

Curry's paradox, Łukasiewicz, and Field

Peter Smith

June 29, 2010

In approaching Ch. 4 of *Saving Truth from Paradox*, it might be helpful first to revisit Curry's original paper, and to revisit Łukasiewicz too, to provide more of the scene-setting that Field doesn't himself fill in. So in §1 I'll say something about Curry, in §2 we'll look at what Łukasiewicz was up to in his original three-valued logic, and in §3 we'll look at the move from a three-valued to a many-valued Łukasiewicz logic.

In §4, I move on to announce a pair of theorems which kill off any prospect of using a Łukasiewicz logic as a general setting for taming paradoxes. In the light of this, we can regard Field's chapter as an effort to rescue something from the wreckage. So in §5 we take an outline look at what Field thinks *can* be rescued, highlight his main result, and wonder very briefly indeed about its significance.

It shouldn't need saying that these remarks are not supposed to be a rounded critical appraisal of Field's chapter! They were originally written very speedily as a short introductory presentation for a seminar – a spur to further discussion, not a final verdict. And to keep the length of the introduction under control, I in particular don't here discuss Field's 'determinately' operator.¹

1 Curry's original paper

Curry published a short note 'The Inconsistency of Certain Formal Logics' in *JSL*, 1942. That paper was originally a sort-of-footnote to a much longer paper on the inconsistency of an early version of the λ -calculus: the point of his additional note is show that we can get a negation-free paradox in that calculus.

Now, the λ -calculus operates on identities between terms, and Curry's original argument works by noting that the particular version is such that, for any term b , there will be a term u such that it is provable that $u = u \rightarrow b$ where ' \rightarrow ' is a conditional-like operator. The argument proceeds, of course, by substituting identicals in identities. What is nowadays called Curry's paradox works not with identities between terms but with biconditionals between propositions. So we need to do just a little work to extract the argument we are going to be interested in from its original setting.

His argument clearly divides into two stages. First (A) Curry gives a very simple deductive schema that can be used whenever we have a standard conditional and a counterpart biconditional: what this schema reveals is that, given standard conditionals, if for *every* ψ , theory T proves *some* biconditional of the form $\varphi \leftrightarrow (\varphi \rightarrow \psi)$, then (iii) T is inconsistent. Then (B) Curry notes two sorts of cases where the schema can be applied to cause trouble.

¹Many thanks to Andrew Bacon and Tim Button for discussions after the seminar which have improved this handout. And having read further into the book, I've also returned to tweak the presentation a little.

(A) Let's say that a connective \rightarrow is a *standard conditional* if it obeys the following rules:

1. (MP) Modus ponens is valid: from φ and $\varphi \rightarrow \psi$ you can infer ψ .
2. (Con) Contraction is valid: from $\varphi \rightarrow (\varphi \rightarrow \psi)$ you can infer $\varphi \rightarrow \psi$.

A word about contraction. Evidently if (i) a conditional allows conditional proof, and (ii) a supposition can be used more than once, then contraction follows. For suppose you are given something of the form $\varphi \rightarrow (\varphi \rightarrow \psi)$ as a premiss: now suppose the relevant φ . Modus ponens using this supposition twice gives ψ . And discharging the supposition by conditional proof gives $\varphi \rightarrow \psi$. So the price of rejecting contraction for a propositional conditional initially seems pretty high! Either (i') we must reject conditional proof (but are we then still dealing with a *conditional*?), or (ii') we must disallow the re-use of suppositions.

Let's further say that \leftrightarrow is the biconditional counterpart of a conditional \rightarrow when \leftrightarrow is governed by the obvious rules:

3. (Bi) From $\phi \rightarrow \psi$ and $\psi \rightarrow \phi$ you can infer $\phi \leftrightarrow \psi$; and conversely, from $\phi \leftrightarrow \psi$ you can infer $\phi \rightarrow \psi$ and also infer $\psi \rightarrow \phi$.

And now Curry notes the following very simple deduction schema, starting from a biconditional of the form $\varphi \leftrightarrow (\varphi \rightarrow \psi)$:

$$\frac{\frac{\varphi \leftrightarrow (\varphi \rightarrow \psi)}{\varphi \rightarrow (\varphi \rightarrow \psi)} \text{Bi}}{\varphi \rightarrow \psi} \text{Con} \quad \frac{\varphi \leftrightarrow (\varphi \rightarrow \psi)}{(\varphi \rightarrow \psi) \rightarrow \varphi} \text{Bi}}{\psi} \text{MP}$$

This simple proof shows, as announced, that if we are working inside a theory T which (i) has a standard conditional \rightarrow (and counterpart biconditional \leftrightarrow), and (ii) where for every ψ , T proves *some* biconditional of the form $\varphi \leftrightarrow (\varphi \rightarrow \psi)$, then (iii) T proves everything and is inconsistent.

(B) Now, as I said, Curry himself was working not with biconditionals but identities; and instead of the (Bi) rules he has substitution of identities (starting from terms of the form $\varphi \rightarrow \varphi$). He then notes that the particular theory which it is his concern to show inconsistent has *two* different methods of constructing, from any b , a provable identity of the kind $u = u \rightarrow b$; one is Russell-like (in effect gets trouble from naive comprehension), one is Epimenides-like (gets trouble from a truth-predicate). Transposing into our biconditional key, that points to two kinds of theoretical frameworks with standard conditionals in which, for every ψ , we will be able to prove some $\varphi \leftrightarrow (\varphi \rightarrow \psi)$, so revealing the frameworks to be inconsistent.

The first framework comprises any theory T_1 which is a set theory with both (i) a standard conditional and (ii) the naive comprehension principle

$$\text{For any monadic } \chi(x), \exists_1 X \forall x(x \in X \leftrightarrow \chi(x)).$$

Choose ψ as you like, and then put $\chi(x) =_{\text{def}} (x \in x \rightarrow \psi)$. Applying comprehension, there is a set we can dub τ such that $\forall x(x \in \tau \leftrightarrow (x \in x \rightarrow \psi))$. So, in particular, $\tau \in \tau \leftrightarrow (\tau \in \tau \rightarrow \psi)$. So put φ for $\tau \in \tau$ and we are done; we've shown that for any ψ you choose, there's a φ such that T_1 proves $\varphi \leftrightarrow (\varphi \rightarrow \psi)$. So such a T_1 is inconsistent.

The second framework comprises any theory T_2 with (i) a standard conditional, (ii) enough arithmetic to establish the usual diagonalization lemma, and (iii) a truth-predicate T such that, within T_2 -theorems, we can always intersubstitute $\mathsf{T}(\ulcorner \varphi \urcorner)$ and φ . Choose ψ as you like, and consider the one-place predicate $(\mathsf{T}(x) \rightarrow \psi)$. The diagonalization lemma assumed in (ii) tells us that for some φ , T_2 proves $\varphi \leftrightarrow (\mathsf{T}(\ulcorner \varphi \urcorner) \rightarrow \psi)$. By (iii), we can substitute φ for $\mathsf{T}(\ulcorner \varphi \urcorner)$, so T_2 proves $\varphi \leftrightarrow (\varphi \rightarrow \psi)$. So again, such a T_2 is inconsistent.

What’s the moral so far? We can’t avoid all paradox just by boxing clever with negation and/or swallowing certain contradictions (and banning the “arguing past contradictions” which produces explosion). As Geach remarked long ago when commenting on a Curry-like argument in his 1955 *Analysis* note ‘On “insolubilia”’ – a title that reminds us that there are medieval antecedents to these discussions – “If we want to retain the naive view of truth, or the naive view of classes . . . , then we must modify the elementary rules of inference relating to ‘if’.”

2 Łukasiewicz’s three-valued logic

I don’t want to suggest that understanding Łukasiewicz’s early work on propositional many-valued logics really helps us understand Field: it doesn’t. But still, it is worth knowing a bit about the background – and we’ll see that a problem in understanding Łukasiewicz carries over to point up an issue in understanding Field.

Łukasiewicz’s early work on propositional many-valued logics was motivated by an antipathy to determinism, and in particular by the thought that the future is genuinely open. So, speaking of the future, some propositions are (already) true, some are (already) false, some are left open, neither yet true nor yet false.

I maintain that there are propositions which are neither true nor false but *indeterminate*. All sentences about future facts which are not yet decided belong to this category. Such sentences are neither true at the present moment, for they have no real correlate, nor are they false, for their denials too have no real correlate.

So, on this view, there are three different statuses that a proposition can have: true, false and indeterminate. Suppose that’s right.

How then should connectives work? We might naturally suppose that if φ is already true, its negation $\neg\varphi$ should be false, and if φ is already false, its negation should be true. While to say it is indeterminate whether φ is to say it is indeterminate whether or not φ – so if φ is indeterminate, $\neg\varphi$ should be indeterminate too.

The natural way of dealing with conjunction is to say that if φ and ψ are both already true, then $\varphi \wedge \psi$ is already true, and if one of them is already false then $\varphi \wedge \psi$ is already false. The remaining cases are where at least one of φ and ψ is indeterminate, and the other is indeterminate or true: in such cases it is still indeterminate whether $\varphi \wedge \psi$ (things might get fixed one way or the other, or maybe never get fixed).

And the natural way of dealing with disjunction is dual. If φ and ψ are both already false, then $\varphi \vee \psi$ is already false, and if one of them is already true then $\varphi \vee \psi$ is already true. The remaining cases are where at least one of φ and ψ is indeterminate, and the other is indeterminate or false: in such cases it is still indeterminate whether $\varphi \vee \psi$.

Note that if φ has the neither-true-nor-false status, then (on our treatment) $\neg\varphi$ has the same status, and hence $\varphi \vee \neg\varphi$ is not yet true but indeterminate. So the law of

excluded middle fails.

We can set all that out in tabular form in the entirely obvious way:

\wedge	T	I	F	\vee	T	I	F	\neg	
T	T	I	F	T	T	T	T	T	F
I	I	I	F	I	T	I	I	I	I
F	F	F	F	F	T	I	F	F	T

We might casually call these ‘three-valued’ tables – indeed, that’s the standard label! But this is potentially misleading. We could equally well say – indeed, maybe it would be better to say – that there are still just two truth-values that a proposition can take, truth and falsity: what we are now explicitly recognizing is the possibility that a proposition might not (yet) have a determinate value.

And what about a conditional? If we are in the game of looking for something which is ‘value’-functional and which coincides with the classical truth-functional conditional when we dealing with definite values (true, false) then we have to start the table like this:

\rightarrow	T	I	F
T	T	?	F
I	?	?	?
F	T	?	T

where the left column gives the value of the antecedent and the top row the value of the consequent.

Now, we might complete the left column by reflecting that if the consequent of $\varphi \rightarrow \psi$ is already true, then that settles it: ψ holds whatever the future brings with respect to φ (if it brings anything at all!), and in particular ψ still holds even if φ . And if we want our conditional to allow contraposition, we’ll need a corresponding T in the middle of the bottom row too.

\rightarrow	T	I	F
T	T	?	F
I	T	?	?
F	T	T	T

Further, we can complete the top row by reflecting that if φ is true, and ψ is indeterminate, then there’s all to play for with $\varphi \rightarrow \psi$ – it is as yet indeterminate and could go either way when ψ ’s value gets settled. Similarly for the right column. So we get

\rightarrow	T	I	F
T	T	I	F
I	T	?	I
F	T	T	T

So far, then, that’s all pretty natural.

But now, Łukasiewicz arguably goes off message. You’d expect him to say that, if at the moment it is indeterminate whether φ and indeterminate whether ψ , nothing (yet) settles whether $\varphi \rightarrow \psi$. But in fact he completes the table like this:

\rightarrow	T	I	F
T	T	I	F
I	T	T	I
F	T	T	T

He isn't exactly clear about *why* he makes this choice. He says in 1930 that "the desired equations I obtained on the basis of detailed considerations, which were more or less plausible to me". But it isn't easy to find explicit considerations (though I've only done limited homework). So what's going on?

Well, note that in the alternative table with 'I' in the central position in the matrix, there will be no tautologies in the logic at all (meaning now a proposition that always takes the value T, whether its atomic components take T, F or indeed I). For on that alternative table, complex propositions built up entirely from constituents taking the 'value' I will always take the 'value' I. This, perhaps, Łukasiewicz took as step too far away from classical ideas: after all (you might think) whatever φ 's status, doesn't it just *have* to be the case – and so is true already – that *if* φ , then φ ?

Two points about this, however. (i) While bandying intuitions, you might have thought that it also just can't be the case that if φ , then $\neg\varphi$ without providing us a reductio of φ which shows it to be false. Yet on Łukasiewicz's table, $\varphi \rightarrow \neg\varphi$ is *true* even if φ is I. (ii) Maybe there's a deeper issue here. Sure, the *inference* φ therefore φ indeed has to be valid; but once we move away from the classical framework, perhaps we mustn't run together validity of inference with truth of conditionals (incidentally, confusion about validity and truth of conditionals famously underwrites Łukasiewicz's treatment of the Aristotelian syllogism, though that is another story).

Anyway, whatever the rationale, Łukasiewicz settles on the conditional table as given. *And that choice makes the conditional non-standard.* Take valid inference in this context still to be defined as before: if the premisses are (already) true, the conclusion must be true too. Then (MP) remains reliable: if φ is T, and $\varphi \rightarrow \psi$ is T we have to be at the top left hand corner of the value matrix, and ψ is true. But (Con) fails. For suppose φ is I and ψ is F. Then $\varphi \rightarrow \psi$ remains I, and so we read off the value of $\varphi \rightarrow (\varphi \rightarrow \psi)$ from Łukasiewicz's deviant(?) filling of the central slot to get T. Hence (Con) can take us from a T premiss $\varphi \rightarrow (\varphi \rightarrow \psi)$ to a non-T conclusion $\varphi \rightarrow \psi$ and so isn't valid on Łukasiewicz's understanding of the value-functional conditional in his three-valued setting.

3 From three values to lots of values

We've seen, then, that we can purchase off the shelf a conditional which is non-standard, where the contraction principle (Con) fails, and so Curry paradox argument will fail for a theory using this conditional. The conditional lives in a three-valued logic; and true, the logic's original motivation for its conditional is more than a bit suspect. But ok, now we've got the tables available, we can perhaps 'borrow' them for a logic of paradox and suggest a different motivation. After all, we might reasonably suppose that the lesson of the paradoxes is that some claims are neither T nor F but Indeterminate.

Is there any mileage in this idea? Given the Łukasiewicz tables, a counterpart biconditional will be true just when both sides take the same value. Now in the framework with naive truth and a diagonalization lemma for its biconditional, for every ψ , there will be a φ such that it comes out true that $\varphi \leftrightarrow (\varphi \rightarrow \psi)$. Suppose α is false, then by

inspection of the tables, we need a wff β such that β is Indeterminate for $\beta \leftrightarrow (\beta \rightarrow \alpha)$ to come out true. So then there must be an γ such that $\gamma \leftrightarrow (\gamma \rightarrow \beta)$ comes out true, with β Indeterminate. But this time, inspection of the tables shows that this is impossible in Łukasiewicz's logic. Trouble!

I suppose some might already get off the bus at this point, and jettison the tentative idea of using the dodgy Łukasiewicz conditional in a paradox-avoiding theory, since it falls badly at the first hurdle. But other logicians are made of sterner stuff, pick themselves up, and press on. Since we get into trouble with three 'values', what we need, the story goes, is to recognize *more* values! How do we do the trick technically? And how are we to conceive of these extra values?

To ease our technical path, let's change our labelling of the various 'values' involved in the original, Mark I, three-valued logic. Let's now use '1' for T, '0' for F, and ' $\frac{1}{2}$ ' for Indeterminate, neither-true-nor-false. And we'll write $|\varphi|$ for the value of φ . Then, rather neatly, we have in the three-valued case that, for all φ, ψ ,

1. $|\neg\varphi| = 1 - |\varphi|$
2. $|\varphi \wedge \psi| = \min(|\varphi|, |\psi|)$
3. $|\varphi \vee \psi| = \max(|\varphi|, |\psi|)$
4. $|\varphi \rightarrow \psi| = \min(1, 1 - |\varphi| + |\psi|)$

These, of course, aren't the only equations we could read off the three-valued table, but they are the most natural. And they evidently can be carried over to a framework where we allow more than three values.

But what are we to make of the suggestion that we should use more than three values? As I remarked before, in the original three-valued framework it might be better to say that there are still just two values that a proposition can take, truth and falsity: we are simply explicitly marking the (supposed) possibility that a proposition might not (yet) get to determinately have one of those value. Indeterminate is not really an intermediate *value*, so much as a lack of value: but then what sense does it make to suppose we should interpose more of indeterminacies between truth and falsehood? Frankly I have no idea!

Field in his §4.3 briefly addresses this issue, but almost all he says at this point is that "Just as $|A| = 1$ suffices for the truth of A , $|A| \leq |B|$ suffices for the conditional that if A is true, so is B ." But that merely tells us what the table for a Łukasiewicz conditional *says*: it does nothing to illuminate the *nature* of the intermediate values.² It seems, however, that 'values' for Field are just apparatus involved in enabling us to *specify* which inferences are valid: they do not in any way *elucidate* their validity. Don't we want more from a genuine 'semantics' than an algebraic device for cataloguing validities?

However, let that pass for the moment. Let's now consider on its own terms the prospects for a many-valued Łukasiewicz logic with its non-standard conditional as a logic for preserving a naive truth-theory. By repetition of the argument which showed

²Field does add,

In any case, the semantics doesn't commit to any value that represents the division between the true and the untrue; a large part of the "fuzziness" of 'true' can be thought of as a "fuzziness" as to which values are sufficiently high to count as truth.

Eh? Where on earth has the idea of some fuzziness in *true* come from? At this stage, this is just unexplained arm-waving on Field's part.

that a three-valued logic won't do, no finite-valued logic will do either. So we need to go infinite valued. There would initially seem to be choices here: we could (as far as arguments so far tell us) take the values to be rationals q where $0 \leq q \leq 1$ and q has the form $k/2^n$; or all rationals in the interval; or all real. Or come to that, all surreals! What's to chose?

Field plumps for a continuum-valued logic, he says "to be safe", but in fact the choice is going to matter for the official proof of his ensuing main result. What's the prospect of getting a paradox-avoiding framework based on such a continuum-valued logic?

4 The Restall and Hájek/Paris/Shepherdson theorems

In 1992, Greg Restall wrote an elegant paper 'Arithmetic and truth in Łukasiewicz's infinitely value logic'.³ The headline news is this.

Extend the continuum-valued logic in the obvious way to deal with quantifiers. (So if we are dealing with arithmetic, so that quantifiers run over the numbers, then the idea is that $|\forall x\varphi(x)|$ comes out as the infimum of the values of $|\varphi(n)|$. Dually for existentials). Then take a natural axiomatization of first-order Peano arithmetic with an induction rule. Then the set of arithmetic consequences of these axioms is the classical set because excluded middle can be shown to hold for the arithmetical sentences built up from arithmetical identities using connectives and quantifiers. Now add a truth-predicate T and all the T biconditionals for this predicate. In a classical setting we'd be in trouble: diagonalization would give us a Liar sentence λ , such that $\lambda \leftrightarrow \neg T(\ulcorner \lambda \urcorner)$, in flat classical contradiction with the instance of the T -schema for λ . Łukasiewicz's hobbled conditional preserves from contradiction. But Restall shows that things are almost as bad: we don't have inconsistency, we have ω -inconsistency! So adding the truth-predicate T to PA with a Łukasiewicz logic kills off arithmetic-as-we-intended-it!

That's bad. And in *JSL* 2000, Hájek, Paris and Shepherdson show there's worse: if we allow in addition that induction holds not just for one-place predicates of the original language of PA but for predicates of the whole language, possibly embedding T , and also want T to commute with connectives (a step to having φ and $T(\ulcorner \varphi \urcorner)$ intersubstitute), then the resulting theory is inconsistent.

We won't prove the theorems, but as these things go they are straightforward.

5 What can we save from the wreckage?

Even if you stuck with it earlier, you might think at *this* point we should certainly give up on the idea of using a Łukasiewicz logic to get a non-standard conditional as a setting for a naive truth theory!

Well, you could blunt the force of the second theorem I just mentioned by arguing that we shouldn't use a truth-predicate in induction. Since the second order axiom applies to any genuine property, this is to say that the truth-predicate doesn't express a genuine property – which is what a disquotationalist *should* say. But Restall's result, you might well think, is bad enough.

And after all, remember, (a) the Łukasiewicz conditional hasn't yet been particularly well motivated (Field does nothing in this chapter to defend it or explain its supposed intrinsic attractions), (b) we haven't yet the foggiest idea about what the 'intermediate'

³See <http://consequently.org/papers/arithluk.pdf>

values in an infinite-valued Lukasiewicz logic might signify (if we are after more than an algebraic device for cataloguing validities), and (c) we’ve now a knock-down demonstration that it doesn’t work in even the simplest case of adding a truth-predicate to arithmetic. (Do the words ‘degenerating research programme’ spring to mind?)

But Field still wants to rescue something from the wreckage. He observes that the problematic results require that the Lukasiewicz logic is extended to cover quantifiers in the natural way. Suppose, then, that we stick inside a quantifier-free framework (lacking definite descriptions too). If we have some suitable apparatus for self-reference, predicates and identity, we can *still* get some paradoxes.

Field now proves the key result of this chapter. Unfortunately, his usual poor standards of exposition apply. Here’s the opening of his technical appendix where you’d at least expect him to be trying to be accurate:

Recall that S is a language without quantifiers or definite descriptions, but has a primitive device of self-reference. The theorem ... is: Let M be any classical ω -model of the ‘True’-free part of the language S

But languages don’t have models; it’s *theories* that do.

In the main text, Field at least gets round to mentioning a theory T built in the language S before stating his theorem, but the theorem as he states it falls back to talking about models of S rather than of T . But clearing things up, here’s the idea.

Suppose we start with a first-order language S , which is quantifier-free and also description-free, and which has some protosyntactic apparatus – names for each symbol of its alphabet, a binary symbol for concatenation, a *name* function that maps expressions to their structural-descriptions, and a *substitution* function which maps a formula with the free variable ‘ x ’ and a name e to the result of substituting e for the variable.

Let M be a classical ω -model [not for S but] for some set of S -sentences. Let T be the set of sentences true in that model.

Question: Can we consistently augment T to T^+ , by expanding our theory’s language with a truth-predicate T , and adding the naive axioms for truth, i.e. all instances of the T -schema, and closing under the principle that $\mathsf{T}(\langle\varphi\rangle)$ intersubstitutes for φ (where $\langle\varphi\rangle$ is the structural description for φ)?

Answer: Yes! There is a continuum-valued model M^+ for the full T^+ in which

1. $\mathsf{T}(\langle\varphi\rangle)$ and φ have the same value in M^+ , for any sentence φ in the language (including those that contain ‘ T ’), and
2. If φ is a sentence not containing ‘ T ’, it has value 1 in M^+ if it is true in M and 0 if it is false in M .

Ok, that is cute and original to Field (note, the proof as stated does call on the assumption that there is a continuum of values). But what does it show us?

It shows that by radically hobbling our expressive resources, we *can* produce some *consistent* quantifier-free theories, where the truth-predicate behaves naively and where we can construct Liar-style sentences and Curry-style sentences and the like without disaster (they just get an intermediate value), while all the non-semantic stuff stays unaffected. Nice!

But the gloss that Field puts on this is *very* extravagant: he says “It is worth noting that the [theorem] covers an immense variety of standard paradoxes”. Well, it covers the quantifier-free analogues of a variety of standard paradoxes. But if the solution can’t be

carried over to their immediate next-door neighbours, to allow us e.g. to add a naive truth-predicate to (quantified) arithmetic, then there's an immense variety of standard paradoxes it doesn't cover. So what does that show us?

Compare: if a strategy that seems to deal with the Liar is immediately in trouble with the Strengthened Liar, most logicians think the strategy is in trouble! The same sort of thought seems to apply in spades here – if a strategy that seems to deal with the Liar for the special case of formulations in a hobbled quantifier free-language blows up as soon as we try to apply it to Liars in a quantified language, you'd have thought that the strategy is hopeless. What Field says, however, is:

The clear inadequacy of the continuum-valued semantics for languages with quantifiers should not blind us to its virtues in a quantifier-free context. Indeed, one might well hope that some simple modification of it would work for languages with quantifiers. In fact, this does not seem to be the case: major revisions in the approach seem to be required. Even so, some of the features of the quantifier-free continuum-valued approach will in fact carry over to more adequate solutions.

Well, we will just have to see