mu \geq \nu$ so that if, say, A alternates its value unboundedly in the ordinals, it does so unboundedly below μ in particular. (This already shows that the class of acceptable points is unbounded in the class of all ordinals.) But further, if H_μ is the set of sets of hereditary cardinality less than μ, then H_μ is a ZFC^- model. Now by the Löwenheim-Skolem theorem let $X \prec H$ be a countable elementary submodel of H with $\mathcal{M} \in X$. Let N be a transitive set isomorphic to X. (Such an N exists by the Mostowski collapsing Lemma.) Let $\Delta < \omega_1$ be the ordinal height of N. Then Δ is an acceptable point: one may easily verify that $|A|_\Delta = |A|_\mu$ for any A.

By modifying the argument one may show that there is in fact a closed and unbounded set $E \subseteq \omega_1$ of acceptable ordinals.

Second demonstration: One could directly show that the construction, and argument that there are acceptable points, can be performed not just in ZFC^- but in much weaker theories, for example, Kripke-Platek (KP) augmented with a Σ_2-Separation axiom, or even in second order number theory with Π^1_3 Comprehension. (We aren't advocating actually *doing* this; that it could be done will follow from the third demonstration). If this is granted, then we can find countable models of such theories for which the construction is absolute. That it *cannot* be done in much weaker theories is also a concomitant of the theorem.

Third demonstration: This contains the heart of this note. We shall define an ordinal ζ. To see what this is we make the following observations. The Σ_2 nature of the defining clauses of $\|A\|$, together with the absolute nature of the construction, hint not only (a) that the definition of the function $\|\cdot\|$ can be construed as a valuation coming from a quasi-inductive definition but (b) that the first repeat point, that is the first acceptable point, is ζ - this latter ordinal has been characterised as the first Σ_2-extendible ordinal ((Burgess, 1986) Sect.14), that is the first ordinal ζ so that L_ζ has a proper Σ_2-elementary end extension.[4] Equivalently it is also the first place where all Infinite Time Turing machine (ITTM) computations have either halted or have entered an infinite loop ((Welch, 2000) Thm. 2.1). In the next theorem Σ is the periodicity, or ordinal length, of that loop, or, equivalently, the ordinal height of that smallest Σ_2 end-extension of L_ζ. We shall prove:

THEOREM 2.1. *Over the ground model $\mathcal{M} = \mathbb{N}$ of arithmetic, the first acceptable point $\Delta_0 = \zeta$. The class of acceptable points above ζ is: $\{\Sigma.\rho \mid \rho \in On\}$.*

To state explicitly at the outset, we shall reserve the word *recursive* to have its usual standard meaning throughout this note. The notion of $P \leq_T Q$ "P is *recursive in* Q", for sets of integers P, Q, says simply that χ_P, the characteristic function of P, is recursive in an oracle for Q; alternatively put, there is an index $e \in \omega$ so that χ_P is the e'th function recursive in Q: $\chi_P = \{e\}^Q$. To say that two sets P, Q are *recursively isomorphic* (written "$P \equiv Q$") is to say that there is a total recursive *bijection* $F : \mathbb{N} \longleftrightarrow \mathbb{N}$ with $\forall n\; n \in P \iff F(n) \in Q$. For such P, Q membership questions about P can be converted by a pencil and paper algorithm to membership questions about Q and conversely. For most mathematical purposes then, these sets are seen to be isomorphic if not identical. A weaker notion, written $P \leq_1 Q$, is that "P is *(1-1) reducible to* Q", that is, there is a recursive *injection* $F : \mathbb{N} \longrightarrow \mathbb{N}$ with $\forall n\; n \in P \iff F(n) \in Q$. Here we could only paraphrase by saying that membership questions ?$n \in P$? can be converted into the questions ?$F(n) \in Q$? again *via* such an algorithm. Hence in this situation, as a slogan: "Q is at least as complicated as P". An effective version of the proof of the classical Cantor-Schröder-Bernstein theorem due to Myhill ((Rogers, 1967) Theorem (VI)) shows that $P \equiv Q$ iff $P \leq_1 Q$ & $Q \leq_1 P$.

We introduce some further definitions.

DEFINITION 2.1. *For $x \in L$, we let $\rho_L(x) =_{df}$ the least α such that $x \in L_{\alpha+1}$.*

For E a class of ordinals, let E^* denote E together with the set of its limit points. E^* is thus the *closure* of E.

DEFINITION 2.2. *(i) Let* $ADM = \{\alpha \mid \langle L_\alpha, \in\rangle \models KP\}$.
(ii) Let $\langle \tau_\iota \mid \iota < \omega_1\rangle$ *enumerate* $ADM^* \cap \omega_1$.

Here "KP" denotes Kripke-Platek set theory in the language $\mathcal{L}_{\dot\in}$. If $\alpha \in ADM$ it is called *admissible*. Then $\tau_0 = \omega, \tau_1 = \omega_{1ck}$, where the latter is the first non-recursive ordinal. Note that ADM is not closed: τ_ω is not admissible.

We shall make use of the following facts:

LEMMA 2.1. *(i)* $\zeta = \tau_\zeta;\ \Sigma = \tau_\Sigma$.
(ii) $\forall \mu < \Sigma\, \neg\exists \nu < \mu \quad L_\nu \prec_{\Sigma_2} L_\mu$,

((i) that $\tau_\xi = \xi$ is true for any Σ_2-extendible ordinal ξ, as for any such ξ L_ξ can be shown to be a model of Σ_2-KP which is moreover a union of such; ξ is thus a limit of "Σ_2-admissible ordinals" even; this can be deduced from the reflection properties enjoyed by Σ_2-extendibles. (ii) uses, or rather is just a restatement of, the leastness assumption on ζ together with that of Σ being the height of the least Σ_2 end-extension of L_ζ.

DEFINITION 2.3. *Let* $C \subseteq \omega$. *We let* $A^{1,C} =_{df}$ *the complete* Σ_1*-theory of* $\langle L_{\omega^C_{1ck}}[C], \in, C\rangle$.

$A^{1,C}$ may thus be identified with a set of integers coding Σ_1 sentences in the language $\mathcal{L}_{\dot\in,\dot C}$. This set has itself a Σ_1-definition over the structure $\langle L_{\omega^C_{1ck}}[C], \in, C\rangle$, which is uniform in C (that is, the defining formula is independent of the choice of set C). Here ω^C_{1ck} is the first ordinal not recursive in C (equivalently: not arithmetic in C) and is the first ordinal which is admissible in the predicate C. (We shall henceforth drop the "ck" when considering the first ordinal not recursive in some C; we shall keep it for the first non-recursive ordinal ω_{1ck}; the first uncountable cardinal is thus $\omega_{1,}$ to avoid any confusion.)

DEFINITION 2.4. *For* $\iota < \omega_1$ *let* $A^2_\iota = A_\iota =_{df}$ *the complete* Σ_2*-theory of* $\langle L_{\tau_\iota}, \in\rangle$.

It is thus $\widetilde{O} =_{df} A_\zeta$ which we wish to show is recursively isomorphic to D.

PROPOSITION 2.2. *(i)* $\forall \iota < \Sigma\ \rho_L(A_\iota) = \tau_\iota$;
(ii) $A_\Sigma = A_\zeta$ *and hence* $\rho_L(A_\Sigma) = \zeta = \tau_\zeta$;
(iii) $\forall \iota < \Sigma\ \omega_1^{A_\iota} = \tau_{\iota+1}$.

Proof: (i): Assume $\iota < \Sigma$ and let $\tau = \tau_\iota$; as $\tau < \Sigma$ for no ξ do we have $L_\xi \prec_{\Sigma_2} L_\tau$. We say that a set $x \in L_\tau$ is (*implicitly*) Σ_2*-definable over* L_τ if there is a Σ_2 formula $\varphi(v_0)$ so that in L_τ x is the unique set so that $\varphi[x]_{L_\tau}$. We let $\lambda =_{df} \sup\{\gamma \mid \gamma$ is Σ_2-definable over $L_\tau\}$. Then $\lambda \leq \tau$. We shall claim that $\lambda = \tau$. But first we note

(1) Any $\gamma < \lambda$ is Σ_2-definable over L_τ.

For, if $\delta < \lambda$ is Σ_2-definable over L_τ, then note that any $\gamma < \delta$ is actually Σ_1-definable from δ: this is because for any $\tau \leq \Sigma$ we have that $L_\tau \models$"$\forall\alpha\exists f : \omega \longrightarrow \alpha$ with f surjective". (This latter assertion holds, since if $L_\tau \models$ "η is uncountable" we should be able to prove that there are even $\mu < \nu < \eta$ with $L_\mu \prec L_\nu$ even which would contradict our leastness assumption on Σ.) Hence, given a $\gamma < \delta$, the global wellordering of L is Σ_1-definable and so the L-least such surjective map f onto δ is Σ_1-definable in L_λ from δ. Thus if $f(n) = \gamma$ then in turn γ is Σ_1-definable from δ. Putting this together with the Σ_2-definition of δ yields a Σ_2-definition of γ. QED(1)

Let A_λ be the complete Σ_2-theory of $\langle L_\lambda, \in\rangle$.

(2) $A_\iota \subseteq A_\lambda$

Suppose $\sigma \in A_\iota$. As σ is of the form $\exists u\forall v\psi(u, v)$ for some ψ with all quantifiers bounded, then for some γ we have that

$$L_\gamma \models \text{“}\sigma \wedge \forall\eta\exists\eta' > \eta(L_{\eta'} \models \neg\sigma)\text{” \& } \forall\gamma' \geq \gamma L_{\gamma'} \models \sigma.$$

(γ is thus the least point on from which σ is seen to be true.) But now the above yields "$g = \gamma$" as Π_1 over L_τ. Hence γ is Σ_2-definable over L_τ and thus $\gamma < \lambda$. We conclude by the form of the above statement, since $\lambda > \gamma$, that $\sigma \in A_\lambda$.

QED(2)

(3) $L_\lambda \prec_{\Sigma_1} L_\tau$.

This is trivial if $\lambda = \tau$ so assume otherwise. If $\xi < \lambda$ and $\varphi(\xi)$ is a Σ_1 assertion about ξ that is true in L_τ but *not* in L_λ then there is first ordinal $\gamma < \tau$ which sees a witness for the existential quantifier of φ: thus $L_\gamma \models \varphi(\xi)$. We thus have (where ψ is the Σ_2 defining formula for ξ):

"$\exists x(\psi(x) \wedge g$ is the least ordinal so that $L_g \models \varphi(x)$".

This is a Σ_2-definition over L_τ of γ. Hence $\gamma < \lambda$. Hence (3) holds by the upwards persistence of Σ_1 formulae to any larger model. QED(3)

But (3) immediately implies (by the upwards persistence of Σ_2 formulae between Σ_1 elementary substructures):

(4) $A_\iota \supseteq A_\lambda$.

Hence with (2) we have equality here. But $\lambda < \tau$ will imply that $L_\lambda \prec_{\Sigma_2} L_\tau$ which contradicts the leastness of (ζ, Σ). To see this last point, let $\varphi(\xi) \equiv \exists u\forall v\psi(u, v, \xi)$ be a Σ_2 assertion about $\xi < \lambda$ which holds in L_τ. As $\xi < \lambda$ it has a uniquely defining Σ_2 formula $\theta(v_0)$ which is only true in L_τ of ξ itself. Hence L_τ is a model of the following:

"$\exists x \exists u[\theta(x) \wedge \forall v \psi(u, v, x)$"

This is a Σ_2 sentence and so is in $A_\iota = A_\lambda$. It thus holds in L_λ. But note that $\theta(\xi)$ holds in L_λ uniquely also: (a) if $L_\lambda \models \exists \xi_0 \exists \xi_1(\xi_0 \neq \xi_1 \wedge \theta(\xi_0) \wedge \theta(\xi_1))$ then this is a sentence in A_λ which fails to hold in A_ι - a contradiction; so (b) suppose $L_\lambda \models \theta(\xi_0) \wedge \xi_0 \neq \xi$. By (3) we should again have by the upwards persistence of the Σ_2 formula $\theta(\xi_0)$ that $L_\tau \models \theta(\xi_0)$ also a contradiction.

Hence $\lambda = \tau = \tau_\iota$ and $\rho(A_\iota) \geq \tau$. QED(i)

For (ii): here Σ is the very least point where there is some $\zeta < \Sigma$ with $L_\zeta \prec_{\Sigma_2} L_\Sigma$. Further these two levels of the L-hierarchy have identical Σ_2-theories, *i.e.*, $A_\Sigma = A_\zeta$. Hence $\rho_L(A_\Sigma) = \tau_\zeta$. That $\tau_\zeta = \zeta$ is Lemma 2.1 (i) above.

For (iii): again let $\tau = \tau_\iota$; we saw above in (i) that every $x \in L_\tau$ has a Σ_2 definition. A_ι itself is thus a complete Σ_2-truth set for L_τ (*i.e.* it contains a record of all elementary $\in$ and $=$ facts about L_τ) and arithmetically in A_ι we can decode a wellordering $w \subseteq \omega \times \omega$ of order type τ. Hence $\omega_1^{A_\iota} > \tau_\iota$. However for any $B \subseteq \mathbb{N}, B \in L_{\tau_{\iota+1}}$, we have that $\omega_1^B \leq \tau_{\iota+1}$. In particular $\omega_1^{A_\iota} \leq \tau_{\iota+1}$. QED

The Σ_1-Separation scheme (see, e.g. (Barwise, 1975) Sect. I.9) is the infinite set of axioms which are the universal closure of axioms of the form:

$\exists u \forall v(v \in u \longleftrightarrow v \in w \wedge \varphi(v))$

where $\varphi(v)$ is any Σ_1-formula (possibly with other set variables). For us, what matters is the following relatively easily proven fact (again see (Barwise, 1975), V 6.3 & 7.12):

PROPOSITION 2.3. *Suppose $\beta > \omega$. $L_\beta \models \Sigma_1$-Separation if and only if $\{\gamma < \beta \mid L_\gamma \prec_{\Sigma_1} L_\beta\}$ is unbounded in β. If $L_\beta \models \Sigma_1$-Separation then $\beta \in ADM \cap ADM^*$.*

PROPOSITION 2.4. (cf.(Welch, 2000) Thm 2.1) $L_\zeta \models \Sigma_1$-*Separation.* [5]

We shall use the following notation. Let $C_\alpha = (C_{\alpha,T}, C_{\alpha,F})$ where $C_{0,T} = C_{0,F} = \varnothing$. Then for $\alpha > 0$ set:

$C_{\alpha,T} = \{\top \longrightarrow A : |\top \longrightarrow A|_\alpha = 1\}$;

$C_{\alpha,F} = \{A \longrightarrow \bot : |A \longrightarrow \bot|_\alpha = 1\}$.

In the above we have omitted gödel corners, but we assume that $C_\alpha \subseteq \mathbb{N}$, after some suitable recursive coding, and pairing, of $C_{\alpha,T}$ with $C_{\alpha,F}$. Note that if we set, for $\alpha \geq 0$:

$C^+_\alpha = (C^+_{\alpha,T}, C^+_{\alpha,F})$ where we have taken:

$C^+_{\alpha,T} = \{A : |A|_\alpha = 1\}$;
$C^+_{\alpha,F} = \{A : |A|_\alpha = 0\}$,

then, for example, $C_{\alpha+1,T}$ is little different from $C^+_{\alpha,T}$, as the former is simply the set of members of the latter with "$\top \longrightarrow$" inserted before them (and similarly for $C_{\alpha+1,F}$ and $C^+_{\alpha,F}$ *m.m.*). Note that if ξ is acceptable in Field's sense, then $C_{\xi,T} = \{\top \longrightarrow A \mid A \in D\}$ and $C_{\xi,F} = \{A \longrightarrow \bot \mid \neg A \in D\}$. Hence C_ξ is (1-1) reducible to D.

LEMMA 2.2. *For $\iota < \Sigma$ (i) A_ι is arithmetic in C_ι; (ii) $\omega_1^{C_\iota} = \tau_{\iota+1}$; (iii) $\rho_L(C_\iota) = \tau_\iota (\iota > 0)$; (iv) if $L_{\tau_\iota} \models \Sigma_1$-Separation, then $A_\iota \leq_1 C_\iota$; moreover this latter clause is uniform in ι for such models of Σ_1-Separation.*

By "uniform in ι" in (iv) above, we mean that the (1-1) recursive function $f{:}\mathbb{N} \to \mathbb{N}$ that effects the reduction of A_ι to C_ι is the same for each such $\iota < \Sigma$ where $L_{\tau_\iota} \models \Sigma_1$-Separation. Taking $\iota = \zeta$ shows that $\widetilde{O} = A_\zeta \leq_1 C_\zeta \leq_1 D$. By Proposition 2.3 the hypothesis of (iv) is false if ι is a successor ordinal (and is also for many limit ordinals).

Proof of Lemma 2.2: We have mentioned the C^+_α versions decorated with the plus sign, as this allows the following formulation.

PROPOSITION 2.5. *(Kripke) C^+_α is a complete Π^{1,C_α}_1 set of integers, (and hence by the above remark, so is $C_{\alpha+1}$).*

This was proven by Kripke in the case when $\alpha = 0$, for his original theory using strong Kleene. (A proof is in (Burgess, 1986).) By Field's definition, each C^+_α is constructed as the Kripkean fixed point using this scheme, with the distribution of semantic values $0, 1/2, 1$ assigned to the various sentences involving the $\longrightarrow$ operator. The proposition above then, is simply the relativisation of Kripke's result to the "starting distribution" coded into C_α. Once we have C^+_α is a complete Π^{1,C_α}_1 set of integers, we may apply a result that identifies such with the Σ_1-theory of the least admissible set over C_α :

PROPOSITION 2.6. *If $D \subseteq \omega$, and D^+ is a complete $\Pi^{1,D}_1$ set of integers, then $D^+ \equiv \Sigma_1$-Theory of $\langle L_{\omega_1^D}[D], \in, D\rangle$; in fact there is a recursive bijection $g : \mathbb{N} \longleftrightarrow \mathbb{N}$, so that uniformly for all D,*

$$n \in D^+ \Longleftrightarrow g(n) \in \Sigma_1 \cap Sent \wedge \langle L_{\omega_1^D}[D], \in, D\rangle \models g(n).$$

In particular: $C^+_\alpha \equiv \Sigma_1$-Theory of $\langle L_{\omega_1^{C_\alpha}}[C_\alpha], \in, C_\alpha\rangle$.

The proposition is essentially obtained from the relativisations of the theorems of Kleene and Spector, (see, *e.g.*, (Rogers, 1967), Theorems XLI &

XLIV) that classify the Π^1_1 predicates on integers as those Σ_1-definable over HYP (the latter the class of hyperarithmetic reals, together with the fact that the reals of $L_{\omega_{1ck}}$ are those of HYP).

We now proceed to the proof of the clauses (i)-(iv) of the lemma. First observe that (ii) follows from (i) and (iii) using Proposition 2.2(iii): first we note that in general, if B is arithmetic in C, then $\omega_1^B \leq \omega_1^C$. Thus $\omega_1^{C_\iota} \geq \tau_{\iota+1}$ (using (i) and Prop 2.2 (iii)). Further, using (iii), as $\rho(C_\iota) = \tau_\iota$, $\omega_1^{C_\iota} \leq \tau_{\iota+1}$. Hence (ii) holds.

We proceed to prove (i),(iii),(iv) by induction on $\iota < \Sigma$. For $\iota = 0$ they either hold vacuously or are trivial. Suppose now $\iota = \eta + 1$, and (i) to (iv) hold with η in place of ι. A_ι is the complete Σ_2-theory of $\langle L_{\tau_\iota}, \in\rangle$.

But note that $\tau_{\eta+1} = \tau_\iota$ is the next admissible ordinal above τ_η and hence is the height of the smallest admissible set containing C_η as an element: $\tau_\iota = \omega_1^{C_\eta}$ (using for (ii) the inductive hypothesis.) As sets, $L_{\omega_1^{C_\eta}}[C_\eta]$ = $L_{\omega_1^{C_\eta}}$ (because A_η is definable over L_{τ_η} and C_η is arithmetic in A_η). Using our nomenclature above:

$A^{1,C_\eta} = \Sigma_1$-Th($\langle L_{\omega_1^{C_\eta}}[C_\eta], \in, C_\eta\rangle$)= Σ_1-Th($\langle L_{\tau_\iota}, \in, C_\eta\rangle$).

However the Σ_2-theory of this structure is simply obtained: it is (recursively isomorphic to) the Turing jump of A^{1,C_η}. Thus:

$(A^{1,C_\eta})' \equiv \Sigma_2$-Th($\langle L_{\tau_\iota}, \in, C_\eta\rangle$.

However A_ι is the "pure $\dot{C}$-free" part of this theory, namely:

$A_\iota = \Sigma_2$-Th($\langle L_{\tau_\iota}, \in, C_\eta\rangle \cap \mathcal{L}_{\dot{\in}}$.

Putting this together one gets: $A_\iota \leq_T (A^{1,C_\eta})'$, and the latter is arithmetic in $A^{1,C_\eta} \equiv C^+_\eta \equiv C_{\eta+1} = C_\iota$ (where the last sentence of Proposition 2.6 is used for the first (1-1) equivalence here). Hence (i) holds for $\iota = \eta + 1$, and we even have that A_ι is recursive in the Turing jump of C_ι. We have seen that A^{1,C_η} is definable over $\langle L_{\tau_\iota}, \in\rangle$ (using the parameter C_η) which is a member of L_{τ_ι}; hence $\rho(A^{1,C_\eta}) \leq \tau_\iota$. But we have just asserted that $A^{1,C_\eta} \equiv C_\iota$. Hence $\rho(C_\iota) \leq \tau_\iota$ as well. But $\iota < \Sigma$, so we cannot have then $\rho(C_\iota) < \tau_\iota$, as otherwise we should also have $\rho(A_\iota) < \tau_\iota$, and this would contradict Proposition 2.2(i). This proves (iii) for ι. (iv) is vacuous for successor ι.

We now suppose that $\iota < \Sigma$ is a limit, and the inductive hypotheses (i)-(iv) hold for $\eta < \iota$. As $\iota < \Sigma$ we shall have that L_{τ_ι} is the Σ_2-skolem hull of $\varnothing$ inside L_{τ_ι} (by using the argument of Proposition 2.2 again); that is, $\forall x \in L_{\tau_\iota} \exists k L_{\tau_\iota} \models$"$\varphi^2_k[x]$" where $\langle \varphi^2_k | k \in \omega\rangle$ is the above recursive enumeration of all Σ_2-formulae of the language $\mathcal{L}_{\dot{\in}}$ with v_0 the single free variable. Thus we have that any $x \in L_{\tau_\iota}$ is implicitly Σ_2-definable for some k as $\varphi^2_k[x]$. We shall assume that the form of such formulae can be expressed as :

$(*)\ \varphi^2_k(v_0) \Longleftrightarrow \exists u \forall v \psi_k(u, v, v_0)$

for some Σ_0 ψ_k.

We divide into cases depending on whether L_{τ_ι} satisfies the Σ_1-Separation scheme or not. In the former we shall have a uniform (1-1) recursive way of recovering A_ι from C_ι; in the latter case it is the presence of parameters in the argument that seemingly prevent a uniform reduction.

Case 1 ι is a limit ordinal, but $L_{\tau_\iota} \not\models \Sigma_1$-Separation.

Let $\langle \varphi_i | i < \omega \rangle$ be a recursive enumeration of all Σ_1 formulae of one free variable of the language $\mathcal{L}_{\dot{\in}}$. The case hypothesis assures us there is a "Σ_1-parameter" $p \in L_{\tau_\iota}$, so that L_{τ_ι} is the Σ_1-skolem hull of the singleton set $\{p\}$[6]. By the comment on Σ_2-skolem hulls above, there is k so that φ^2_k uniquely defines p. Moreover, by the form at $(*)$ of the formulae φ^2_n, there is α_0 so that for all $\alpha < \iota, \alpha_0 < \alpha \longrightarrow L_{\tau_\alpha} \models$"$\varphi^2_k[p]$"; that is p is named in the same way in the structures L_{τ_α}. [7] We then have the following chain of equivalences.

$l \in A^1_{\tau_\iota} \Longleftrightarrow_{df} L_{\tau_\iota} \models \varphi_l[p]$

$\Longleftrightarrow \exists \beta_0 > \alpha_0 \forall \beta > \beta_0 \;\; L_{\tau_{\beta+1}} \models$"$L_{\tau_\beta} \models \varphi_l[p]$"

$\Longleftrightarrow \exists \beta_0 > \alpha_0 \forall \beta > \beta_0 \langle L_{\tau_{\beta+1}}[C_\beta], \in, C_\beta \rangle \models$"$\exists \overline{p} \; (\varphi^2_k[\overline{p}] \wedge L_{\rho(C_\beta)} \models \varphi_l[\overline{p}])$"

(using in the last equivalence the inductive hypothesis (iii)); the statement in the last line's quotation marks here is actually a Σ_1 sentence in $\mathcal{L}_{\dot{\in},\dot{C}}$ about k, l; so let us call this $\sigma(k,l)$. Note that the map $(k,l) \mapsto \sigma(k,l)$ can be assumed recursive. Using the Proposition 2.6 the last equivalence becomes:

$\Longleftrightarrow \exists \beta_0 > \alpha_0 \forall \beta > \beta_0 \;\; g^{-1}(\sigma(k,l)) \in C^+_\beta$

$\Longleftrightarrow g^{-1}(\sigma(k,l)) \in C_\iota$.

We thus have $A^1_{\tau_\iota} \leq_1 C_\iota$. This implies that A_ι is recursive in $(A^1_{\tau_\iota})' \leq_T C'_\iota$. Hence (i) holds: A_ι is arithmetic in C_ι. This leaves (iii): $\rho(C_\iota) \not< \tau_\iota$ (for otherwise $\rho(C_\iota) < \tau_\alpha$ for an $\alpha < \iota$, hence $A_\iota \in L_{\tau_\iota}$; this is a contradiction as $A^1_\iota \notin L_{\tau_\iota}$). However C_ι is $\Sigma_2(L_{\tau_\iota})$. Hence $\rho(C_\iota) = \tau_\iota$ as required.

Case 2 ι is a limit ordinal, and $L_{\tau_\iota} \models \Sigma_1$-Separation.

Now there is no such parameter p as in *Case 1*. Let $\langle \varphi^2_l \mid l \in \omega \rangle$ enumerate all Σ_2-sentences of $\mathcal{L}_{\dot{\in}}$.

$l \in A_\iota \Longleftrightarrow_{df} L_{\tau_\iota} \models \varphi^2_l$

$\Longleftrightarrow \exists \alpha_0 \forall \beta \geq \alpha_0 \;\; L_{\tau_\beta} \models \varphi^2_l$

We should justify this last equivalence: suppose φ^2_l is $\exists u \forall v \psi_l(u,v)$. ($\Longrightarrow$) If z is such that $L_{\tau_\iota} \models \forall v \psi_l(z,v)$, then we may choose $\alpha_0 < \iota$ with $z \in L_{\tau_{\alpha_0}}$. Then $\forall \beta \geq \alpha_0 \;\; L_{\tau_\beta} \models \forall v \psi_l(z,v)$. ($\Longleftarrow$) Let α_0 be as hypothesized. By Proposition 2.3, find $\gamma < \iota$ with $\alpha_0 \leq \gamma$ and $L_{\tau_\gamma} \prec_{\Sigma_1} L_{\tau_\iota}$. Now if $L_{\tau_\gamma} \models \forall v \psi_l(z,v)$, this is a Π_1 sentence about z that persists upwards to L_{τ_ι}.[8]

Now, using similar reasoning:

$\Longleftrightarrow \exists \alpha_0 \forall \beta \geq \alpha_0 \;\; L_{\tau_{\beta+1}} \models$"$L_{\tau_\beta} \models \varphi^2_l$"

$\Longleftrightarrow \exists \alpha_0 \forall \beta \geq \alpha_0 \;\; \langle L_{\tau_{\beta+1}}[C_\beta], \in, C_\beta \rangle \models$"$L_{\rho(C_\beta)} \models \varphi^2_l$"

(again using in the last equivalence the inductive hypothesis (iii) as before); now the sentence in quotes depends only on l, so let us call it $\sigma(l)$, again with $l \mapsto \sigma(l)$ assumed recursive. For the same reasons, we have the last equivalence:

$$\iff \exists\alpha_0\forall\beta > \alpha_0 \;\; g^{-1}(\sigma(l)) \in C^+_\beta$$
$$\iff g^{-1}(\sigma(l)) \in C_\iota.$$

Thus $A_\iota \leq_1 C_\iota$ and the reduction here $(g^{-1} \circ \sigma)$ is independent of ι.

QED (Lemma 2)

Proof of Theorem 3. As stated at the beginning of this "*third demonstration*", we could consider running Field's construction inside L, indeed inside L_Σ. The defining clauses of his semantic value assignments form a Σ_2-recursion. We note then:

(1) $|A|_\zeta$=$|A|_\Sigma$ for any sentence A of $\mathcal{L}^+$. Hence $C_\zeta = C_\Sigma$.

Proof: Immediate from the fact that there is a uniform Σ_2-recursion in the sentence variable A that defines these values over any L_λ, if $\lambda = \tau_\lambda$, say, and then $|A|_\lambda$ will be Σ_2-definable over L_λ. Focussing on L_ζ, and L_Σ respectively, we have that $L_\zeta \prec_{\Sigma_2} L_\Sigma$, and the result then follows.

QED(1)

(2) If $|A|_\zeta$= 0 (resp. 1) then $\forall\alpha \in (\zeta, \Sigma)$ $|A|_\alpha$= 0 (resp. 1).

Proof: Suppose this were false for some particular A with, say, $|A|_\zeta$= 0, but $|A|_\alpha \neq$ 0. Suppose $\beta_0 < \zeta$ is such that $\forall\beta \in (\beta_0, \zeta)|A|_\beta$= 0. By supposition $L_\Sigma \models \exists\alpha > \beta_0|A|_\alpha \neq$ 0. The latter is a Σ_1 sentence about A and β_0 and will reflect down to L_ζ (as all Σ_2-sentences about objects in L_ζ do). Hence $L_\zeta \models \exists\alpha > \beta_0|A|_\alpha \neq$ 0. However this contradicts our choice of $\beta_{0.}$ QED(2)

However (1) implies that the distribution of semantic values according to Field's clauses are exactly the same at stage ζ and at stage Σ. Hence they will be identical at stage $\zeta + 1$ and $\Sigma + 1$ and then at stages Σ and $\Sigma + \Sigma$, etc., etc. We thus shall have a "periodicity" of length Σ in these "snapshots" of semantic values. In short for any sentence A we have that $|A|_\zeta = |A|_{\Sigma.\rho}$ for any $\rho > 0$ (note that Σ is additively closed, so $\zeta + \Sigma = \Sigma$). The force of (2) is that any 0,1 semantic value assigned at stage ζ to a sentence A will remain assigned *at all later stages*. As (1) shows any sentence which has an unstable value at stage ζ will also do so at Σ and indeed by repetition as above, at any stage $\Sigma.\rho > \zeta$; we deduce that ζ is an acceptable point.

(3) ζ is the least acceptable point.

Proof: To summarise, we have shown that there is a (1-1) recursive function $h : \mathbb{N} \Longrightarrow \mathbb{N}$ so that $A \in D$ iff $h(\ulcorner A \urcorner) \in C_\zeta$ iff $h(\ulcorner A \urcorner) \in C_\xi$ where ξ is any acceptable point. Suppose for a contradiction that $\xi < \zeta$ was the least acceptable point.

We claim there is a $\theta < \zeta$ which is an acceptable point and further $L_\theta \models \Sigma_1$-Separation. This is because $L_\Sigma \models$*"There exists $\theta > \xi$ with $C_\theta = C_\xi$ and $L_\theta \models \Sigma_1$-Separation"*. ($\zeta$ is such). Hence there is, by Σ_1-reflection such a θ below ζ. However then we have that $C_\theta = C_\zeta$. This is a contradiction to Proposition 2.2 (i), since, as we have seen, A_ζ is recursively obtainable from $C_\theta = C_\zeta$, we should have that $\rho(A_\zeta) < \zeta$. QED(3)

The sequence of sets of assignment values will then reappear with a periodicity Σ, as required for the theorem. QED(Theorem 3)

Proof of Theorem 2 Lemma 2 shows that $\widetilde{O} \leq_1 D$. However the argument at (1) above shows that there is a Σ_2 formula $\psi(v_0) \in \mathcal{L}_{\dot{\in}}$ so that $A \in D$ if and only if $\psi(A)$ is true at L_ζ (equivalently L_Σ). Hence $D \leq_1 \widetilde{O}$. The result follows by Myhill's Theorem - the effective Cantor-Schröder-Bernstein Theorem mentioned above.

QED(Theorem 2)

3. Further equivalences.

We discuss some further equivalences. We state the result first before going on to discuss the sets S, Q being introduced.

Fact 1 [9] *The following are recursively isomorphic:*
(i) S ; (ii) A_ζ ; (iii) Q ; (iv) H.

We introduce some nomenclature (taken from (Hamkins and Lewis, 2000)) in order to talk about the class of *infinite time turing machine* (ITTM) computations. This will be the straightforward generalisation of that from ordinary recursion theory. We let $\langle P_e | e \in \omega \rangle$ be an enumeration of all ITTM programs (which are no different from ordinary Turing programs, apart from the addition of a single new "*limit state*", q_L, to which the machine enters at limit stages of time). We use the notation that "$P_e(k) \downarrow l$" to mean the computation halts with output tape containing the string $l \in 2^\omega$. "$P_e(k) \mid 0$" means that the computation halts, or eventually enters an infinite loop, but that 0 (the infinite string of zeros) remains on the output tape (even if the machine has not formally halted). Note there is, as expected, a third class of computations: it may be that the e'th program on input k loops without any stable contents on its output tape, we denote this as usual : "$P_e(k) \uparrow$". The definition of S to follow is a generalisation of *Turing jump.*

DEFINITION 3.1. $S =_{df} \{e : P_e(0) \mid 0\}$

A perhaps more understandable version of any of the sets we have been considering is the set Q, the *complete arithmetic quasi-inductive* set of integers. To see what Q is we follow Burgess (Burgess, 1986) (who distilled this definition from that of a Herzberger revision sequence), and consider the following.

Let $\Gamma : \mathcal{P}(\omega) \to \mathcal{P}(\omega)$ be any arithmetic operator (that is "$n \in \Gamma(X)$" is arithmetic; we emphasise that Γ need be neither monotone nor progressive). We define the following iterates of $\Gamma : \Gamma_0(X) = \Gamma(X); \Gamma_{\alpha+1}(X) = \Gamma(\Gamma_\alpha(X)); \Gamma_\lambda(X) = \liminf_{\alpha \to \lambda} \Gamma_\alpha(X) = \cup_{\alpha<\lambda} \cap_{\lambda>\beta>\alpha} \Gamma_\beta(X)$. We say that $Y \subseteq \omega$ is *arithmetically quasi-inductive* (AQI) if for some such Γ, Y is (1-1) reducible via a recursive function to $\Gamma_{On}(\varnothing)$. The *complete arithmetically quasi-inductive set* Q, to which all others are similarly reducible, has then also the various recursively isomorphic characterisations stated in Fact 1 above. The list of recursively isomorphic sets displayed in the Fact above all arise from definitions or procedures that can be seen to fall under the scheme of being arithmetically quasi-inductive. It can be shown that the set of stable truths of a Herzberger revision sequence, and Field's set of ultimate truths, can be obtained *via* an arithmetic quasi-inductive definition.[10] Each ITTM program can be thought of as producing a set of 0/1 values on an infinite tape according to a *recursive* quasi-inductive definition. Moreover, by consideration of a universal ITTM program, one may see that the complete arithmetical quasi-inductive set is in fact recursively quasi-inductive (and hence so are $D, H \ldots$ etc.) A further example of an occurrence of this scheme appears in computer science: see (Kreutzer, 2002) where the author seeks to separate out various fixed point logics on finite structures, by considering essentially the same kind of infinite quasi-inductive process.

COROLLARY 3.1. *Let $\langle A_n | n \in \omega \rangle$ be a recursive enumeration of all sentences of $\mathcal{L}^+$. Let $B \subseteq \mathbb{N}$ be any arithmetical quasi-inductive set. Then there is a recursive (1-1) function f so that so that $B = \{n : A_{f(n)} \in D\}$.*

Perhaps unsurprisingly:

COROLLARY 3.2. *$\models_{LCC}$ is not axiomatisable.*

The last part of the Remark (ii) - that certain theories are insufficient for the conservativeness result of Field's Sect.6 - is only an observation that is a consequence of the fact that such theories cannot prove the existence of acceptable points, in particular of ζ. The least β-model of $\mathbf{\Delta^1_3}$-CA_0 occurs as the set of the reals of some level L_γ of the constructible hierarchy, for an ordinal γ that is much smaller than ζ. Hence we can find such a model containing no code for a wellordering along which we can perform an arithmetic quasi-inductive definition and reach a point in the ordering where we have a repetition in the values produced. In the terms here, such orderings have no "acceptable" points.

4. Revenge-immunity

We discuss here some aspects of Field's claim to have produced a "revenge-immune" solution to the semantic paradoxes.[11] This claim itself seems at first sight paradoxical: surely any system based on a sufficiently strong theory to enable Gödel coding of the syntax and diagonalisation arguments to be performed, will have some semantic predicate of the system that allows one to "diagonalise out" of the system and thus either obtain a self-referential sentence concerning that semantic notion leading to a contradiction, or else one concludes that the semantic notion is not "representable" within the system?

Indeed Field's system is no different (indeed cannot be) from others in this regard if we maintain a classical metatheory. The relevant semantic notion here is that of *ultimate truth*. Indeed he points out in Section 6 that (again assuming a classical metatheory) we may formulate a predicate $D^*(A)$ which is defined to have value 1 if $\|A\| = 1$, and 0 otherwise, then this cannot be semantically represented in his system *via* a formula in the language, for otherwise we should be able to produce a self-referential sentence L_* equivalent to $\neg D^*(\text{True}(\langle L_* \rangle))$ in the usual way, and thus a contradiction.

There is demonstrably within his system a hierarchy of 'determinately true' operators. Moreover such can be used to give an account within the system of how the traditional Liar fails to gain a truth value. Indeed he gives a hierarchy of such operators $\langle D^\alpha | \alpha < \lambda_0 \rangle$ that can deal with more and more complicated super-liars.

> Third, and probably most important, the new conditional can be used to show that the theory is not subject to "revenge problems."
>
> More fully, the addition of the new conditional operator to the language allows for the definition of a natural "determinately operator", so that we can consistently handle "extended paradoxes", such as sentences that assert of themselves that they are not determinately true." (Field, 2003) p140.

The determinately true operators do go some way to helping Kripke out from his dilemma that Liar sentences are not classified as untrue in his minimal fixed points, whilst at the same time we are precluded from expressing this in the object language. Field with his operator D effortlessly can form such object language expressions. We note below in the next subsection that it is not the new conditional that allows for a notion of "determinately true" - this can be captured using the older notions of a Herzbergerian Revision sequence (although in such sequences there is no "conditional" which allows for a naive theory of truth to be stated using them).

He then argues that the reason D^* is not semantically representable by any formula in the language, is not because of any limitation of expression on the part of the object language (which includes arithmetic) or that of the metalanguage and theory in which his theorems are proven (that of set theory, ZFC). We take no issue then with the claim on p166 (op.cit) which states *"there is no need to use a broader classical metalanguage to do the semantics; we can use L itself"*. In general it seems then, to be a wrong turn down a blind alley, to argue that since the predicate indicating "has semantic value 1" cannot be representable in such a semantic theory of truth, that therefore one must broaden the meta*language*. Of course we can beef up the meta*theory* and use a construction dependent on a longer initial segment of the ordinals to capture a more complicated notion of ultimate truth by proceeding in a way that goes further up the L-hierarchy. Just as Kripke's least fixed point theories (supervaluational, strong Kleene or otherwise) can be established over a model, say that of the standard natural numbers, using a relatively weak classical metatheory, so we might regard *this particular construction* of Field's as pursued in a much stronger classical metatheory (and necessarily so as we have shown).

(Indeed one may analyse precisely what predicates or operators are so representable in *this particular* semantical construction over the standard model of arithmetic, and one obtains a precise answer: those predicates that are both AQI and co-AQI. As D^* is a *complete* AQI set in effect, it is universal for all AQI sets, and by a standard diagonalisation argument cannot itself be both AQI and co-AQI. Hence, as we already knew, it cannot be representable.)

I emphasize 'this construction' because I take it that Field only wishes to show that his logical system is consistent. The question remains open as to whether a consistency proof could be established using a weaker classical metatheory. However it would appear that any argument for "revenge immunity" can not be done within a *classical* metatheory for the reasons outlined at the beginning of this section. If it could be shown that a consistency proof *required* the use of a strong theory - and I think showing this might be rather difficult - then this could be construed as a criticism of the theory, in the way that, say, principles that can 'only' be shown consistent relative to large cardinals in set theory, get criticized. In the context of 'truth' for the standard model of the natural numbers, the second order theory $\mathbf{\Delta^1_3}$-CA_0 might well be said to have the status of a 'large cardinal theory'. It is a very strong fragment of second order number theory, well beyond anything one might call "predicative" in Feferman's sense, and beyond current proof theoretical techniques of ordinal analysis. (Another way to put this is to say that the least acceptable point Δ - here called ζ, is extremely large.) It is hard to think of any theorem in mathematics, in particular in *analysis* other than Borel Determinacy results, that require this strength in the background theory. If no simpler consistency result can be found for the theory LCC (using a classical metatheory) then one is left wondering as to the status of "truth for the natural numbers" if one wishes to adopt this system. That is so independently of the stance one takes on the "revenge-immunity" part of the theory. On the other hand it may well be that by abandoning classical reasoning in the metatheory some other argument may allow a simpler justification for LCC, for those brave enough to take that leap. Then one may have Field's "revenge immunity" if one is prepared to have such a weak non-classical meta-theory.

However, that aside, and returning to Field's construction under discussion, we might imagine after following it through, that we no longer have the feeling, as Kripke seemingly does at the paragraph towards the end of his paper (Kripke, 1975), that the 'ghost of the Tarski hierarchy is still with us.' Kripke might have felt that Liar sentences are *not true* in the object language, '... but we are precluded from saying this is the object language...' What the semantical notion of ultimate truth, or D^*, does is transport us to a situation where the assertion '$D^*(A) \neq 1$' is so far removed from ordinary 'untruth' about A that we may be uninclined to worry that, residually, L_* ought to be, or is 'nevertheless', untrue. However exporting our possibly self-referential notions away to these remoter regions, or complicating them up, so that they no longer enter the circle of our concerns about basic, immediate, common or garden 'truth' seems not to be a clear way through to a genuine revenge-immune solution. The ghost, although thinner and ghostlier, is still perceptible. We seem to need the failure of the Law of Excluded Middle within a non-classical metatheory, (so that we cannot make claims about any ultimate D^*'s bivalency). In which case though, presumably much is up for grabs...

5. Endnote: 'Determinately true' predicates in a herzbergerian setting.

Whilst Field uses the '$\longrightarrow$' operator in his consistency proof construction to capture some of the notions of a conditional, its most striking feature is that it allows one to capture some notion of the 'history' of the semantical values of a sentence within the object language. (i) $\top \longrightarrow A$, when evaluated at a successor stage $\alpha = \gamma + 1$, tells you whether A was true or not at stage γ; (ii) when at a limit α, whether or not A was true on a 'tail' of the ordinals below α. His operator $D(A)$ which is $(\top \longrightarrow A) \wedge A$, if (i) it has value 1 at stage $\gamma + 1$ says "A was true at the previous stage and it is so now"; (ii) for limit stages this has the effect of "A has recently been always true and is so now" (with a little fanciful interpretation of stages as being that of time.)

However, note how we can get this effect in Herzberger's two-valued setting: $\top \longrightarrow A$ is no more than "Tr(A)" and $D(A)$ becomes simply "$Tr(A) \wedge A$". The effects of (i) and (ii) are thus there also in the Herzberger system. Field then iterates these definitions producing: $D^{\alpha+1}(A)$ as $D(D^{\alpha}(A))$, and for suitable recursive limit ordinals, and in a suitable notation system, $D^{\lambda}(A)$ as well. One can mimic this process exactly within Herzberger's system also. In short: one does not need a new conditional in order to have a "determinately true" operator in order to handle "extended paradoxes" of this kind within the object language. In other words the description of the ' "singular" status of the Liar and the Truth-teller,' as Field puts it (Field, 2003) p.157, as well as the extended variants, can be equally well handled within Herzberger.

Notes

[1] In brief, this is second order number theory with a set induction scheme, but with comprehension restricted to formulae expressible in a Δ^1_3 fashion.

[2] In a nutshell: we are working in a part of the constructible hierarchy, where each level has a Σ_1-truth set that is definable using, in general, a necessary single infinite ordinal parameter. This parameter has a Σ_2-definition over L_α but not a uniform one independent of α. For those L_α which are models of Σ_1-Separation, there is in any case no such parameter from which all sets are Σ_1-definable. However for these latter type of L_α, we do in fact have a uniform *recursive* way of recovering a Σ_2 truth set from the Fieldian set of locally stable truths C_α, as this is the "liminf" of the previous Σ_2-truth sets for the models L_β $(\beta < \alpha)$.
Note: Since writing this, Sy Friedman has pointed out that we can uniformly treat *both* types of stages with a marginally more complicated reduction procedure; one then has for models L_λ for limit λ a way of finding the Σ_2-truth sets recursively in the *Turing jump* of the "liminf" truth sets. We have not re-written this into the above, as it would only involve further technicalities, without bringing forth any great perspiciences.

[3] One might further add that a similar situation pertains to the least Kripkean fixed points under the various schemes, such as weak- and strong-Kleene, and the supervaluation scheme: such sets are all recursively isomorphic - without there being any obvious recursive "translation" mechanism.

[4] This means that for some $\xi > \zeta$ we have that $L_\zeta \prec_{\Sigma_2} L_\xi$, the subscript Σ_2 indicating that Σ_2 formulae with parameters allowed from L_ζ have the same truth value in both structures.

[5] Actually the theorem cited shoes that L_ζ is a model of the Δ_2-Comprehension, and Σ_2-Collection Schemes.

[6] The existence of such a parameter can be established as follows: suppose there were no such p. Then the Σ_1-skolem hull of $\varnothing$ in L_{τ_ι}, is not all of L_{τ_ι}; one can argue that this hull is transitive, and in fact is thus L_{α_0} for some $\alpha_0 < \tau_\iota$. Hence $L_{\alpha_0} \prec_{\Sigma_1} L_{\tau_\iota}$. Now consider the Σ_1-skolem hull of $\{\alpha_0\}$ in L_{τ_ι}; again by supposition this is not all of L_{τ_ι}; it is transitive and hence is some $L_{\alpha_1} \prec_{\Sigma_1} L_{\tau_\iota}$, with $\alpha_0 < \alpha_1 < \tau_\iota$. Continuing in this way we could create an infinite chain of such substructures of ordinal height $\alpha_n < \tau_\iota$. This would contradict Proposition 2.3. In such a situation as here there is a "standard Σ_1-parameter" $p = p^1_{\tau_\iota} \in L_{\tau_\iota}$, so that L_{τ_ι} is the Σ_1-skolem hull of the singleton set $\{p\}$. Such a parameter is least in a

particular wellordering on finite sequences of ordinals below τ_ι. It can be shown that in our situation this standard parameter is a single ordinal $\pi < \tau_\iota$, (perhaps just 0). For a fuller account of such parameters and hulls see (Devlin, 1984) IV.

[7] The point here is that if we suppose that $L_{\tau_\iota} \models$"$\varphi^2_k[p]$ " where $\varphi^2_k[p] \iff \exists u \forall v \psi_k(u, v, p)$, then if $L_{\tau_\iota} \models \forall v \psi_k(z, v, p)$ for some particular z we may take α_0 sufficiently large so that $z \in L_{\tau_{\alpha_0}}$. Then for any $\alpha > \alpha_0$ we shall have $L_{\tau_\alpha} \models \forall v \psi_k(z, v, p)$ (as the latter is Π_1, *i.e.* a universal statement). Thus p is named in the same way by φ^2_k in all such structures. Note that p will have other, different, Σ_2 definitions in each such structure, but they will not necessarily persist upwards to L_{τ_ι}.

[8] Note how this argument can fail if L_{τ_ι} is not a model of Σ_1-Separation: take $\iota = \omega$, and φ^2_l as "there is a largest admissible ordinal". Then for all $n \in \omega$, $L_{\tau_n} \models \varphi^2_l$, but this sentence is false in L_{τ_ω} (and moreover $g^{-1}(\sigma(l)) \in C_\omega$).

[9] The equivalence of (i) and (ii) is (Welch, 2000) Thm 2.6; whilst working on this latter paper, we were unaware of the very concrete links with the much earlier work of Burgess in (Burgess, 1986). There he explicitly states the equivalence of (iii) and (iv) as Prop.13.1; but the equivalence of (ii) and (iii) is essentially his Theorem 14.1 there too.

[10] This is done directly for the Herzberger sequence with a null starting hypothesis in (Burgess, 1986).

[11] Since this article was written Field has produced an extended explication of the claim to have produced such a solution ((Field, 2008)) - hence we have not considered that here. A very detailed reply to Field on this topic is the recent (Leitgeb, 2006).

References

Barwise, K.: 1975, *Admissible Sets and Structures*, Perspectives in Mathematical Logic. Springer Verlag.

Burgess, J.: 1986, 'The truth is never simple'. *Journal for Symbolic Logic* **51**(3), 663–681.

Devlin, K.: 1984, *Constructibility*, Perspectives in Mathematical Logic. Berlin, Heidelberg: Springer Verlag.

Field, H.: 2003, 'A revenge-immune solution to the semantic paradoxes'. *Journal of Philosophical Logic* **32**(3), 139–177.

Field, H.: 2008, 'Solving the Paradoxes, escaping revenge'. In: J. Beall (ed.): *The Revenge of the Liar*. Oxford: O.U.P.

Gupta, A. and N. Belnap: 1993, *The revision theory of truth*. Cambridge: M.I.T. Press.

Hamkins, J. D. and A. Lewis: 2000, 'Infinite Time Turing Machines'. *Journal of Symbolic Logic* **65**(2), 567–604.

Herzberger, H.: 1982a, 'Naive Semantics and the Liar Paradox'. *Journal of Philosophy* **79**, 479–497.

Herzberger, H.: 1982b, 'Notes on naive semantics'. *Journal of Philosophical Logic* **11**, 61–102.

Kreutzer, S.: 2002, 'Partial fixed point logic on infinite structures'. In: *Annual Conference of the European Association for Computer Science Logic (CSL)*, Vol. 2471 of *Lecture Notes in Computer Science*. Springer.

Kripke, S.: 1975, 'Outline of a theory of truth'. *Journal of Philosophy* **72**, 690–716.

Leitgeb, H.: 2006, 'On Field's "Solving the Paradoxes, escaping revenge"'. *APA Proceedings*.

McGee, V.: 1991, *Truth, Vagueness, and Paradox: An essay on the Logic of Truth*. Hackett.

Moschovakis, Y.: 1974, *Elementary Induction on Abstract structures*, Vol. 77 of *Studies in Logic series*. Amsterdam: North-Holland.

Rogers, H.: 1967, *Recursive Function Theory*, Higher Mathematics. McGraw.

Simpson, S.: 1999, *Subsystems of second order arithmetic*, Perspectives in Mathematical Logic. Springer.

Welch, P. D.: 2000, 'Eventually Infinite Time Turing Degrees: infinite time decidable reals'. *Journal for Symbolic Logic* **65**(3), 1193–1203.