

A Debate on Hyperintensionalism*

Kit Fine (NYU)

Timothy Williamson (Oxford)

Opening Speeches

KIT FINE

It's a very great pleasure to be here. I should begin by saying that I'm a great fan of intensionalism. The first 20 years of my career were devoted, actually, to philosophical and technical problems in modal logic. But over time, I became disillusioned. I thought that, good as it was, it wasn't good enough.

Now, this is a vast topic and so, clearly, in 15 minutes it's very limited what I can do. What I will do is explain how some of the reasons philosophers on the other side might have had for not liking hyperintensionalism are not actually good reasons. These not-so-good reasons, I'm not attributing to Tim, of course. But I do think that they are perhaps in the back of mind for many people who are disinclined to go for a hyperintensionalist approach. In fact, they say things that suggest as much.

I should begin by saying what I mean by intensionalism and hyperintensionalism. Actually, what I want to do is to distinguish – and it's really central to my argument – between two forms of intensionalism: intensionalism in regard to reality, and intensionalism in regard to language.

So, what is intensionalism in regard to the world? It says things like: if two facts necessarily co-obtain, they're the same fact – if there's no modal difference in when the facts obtain, then there's no difference at all in the facts themselves; similarly for properties – if two properties are necessarily co-extensive, then those properties are the same. If we think of these facts or properties as occupants of the world, what we're saying is that reality is intensional.

We can also put it in terms of possible worlds rather than in terms of modality. We can say that: if two facts obtain in the same possible worlds, then they're the same; if two properties have the same extension in each possible world, then they are the same. So for that reason, I'm going to call this a *worldly* conception of reality. Intensionalists have this worldly conception: no distinctions in reality without a modal distinction.

*Hosted March 3rd, 2026 by the Oxford Socratic Society. Transcribed by Oak Hu and published by [The Philosophers' Magazine](#) with permission of the hosts and the authors.

So that's intensionalism in regard to reality. What's intensionalism in regard to language? Well, take two sentences and suppose that they're necessarily equivalent: necessarily one is true just in case the other is true. Then the intensionalist wants to say that there's no semantic difference between these two sentences.

Now, what does that mean, 'no semantic difference'? At least one *implication* of there being no semantic difference is this: that, in a well-constructed language, if we substituted the one sentence for the other in a given context, the two resulting sentences would be necessarily equivalent. So in that sense, we cannot, within a well-constructed language, distinguish the behavior of these two sentences. If we use one in a construction, or we use the other, the resulting constructions will be necessarily equivalent. Similarly for predicates. So if predicates are necessarily true of the same things, we can always substitute one for the other *salva* not just *veritate*, but *salva* modal status.

I think – and this is an error I'm going to try to expose – it's often been thought that if your concern is to construct a language that describes reality, and if you have an intensionalist conception of reality, then you should have an intensionalist conception of the language by which that reality is described. You might think: how come you make distinctions in the language that don't bear on reality? The way you can put it is: if reality is worldly, then language should be worldly, too.

What that means is any language that seems to be attempting to describe the world should be intensional. That applies, I take it, to many constructions, like the counterfactual construction. When we say 'if we were to strike the match, it would light', this is something about the world. Or again, you might think that metaphysics should principally be concerned to describe the world. So this means that the language we use in making metaphysical claims should also be intensional, according to this view.

Now, I do not accept an intensional conception even of reality. I don't think reality is worldly – it's unworldly, which is clearly a better thing. But for the sake of argument, I'm going to accept that reality is worldly; I'm going to go along with the suggestion, for example, that facts that necessarily co-obtain are the same.

What I'm going to try to argue is: that doesn't mean the language which you use to describe reality should also be worldly. Intensionality in regard to reality doesn't demand intensionality in regard to language.

The issue turns on how language connects with reality. We've got a sentence; suppose it's true, and something makes it true. What makes it true? There are two views you can have. One is what you might call a *one-to-one* conception, and the other I'm going to call the *one-to-many*

conception. On the one-to-one conception you can say, given any true sentence, there's just going to be one thing that does make – or is even capable of making – the sentence true. Whereas on the other view we want to say: no, I don't think there's a one-to-one correspondence here; it's a one-to-many correspondence – given a sentence, there can be many different things in the world that make it true.

I can illustrate this with a disjunction. I say 'it's rainy or windy'. On the one-to-one conception, the truth of that sentence just relates to the worldly fact that it's rainy or windy – that single fact. That's the correspondence in that case. But on the one-to-many conception, that sentence – at least at the simplest level – is capable of relating to two facts. If it's true, it could be true either in virtue of the fact that it's rainy, or in virtue of the fact that it's windy. So the correspondence here now is not between the single sentence and the single fact, but the single sentence and many facts – the many facts that are capable of making it true.

Now, suppose you have the one-to-many conception of how language relates to reality, and we're granting that reality is worldly. Then, there's room to have a conception of language which is not intensional. Again, I emphasize that this language is describing reality! It's describing reality in the worldly sense of the term.

Let me introduce you to Johnny. He likes ice cream; he also likes chocolate. So look at the sentence 'Johnny is having ice cream'. Now look at the sentence, 'Johnny's having ice cream, or Johnny's having ice cream and chocolate'. One's of the form 'P', the other's of the form 'P, or P and Q'. These two sentences are necessarily equivalent. So, in an intensional language, we wouldn't discriminate between them. But on the one-to-many conception, we can distinguish between them because we want to say: look, the sentence 'Johnny is having ice cream' simply corresponds to a single fact; but 'Johnny's having ice cream, or Johnny's having ice cream and chocolate' corresponds to two possible facts – one is his just having ice cream, the other is his having ice cream and chocolate.

I hasten to add that when we have the 'P' and the 'P, or P and Q', in both cases we're describing the worldly reality. But we're connecting the language to that worldly reality in a different way than under the one-to-one correspondence.

What I've said is not that we *should* have a non-intensional language to describe a worldly reality, but that this should be on the cards: even if someone thinks that reality is worldly, they should be open to having a hyperintensional language – one that distinguishes, for example, between 'P' and 'P, or P and Q'. Now, there's a question that is: well, okay, I'm saying this should be an option but should we do it?

It depends what your purposes are. One purpose is to understand natural language. How is natural language working? Especially those parts of the language which we think are related directly to the world. But another question has to do with formal languages. You might want to construct or regiment natural language for very many different purposes.

I think that there's something to be said in both of these cases for adopting a hyperintensionalist stance. I'm going to have to be very brief here, but go back to Johnny. I say to Johnny, 'You may have ice cream'. Johnny looks to Tim and says 'Look, having ice cream is the same as having ice cream or ice cream and chocolate, right? They're classically equivalent. So when Kit said that, then it's the same as saying that I may have ice cream or chocolate and ice cream. If you give me permission to do X or Y, then I'm allowed to do either. So I'm allowed to have chocolate and ice cream'. Of course, we can substitute anything else for chocolate here.

We might say against Johnny: 'No! Don't think that those two sentences – 'P' and 'P, or P and Q' – are semantically equivalent'. So we have here a very simple explanation of why this reasoning of Johnny breaks down. I'm not saying there isn't an argument on the other side – this is a big issue – but all I want to say is there are some very straightforward explanations we can give of certain linguistic phenomena in natural language.

I'm just going to briefly discuss this issue of regimentation of formal language. I'm interested, as is Tim, in formalizing and regimenting our discourse concerning counterfactuals. Now there's an amazing fact about the present approach, which is this.

There are two things we might want of the logic of counterfactuals. One is the freedom to express embedded counterfactuals: counterfactuals of the form 'if P, then if Q were to be the case, then R would be the case' – I'm talking about embedding to the right. Another thing we might want is that every counterfactual sentence, even if it involves embedding, should be equivalent to an unembedded counterfactual. So when I say 'if P were to be the case, then if Q were to be the case, then R would be the case', I want that to be equivalent to a straight counterfactual, no embedding. The obvious candidate is 'if P and Q were to be the case, then R would be the case'. Because, after all, our practical interest in counterfactuals when making decisions is in these straight, unembedded counterfactuals. But it's still useful to be able to express these embedded counterfactuals. When someone states a counterfactual, and I say 'but what if this were the case, then what would the status of your counterfactual be?' – I want the freedom to be able to say things like that.

Now, if you adopt an intensional approach, this is disastrous. You can't

do it. You can't have both the freedom to express embedded counterfactuals and the equivalence to unembedded counterfactuals.

But under this alternative hyperintensional approach, you can do this. People often say: look, hyperintensionalism is so complicated – you have this complicated logic. But in some ways, the complication in logic can earn you a kind of simplification in other aspects of your logic. Thank you.

TIMOTHY WILLIAMSON

I'm going to be defending an intensionalist approach: one that's modeled on the kind of framework of possible worlds that you get in modal logic, and that's being used both in order to understand what properties, relations, states of affairs and so on are, and as a framework for the semantics of the language.

So it's saying, for example, all that matters about, let's say, a property is which things have it in which worlds, and if you have properties that are the same in that respect, then they're just the same property.

I'm not so concerned about exactly what the setup of the language is because, for example, it's perfectly fine to have a language in which you can have quotations: you can form a quotation of a word as a way of referring to that word, and you can do that even though your aim is to describe reality or the world, because language is *part* of that, and you're using these quotations to talk about that part of the world.

But then quotation is not intensional, because if you substitute, for example, two things which are necessarily equivalent but aren't in exactly the same words, then the quotations will refer to different things. So I'm not going to be bothering about those. But I will be talking about the things that are not so obviously just specific to language.

The great advantage of this framework is that it's very simple; it's very well understood; and it's known to be consistent – whereas when you go hyperintensional, all bets are off. And if you go for a really, really full-blooded hyperintensionalism, then you're making so many distinctions that you actually get a kind of paradox known as the Russell–Myhill paradox – which is related to Russell's more famous paradox – but it's to do with propositions. That's not a killer by itself for hyperintensionalism because you can introduce all kinds of distinctions and restrictions to prevent the Russell–Myhill paradox from occurring, but you do have to complicate your theory quite a lot in order to do that, and reduce its power.

So that's one kind of advantage of intensionalism – Kit was mentioning the specific case of counterfactuals, but I'm not going to say anything

about that right now. Another one is that we really want to avoid projecting features of language onto the world. Let me explain how that can easily happen in the kind of arguments that people give for hyperintensionalism – this isn't specifically to do with the arguments Kit was giving just now.

For example, when people talk about truth, they sometimes say things like – this is actually an idea that goes back to Aristotle, so it has a good lineage – ‘the statement that grass is green is true because grass is green: that's correct; but it's not the case that grass is green because the statement that grass is green is true: that seems to be getting the information the wrong way round’.

That's a case where you have one ‘because’ statement which seems to be true, which is, roughly speaking, explaining truth in terms of how things are in the world; and another one that seems to be false.

That, if correct, would be an example for the hyperintensionalist case, because it would mean that ‘because’ was working in a hyperintensional way, because if you take the two statements on either side of the ‘because’, one is just ‘grass is green’, the other is ‘the statement that grass is green is true’; and those are actually necessarily equivalent, and so they would be expressing the same proposition.

On an intensionalist view, it shouldn't matter which order you put them in, in relation to the ‘because’, because it's actually the same proposition that is being expressed, assuming that when we use the word ‘because’, we're not talking about the words, but about the proposition.

Now, that seems actually like a pretty good argument for hyperintensionalism. But the trouble is that there are other arguments in a very similar spirit that aren't going to work, and where something is going wrong.

For example, take the following: somebody says ‘furze is as prickly as gorse because furze is gorse’. That seems a perfectly reasonable explanation; that seems like a true ‘because’ statement.

But if you were to say ‘furze is as prickly as gorse because furze is furze’, that seems like a complete *non-sequitur* and totally unhelpful: surely that can't be why furze is as prickly as gorse, just that furze is furze.

So it's very tempting to say, well, the second statement is false; it's not a correct ‘because’ claim. But, in fact, the words ‘furze’ and ‘gorse’ are just synonyms of each other. They're just different words for the same thing. There is no difference in meaning between them at all. It's just that they're different words. You could do the same trick with any pair of synonyms. So it can't really be that – given that the semantics is compositional, that the meaning of the whole sentence is built up from the meaning of the parts – there is any difference in meaning between these two sentences.

There's just this very superficial difference at the level of the words being used.

In that case, the judgments that seem to lead to hyperintensionalism can't be right, because they're assigning different truth values to sentences which have the same meaning.

If one thinks about what's going on – why are we misjudging the truth values of these sentences? – here's a natural explanation, which is that we're using a certain kind of heuristic, which is kind of rule of thumb: a quick, easy way that we can judge things, but isn't guaranteed to be 100% reliable. The kind of rule of thumb that we're using is to judge the truth value of a statement of the form 'A because B' by how good the sentence 'B' is as an answer to the question 'why A?'

The first sentence – 'furze is as prickly as gorse because furze is gorse' – corresponds to the kind of question-answer pair where somebody asks 'why is furze as prickly as gorse?', and then somebody tells them 'furze is gorse'. That seems totally fine: that's exactly the answer they need.

Whereas the second sentence – 'furze is as prickly as gorse because furze is furze' – corresponds to an exchange where somebody asks the same question – 'why is furze as prickly as gorse?' – and then the answer they're given is 'oh, well furze is furze'. That's a lousy answer. It's not in the least helpful to the person who's asked the question. Because it's a lousy answer to the question, we judge that the corresponding 'because' statement is false.

But that's just a case where what we're doing is projecting the explanatory value of the answer onto the truth value of the corresponding 'because' claim. Explanatory value is sensitive to all kinds of things other than the actual content or meaning of the sentences: it's sensitive to how perspicuously things were said, and so on.

So that's the sort of way in which, if we just rely on our unreflective judgments, we're vulnerable to the heuristics that – probably without realizing it – we're relying on to make these judgments, and those heuristics are not 100% reliable. In effect, in a case like this, what's going on is that what is fundamentally just the difference between two words is being, as it were, projected onto reality. We're thinking that one of these 'because' statements is correctly describing the world, while the other isn't, when, in fact, there's no neat difference in meaning between them at all.

That is the sort of thing that intensionalism doesn't allow, but it's behind many of the kinds of examples that have inspired people to be hyperintensionalists. There's a severe danger within hyperintensionalism of just, roughly speaking, projecting the structure of our language onto the structure of our world.

One other example: there's a tendency for hyperintensionalists to talk about aboutness as a feature not just of sentences or discourse, but of the propositions and properties and, as it were, the things out there that the words are supposed to be expressing.

Here's an example. Somebody might say that the proposition that Mary is either guilty or not guilty – which is supposed to be the thing expressed by the sentence – is about Mary; whereas if you take the proposition that John is either guilty or not guilty, that one is about John and not about Mary. If you say that, that means that they have to be different propositions: one of them is about Mary but not John, the other is about John but not Mary.

But by intensionalist standards, they're the very same proposition – I'm assuming now classical logic – because they're both just logical truths. They're both true in all possible worlds, and that's enough for them to be the same proposition.

In a case like that, we should be very wary of thinking of aboutness as a feature of the states of affairs and properties and relations that we're talking about, because aboutness is fundamentally a property of discourse: 'What are you talking about?'

This sort of hyperintensional urge is getting people to judge that there's a difference in the states of affairs because they differ in aboutness, which is projecting the linguistic property of being about a certain subject matter onto the things that the words are expressing. That's the sort of temptation that we need to guard against. Intensionalism is an extremely good way of guarding against that kind of projection of linguistic properties onto the world.

Rebuttals

KIT FINE

I'll work backwards. Let's talk about Mary and John being guilty. One sentence is 'Mary is guilty or not guilty', the other is 'John is guilty or not guilty'. Now, according to Tim, thinking of one being about John, the other being about Mary, involves some sort of projection of language onto reality. However, this is not how it looks from the point of view I was trying to lay out before. Just take the actual disjuncts. 'Mary is guilty', it's just made true by a single fact, namely, Mary's being guilty. If you like, think of the fact as a true proposition, it doesn't really matter. We can, in that particular case, identify Mary as a subject of that proposition, I take

it – there's not the same difficulty in saying that that proposition is about Mary. Similarly for Mary not being guilty. So, if this proposition that Mary is guilty or not guilty can be made true these different ways, and these two different ways are in an unproblematic way about Mary, not about John, then we can say that the proposition is about Mary. So, again, it's a matter of whether you have the one-to-one correspondence, we just think of the sentence as relating to this single proposition, or of the sentences relating to different facts.

Now, Tim gave the example of 'because' statements, and again, he thinks that involves projection of language. What I want to say is: the kind of case I was considering can't properly be seen as involving projection of language. When we project language, we confuse how something is said with what is said – that's the confusion. That's not what's happening with my example of 'P' versus 'P, or P and Q'. You could think of it that way, but you don't have to think of it this way. These two sentences are relating to different facts, facts in the world. So this just has to do with the connection between language and the facts. There's simply no reason to think in this particular case that anything like projection is going on. I mean, that's a possible hypothesis, but you don't have to think that. I think for this reason that the case is fundamentally different.

'Furze' and 'gorse'. This is complicated. I am a semantic relationist. I believe that two expressions can have the same meaning. I've got 'furze' and 'gorse', on the one hand – I'm talking about the words – and we've got 'furze' and 'furze', on the other hand. So there's 'furze' and 'gorse', that pair of expressions; there's 'furze' and 'furze', that other pair of expressions. 'Furze' means the same as 'furze'; 'gorse' means the same as 'furze'. I grant that, yes! Does 'furze' and 'gorse' mean the same as 'furze' and 'furze'? I say no, because there are semantical relations between pairs of expressions that cannot be understood in terms of the meanings of those expressions. With 'furze' and 'furze', it's a semantic requirement that they be coextensive. There's no semantic requirement that 'furze' and 'gorse' be coextensive, even though in fact they mean the same. The semantic facts behave this way it seems to me. Anyway, this is a possible view, okay? There are these relational differences in meaning, not reducible to intrinsic aspects. And that would be enough to dispel the parallel.

I'll say a little bit more about this case. I'm very interested in having a properly regimented language, so we can express clearly what we want to say. I'm saying in this particular case of 'furze' and 'gorse', we're using a language in which there are these irreducible semantic relations. I don't like those languages for formal purposes. So I would not want to actually have a logic of 'because', or ground, that allowed this kind of case. What

one can do, it seems to me, in order to be clear about these matters, is just factor out how it is that language might be interfering with what you're saying. So I can say: take the thing X that is furze, take the thing Y that is gorse. Is X as prickly as Y because X is identical to Y? I'm taking those things, actually the properties, whatever they are. Well, okay. Presumably not, because they're the same thing, and so you're saying something's the same as itself because it's the same as itself. So we can say: look, this is not a good example of the kind of thing we have in mind when we're using language in this transparent way – when we're making clear we're talking about these things, and how we're talking about these things is not coming into it.

The truth case, where you say the proposition that it's raining is true because it's raining: that can't be cleaned up in that way. If someone said, look, I need to clean up the 'furze'-'gorse' example, fine! Let it be cleaned up; let's abandon it. The same considerations do not apply to this other case. I say: take the proposition P that it is raining; then P is true because P. That seems fine. So there's still room for someone to say: look, I accept this 'because' statement involving truth, even though I'm willing to jettison the other statement – because *there*, not here, there is a danger of not properly focusing on what we're talking about rather than the way we're talking about it.

TIMOTHY WILLIAMSON

I'll talk about a couple of things in what Kit has been saying. One of them is this idea of 'making true' – which he was using in his original presentation, and also now – as it were, what facts make some statement true.

There are a small number of cases where that idea makes fairly straightforward sense, or at least seems to. Those include disjunctions being made true by their disjuncts; conjunctions being made true by their conjuncts together; and existential claims like 'there is a pen on the table' being made true by their instances.

This is quite an old idea. You can already find Bertrand Russell talking about it, more than a century ago, in his lectures on logical atomism. The difficulty with it is to generalize it. One point at which the problem comes up is just as soon as you look at negative statements: what makes it true that it is not snowing? Russell was worried: is there such a thing as a negative fact out there in the world to make it true that it's not snowing; or is there some other fact which does obtain, which is positive, but is somehow incompatible with its snowing? But then you also have to postulate this relation of incompatibility.

You also get huge problems with universal claims. Take a negative one, like 'there are no dragons'. We're talking about the whole universe, maybe there are no dragons in the universe. The difficulty is, well, what makes that true? It seems it's hard to say what makes it true other than just the fact that there are no dragons. But then this idea of things being made true by something more basic isn't even working for universal claims.

The difficulty has been that it's an idea which seems to work for a very small number of initial cases, and then Russell didn't know how to generalise it, and I think that people who do truthmaker theory – the theory of making true – haven't really solved any of these problems. It's an idea which doesn't properly generalise. We should be very sceptical of it, because it seems to be just exploiting the features of a very small number of initial cases.

I should say something about the issue of embedded counterfactuals. If you have an intensional approach to counterfactual conditionals, as I do, it's not that there is any special problem in having counterfactual conditional sentences, with one embedded in another, and getting them to behave in sensible ways.

What we do have is a difficulty in expressing them in a straightforward way within natural language. I think there is a reason for that, which is that natural language – and this is not just English, but seems to be a cross-linguistic feature – expresses counterfactuality by what linguists call the fake past, which is a use of the past tense not really to be talking about the past time, but to be just distancing yourself from a possibility that is non-actual. You can see that in English with the word 'would', which is the standard way of expressing counterfactual claims. 'Would', there, is in fact the past tense of 'will', and you can see that there are quite straightforward uses of 'would' just in a past way.

For example, if you're talking about Napoleon going into exile after the battle of Waterloo, his ship sailing off, somebody at the time could say 'he will never see France again', and then a historian can say, in the past tense, 'he would never see France again', which is just the past tense of 'will'. We're using the same linguistic construction, but we're using it to talk about counterfactuality – that's a very widespread phenomenon in language. But it's difficult because embedding the past within the past isn't a straightforward thing to do. Of course, we've got the pluperfect, and so on, but it gets really messy. Giving clear expressions of embedded counterfactuals in natural language is quite tricky; but, the semantics themselves – the intensional semantics – make perfectly good sense of these. It is really just a problem with expressing them using the somewhat restricted apparatus of natural language.

Free Discussion

KIT FINE

Just about the last stuff on counterfactuals. I was actually interested in – I said it very briefly – embedding on the right. I wasn't interested in when the antecedent of the counterfactual was itself a counterfactual, I was interested in the case when the consequent of a counterfactual was a counterfactual. And, of course, there's no difficulty in saying that if Mary were to be at the party then, if Bill were also to be at the party then the party would be a failure. So we have no special difficulty understanding or saying those things.

My point was simply that there are two desiderata you might want to satisfy, which can be satisfied on the kind of semantic approach I was indicating – by the way, it's called truthmaker semantics, I didn't give it a name. The two desiderata can be satisfied under truth-making semantics, which has a non-classical logic, but cannot be satisfied under an intensional view. One is a kind of normal form requirement that every counterfactual – I'm not allowing embedding on the left, only on the right – should be equivalent, even if it involves all this embedding, to one that doesn't involve any embedding. So, for example, in the case of a counterfactual of the form 'if P would be the case, then if Q would be the case, then R would be the case', the thought would be: that's equivalent to 'if P and Q were to be the case, then R would be the case'. So, one requirement is that every complex counterfactual statement should be equivalent to a normal form in which there's no embedding; and secondly, that we do allow embedding on the right. Of course, if you don't allow it on the right the normal form theorem is trivial, but we do allow arbitrary embeddings on the right.

I was just amazed by this: that you can get that result within a truthmaker framework. I would have thought these are two very desirable things. First of all, our interest for the most part, especially when it comes to practical decision making, is just in these unembedded counterfactuals. But we also want to be able to embed on the right. Someone is wondering what to do and they say 'if this were the case, ...', you know, they go through this counterfactual; and I say 'but what if something else were the case?' – that is, I'm asking them now: look, in each of your counterfactuals, what would the status be if this particular antecedent were to hold? It's good to have that kind of flexibility. A similar point holds for conditional imperative statements as well: I say 'if it's raining, then, if it's windy, then take an umbrella and wear a coat' or something like that. So

the consequent of one conditional imperative can be another conditional imperative – a similar point holds there: that you can get this kind of simplification of these things.

The point really was that, forgetting about natural language, if we're trying to construct certain formal languages with certain desirable features, there may be something to be said for adopting an alternative to the intensional framework. There may be certain desiderata you want to satisfy which cannot be satisfied if you adopt it. That was simply my point.

TIMOTHY WILLIAMSON

I'm not sure why you're so confident of this constraint that we should be able to reduce the embedded counterfactual – I mean embedded on the right – to an unembedded one. It seems to me that you can say things like 'if this meeting hadn't been organised, then it would have been the case that if it had been organised, it would have happened more or less as it did'. The thing is, if you try to reduce that, then you just get an inconsistent antecedent, and so it's kind of trivial, whereas the original thing was non-trivial. It seems to me that there's a good reading of the original embedded one on which it might well be false, and so it's non-trivial.

I think what's going on is that we're making suppositions within suppositions, but it isn't compulsory to just pile up these suppositions together. We can make a supposition, and then within that, we can make another counterfactual supposition and then just evaluate that under the original one. We're not assuming that, as it were, the initial antecedent and the embedded antecedent will have to be true together. It seems to me those readings are available; the intensional approach doesn't prevent that from happening. Of course, in principle, it may be that we have some kind of default where we're assuming that this is to be interpreted with the antecedents piling up, but that isn't compulsory.

KIT FINE

Yes. Look, I think that other reading is available. My concern was not actually with natural language, it was with a regimented language designed to serve certain purposes. Maybe I could use a certain analogy here. Those of you who know classical propositional logic know that every formula of classical propositional logic has a disjunctive normal form. If the sentence letters are 'P', 'Q', 'R', then 'either P and Q and R, or not P and Q and R', or blah blah blah, so you have a disjunction of that sort. That's a wonderful result because, in a way, that disjunctive normal form perspicuously

expresses what the original formula was expressing.

But still, we don't want to just confine our attention to disjunctive normal forms. If this is the perspicuous way of saying it, why don't we just use disjunctive normal forms? Well, you assert a disjunctive normal form; I may want to deny what you say! Oh, well, what's the disjunctive normal form of that? That's not a straightforward question. So, we want the expressive freedom that arises from allowing these more complicated constructions. But we also find it desirable that there should be a perspicuous representation of what's being said.

So the thought is that we would like there to be a similar normal form theorem in the case of counterfactuals. Basically, we want to know what would happen if this were to happen. How would things be if this were to be the case, how would things be if that were to be the case, and so on and so forth. That is what corresponds to 'P' and 'Q' and 'R' or what have you. We then want every complex counterfactual statement to be equivalent to a disjunction of statements stating how things would be under different circumstances.

Our only interest at the end of the day might be in these unembedded counterfactuals – what would happen if this were to be so, where what would happen is itself not a counterfactual. But we might also want to have the expressive freedom to construct these arbitrary counterfactuals. So my point was simply – and I think this is an extraordinary result – that truth-making semantics allows you to do that, whereas it's not possible under an intensional approach. Now, of course, what that means is this reading that Tim was talking about would not be available. But it is not clear that that would be a loss.

TIMOTHY WILLIAMSON

From my perspective, the counterfactuals are simply modal statements. What the 'would' is expressing is a kind of very very localized sort of necessity, simply in the sense of 'in all relevant worlds' or something like that, where the relevance is contextually restricted. So when you're embedding the counterfactuals, you're embedding one necessity operator within another. As Kit very well knows, generally speaking when you're doing that the sentences that you get will not be equivalent to unembedded modal sentences, because as you pile up the modalities, as it were, each layer of modality may take you to further-off possible worlds. So the kind of normal form that Kit is talking about is available in a few modal logics, but very, very few. So it's an extremely special case, and I don't think we should really expect, as it were, the language somehow to be

designed in order to provide for this very, very special case.

KIT FINE

Since we're on this topic, why don't I just discuss the case of conditional imperatives as well? So, I say 'shut the door', and I say 'if there's a draft, shut the door'; so that's a conditional imperative. As I said, the consequent of a conditional imperative can itself be a conditional imperative. So I say 'if it's drafty, then if you're feeling cold, shut the door'. Now, when it comes to imperatives, a very natural thought, I think, is that we should be able to express the content of an imperative in terms of a disjunction of plans.

What I mean by a plan is something that says what you should do in different circumstances: if this is the case, then do this; if that's the case, then do that. That's a plan. A very natural thought is: well, when you have a conditional imperative, the whole point is to provide someone with an option of different plans. So every imperative should be equivalent to a disjunction of these plan imperatives of the form 'if this is the case, do this', blah blah blah blah. This is very analogous to the counterfactual case. But we might also want to allow, as I said, these embedded conditional imperatives. They would arise very naturally: someone may have a plan, and I might want to say, well, look, let's conditionalize this plan: if *this* is the case, then adopt *this* plan; if *that's* the case, then adopt *that* plan; and so on and so forth.

We'd like to have that expressive resource, so that we can actually conditionalize on the plans. So there are two desiderata here: we want to be able to conditionalize plans; and we want to have every imperative equivalent to a disjunction of plans. As I said, if you adopt a classical view, an intensionalist view, these desiderata can't be met, whereas they can be met in a non-classical or hyperintensional setting. This case is just meant to be illustrative; it's meant to show that when you're trying to regiment a certain discourse, there may be certain desirable features that you're after, which can only be realized by adopting a hyperintensional stance. And another moral – I think I mentioned this – is that people often complain that the hyperintensional stance is so complicated, classical logic is so beautifully simple. But there's a way in which a hyperintensional stance can be simpler than a classical approach, because you can actually have these normal form theorems – you can actually show that expressions can be simplified – which you can't do in a classical approach. Again, I think these are quite remarkable results. I, myself, didn't expect them to hold. Anyway, that was simply my point.

TIMOTHY WILLIAMSON

I'm not convinced that imperatives really need to be understood in terms of plans. They're also different from counterfactuals in that they don't have a modal element; they are indicative. Of course, on my view, if there isn't a modal element there, then you're just dealing with material conditionals, for which the 'if P, then if Q, then R' just is equivalent to 'if P and Q, then R', and I would be aiming to understand imperatives in such a way that you were just using the material conditional and that then you could get such equivalences in that way. Of course, there are all sorts of apparent obstacles to giving just an ordinary, truth-functional, material reading of the conditional, but I think they can be overcome. So that's the kind of strategy that I would be using in those cases.

Maybe I can ask something of Kit. So this is, as it were, a slightly more meta question. In the case of intensionalism, there was really a very dramatic revolution in philosophy, where right about 1960, Kripke came up with this new form of semantics – I mean, it wasn't that nobody had ever thought of possible worlds, but he found a much better way of doing the semantics in terms of that than people had been kind of messing around with before. This immediately solved a bunch of technical problems, but it was also something that philosophers – well, Kripke – could latch onto. Up to that point people had been thinking that when you're talking about modal operators, you're just talking about something like logical necessity and logical possibility; but Kripke found a way of doing things where the possible worlds were actually in some ways modeled on times, and therefore had a more metaphysical element.

So within a very few years, philosophy had been transformed – I mean of course analytic philosophy – in the sorts of framework that it was using. This possible worlds framework – with slight variations, but basically that one – became the framework in which an enormous number of philosophers were working. People were playing all sorts of variations on that theme, but it was fundamentally clear what the overall structures looked like; whereas with hyperintensionalism there's been much less in the way of a new, systematic, mathematically powerful framework.

Of course, Kit has done more than anybody else to provide hyperintensionalism with some kinds of systematic and formally worked-out framework; but there's a huge variety of different systems and approaches out there, and I wonder if you think that there's any sign of convergence amongst all these different versions of hyperintensionalism, because lots of people are not working in terms of truthmaker semantics at all.

KIT FINE

Well, I'd like to think that there would be convergence; that's at least a sign that one's onto something. It may be that we still are awaiting our Kripke; but before Kripke, right, as you said, the subject was a mess. People had many different ways of understanding these systems; there was nothing very systematic about what they did. So I agree that Kripke's work somehow solidified the field in the way that hasn't yet happened with these hyperintensional approaches. I would agree with that, and I also agree that it would be desirable if that did happen; but compared to modal logic, it's early days.

TIMOTHY WILLIAMSON

Yes, but I'm inclined to say that if there was this potentially unifying system that would clarify everything in the way Kripke's work did, you'd already be there!

KIT FINE

That's a backhanded compliment!

TIMOTHY WILLIAMSON

Would it not be an idea that we just throw it out into the audience? I don't feel that the two of us have to be making life miserable for each other, as a condition of the event.

Audience Questions

STUDENT 1

Just a question for Kit Fine. Is there a normal form in general for 'if it were that P, then if it were that not P, then it would be that Q'?

KIT FINE

Well, you just conjoin. You've got 'if P were to be the case then if Q were to be the case then R would be the case'. Now we take that to be equivalent to 'if P and Q were to be the case, then R would be the case'. So, in the case when 'Q' is 'not P' – this is the case that Tim was discussing – then of course you would end up with 'if P and not P were to be the case, then R

would be the case', which is actually, I think, a way of reading the English. But, as Tim pointed out, there's another way of reading it. So, my way of doing things is not getting at the other reading. But as I said, my interest wasn't in natural language. My interest was in the kind of things we're trying to say when we use counterfactuals but in a regimented way, which might depart from some of the ways in which we use natural language – just as classical first-order logic and its treatment of the quantifiers and the connectives departs from natural language, I would say.

STUDENT 2

I was wondering if you could return to the point of the debate earlier about the difficulty of extending the truthmaker framework to negation and universal claims, and I think Tim brought up a quite interesting point that we didn't really get into very much.

KIT FINE

Yeah, well let me make two points. One is, I think of truthmaker semantics as a semantics, not as a metaphysics. It may be that Russell was interested in metaphysics, in trying to understand ultimately how things work. So I'm perfectly happy with the thought that there's a so-called state space which tells you what the facts are, and I'm perfectly happy with the idea that that is to some extent a contextual matter – what it is we regard as the truthmakers is to some extent a contextual matter. It depends on the discourse: if you're an Eskimo, what will make it true that something is white? Would it be this white or that white? The rest of us, we're just happy with white. You might also just be happy with – you ask, what makes it true it's not raining? – you may just be happy with: it's not raining, that negative fact. You think, oh my God, negative fact in the world? When I'm doing semantics, I'm not necessarily interested in the world. So, that would be what I could say.

Now, when it comes to the negative existential 'there are no dragons' – there's a question that's actually how you regiment this – but one thing you could say is that you go through everything in the universe, and, in each case, there's the fact that it's not a dragon, and then there's also the fact that these are all the things that there are. So it's that big thing: 'this is not a dragon, that's not a dragon', blah blah blah. Or maybe, if it's not being a dragon, it's a matter of some positive fact, but anyway, just stick with this, the thing not being a dragon. So the truthmaker would be: this is not a dragon, that's not a dragon, and these various things are all

the things that there are. That would be the truthmaker, at least on one standard approach. I actually have a highly unorthodox approach, where I have these generic truthmakers. They're not in the world, but I think they do the job of explaining what it is that's making these things true.

I should mention here that there's a lack of fit between what I was saying earlier, about having an intensional conception of reality. As I said, for the purposes of doing semantics, I'm not even sure that the semantics should directly relate the languages to the world. But anyway, that's a separate issue.

TIMOTHY WILLIAMSON

Could I just make a couple of comments about that? One of them is just that I think Russell simply was not very clear about when he was doing semantics and when he was doing metaphysics. As we would now put it, he was pretty wobbly on the use-mention distinction, and I think people used the word 'logic' for something that was sort of semantics and sort of metaphysics.

But on the specific thing about going through some inventory of the contents of the universe and saying 'that's not a dragon', 'that's not a dragon', and then finally saying 'and there's nothing else': I think Russell did actually consider a view like that, and what worried him was that the 'and nothing else' was just another negative existential of the very same kind that you're dealing with. It's saying that it's not the case that there is something which isn't on this list. So there's a kind of circularity in giving a truthmaker of that kind, because it doesn't really seem that the negative fact – that there is not something which is not on the list – should be anything special; it's just one more negative existential. I think that was one of the alternatives that he considered and rejected for circularity – and again, he wasn't really able to find anything that would satisfy him.

But I think the difficulty is that what gives a nice initial motivation to the truthmaker idea is that we have these kinds of non-trivial example, as with disjunctions and conjunctions and positive existentials. Because if they'd just gone around saying: look, we need to talk about what makes statements true, and the proposition that grass is green is true because of the fact that grass is green, the proposition that there are more than twenty people in the room is true because of the fact that there are more than twenty people in the room, and it just went on like that; I think the audience would be wondering, what's the point of all of this? We're not really getting any bang for our buck. So it's the initial non-trivial examples that give us a sense that we're getting some real action with this theory.

But if it turns out that the non-trivial examples are just a sort of special case, then that undermines the initial promise of the theory, because what we were hoping for was that it could deliver by producing similar kinds of non-trivial truthmakers for these other statements that were going down towards the simpler and simpler facts or truthmakers, whatever you call them. If we're not getting that, then it makes it look as though the original examples were just a false promise.

KIT FINE

An *ad hominem* argument just occurred to me, which is this. Consider 'everything is not a man or mortal'. You might ask: what makes that true? Each thing, either not being a man, or its being mortal – or perhaps a man and a mortal – and these things being all the things there are. That last thing was the thing that was worrying us.

But if we're necessitists, we don't need to say these things are all the things there are, because this is something that holds as a matter of necessity. So we don't need this further fact if we're necessitists. Under truthmaker semantics for a constant domain, what makes a universal statement true is the same as what makes the conjunction of all of its instances true. So the semantics for a universal quantification is no more problematic than the semantics for conjunction.

TIMOTHY WILLIAMSON

I think it's easier for an intensionalist like me to say that than for a hyperintensionalist. I'm speaking on the basis of very faint memories of his lectures on logical atomism, but I think Russell wasn't thinking 'oh, it could have been a different set of things'. I think he was looking for some kind of perspicuous explanation of why this statement was true. Just as, if you don't say anything at all about all these things being an exhaustive inventory of the contents of the universe, then there's something lacking in your explanation, because there's a gap in it. Although a hard-line, coarse-grained intensionalist like me isn't going to be moved by that, I would expect a hyperintensionalist who wants the relations to be a bit more perspicuous than that to feel that there was something lacking in this kind of truthmaking – because something that was super-relevant to the truthmaking just hadn't been put on the table.

KIT FINE

Well, you've answered my *ad hominem* argument with another *ad hominem* argument.

TIMOTHY WILLIAMSON

Yes.

KIT FINE

But what I can say is there is a version of truthmaker semantics on which this isn't going to be a problem, namely that we don't insist upon this explanatory demand: as long as the truthmaker necessitates what is made true, that should be enough, and this further explanatory demand doesn't have to be met. So there would be a version of truthmaker semantics where this isn't a problem.

TIMOTHY WILLIAMSON

Yeah, but it doesn't seem very much in the spirit of the way that you've developed it, because in developing truthmaker semantics you've insisted that this is a non-modal theory, and that these sorts of thing that are going on in the truthmaker framework are just not to do with modality. So to appeal to modality in terms of the necessitation at that point – of course you're better placed than I am to say what's in the spirit of your view – doesn't seem very much in the spirit of your view as you've expounded it.

KIT FINE

I'm trying to get you to meet me halfway!

TIMOTHY WILLIAMSON

A politician – it might have been Aneurin Bevan – said: we all know what happens to the man who takes a position in the middle of the road – he gets run over.

STUDENT 3

You talked about this point towards the end of the debate about the sort of Kripke of hyperintensional logics, and I'm wondering what sorts of char-

acteristics would the work of such a figure, hypothetically speaking, be expected to have, or what sort of goals would it have to meet in order for it to sort of achieve something like what Kripke achieved?

TIMOTHY WILLIAMSON

Maybe this future Kripke is somewhere in the audience.

KIT FINE

Kripke gave something like a canonical formulation; as Tim said, people had already moved in that direction without quite getting there. It just seemed right, the framework he developed. But it's hard to say in the abstract what would be required.

Also, I think it's not just a question of it being a nice theory, but seeing that there's this range of applications as well. So there's a kind of systematicity and unity; but I don't know what exactly that would mean.

TIMOTHY WILLIAMSON

One of the nice features of Kripke's intensional framework is that it has had lots and lots of applications outside philosophy: in mathematics, in computer science, in linguistics, in theoretical economics through epistemic logic and so on. So maybe you could say something about what you think are the most promising applications of the hyperintensional framework outside philosophy.

KIT FINE

Well, actually, there have been a significant number of applications. People are using that formalism to develop various AI products. I'm afraid this is not something I really know much about, but people in AI have found it useful. Just today, I came across an article, I can't remember its title, but it had to do with the safety of AI procedures. Indeed, for me, one of the principal motivations behind truthmaker semantics was something called the frame problem in AI, but I don't know whether anyone's pursued that particular direction. Also, a fair number of people are now applying truthmaker semantics within linguistics. In fact, one of my colleagues in the linguistics department was an advocate of the possible world semantics and has now come over to the other side, the dark side. So it is catching on within linguistics.

I myself would like to think that there should be significant applications to confirmation theory and that sort of thing, and I know a number of people who have worked in that direction, but I can't speak for how good the work is. There certainly has been a gathering of momentum in the applications that have been made.

TIMOTHY WILLIAMSON

One thing to say about confirmation theory is that most of the work in it is done within a probabilistic framework. I'm not sure if it's obvious, but it is in fact the case that the state spaces that probability theory is based on, mathematically, are pretty much exactly the same as Kripke's possible worlds framework. The terminology is all different: for a start, in possible world semantics, the set of worlds is an uppercase 'W', and in the probability case you have to look for an uppercase ' Ω '; and then the probability theorists talk about outcomes rather than worlds, but basically the formal properties are exactly the same. In possible world semantics, sets of worlds are called propositions, whereas in probability theory they're called events; but again, it's basically formally exactly the same.

In a way it's not a surprise because Kripke was partly inspired to his possible world semantics – he told me – by looking at the kinds of state space you get in natural science and physics, for example – sometimes called phase spaces, sometimes called dynamical systems, depending on the extra structure that they have. That's a kind of very standard approach in the natural sciences; it's also the approach that probability theory, as a mathematically well-developed theory, is based on. So Kripke was actually just applying the same framework to modal logic. It's not some kind of weirdo science-fictionish thing about possible worlds as some kind of strange alternative realities or anything like that; it's in fact an application of a very well-developed approach that you get in lots of the natural sciences – and, as I said, it's what is going on in probability theory.

The kind of work that I've seen on confirmation theory has mainly been within a probabilistic framework, and so it's just, as it were, more intensional activity. There may be room for something else, but at least the default in that area, I think, is still an intensional approach.

STUDENT 4

This question is sort of in the spirit of Professor Williamson's last question: what do you both think are the very most important issues, theoretical and practical, that are likely to turn on this dispute?

KIT FINE

There are lots of small issues, so it's hard to say. Take something like the atomic theory of gases. It took some time for that to be accepted, and it was an incremental process. It just turned out other views didn't work; the view did work and it was a beautiful theory.

It's somewhat similar. How is the issue to be resolved? I think the applications are very significant in this regard: to the extent that one can show that this theory actually is good for something, any reasons one might have to doubt it would seem like being philosophic reasons in a bad sense of the term. If it can really do some important work – I mean through the applications, it could be linguistic, or what have you – that would be the best evidence it seems to me that we can have that there was something worthwhile there.

TIMOTHY WILLIAMSON

I agree with that. One thing I would say is that what you're looking for in a framework – which is what intensionalism is for a lot of work – is something settled and robust, so that you don't have to keep making it up as you go along; you just have this well-defined framework which is understood, and then you can apply it. You're just using it, you're not tweaking it; so it has its own autonomy. That's what you want from a framework: to take it out of the dispute, so it's just something there that we can use without having to reinvent it every time we use it, or to make extensions of it – we've got something that is general enough that we're genuinely applying it rather than inventing some new theory which is kind of similar to something else. One advantage of the intensional framework has been that you can quite easily use it in that spirit. Of course, it's been generalized in various ways, but it is basically pretty stable and robust.

STUDENT 5

Back to the question of what it is for a sentence to be true, such as the sentence 'grass is green'. I'm just curious what to make of the view that for a sentence to be true is for the predicate to apply to the subject, such as the subject grass, and the subject is among the extension of the predicate.

TIMOTHY WILLIAMSON

You seem to be assuming that every sentence is of subject–predicate form, but that isn't the case. For example, if you take a disjunction, like 'Ei-

ther Mary is guilty or John is guilty', that doesn't have an overall subject-predicate form. One of the ways in which semantics has had to get much more sophisticated in the 20th – and 21st – century is simply that we're much more aware of the fact that we have to deal with a huge range of sentences, which differ from each other a lot in their structure, so that you can't assume in advance that what you're dealing with is a sentence of some particular structure.

It took quite a long time for this to be fully accepted in logic. For example, John Stuart Mill didn't like the idea that you can have conjunctive propositions of the form 'A and B' – he said 'such a sentence is no more a conjunctive proposition than a street is a conjunctive house'. He just hadn't got the kind of generality of the category of proposition-expressing sentences. A lot of the reason why the way people talk about truth now is somewhat different from the way they used to is just appreciation of the fact that we really need a general theory that will work for declarative sentences of all kinds of structure – that we just shouldn't take it for granted in advance that we know in advance what the structure of a given sentence will be.

STUDENT 6

May I take this debate to the arena of normativity? How would either framework apply? From what I understand, you've written recently on intensionalism as applied to problems in normativity, moral realism, etc., and one of the implications of the intensional framework is that you end up with some kind of naturalism. For any necessitated normative thesis, there will be a co-intensional, non-normative...

TIMOTHY WILLIAMSON

You mean, co-intensional.

STUDENT 6

Co-intensional, sorry – non-normative fact. You wouldn't want to fill in the details as to how we should carve normative space by giving it some kind of Canberra Plan analysis of concepts.

TIMOTHY WILLIAMSON

God, no! I don't want to have anything whatsoever to do with the Canberra Plan.

STUDENT 6

Right. But neither would you want to, if you're a naturalist, in your case. . .

TIMOTHY WILLIAMSON

I don't think you've got the position quite right. It's certainly the case that I think there are examples where a predicate formulated using normative expressions expresses the same property as a predicate that doesn't contain normative expressions.

For example, let's assume that torturing babies for fun is necessarily wrong, it's just impossible for it not to be wrong. Then that means that 'torturing babies for fun' and 'wrongly torturing babies for fun' pick out the very same property, but one of them picks it out using a normative expression and the other one doesn't.

But the conclusion I'm drawing from that is not that the underlying property is natural, is non-normative, or something like that; it's simply that a distinction that can be made at the level of linguistic expressions – whether they contain normative material – just doesn't project onto a distinction between the things that they express. It's not part of my agenda that somehow the natural or the non-normative vocabulary is more basic than the normative vocabulary or anything like that; it's just that this normative–non-normative distinction, which we can make at the level of linguistic expressions, shouldn't be applied to what the expressions are picking out in the world because these differences just don't project down. It's not that I have some kind of agenda where I think that let's say, the language of fundamental physics is somehow more basic than anything else. It's more simply a matter of: be careful about whether distinctions are being made at the linguistic level or at the metaphysical level.

STUDENT 6

Right, but in that case then my question is: if the distinction between the normative and the non-normative depends only on the level of our representations – so there are no irreducibly normative facts or some moral. . .

TIMOTHY WILLIAMSON

I'm not working with facts anyway, but you may use true propositions.

STUDENT 6

But how would you carve normativity at the level of language? How would you know which are the true normative predicates, if not by doing some kind of conceptual analysis?

TIMOTHY WILLIAMSON

Definitely not by doing conceptual analysis! I think we have some capacity to recognize which linguistic expressions have some kind of evaluative valence to them, like 'right' and 'wrong', 'good' and 'bad', and so on, and deontic modal expressions like 'permitted' versus 'forbidden' and so on; obviously these are expressions which have certain kinds of connection with action, assessments of action, and so on.

There will be some story along those lines, but I don't think that these expressions can all be somehow analysed and then there will be some particular component in them which is the evaluative one. I don't think those kinds of program of analysis are likely to work out. In that sense, we're not using conceptual analysis, we're just using our ordinary understanding of these expressions and some ability to recognize expressions which have a valence versus ones that don't. Then, of course, the specifically moral ones are a subcategory of that, because lots of forms of evaluation are not moral evaluation.

KIT FINE

If I could just interject my own view on this question? I think it is necessary that torturing babies for fun is wrong. But for me, this is what I call normative necessity, not metaphysical necessity. Intensionalism is always relative to a certain modality; normally people take this to be the metaphysical modality. I want to deny that it's a metaphysical necessity that torturing babies for fun is wrong, even though there's another notion of necessity on which it is necessary. So even if I were an intensionalist with respect to metaphysical necessity, I wouldn't go the way of Tim. But that's just my opinion about the kind of necessity that's involved in these statements.

STUDENT 7

This is perhaps broadening things out a bit, but I'm interested in where you both see the heart of your disagreement lying, because I feel like there's been a number of different areas which have been discussed in which kind of your theories – you see them as being more useful individually in that area. Do you think your disagreement lies in just the fact that you find your theories more useful in specific areas, or do you think it's more fundamental and can be pinpointed? And do you agree on where the disagreement is, or not?

TIMOTHY WILLIAMSON

I think it is in some ways a methodological disagreement, which has to do with the kind of comparative weight that we give to different sorts of criteria in evaluating theories. I'm not sure Kit would agree with this characterization, but I tend to be suspicious of complicating theories. I'm very wary, for example, of just abandoning some simple theory that seems to be working pretty well just because we come across something which feels like a counterexample to it.

I'm as keen on counterexamples as anybody; I have been called a counterexample machine, so I can't be too hostile to them, but I think when we're assessing them, we're using what psychologists would call heuristics, that are quick, easy ways of assessing things, but are not 100% reliable. So when we have a conflict between a simple, so-far well-working theory and a counterexample, I think we should be pretty suspicious of the counterexample. I'm putting a lot of weight on criteria like simplicity and elegance in what's sometimes called the abductive comparison of theories, whereas I think Kit is less averse to complicating things and is more inclined than I am just to take what feel initially like compelling counterexamples at face value and think that these really just are counterexamples.

But if there is such a difference between us, that's just at the methodological level, and then of course there are more specific disagreements between us, e.g., on intensional versus hyperintensional frameworks, which I would be inclined to explain by this. I don't think it's a huge methodological difference between us; my feeling is that we're actually not that far apart on the spectrum of philosophers, where our views and approaches are really not radically different; it's just that when we zero in on something specific – like using an intensional framework versus a hyperintensional framework – then we have some specific disagreements.

KIT FINE

I think you're right about one thing about my own view, wrong about another. I hate complications; I'm much in favor of simple theories as you are; and I think simplicity is a guide to the truth. But you're right in thinking that I'm much less suspicious of intuition than you are. The way I see it, the question is how best to make progress in philosophy.

One way is to pursue a theory and to give it your all, but another is to expose yourself to the complexity of the data and so on and so forth. In the case of science, the first method worked very well. Newton came up with a beautiful, simple theory; it was just amazing. If he'd been bothered about all the complexities and so on and so forth, he wouldn't have got where he did.

That method, there's a certain sense in which it's not *a priori* that this was going to work – maybe in some sense it is *a priori*, but not in any practically meaningful sense. One wasn't to know that something like Newton's theory would be so successful, even though it ignored all the complications. Now, there's a real question as to whether we should, as it were, have the same attitude towards theory that scientists seem to have, and my own inclination is that right now, in the current state of philosophy, this is not the way progress is to be achieved.

My view is very oriented towards very particular intuitions that we have. I'm going to have to exercise judgment about them: is there something here that might be helpful, that might give us a better understanding of something? So I'm much more inclined, as it were, to work upwards from certain intuitions that I feel have something going for them, rather than work top-down from a theory. I hate it when these people have these theories – everything's physical and so on – and then they see everything in that light. It's like utilitarianism: progress in moral philosophy is to be gained not by just assuming a particular very general theory is correct; you need to dig deep into the details of very particular things. I guess that's my attitude right now about metaphysics, and really philosophy in general. We're not ready for big theories. I'm suspicious of anyone who has big theories.

TIMOTHY WILLIAMSON

We've got a lot of very good theories just in logic, which are extremely general.

KIT FINE

Oh, no, I'm talking about – yeah.

TIMOTHY WILLIAMSON

Yeah, but I'm taking it that the philosophical end of logic is part of philosophy. But I don't see such a big difference here between philosophy and natural science. There are some quite illuminating things that one can get from Thomas Kuhn's *Structure of Scientific Revolutions* – which of course became fantastically successful in the '60s because it was, as it were, the hippies' philosophy of science, and people thought that what it was saying was that science is just as irrational as everything else, or something like that. But I think it's in fact a much more sophisticated picture, and part of Kuhn's point is that every scientific paradigm at any point has anomalies. It's pretty obviously true of philosophy, as well, that every theoretical paradigm in philosophy has its own anomalies and people keep working on it despite that.

I think part of Kuhn's model is that what in a sense may be irrationality at the individual level can add up to rationality at the group level. What he describes is a situation in which it's typical of scientists to be more committed to their paradigms than is fully warranted by the evidence. He famously mentions Max Planck supposedly saying that progress in science is measured by funerals, which means basically: people, once they're really invested in a theory, never give up on it, but they do eventually die; where the progress comes from is that the students – who are not so invested already in any one paradigm – make somewhat better judgments about whose research program is more progressive and follow them. That's where the dynamics come from. But Kuhn's point is that having people at an individual level irrationally committed to a given paradigm or research program is a very good way of making sure that its explanatory potential is fully explored.

That holds just as much in philosophy. So it's very important in philosophy that we have both people like me who are fanatically committed to intensionalism, and people like Kit who are fanatically committed to hyperintensionalism, because that's how the potential of these different research programs gets worked out. Each of us is doing our damndest for the theory that we like, and then in the long run that will enable people to make informed decisions about what the capacity and limitations of these different paradigms are. So it's good that we're disagreeing, and that we're both willing to go all-out on our conflicting research programs, because,

at the level of the whole philosophical tradition, that is how progress gets made.

KIT FINE

I wonder if I could just make a further point about how it's often unclear what is the best method to use. I was in Edinburgh in the early 1970s; I worked a lot with people in cognitive science, and they were very concerned with recognition of speech, speech production, and so on. They tried to work this out on first principles – just really trying to analyse how we speak, what patterns of sounds make sense as productions of language. It was a huge failure! You might have thought *a priori*: surely this is the way to go! We speak; my words now sound like sentences of English; they flow the way they should. You might think: surely there are very general principles governing how this is so. This approach turned out not to work; and of course now, with pattern recognition techniques where we don't even understand what the underlying principles are, we're able to do these tasks. As I said, this is an astonishing fact! One might have thought *a priori* that we need to look for first principles. So, I think we need to have a properly humble attitude as to what the correct methodology is.

Philosophy is strewn with attempts to say: this is how we should be pursuing the subject. Look at Descartes, or Kant, and so on. None of us really know how best to pursue the subject. It's very much an open question, just as it was for speech in Edinburgh – or should have been.

Final Words

KIT FINE

One of the things I've been trying to say is that, look, what I've been calling truthmaking semantics isn't just a semantics of natural language; it's also a semantics of formal language. Let me just give one example where I think it has been illuminating. I don't know how many of you know about intuitionistic logic. Intuitionists don't accept the law of excluded middle – P or not P. There have been two main semantics for intuitionistic logic: one in terms of constructions or proofs; and another, by Kripke actually, in terms of conditions under which a statement can be true – think of those conditions as bits of partial information. However, it has been hard to see how these two approaches are related.

There's something I developed, which is a truthmaker semantics for intuitionistic logic, which bridges the gap between the two. It actually does a reasonably good job of explaining how you can go from one to the other. So, I think there's something genuinely illuminating about this; it's creating a kind of synthesis, if you like, between two approaches that otherwise do not appear to relate to one another. And it seems to me that the fact that it can do that is symptomatic of the fact that there's something worthwhile here; there's some useful techniques here that we can perhaps apply to a wide variety of different fields.

TIMOTHY WILLIAMSON

Studying intuitionistic logic can be genuinely illuminating, but what it's illuminating is an approach to mathematics that is inspired by fundamentally, utterly mistaken views about knowledge and about the nature of mathematics. I'm not saying that your truthmaking vision is implicated in those, but what we're doing is exploring a way of thinking that is very interesting, but also philosophically mad.

KIT FINE

I accept that the initial motivation for intuitionistic logic was not a good motivation. Still, there's a lot to be said for intuitionistic logic; it can do a lot for us. What we can think of it as is a formalization of the notion of constructive proof. I might be able to show that a contradiction can be deduced from the fact that everything Fs. So classically I know that there's something that's not an F, but that's not to provide an example of something that's not an F. It's not a constructive proof because it's not telling me what the counterexample is. One thing that intuitionistic logic does is it forces you to provide a constructive proof; and that's something you can accept about the logic without necessarily accepting the kind of epistemological motivation that lies behind it.

TIMOTHY WILLIAMSON

Here's one way in which the difference between the different attitudes to intuitionistic logic comes out. Of course it's clear that there is a difference between constructive and non-constructive proof, and that it merits exploration, and intuitionistic logic can play a role in that – but what really sorts out who's on which side is here: if you think that intuitionistic logic is basically a good tool to use in understanding the difference between construc-

tive and non-constructive proof, then you should be happy to use classical logic and non-constructive proof in the metalanguage, because you're not taking the view that there's anything wrong with non-constructive proof, just that constructive proof has special characteristics of its own.

Some people investigate intuitionistic logic using a classical metalogic because they're not buying into the original intuitionistic ideology. But hardline intuitionists insist on using intuitionistic logic in the metalanguage as well as in the object language. As a result of its restrictions – it's a funny situation – there are nice properties of intuitionistic logic which can be proved if you use classical logic in the metalanguage, but which can't be proved if you use the intuitionistic logic in the metalanguage so that intuitionism in some respects has virtues which can only be recognized by its enemies.