THE REVISION OF LOGICAL FLAWS

Hartry Field
Silver Professor of Philosophy

January 2004

The Revision of Logical Flaws

Can there be rational reasons for revising our most fundamental logical rules (for instance, the rule of inferring from *A* to *A* or *B*; or the rule of inferring from *A or B* and *not A* to *B*)?

I think there can.

Nonetheless, the supposition that there can be rational reasons for such revisions gives rise to a wide variety of puzzles, which have led some philosophers to think that rational revision of logic is impossible. I'll start by briefly mentioning two of them, without going into any detail. (The main body of the paper will be an argument that despite the puzzles, rational revision of logic is possible, and on *this* I *will* provide some details.)

A first puzzle about rational revision of logic is this: any serious evaluation of anything requires logical reasoning; so in evaluating a proposed revision of logic, you need to use logic. But if you use a logic in the evaluation process, won't it automatically rule against all competing logics?

One strategy for responding to this would be to argue that a principle of logic can rationally be revised by an evaluation that uses only other principles of logic. (If that's the best we could do, then if there are parts of logic that must

be used in the evaluation of any principle of logic, *those parts of logic* must be immune from revision.) A better idea, I think, is to try to give an account of reasoning which would allow a logical principle to appear in an evaluation that undermined itself. But to say that is to say very little: an account of just *how* a logical principle can be used in an evaluation that undermines itself is needed. Still, there seems to at least be room for such an account: it would presumably work by appealing to methodological principles that go beyond the logic itself.

There is a variation of the puzzle that is harder. For there is a *prima facie* compelling argument that it is impossible to rationally revise *one's most fundamental methodology*: a methodology that advised us under some circumstances not to follow its own dictate p, but to follow instead the dictate p* of some other methodology, would simply be inconsistent: it would recommend both p and p*. It seems to follow that any consideration for rationally revising a methodology M must use methodological principles not included in M, so that if M is rationally revisable it isn't fundamental. If this argument is right, then the only way to allow for the rational revision of logic is to separate logic from fundamental methodology, in a way that seems to go against the traditional connection between logic and the laws of rational thought.

Despite these worries, there is at least one very strong reason for thinking that logic is rationally revisable. That's what I will discuss in this paper. Puzzles like those alluded to in the previous paragraph are serious, and I hope will be a focus of intensive work in epistemology over the next few years; the goal of

the present paper is simply to argue that they need to be solved, because logic *is* rationally revisable.

I believe that the best rational case for the revision of our logical practices is based on the existence of certain paradoxes. The simplest ones have been known for thousands of years, but have never been adequately resolved. All these paradoxes (not just the simplest ones) have been studied with much greater intensity in the last thirty years than ever before. In my opinion we are finally approaching a resolution of them, a resolution that requires a weakening of standard logical rules.

The very simplest such paradox (essentially due to the Greek Epimenides) is the "Liar paradox." It is based on the fact that there are sentences that assert their own untruth. Here is one way that this might naturally come about.

Suppose I see several sentences written on the blackboard next door, and I can see that they have an unacceptable conclusion; this motivates me to write on my own blackboard "Not all of the sentences on the blackboard in Room 503 are true." There's no doubt that what I've just written is meaningful. It's meaningful even if I've mixed up the room numbers: if, for instance, Room 503 is actually on the floor below and has only true sentences on it, then by my mistake about room numbers I've inadvertently said something false. But suppose that my room mix-up was different: suppose room 503 is the very room

in which I write. And suppose that everything else I've written on my blackboard is uncontroversially true. Given these facts, the sentence I've written in effect asserts its own untruth.

It is common, and harmless, to simplify the discussion of such paradoxes, by imagining sentences that directly assert their own untruths, by containing names of themselves: sentence L (the Liar sentence) is of form

L is not true,

where *L* is a name of that very sentence. Such direct self-reference may seem suspicious. But there is no remotely promising way to deny the possibility of at least indirect self-reference, like that discussed in the previous paragraph. (And anyway, there are paradoxes analogous to the Liar that avoid self-reference entirely.) Once you have the indirect self-reference you get the paradox; direct self-reference just allows it to be discussed more easily.

OK, so why is the Liar sentence (L, i.e. "*L* is not true") paradoxical? Well, suppose L is true. If "Napoleon is short" is true, Napoleon (the thing named by "Napoleon") is short; so if "*L* is not true" is true, L is not true. That is, if L is true, it is not true. The supposition that it's true leads to the conclusion that it is also not true, and thus lands us in a contradiction.

So far, no paradox: it just seems that it is contradictory to suppose L to be true. But the paradox arises because it is also contradictory to suppose L not to be true! For if L is not true, then "*L* is not true" must be true; that is, I must

be true. The supposition that L is not true leads to the conclusion that it is also true, and thus lands us in a contradiction.

In sum: the supposition that L is true leads to contradiction, and so does the supposition that L is not true. We're in trouble either way, and that's the paradox.

What responses are possible? One possibility is to accept the contradiction: L is both true and not true! This position (called *dialetheism*) is difficult to accept. The only way it can even be contemplated is if it is accompanied by a proposal to change standard logic. The reason is that in standard logic, contradictions imply everything: if we accept that L is both true and not true, then according to standard logic we will have to accept that the earth is flat (as well as not flat), and so forth, which is a position too silly to contemplate. [How does the contradiction that L is both true and not true imply that the earth is flat, in standard logic? It's easy: from "L is true" we can derive "Either L is true or the earth is flat"; from this and "L is not true" we can derive that the earth is flat.] The "contradictions can be true" option doesn't make even minimal sense except in the context of a revision of logic.

Is it possible to avoid revising logic while finding a way out of the paradoxes? It is, if we don't count the principles of truth as part of logic. For the paradoxical argument rested not only on logical principles narrowly conceived, but also on the assumption that "'Snow is white' is true" is equivalent to "Snow is white," and similarly for every sentence of our language other than "Snow

is white." We can avoid a revision of logic by revising truth theory instead, and it is usually thought that we should do so. But there are serious obstacles to finding a satisfactory solution of this sort.

If we are to preserve standard logic, it looks like we must either block the argument that "L is true" leads to contradiction or block the argument that "L is not true" leads to contradiction. (Later I'll consider the question of whether there might be a third option within standard logic.) Of these, the second option has proved by far the more popular: for instance, for many people the first reaction to the Liar paradox is to say that L is neither true nor false; but this implies that L is not true, so it requires the second option. But defending this option requires that we assert the untruth of many of our own assertions. For instance, we assert "L is not true"; but that is equivalent to L, so we are asserting L while also asserting it untrue. And this is decidedly odd: why assert what you take not to be true? It can be shown that the oddity is not confined to a few pathological examples: if we develop a theory of truth that accords with this second option, we will have to declare that *the fundamental principles of that very theory of truth* are not true, i.e. we will have to declare untrue the very theory of truth that we are advocating. That does not seem to me a happy outcome.

To be fair, I should note that proponents of the second option often try to evade the difficulty I've just mentioned by supposing that the word "true" is radically ambiguous: we need to suppose that there is $truth_0$, $truth_1$, $truth_2$, and so on. The principles of the theory of $truth_0$, can't be supposed $true_0$, but at

least they're true$_1$; the principles of the theory of truth$_1$, can't be supposed true$_1$, but at least they're true$_2$; and so on. I don't think this proposal can be made satisfactory, but will not discuss it further.

The other obvious option in standard logic that I mentioned — accepting that L is true, and blocking the argument that it leads to contradiction — fares no better, and to my knowledge has never been seriously advocated.

There is however a third option, which at least in some sense manages to preserve standard logic. The idea is to grant that "L is true" leads to contradiction, and also grant that "L is not true" leads to contradiction. But we refuse to conclude that "Either L is true or L is not true" leads to contradiction, and hence we feel free to accept that!

Technically, the inference from the contradictoriness of A and the contradictoriness of B to the contradictoriness of A or B isn't an *inference* of standard logic, it's a *meta-inference*, so many people count this as not a real violation of standard logic. And in some ways it leads to a much more satisfactory theory of truth than the first two options. But to my mind there's something deeply suspicious about it: it seems to me that if A leads to trouble and B leads to trouble, then you're in trouble if you accept A or B, even if you haven't made your mind up which of A and B to accept.

This however naturally suggests another alternative, one which gives up all pretense of keeping standard logic. As with the third option, we grant

that "*L* is true" and "*L* is not true" each lead to contradiction, but this time we accept the natural conclusion that "Either *L* is true or *L* is not true" leads to contradiction. How then do we escape accepting a contradiction? By refusing to assume that either *L* is true or *L* is not true.[1] This is definitely a weakening of standard logic, for in standard logic anything of form "*A or not-A*" is a logical truth — it follows from anything. That's called "the law of excluded middle." The suggestion is that we abandon this basic principle of standard logic.

The idea of doing this is not new, though those who have advocated it have not always clearly distinguished it from the approach that keeps standard logic but says of certain sentences of form *A or not A* that they are not true. (I know that these theories don't *sound* distinct, but they are. The reason they are distinct is that the standard-logic theories give up on the equivalence between

(1) *A or not A*

and

(2) "*A or not A*" is true;

they accept the former while denying the latter. It is just this sort of weirdness that we can avoid by weakening standard logic: weakening the logic allows us to consistently regard (1) and (2) as equivalent to each other, and we can reject both for certain pathological sentences like *L*.) But though the *idea* isn't new, it was never adequately worked out until about two years ago. In particular, until then there were no known logics that allowed us to consistently accept all of the standard postulates governing the notion of truth without endorsing contradictions.

One standard postulate about truth is that a claim of form *True("A")* is always equivalent to the corresponding claim *A*. This postulate is conspicuously violated by theories that keep standard logic, as the previous paragraph illustrates dramatically; but in an extremely influential paper in 1975, Saul Kripke demonstrated that the postulate can be maintained in a logic without excluded middle. But there is a second standard postulate about truth, that claims of form

True("*A*") if and only if *A*

hold or are true. This postulate might seem to follow from the first: after all, claims of form *A if and only if A* are surely true, so it follows from the first postulate that the corresponding claims *True("A") if and only if A* are true. The problem with this reasoning is that while claims of form *A if and only if A* are true in standard logic, they aren't true in the very weak logic that Kripke used!

I think this is a serious defect in Kripke's theory. And I think that what we should hope for is a theory that is like Kripke's but that involves a much richer logic. Unfortunately there are very serious technical obstacles to overcome in carrying this out, and until recently all attempts to do so (e.g. using "fuzzy logic") turned out to be inconsistent. It has often been said in the recent literature that the idea is pretty much hopeless, and that the only real possibility if we are to save the standard postulates about truth is the "dialetheic" one of accepting contradictions like "*L* is true and not true," though in a logic that prevents the derivation from this of absurdities like "The earth is flat."

However, in the last couple of years I've found a couple of ways of carrying out the program of consistently adhering to the standard theory of truth (both postulates of it), without generating contradictions, in a stronger logic than the one Kripke used. The first of these ways was not offered as a serious proposal, just a way of showing (contrary to received skepticism) that it could be done, and giving some idea what logics in which it was done might be like. The second I do take as a serious proposal. But I suspect that other serious proposals are possible, and one thing I look forward to in the next 5 to 10 years is the development of more of them and a debate as to which is best.

Another thing that I expect to see in the next 5 to 10 years is a sustained debate between views of the sort I've been suggesting and other views of the paradoxes: both views within standard logic of the various types I've mentioned, and also "dialetheic" views that accept contradictions but prevent them from spreading everywhere. In the last few years these dialetheic views have been widely discussed; most people find them repellent, but it has turned out to be surprisingly difficult to argue against them. I myself suspect that they inevitably do less well on the paradoxes than views of the type mentioned in the previous paragraph do, and that this is connected with general reasons against dialetheic views; but this needs detailed arguing.

If these are indeed major topics in coming years, then it will vindicate a

claim I made in my opening remarks: rational debate about logical principles is possible, and of importance. (The importance is far greater than I have indicated: a good case can be made that the same weakening of logic that is important for avoiding the paradoxes has great importance elsewhere as well, especially in allowing for the coherence of claims that certain issues are "non-factual." But that's a whole other story.)

The revision of logic to handle the paradoxes will also, I hope, give a nice case study for how the revision of logic is to succeed. Up to now, philosophical discussions of the revision of logic have largely been focused on three examples: "intuitionist logic," "quantum logic" and "relevance logic." These are all extraordinarily poor examples for general discussion, for a variety of reasons, among them being that in none of these cases has there been any serious proposal of what it would be like to adopt it as one's all-purpose logic. In the case of the logics I'm suggesting, there is little if any difficulty in thinking of them as all-purpose logics, so they make excellent case studies for how logical revision is to proceed.

I began the paper with some vague worries about how revision of logic is possible. There is much discussion in the literature about such worries as these, and I expect and hope that in coming years the worries will be sharpened and their investigation deepened. It is my *bet* — I don't know how to fully substantiate it yet — that the worries are based on basic assumptions about epistemology the overturning of which would have revolutionary and liberating

effects not just in the epistemology of logic but in epistemology more generally. It is often the case in philosophy that the solution to a small technical problem ultimately has widespread ramifications elsewhere, and perhaps it isn't unreasonable to suspect that the solution to paradoxes involving such a fundamental notion as truth might turn out to be one example.

---

1 But we also don't accept that it is *not* the case either L is true or L is not true; for on almost any logic, that would imply that L is not true and hence lead to paradox.