CANADIAN JOURNAL OF PHILOSOPHY, 2016
VOL. 46, NOS. 4–5, 542–547
http://dx.doi.org/10.1080/00455091.2016.1205852

Routledge
Taylor & Francis Group

Reply to Bacon, Hawthorne and Uzquiano

Timothy Williamson

Wykeham Professor of Logic, University of Oxford, Oxford, UK

In 'Higher-order free logic and the Prior-Kaplan paradox' (Bacon, Hawthorne, and Uzquiano 2016), Andrew Bacon, John Hawthorne and Gabriel Uzquiano (BHU) explore various strategies for avoiding some disturbing results in higher-order logic derived by Prior (1961), Kaplan (1995) and others, under apparently reasonable assumptions. Specifically, the results involve quantification into sentence position. In particular, they are derivable in systems of higher-order modal logic of the sort defended in *Modal Logic as Metaphysics*, so I cannot just say that they are not my problem.

Typical of the results is this:

Prior's Theorem $\quad Q\forall p(Qp \rightarrow \neg p) \rightarrow \exists p(Qp \,\&\, p) \,\&\, \exists p(Qp \,\&\, \neg p)$

Here *Q* is any sentence operator. For instance, we can read *Q* as 'TW visibly wrote at 8.45 that', with reference to the morning on which I wrote this paragraph. Under this interpretation, roughly paraphrased in English, Prior's theorem says that if TW visibly wrote at 8.45 that whatever TW visibly wrote at 8.45 is not so then something TW visibly wrote at 8.45 is so and something TW visibly wrote at 8.45 is not so, and consequently TW visibly wrote at least two things at 8.45.

The main problem with Prior's Theorem is *not* that it is beyond the remit of logic to give us this sort of information. As I argue in Williamson 2013, logic is defined by its generality, not by its neutrality or uninformativeness. Strong, informative theories are as valuable in logic as they are in any other science. The main problem with Prior's theorem is much worse and more straightforward. It just looks to be false, under various interpretations including the one above. For, as I can assure the reader, a natural description of what happened at 8.45 on the morning I wrote this paragraph is this: TW visibly wrote that whatever TW visibly wrote at 8.45 is not so; TW visibly wrote nothing else (in the relevant sense). Given the natural description, TW visibly wrote only one thing at 8.45, so the consequent of Prior's theorem is false, while its antecedent is true. Thus Prior's theorem looks to make false predictions about easily observable events.

CONTACT Timothy Williamson timothy.williamson@philosophy.ox.ac.uk

On the face of it, the problem has nothing special to do with higher-order *modal* logic. Prior's theorem contains no specifically modal operators, and the reading above on which it appears false was specified in non-modal terms. However, BHU develop modal aspects of the problem in several ways.

First, BHU formulate various plausible-looking possibility claims that can be refuted by reasoning like that in the derivation of Prior's theorem. For instance, if you doubt my description of what *did* happen at 8.45 this morning, but concede that it (metaphysically) *could* have happened that way, your concession is incompatible with the necessitation of Prior's theorem, which is standardly derivable in higher-order modal systems, including mine. BHU use some plausible-looking generalized possibility claims as benchmarks in their search for higher-order modal systems that avoid Prior-like results: one can test systems for their consistency with such claims. However, since the original problem arose without appeal to modality, we may reasonably hope for a non-modal explanation of what went wrong. For understanding the root of the problem, modality looks like one moving part too many, an unhelpful complication and distraction.

The second way in which BHU develop a modal aspect of the problem is by treating quantification into sentence position intensionally, modelling it in terms of quantification over sets of worlds, at least in some of the approaches they consider. That is also how *Modal Logic as Metaphysics* treats quantification into sentence position. However, that too is not essential to the underlying problem, as BHU show. The problem arises even if quantification into sentence position is treated extensionally, in effect as quantification over truth-values, since the consequent of Prior's theorem then implies, for instance, that at 8.45 I visibly wrote two very coarse-grained things that differ in truth-value. Equally, the problem arises even if quantification into sentence position is treated hyperintensionally, in effect as quantification over very fine-grained entities such as Fregean thoughts, since it still seems false that at 8.45 I visibly expressed in writing two Fregean thoughts that differ in truth-value. Thus, the underlying problem is robust on the dimension of fineness of grain in the interpretation of quantification into sentence position. The intensional approach is not to blame, though an adequate diagnosis must no doubt take account of the operative level of individuation.

BHU emphasize a third modal theme with reference to *Modal Logic as Metaphysics*. The first part of their paper constructs a variety of ingenious models to investigate 'Fregean' higher-order modal systems in which, as in the book, one avoids a Russellian hierarchy of ramifications within a given grammatical category. The sentential quantifiers are interpreted as ranging over intensions (in effect, sets of worlds). However, by contrast with the book, the quantifiers do not mandatorily range over *all* intensions. Rather, they are restricted to a proper or improper subset of intensions; BHU restrict the term 'proposition' to intensions in that subset. The effect is to invalidate both the rule of universal instantiation and various comprehension principles, thereby undermining Prior's proof of his theorem and most similar proofs. The result is a higher-order *free* modal logic. But although the domain of quantification is restricted, the restriction is not relativized to worlds. BHU are still doing constant domain semantics, as in the book. Thus sentential quantifier analogues of the Barcan formula and its converse and the necessity of existence are still valid. I derive those principles by universal instantiation, but they hold in BHU's models even though universal instantiation does not. As BHU say, this pulls apart two strands of thought that go together very naturally in the argument of *Modal Logic as Metaphysics*: in their terms, 'First that existence doesn't modally come and go, and second, that existence comes cheap'.

A quick way to check the separation of the two strands is by considering models with only one world. In such models, there is no contingency at all, so trivially existence doesn't modally come and go. But the domain of quantification may still be restricted, for instance to truths, so that existence isn't cheap in BHU's sense: in such models, $\forall p\, p$ is true but $p \,\&\, \neg p$ false, so universal instantiation fails, and $\neg(p \,\&\, \neg p)$ is true but $\exists p\, \neg p$ false, so existential introduction fails.

However, 'existence is cheap' is a slightly misleading slogan for the aspect of the theory defended in *Modal Logic as Metaphysics* that BHU have in mind (although I do not think that they misunderstand the theory). Of course, some philosophers may use the word 'existence' for an 'expensive' property, such as concreteness, that some things have and others lack, but that is irrelevant because the book explicitly avoids using the word 'existence' in its theorizing. The real issue must concern *being something*, on an unrestricted reading of that phrase at whatever order is in play. It is also irrelevant that, on my view, any individual is something in all possible circumstances, and likewise for any proposition, for that is just the claim that 'existence doesn't modally come and go', with which BHU are *contrasting* the claim that 'existence is cheap'. Rather, they seem to have in mind the unrestricted validity at all orders of the standard introduction rule for the unrestricted quantifier ∃ and the standard elimination (instantiation) rule for the unrestricted quantifier ∀, according to the theory. But of course that validity does not mean that the rules can be applied to whatever has the syntactic appearance of an expression of the appropriate category. After all, it is no objection to the standard introduction rule for disjunction (∨) that *A* may be true while $A \vee B$ is not because (we may suppose) although *B* has the syntactic appearance of a sentence, it has never been assigned a meaning, so the disjunction too is semantically defective. Theorizing in any discipline, logic, physics or anything else, is liable to go wrong when one does it in semantically defective terms. To speak for convenience with typical ambiguity: for any grammatical type whatsoever, unless an expression of that type is semantically defective, it has a semantic value of the appropriate kind (perhaps relative to an assignment), and an unrestricted quantifier of the same type ranges over all values of that kind, so the standard quantifier rules preserve truth (when applied to semantically non-defective expressions). None of this prejudges the question of how easy or hard it is for an expression of a

https://doi.org/10.1080/00455091.2016.1205852 Published online by Cambridge University Press

given type to be semantically non-defective. With these clarifications, I will for simplicity continue using BHU's phrase 'existence is cheap' for the claim that the standard quantifier rules are truth-preserving.

In BHU's sense, there are models in which existence doesn't modally come and go but isn't cheap. As they also note, the converse is much harder to make sense of: cheap existence modally coming and going. For if it sometimes modally goes, it is not as cheap as all that. More formally, the characteristic theses of higher-order necessitism are derivable from the standard higher-order rule of universal instantiation (the first-order case is trickier).

Although BHU construct various Fregean models in which the existence of propositions isn't cheap, but doesn't modally come and go, and various Priorian theorems are false, in the end they declare themselves pessimistic about the prospects for this strategy in defusing Prior–Kaplan paradoxes. For too many other plausible assumptions have to be given up in blocking all routes to the paradoxical conclusions. Within the Fregean framework, one can derive versions of the theorems from plausible principles about the closure of the class of propositions under various operations (6.1) or the supervenience of all truths on fundamental truths (6.2), or by making only minimal assumptions about the existence of propositions and consequently ruling out more complex but still apparently possible scenarios (6.3). Moreover, without the standard quantifier rules, higher-order quantifiers are much less useful for enhancing the expressive power of the language (6.4 and 6.5). One would have to be very confident indeed that the Priorian theorems have false readings to be willing to give up so much of higher-order logic in order to vindicate the legitimacy of those readings.

In Section 7 of their paper, BHU consider strategies that reject universal instantiation, and so deny that existence is cheap, within a Russellian framework, ramifying either the quantifiers for a given grammatical category, or the attitude verbs (such as 'write' and 'think') used to specify the paradox-generating readings of the operator *Q*, or both. Once again their conclusions are pessimistic: they provisionally judge the costs to outweigh the benefits.

In concluding, BHU write 'our own inclination is to take Prior's results at face value', so they incline to rejecting principles that conflict with them (Section 8). But, as they say, 'There still remains the question of how to implement that discovery'. They suggest that we may still need to postulate ramifications of subtly different relations associated with a given attitude verb to explain why Prior's results *seem* to fail, but without ramifying the quantifiers or postulating failures of universal instantiation. This direction of research is congenial to my own instincts. I have suggested such a non-logical version of indefinite extensibility in dealing with the semantic and set-theoretic paradoxes (Williamson 1998). The guiding abductive methodology is to keep the overall logical framework simple and strong, while explaining the paradoxes by allowing messier proliferation at the level of less general distinctions. That way one prevents the resultant restrictions from undermining the explanatory power of theories in other disciplines (such as the natural sciences), as they do when logical principles themselves are restricted in response to the paradoxes, and so have to be supplemented with numerous ad hoc assumptions in order to recover their full strength in scientific applications (Williamson forthcoming).

That general methodology is applicable to the case BHU discuss, where someone restricts the rule of universal instantiation for quantification into sentence position by adopting some form of free logic instead. Thus by itself $\forall p\, A$ no longer in general yields its instance $A(B/p)$ (subject as always to the normal syntactic qualifications to prevent free variables in B becoming bound when B is substituted for p in A). To derive the instance in free logic, one also needs an auxiliary premise such as $\exists q(B \equiv q)$ (where q does not occur in B and $\equiv$ is the sentential analogue of identity), to ensure that B expresses a proposition in the domain of quantification.

Now suppose that the background logic for a theory about some topic far from Prior-Kaplan paradoxes employs quantification into sentence position. For instance, that might be the most economical way of understanding some talk of 'conditions' or the like in natural science, since (unlike first-order quantification) it avoids ascribing an apparatus of nominalizing and denominalizing devices to mediate between variables in name position and sentences; for a detailed account of how scientific applications of the mathematical theory of dynamical systems are naturally understood in terms of quantification into sentence position see Williamson 2016. Thus, a scientific explanation of some phenomenon may involve deriving it from various theoretical assumptions that include generalizations of the form $\forall p\, A$. In applying those generalizations, the derivation will typically need to instantiate them. But, as we noted, in the free logical framework each such move generally requires postulating an auxiliary non-logical assumption of the form $\exists q(B \equiv q)$. If such an assumption is justified at all, it will be by further auxiliary assumptions, such as that universal instantiation fails only where intentionality is involved and that the present phenomenon does not involve intentionality — odd-looking assumptions for an explanation in physics! Thus, restricting universal instantiation to block Prior-Kaplan paradoxes has unfortunate repercussions in all sorts of cases far from those paradoxes. By contrast, ramifying the attitude verbs is far more local in its effects, since those verbs are absent from most theories in the natural sciences.

Of course, such general methodological remarks alone do not satisfyingly explain away the apparent falsifications of predictions made by Prior's theorem about observable events. I will not attempt to provide any such explanation here. However, I do suggest that an appropriate setting for such an explanation is likely to be an externalist theory of content, on which what, if anything, one has an attitude towards depends on many factors and preconditions that are utterly non-obvious to both oneself and others; they can only be uncovered through painstaking theoretical investigation. Such a setting helps reduce the shock of Prior's results, although we are still very far from completely removing their sting.

References

Bacon, Andrew, John Hawthorne, and Gabriel Uzquiano. 2016: "Higher-order Free Logic and the Prior-Kaplan Paradox." *Canadian Journal of Philosophy* 46 (4–5): 493–541.

Kaplan, David. 1995. "A Problem in Possible Worlds Semantics." In *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*, edited by W. Sinnott-Armstrong, D. Raffman, and N. Asher, 41–52. Cambridge: Cambridge University Press.

Prior, A. N. 1961. "On a Family of Paradoxes." *Notre Dame Journal of Formal Logic* 2: 16–32.

Williamson, Timothy. 1998. "Indefinite Extensibility." *Grazer Philosophische Studien* 55: 1–24.

Williamson, Timothy. 2013. *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

Williamson, Timothy. 2016. "Modal Science." *Canadian Journal of Philosophy* 46 (4–5): 453–492.

Williamson, Timothy. forthcoming. "Semantic Paradoxes and Abductive Methodology." In *The Relevance of the Liar*, edited by B. Armour-Garb. Oxford: Oxford University Press.