

FIXED POINTS FOR CONSEQUENCE RELATIONS

TOBY MEADOWS*

ABSTRACT

This paper provides a way of dealing with paradoxes associated with consequence relations via a generalisation of Kripke's fixed point technique. In particular, we focus on Beall and Murzi's paradox, although the framework outlined should have more general application. We first attempt to locate this problem among the existing corpus of semantic paradoxes. We then examine two canonical approaches to the issue and conclude with an inductive construction which, in some fashion, goes beyond those approaches.

1. Introduction

In [forthcoming], Beall and Murzi offer up a fresh semantic paradox. Rather than focusing on truth or necessity, Beall and Murzi investigate what would happen if we attempted to introduce a validity predicate, *Val*, into a language with sufficient expressive resources to talk about its own syntax. The validity predicate, *Val*, is intended to represent a primitive conception of validity. We thus write $Val(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$ to mean that ψ , the sentence coded by the object $\ulcorner\psi\urcorner$, is a consequence of ϕ , the sentence coded by $\ulcorner\phi\urcorner$. Since we are taking the notion as primitive, we have no need to define it, although we may nonetheless constrain it. Beall and Murzi propose two constraints, which together have disastrous consequences. The first

(VP) $\text{If } \phi \vdash \psi, \text{ then } \vdash Val(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$

is intended to capture the idea that if we may prove ψ from the assumption ϕ , then we should also be able to prove that ϕ is a valid consequence of ψ . This seems very reasonable. The second principle

(VD) $\phi, Val(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner) \vdash \psi$

is a natural generalisation of *modus ponens*: if we assume ϕ and we assume that ψ is a valid consequence of ϕ , then surely we should be able

* I would like to thank Dave Ripley for his invaluable assistance in improving this paper.

to conclude that ψ . However we see that these seemingly innocent assumptions render any system with such a validity predicate inconsistent as follows. We shall assume that we have sufficient power to prove something like the diagonal lemma, and as such that there is a sentence π where:

$$\vdash \pi \leftrightarrow Val(' \pi ', ' \perp ').$$

We then demonstrate that $\vdash Val(' \pi ', ' \perp ').$

$$\frac{\pi^{(1)} \quad \frac{\pi \rightarrow Val(' \pi ', ' \perp ')}{Val(' \pi ', ' \perp ')} Thm}{Val(' \pi ', ' \perp ')} VD$$

$$\frac{\perp}{Val(' \pi ', ' \perp ')} (1) - VP$$

And from here inconsistency follows easily.

$$\frac{\frac{Val(' \pi ', ' \perp ')}{\pi} Thm \quad \frac{Val(' \pi ', ' \perp ') \rightarrow \pi}{Val(' \pi ', ' \perp ')} Thm}{\perp} VD$$

This paradox is the focus of this paper. Our goal is to use a generalisation of Kripke's fixed point technique for truth to provide a consequence relation capable of supporting both *VD* and *VP*. However, we shall first concentrate our attention on the following two tasks:

- (1) isolating the Beall-Murzi paradox within the contemporary landscape of semantic paradoxes; and
- (2) investigating two different approaches that have been taken to this or similar paradoxes and which are already present in the literature.

We shall begin by looking at the prototypical semantic predicate, truth, and two of the simplest paradoxical sentences that can be formed from it. We then move to the more restrictive case of necessity. This will help us to see how the Beall-Murzi paradox works in more well-worn contexts. Moreover, we shall see that, as opposed to truth-based paradoxes, there is less room to move with validity: the Beall-Murzi paradox is a special case. We shall use classical logic unless stated otherwise, and we shall assume the full complement of connectives.¹

1.1. Two liar sentences and four ways of proving each of them. Let us expand our background language with a truth predicate, *T*, and let us

¹ It is worth noting that Beall and Murzi elegantly avoid the need for negation in their framing of the paradox; however, we incorporate it here for the purposes of comparison.

assume, naturally enough, that the following schema holds for all sentences ϕ in our language.

$$(Ax) \quad \vdash \phi \leftrightarrow T \ulcorner \phi \urcorner.$$

We look at how we may demonstrate inconsistency using either an untruth-teller or a truly-not-teller.

1.1.1. *The untruth-teller.* Using some version of the diagonal lemma, we assume that λ is such that

$$\vdash \lambda \leftrightarrow \neg T \ulcorner \lambda \urcorner.$$

We shall call λ an *untruth-teller*. For our first proof we start by assuming that $T \ulcorner \lambda \urcorner$.

Example 1.

$$\frac{\frac{T \ulcorner \lambda \urcorner^{(1)} \quad \overline{T \ulcorner \lambda \urcorner \rightarrow \lambda} Ax}}{\lambda} \quad \frac{T \ulcorner \lambda \urcorner^{(1)} \quad \overline{T \ulcorner \lambda \urcorner \rightarrow \neg \lambda} Thm}}{\neg \lambda}}{\frac{\perp}{\neg T \ulcorner \lambda \urcorner}} (1)$$

Thus we see that $\vdash \neg T \ulcorner \lambda \urcorner$; and so

$$\frac{\frac{\overline{\neg T \ulcorner \lambda \urcorner} Thm} \quad \frac{\overline{\neg T \ulcorner \lambda \urcorner \rightarrow \neg \lambda} Ax}}{\neg \lambda} \quad \frac{\overline{\neg T \ulcorner \lambda \urcorner} Thm} \quad \frac{\overline{\neg T \ulcorner \lambda \urcorner \rightarrow \lambda} Thm}}{\lambda}}{\perp}$$

And so the system is inconsistent.

But we can just as easily show the inconsistency by assuming $\neg T \ulcorner \lambda \urcorner$.

Example 2.

$$\frac{\frac{\overline{\neg T \ulcorner \lambda \urcorner^{(1)}} \quad \overline{\neg T \ulcorner \lambda \urcorner \rightarrow \lambda} Thm}}{\lambda} \quad \frac{\overline{\neg T \ulcorner \lambda \urcorner^{(1)}} \quad \overline{\neg T \ulcorner \lambda \urcorner \rightarrow \neg \lambda} Ax}}{\neg \lambda}}{\frac{\perp}{T \ulcorner \lambda \urcorner}} (1)$$

Thus $\vdash \neg T \ulcorner \lambda \urcorner$.

$$\frac{\frac{\overline{T \ulcorner \lambda \urcorner} Thm} \quad \frac{\overline{T \ulcorner \lambda \urcorner \rightarrow \neg \lambda} Thm}}{\neg \lambda} \quad \frac{\overline{T \ulcorner \lambda \urcorner} Thm} \quad \frac{\overline{T \ulcorner \lambda \urcorner \rightarrow \lambda} Ax}}{\lambda}}{\perp}$$

Or we can assume λ .

Example 3.

$$\frac{\lambda^{(1)} \quad \frac{\overline{\lambda \rightarrow \neg T'\lambda'}}{T'\lambda'} Thm \quad \overline{\neg T'\lambda' \rightarrow \neg \lambda} Ax}{\frac{\overline{\neg T'\lambda'}}{\neg \lambda} Thm} \quad \frac{\perp}{\neg \lambda} (1)$$

Thus $\vdash \neg \lambda$.

$$\frac{\overline{\neg \lambda} Thm \quad \frac{\overline{\neg \lambda} Thm \quad \frac{\overline{\neg \lambda \rightarrow T'\lambda'}}{T'\lambda'} Thm}{\lambda} \quad \overline{T'\lambda' \rightarrow \lambda} Ax}{\perp}$$

And finally, we can also prove it by assuming $\neg \lambda$.

Example 4.

$$\frac{\neg \lambda^{(1)} \quad \frac{\overline{\neg \lambda \rightarrow T'\lambda'}}{T'\lambda'} Thm \quad \overline{T'\lambda' \rightarrow \lambda} Ax}{\frac{\overline{\neg \lambda^{(1)}}{\lambda} Thm} \quad \frac{\perp}{\lambda} (1)}$$

Thus $\vdash \lambda$.

$$\frac{\overline{\lambda} Thm \quad \frac{\overline{\lambda} Thm \quad \frac{\overline{\lambda \rightarrow \neg T'\lambda'}}{\neg T'\lambda'} Ax \quad \overline{\neg T'\lambda' \rightarrow \neg \lambda} Thm}{\neg \lambda} Thm}{\perp}$$

So there are at least four different ways of getting to inconsistency using the untruth-teller λ . We may show much the same thing for the truly-not-teller.

1.1.2. *The truly-not-teller.* Now let v be such that $\vdash v \leftrightarrow T'\neg v$. We shall call this sentence the *truly-not-teller*. We may show that the system with Ax is inconsistent by assuming v as follows:

Example 5.

$$\frac{\frac{\frac{\nu^{(1)}}{\nu \rightarrow T^{\ulcorner \neg \nu \urcorner}} Thm}{T^{\ulcorner \neg \nu \urcorner}}}{\nu^{(1)}} \quad \frac{\frac{\frac{\overline{T^{\ulcorner \neg \nu \urcorner} \rightarrow \neg \nu}}{T^{\ulcorner \neg \nu \urcorner} \rightarrow \neg \nu} Ax}{\neg \nu}}{\perp} (1)$$

Thus $\vdash \neg \nu$.

$$\frac{\frac{\frac{\frac{\overline{\neg \nu} Thm}{\neg \nu} Thm}{\neg T^{\ulcorner \neg \nu \urcorner}}}{\neg T^{\ulcorner \neg \nu \urcorner}} \quad \frac{\frac{\frac{\overline{\neg \nu \rightarrow \neg T^{\ulcorner \neg \nu \urcorner}} Thm}{\neg \nu \rightarrow \neg T^{\ulcorner \neg \nu \urcorner}}}{\neg T^{\ulcorner \neg \nu \urcorner} \rightarrow \neg \neg \nu} Ax}{\neg \neg \nu}}{\perp}$$

Similarly, it is easy to see that the inconsistency can be also proven by assuming either $\neg \nu$, $T^{\ulcorner \neg \nu \urcorner}$ or $\neg T^{\ulcorner \neg \nu \urcorner}$. It is almost difficult to not prove the inconsistency using a truth predicate satisfying Ax .

We now provide a definition which will allow us to broadly characterise these different proofs according to the form of the diagonal sentence and the initial assumption made in the proof.

Definition 6. Suppose we have a sentence ϕ given by the diagonal such that:

$$\vdash \phi \leftrightarrow \psi^{\ulcorner \phi \urcorner}$$

for some formula $\psi(x)$. We shall call ϕ the *LHS* and $\psi^{\ulcorner \phi \urcorner}$ the *RHS* of the diagonal. Let us say that a sentence ϕ for a semantic predicate P is in *forward-negation form* if we may establish that:

$$\vdash \phi \leftrightarrow \neg P^{\ulcorner \phi \urcorner}.$$

Let us say that a diagonal sentence ϕ for a predicate P is in *reverse negation form* if it is the case that:

$$\vdash \phi \leftrightarrow P^{\ulcorner \neg \phi \urcorner}.$$

We may then summarise the examples above by saying that for a truth predicate satisfying Ax , we may demonstrate inconsistency using a diagonal sentence in either forward negation or reverse negation form; and in either form we may complete the proof from an opening assumption of either *LHS* or *RHS*.

1.2. Two forms of Montague's paradox each with two forms of proof.

However, when we come to consider necessity, matters become more restrictive. We consider here Montague's not-necessary teller and a variation of it: the necessarily-not-teller [Montague, 1963]. We now remove the truth predicate from the language and replace it with a necessity predicate \Box . Ax is too strong to be taken seriously with respect to necessity. The fact that ϕ is true does not mean that ϕ is necessary; i.e., $\Box\phi$. However, if we could prove that ϕ , then we might think that $\Box\phi$ was true. We thus weaken Ax accordingly into the following two principles.

(Nec) If $\vdash \phi$, then $\vdash \Box\phi$.

(T) $\vdash \Box\phi \rightarrow \phi$.

1.2.1. *The not-necessarily-teller.* We first consider the forward negation form. Using our diagonal lemma, let μ be such that $\vdash \mu \leftrightarrow \neg\Box\mu$. We call this the *not-necessarily-teller*.

Example 7.

$$\frac{\frac{\Box\mu^{(1)} \quad \overline{\Box\mu \rightarrow \mu} \quad T}{\mu}}{\frac{\Box\mu^{(1)} \quad \overline{\Box\mu \rightarrow \neg\mu} \quad Thm}{\neg\mu} \quad (1)}{\frac{\perp}{\neg\Box\mu}}$$

Hence $\vdash \neg\Box\mu$; and thus,

$$\frac{\frac{\overline{\neg\Box\mu} \quad Thm}{\frac{\mu}{\Box\mu} \quad Nec} \quad \frac{\overline{\neg\Box\mu \rightarrow \mu} \quad Thm}{\neg\Box\mu} \quad Thm}{\perp}$$

And we may also assume that $\neg\mu$ at the beginning of the proof.

Example 8.

$$\frac{\frac{\neg\mu^{(1)} \quad \overline{\neg\mu \rightarrow \Box\mu} \quad Thm}{\Box\mu} \quad Thm}{\frac{\overline{\Box\mu \rightarrow \mu} \quad T}{\mu}} \quad T}{\frac{\perp}{\mu} \quad (1)} \quad \neg\mu^{(1)}$$

Thus $\vdash \mu$.

$$\frac{\frac{\frac{\mu}{\mu} \text{Thm} \quad \frac{\frac{\overline{\mu} \text{Thm}}{\Box \ulcorner \mu \urcorner} \text{Nec}}{\Box \ulcorner \mu \urcorner \rightarrow \neg \mu} \text{Thm}}{\neg \mu}}{\perp}}$$

However there is no way of establishing the inconsistency via a proof assuming either μ or $\neg\mu$.

1.2.2. *The necessarily-not-teller.* Finally, let ρ be such that $\vdash \rho \leftrightarrow \Box \ulcorner \neg \rho \urcorner$. We shall call this the *necessarily-not teller*.

Example 9.

$$\frac{\frac{\rho^{(1)} \quad \frac{\rho \rightarrow \Box \ulcorner \neg \rho \urcorner} \text{Thm}}{\Box \ulcorner \neg \rho \urcorner} \quad \frac{\Box \ulcorner \neg \rho \urcorner \rightarrow \neg \rho} \text{T}}{\neg \rho}}{\frac{\perp}{\neg \rho} (1)}$$

Thus $\vdash \neg \rho$.

$$\frac{\frac{\frac{\overline{\neg \rho} \text{Thm}}{\Box \ulcorner \neg \rho \urcorner} \text{Nec}}{\Box \ulcorner \neg \rho \urcorner \rightarrow \rho} \text{Thm}}{\rho}}{\perp}$$

We may also demonstrate the inconsistency by starting with the assumption $\Box \ulcorner \neg \rho \urcorner$ in much the same way as Example 7. The critical restriction is, of course, that necessitation can only be deployed in the second phase of the proof.

1.3. Discussion. The *Val* predicate used in the sentence π for the Beall Murzi paradox takes two arguments as opposed to one. However, it is not difficult to see that π bears a close relationship with ρ , the necessarily-not-teller. The Beall-Murzi sentence does not make use of negation; however, the use of \perp clearly plays a similar role in the usual interpretation of validity which is played out here. Moreover, we see that both sentences rely on embedding the effect of negation within the scope of the semantic predicate.

We may summarise the discussion above in the following table. A “✓” indicates that we can demonstrate inconsistency in the indicated form from the indicated opening assumption. A “×” indicates that we cannot. We let ζ where $\zeta \leftrightarrow \neg \text{Val}(\ulcorner \top \urcorner, \ulcorner \zeta \urcorner)$ be the forward negation form of the validity paradox.

Form	Opening Assumption	Truth	Necessity	Validity
FN	LHS	✓	×	×
	RHS	✓	×	×
	\neg LHS	✓	✓	\times^B
	\neg RHS	✓	✓	\times^B
RN	LHS	✓	✓	✓
	RHS	✓	✓	\times^A
	\neg LHS	✓	×	×
	\neg RHS	✓	×	×

- A. This may be simply recovered if we add the following natural principle for negation: $(\neg I) \vdash Val(\phi, \perp) \rightarrow \neg\phi$.
- B. These may be recovered by adding the following principle to constrain validity: $(Taut) \vdash Val(\top, \phi) \rightarrow \phi$.

2. Addressing the problem

This seems like a sad state of affairs, although perhaps not altogether unexpected. In this section, we look at a couple of ways of responding to it as a prelude to our fixed point construction. The natural thing to think is that one of (VP) or (VD) must go. Toward making such a decision, we might first try to make things more specific. Underlying each of the paradoxical arguments above is the ability to talk about expressions as objects, so some kind of syntax theory is in order. Let us assume, for now and for convenience, that that this theory is PA . The arithmetic vocabulary is $\{0, 1, +, \times\}$; we assume that $=$ is part of the logical vocabulary. An arithmetic sentence is one that involves only arithmetic (and of course, logical) vocabulary. We shall not, however, assume that our proof theory is restricted to arithmetic expressions: we shall admit expressions from any language that our coding system can accommodate.

Let us write $\phi \vdash_{PA} \psi$ to mean that ψ is derivable from ϕ from the axioms of PA and classical logic. Thus, we shall have:

Proposition 10. (i) If ϕ is arithmetic, then $\vdash_{PA} \phi$ iff ϕ is a theorem of PA .

(ii) If ϕ and ψ make no use of arithmetic vocabulary, then if $\phi \vdash \psi$, then $\phi \vdash_{PA} \psi$ (where \vdash is the ordinary derivability relation of first order logic).²

² Given ϕ and ψ do not involve arithmetic vocabulary, it is not the case, however, that: if $\phi \vdash_{PA} \psi$, then $\phi \vdash \psi$. For a counterexample, we have $\vdash_{PA} \exists x \exists y (x \neq y)$ while, $\nVdash \exists x \exists y (x \neq y)$ — we assume identity is a logical relation. This is arguably undesirable. We shall address this issue in the final system (see Section 3.1.1) by adding a relation symbol N whose intended interpretation is just the natural numbers. A similar fix could be applied here, but

This should give us sufficient power to say the things we need to say about our syntax. Matters will, of course, become more subtle for sentences which are in the overlapping area involving both arithmetic and non-arithmetic vocabulary.

We may then consider a language \mathcal{L} expanded with a validity predicate Val . In this (quite natural) context, (VP) and (VD) could be formulated as follows:

(VP_{PA}) If $\phi \vdash_{PA} \psi$, then $\vdash_{PA} Val(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$.

(VD_{PA}) $\phi, Val(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner) \vdash_{PA} \psi$.

The argument of the previous section then shows that, assuming we retain classical logic, one of these two must be sacrificed. However, at this point, we do not actually know that much about the property the Val predicate is representing. Informally, we know it is intended to represent validity or consequence, but what does that come to here? We look at two possible options: one proposed by Ketland; and another which generalises McGee's approach to necessity [Ketland, 2012, McGee, 1991].

2.1. Ketland's Approach. So the first and arguably most natural thing to try is taking Val at face value. We are considering the language of first order logic. We know what a consequence is in that language. Moreover, we have a completeness theorem which tells us that if ϕ is a *consequence* of ψ , then ϕ is *derivable from* ψ in any one of the myriad of proof systems for first order logic on the market. So let us define Val as follows:

Definition 11. We write $Val_K(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$ if ψ is derivable from the assumption of ϕ using some sound and complete proof procedure for first order logic (i.e., $\phi \vdash \psi$).

So this seems like a promising precisification: we are just lifting our definition of Val from the derivability relation of first order logic.

Proposition 12. (i) If ϕ and ψ do not involve not arithmetic vocabulary or the $=$ symbol, then $\phi \vdash_{PA} \psi$ iff $\vdash_{PA} Val_K(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$.³

the problem is not pressing for the current discussion and its application will only make the presentation much more complicated.

It may also be worth remarking on the distinction between a sentence that is not arithmetic and a sentence which does not involve arithmetic vocabulary. A sentence is *not arithmetic* if it makes use of non-logical vocabulary outside $\{0, 1, +, \times\}$. On the other hand, a sentence *does not involve arithmetic vocabulary* if it makes no use of $\{0, 1, +, \times\}$. Thus, sentences which only use the identity predicate are arithmetic, but do not involve arithmetic vocabulary.

³ By also excluding $=$, we avoid the problem noted in Proposition 10, but this also causes a significant restriction on VP_{PA} . This can be tidied up, but it is messy and we shall see a way of doing it in Section 3.1.1.

So there is a tight relationship between provability in the system PA and Val_K . Moreover, we may also show that:

Proposition 13. *For all sentences ϕ, ψ , it is the case that $\phi, Val_K(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner) \vdash_{PA} \psi$.*

Proof. By a simple generalisation of the argument in [Hájek and Pudlák, 1998], we see that $\vdash_{PA} Val_K(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner) \rightarrow (\phi \rightarrow \psi)$ for all ϕ, ψ . The proposition then follows from some propositional logic and the deduction theorem.

Remark. We observe that different choices in background syntax theory may affect this result.

Now Proposition 13 tells us VD_{PA} is satisfied under this precisification of validity. We also observe that, as construed, the Val_K predicate is definable in the language of PA and thus, since we are using classical logic, we must (assuming the consistency of PA) have sacrificed VP_{PA} . In fact, it is easy to give a counterexample as in Fact 15. Moreover, we see that Val_K has some counter-intuitive features.

Proposition 14. (i) $\vdash_{PA} Val_K(\ulcorner 0 = 0 \urcorner, \ulcorner 0 = 0 \urcorner)$;
(ii) $\nvdash_{PA} Val_K(\ulcorner 0 = 0 \urcorner, \ulcorner 1 + 1 = 2 \urcorner)$.

Since Val_K does not privilege arithmetic vocabulary, there is nothing stopping there being a model in which the arithmetic vocabulary is interpreted in such a way that $1 + 1 \neq 2$. Thus, the VP_{PA} will fail in such a situation. From this, a counter-example to VP is easy to construct.

Fact 15. $0 = 0 \vdash_{PA} 1 + 1 = 2$ but $\nvdash_{PA} Val_K(\ulcorner 0 = 0 \urcorner, \ulcorner 1 + 1 = 2 \urcorner)$.

2.1.1. *Rationale behind sacrificing VP_{PA} .* So now we can see where things stand on this proposal, can we provide also a satisfying philosophical explanation of why this should be the case? We discuss a way of doing this and then suggest a way in which this is not the end of the story. One of first things we might note about the presentation of our proof system above is that we have, so to speak, smuggled PA in as a subscript, when *really* we ought to think of its use as further assumptions upon which the derivation depends. So we should be writing

$$PA, \phi \vdash \psi$$

rather than

$$\phi \vdash_{PA} \psi.$$

We should not be thinking of ourselves as using a special proof theory which has PA as axioms; rather, we are just using first order logic with

some extra assumptions. This then puts *Val* perfectly in accord with our proof system \vdash , which seems like a good thing.

Now let us consider VP_{PA} in this context. It now looks like this:

(VP_{PA}^\dagger) If $PA, \phi \vdash \psi$, then $PA \vdash Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner)$.

But this does not look so much like the sort of principle we ought to countenance. In fact it could seem wrong, as the following more general principle illustrates.

(GP) If $\chi, \phi \vdash \psi$, then $\chi \vdash Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner)$.

This does not fit our intuitive gloss for *Val* at all. Suppose I had derived ψ from both χ and ϕ as assumptions. Then just because I am assuming χ , there is no reason to think, because of this, that I could derive ψ from ϕ .

Now we see that VP_{PA}^\dagger is essentially an instance of *GP*, so perhaps there is something fishy about it. I do not want to suggest that this is a proof of any kind — we already have one of those — but what we might have here is a philosophical motivation to see why our temptation to accept VP_{PA}^\dagger was misguided in the first place.

As I understand it, this is the motivation behind Ketland's response [Ketland, 2012].⁴ However, I think there is perhaps more to say. While VP_{PA}^\dagger is a set of instances of the problematic *GP*, it may be that each of these instances is actually okay — or at least in accord with our guiding motivations for the validity predicate. Perhaps we should not be treating *PA* as a set of assumptions like any other. Rather, since we require our syntax theory in order to even formulate these problems, perhaps we should take it that these axioms are not mere assumptions, but are essential: true even. Thus, we should not then consider cases where the axioms of the syntax theory are invalid. It is this kind of idea that may motivate us, so to speak, to bury *PA* as a subscript once more.

But there is a problem with this: if we make this move, then our proof system \vdash_{PA} is no longer in accord with our validity predicate *Val*. *Val* is just intended to give us the valid consequences of first order logic, but \vdash_{PA} also gives us the theorems of *PA*. This point motivates our next approach to precisifying validity.

2.2. Something like McGee. Rather than interpreting *Val* to mean first order derivability, we now interpret it to mean derivability in the system \vdash_{PA} . So we shall say:

⁴ We note that in [2012], Ketland uses a one-place predicate *Val* instead of the two place one used here and in [Beall and Murzi, forthcoming]. For his purposes, the difference is largely cosmetic. However, as we shall see in Section 3, it makes a significant and interesting difference to the resultant consequence relation in the fixed point construction. Thus we retain the two place version in this discussion.

Definition 16. $Val_{PA}(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$ if ψ is derivable in first order logic with PA as axioms (i.e., $\phi \vdash_{PA} \psi$).

Intuitively speaking, the idea is to admit the axioms of PA as *primitive*, since we are only interested in those cases where our syntax theory does what it is supposed to. Moreover, it is easy to see that the Val_{PA} is just a simple generalisation of Gödel's provability predicate for PA , $B_{PA}(x)$, and is thus representable by a Σ_1^0 formula of arithmetic. We then see that the following version of VP is obtained.

Proposition 17. *If $\phi \vdash_{PA} \psi$, then $\vdash_{PA} Val_{PA}(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner)$.*

Proof. This is a simple generalisation of the proof that if $\vdash_{PA} \phi$ then $\vdash_{PA} B_{PA} \ulcorner\phi\urcorner$ (see e.g., Boolos [1979]). \square

Now the Beall-Murzi paradox then tells us, since we are using classical logic, that we must abandon VD_{PA} ; i.e.,

Proposition 18. *There exists ϕ, ψ such that $\phi, Val_{PA}(\ulcorner\phi\urcorner, \ulcorner\psi\urcorner) \not\vdash_{PA} \psi$.*

Proof. We assume that PA is consistent. Using the diagonal lemma, let γ be such that

$$(2.1) \quad \vdash_{PA} \gamma \leftrightarrow \neg Val_{PA}(\ulcorner 0 = 0 \urcorner, \ulcorner \gamma \urcorner).$$

We then assume that $0 = 0, Val_{PA}(\ulcorner 0 = 0 \urcorner, \ulcorner \gamma \urcorner) \vdash_{PA} \gamma$; thus

$$(2.2) \quad \vdash_{PA} Val_{PA}(\ulcorner 0 = 0 \urcorner, \ulcorner \gamma \urcorner) \rightarrow \gamma.$$

Under this assumption, we work within PA and suppose that $\neg \gamma$. Then $Val_{PA}(\ulcorner 0 = 0 \urcorner, \ulcorner \gamma \urcorner)$ by (2.1); and so γ by (2.2), which is a contradiction, so γ . But then by Proposition 17, we get $Val_{PA}(\ulcorner 0 = 0 \urcorner, \ulcorner \gamma \urcorner)$; and by (2.1) $\neg \gamma$ and by (2.2) γ : contradiction. Thus our assumption is incorrect: $0 = 0, Val_{PA}(\ulcorner 0 = 0 \urcorner, \ulcorner \gamma \urcorner) \not\vdash_{PA} \gamma$. \square

Remark 19. This is obviously just a version of Example 7 where we use Val instead of \square and we assume consistency rather than attempt to refute it.

2.2.1. *Why abandon VD_{PA} ?* Perhaps there is less of a philosophically satisfying answer to why we should reject VD_{PA} . However, it is such an obvious and simple generalisation of Gödelian reasoning that perhaps we should be inured to such results. In [1991], McGee takes up much the same approach to the addition of a necessity predicate. He opts for the simplicity of a provability $B_{PA}(x)$ predicate for PA to represent necessity and thus ends up sacrificing the principle T while retaining Nec . This is analogous to a rejection of VD_{PA} while retaining VP_{PA} .

In contrast, Skyrms and Leitgeb have adopted approaches which retain principle *T* while sacrificing *Nec* [Skyrms, 1978, Leitgeb, 2008]. These approaches rely on fixed point construction which are similar to those of Kripke [1975]. They provide the inspiration for the construction in the following section.

2.3. Analysis. The basic lesson from all this is that we seem to have a forced choice between VD_{PA} and VP_{PA} .

- (1) If we use first order derivability (\vdash), we get VD_{PA} but sacrifice VP_{PA} .
- (2) If we use derivability over *PA* (\vdash_{PA}) we get VP_{PA} but lost VD_{PA} .

Our goal in this next section is to show how one might attempt to retain both *V* and *VP*.

3. Building the fixed point

Our goal is to construct a model in which we may support both principles:

(VP) If $\phi \vdash \psi$, then $\vdash Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner)$.

(VD) $\phi, Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \vdash \psi$.

We shall attempt to do this in as simple a fashion as possible. The solution provided is thus more of a prototype than a finished product — it could be improved upon, as we shall remark in the next section. However, the construction presented tackles the fundamental problem and a number of hurdles that any similar construction would need to address. The basic idea is to generalise Kripke's fixed point construction for a truth predicate to the current problem of validity [Kripke, 1975]. The result is less aesthetically pleasing, but nonetheless promising. In contrast to the previous section, we shall characterise validity semantically or model theoretically. The essential idea of the construction may be sketched as follows.

We start with an ordinary language \mathcal{L} to which we want to add the validity predicate, *Val*. We then consider the set of all countable models \mathcal{M} of \mathcal{L} : a canonical representative of each isomorphism-type would suffice.⁵ These models form the basis of the construction. At this stage, we have not specified anything about the content of the validity predicate.

The main difference from Kripke's truth construction is that rather than building a fixed point over a particular model of \mathcal{L} , we are building the fixed point over (practically) all of them. In order to talk about the syntax

⁵ The downward Löwenheim-Skolem theorem tells us that we only require countable models.

of the model, we expand our language \mathcal{L} with the language of arithmetic $\mathcal{L}_{Ar} = \{0, 1, +, \times\}$. We then expand and (possibly) extend each of the models to accommodate the standard interpretation of the arithmetic vocabulary. We want this vocabulary's interpretation to remain fixed across all of the models, while the rest of the vocabulary is re-interpreted. We denoted the set of these models by Mod . Finally, we expand the language with the two-place relation symbol Val . Call the resultant language \mathcal{L}_V . We form the sentences of \mathcal{L}_V , abbreviated $Sent_{\mathcal{L}_V}$, in the usual way. We shall assume that we have a simple arithmetic coding system $\ulcorner \cdot \urcorner: Sent_{\mathcal{L}_V} \rightarrow \omega$.

We then commence the inductive construction that will provide an extension for the validity predicate. We ground the construction by putting every instance of VD into the interpretation of the validity predicate. At the first level we then add to the extension of the Val predicate all those pairs of codes $\langle \phi, \psi \rangle$ such that every model says that either ϕ is false or ψ is true. For example, $\langle \phi \wedge \neg \phi, \phi \rangle$ will get into Γ_1^+ (provided ϕ does not use the Val predicate). Moreover, we take all of the pairs from Γ_0^+ (i.e., VD instances) into Γ_1^+ as we have already accepted them. We also form an anti-extension for Val in an analogous way. We write $\Gamma_1 = (\Gamma_1^+, \Gamma_1^-)$ to indicate this. At the next level we use Γ_1 as a parameter to figure out which sentences ψ follow from ϕ given that we already know that all the pairs in Γ_1^+ are valid and those in Γ_1^- are not. By adding these pairs to Γ_1^+ and Γ_2^+ , this allows us to calculate $\Gamma_2 = (\Gamma_2^+, \Gamma_2^-)$. Essentially, this next step allows us to grab the sentences which have two iterations of the Val predicate. For example, $\langle \phi, Val(\ulcorner \phi \wedge \neg \phi \urcorner, \ulcorner \phi \urcorner) \rangle$ would get into Γ_2^+ since $\langle \phi \wedge \neg \phi, \phi \rangle$ was in Γ_1^+ . We then continue on in this fashion. We shall demonstrate that the construction is non-decreasing and well-defined. Then by a standard cardinality argument, we claim that the construction reaches a fixed point. This will be the intended extension of the validity predicate. Finally, we show that our principles are satisfied in the model.

3.1. The construction

3.1.1. *The basis of models.* As noted in Section 2, the relation \vdash_{PA} demands that there is more than one object. This is not the sort of thing we would ordinarily think of as a logical truth. However, since this framework requires us to take our syntax theory with us in all models, we might think this is a necessary cost. Nonetheless, it would be nice if we could speak about the things that were not numbers and make no such demands with regard to them. Thus, while we are clearly going to have more than one natural number, there might not be more than one object that is not a natural number. We shall define our basis of models, Mod , to accommodate this intuition.

Our basic idea is to separate the domain $|M|$ of a model M into two parts: the natural number part and the rest of the objects. Let us call them the *number domain* and the *concrete domain*; and call their union the *total domain*. We want the consequence relation over the concrete domain to operate in the same manner as ordinary first order logic, so we shall demand that it contain at least one object. Moreover, we shall satisfy ourselves with merely countable concrete domains, since the downward Löwenheim-Skolem theorem tells us that we will not get anything new by considering larger domains.

Now if we make this move, there is a slight technical niggle. The arithmetic functions over the number domain will no longer take arbitrary objects from the total domain as arguments and thus fail to be properly defined. There are a couple of options: we could adapt our model theory to deal with partial functions; or we could use a relational signature for the arithmetic vocabulary. The former adds unnecessary complexity to the presentation, so we take up the latter. So assuming our arithmetic vocabulary was $\{0, 1, +, \times\}$, we replace the functions $+$ and \times by three-place relation symbols, P and M where

$$m + n = k \text{ iff } P(m, n, k)$$

and

$$m \times n = k \text{ iff } M(m, n, k).$$

So from now on, let $\mathcal{L}_{Ar} = \{0, 1, P, M\}$. Moreover, we shall also restrict our non-arithmetic vocabulary to relational signatures. We may then define formulae, using the arithmetic vocabulary, which allows us to determine if an object is in the number domain or the concrete domain. We might say that

$$N(x) \leftrightarrow_{Def} P(0, x, x)$$

and

$$C(x) \leftrightarrow_{Def} \neg N(x).$$

By restricting quantification to N and only using arithmetic vocabulary we may define *purely arithmetic sentences*; and *purely concrete sentences* in a similar fashion. We may now define our basis of models, Mod , as follows:

Definition 20. Take an arbitrary language \mathcal{L} . Let Mod_α be the set of all models M of $\mathcal{L} \cup \mathcal{L}_{Ar}$ where: the domain of M is $\omega + \alpha$; and \mathcal{L}_{Ar} is interpreted over ω in the standard way. Let

$$Mod = \bigcup_{1 \leq \alpha < \omega + 1} Mod_\alpha.$$

This gives us every way of interpreting the vocabulary of \mathcal{L} over every concrete domain of some countable cardinality ≥ 1 .⁶

3.1.2. A strong Kleene valuation function. We now define a semantic evaluation function which takes an extension/anti-extension pair $\Phi = \langle \Phi^+, \Phi^- \rangle$, a model $M \in Mod$ and a sentence ϕ from \mathcal{L}_V and returns a value in $\{0, 1\}$ if we have sufficient information to do so. We use a strong Kleene scheme to determine sufficiency. To save some space, we avoid satisfaction sequences and gear our valuation to a language which expands \mathcal{L}_V by a set of constant symbols for every element of the domain [Smullyan, 1968, Hodges, 1997, Priest, 2008]. Thus for a set of constant symbols C , we write $\mathcal{L}_V(C)$ for that expansion. We shall content ourselves to using the objects from the domain themselves as these constant symbols. We let \bar{m} stand for a tuple of elements from M 's domain $|M|$.

Definition 21. Given $\Phi = \langle \Phi^+, \Phi^- \rangle$ and $M \in Mod$ we let $v_{\cdot, \cdot}(\cdot) : \mathcal{P}(Sent_{\mathcal{L}_V})^2 \times Mod \times Sent_{\mathcal{L}_V} \rightarrow 2$ be a partial function which evaluates sentences $\phi \in Sent_{\mathcal{L}_V(|M|)}$ using an extension/anti-extension pair and a model as parameters.

$v_{\Phi, M}(\phi) = 1$ iff

- $\phi := R\bar{m}$ and $\bar{m} \in R^M$ where $R \in \mathcal{L}$;
- $\phi := R\bar{m}$ and $\bar{m} \in R^N$ where $R \in \mathcal{L}_{Ar}$;⁷
- $\phi := Val(m_1, m_2)$ and $m_1 = \ulcorner \psi \urcorner$ and $m_2 = \ulcorner \chi \urcorner$ and $\langle \psi, \chi \rangle \in \Phi^+$;
- $\phi := \neg \psi$ and $v_{\Phi, M}(\psi) = 0$;
- $\phi := \psi \wedge \chi$ and $v_{\Phi, M}(\psi) = 1$ and $v_{\Phi, M}(\chi) = 1$; or
- $\phi := \forall x \psi(x)$ and $\forall m \in |M|, v_{\Phi, M}(\psi(m/x)) = 1$

and $v_{\Phi, M}(\phi) = 0$ iff

- $\phi := R\bar{m}$ and $\bar{m} \notin R^M$ where $R \in \mathcal{L}$;
- $\phi := R\bar{m}$ and $\bar{m} \notin R^N$ where $R \in \mathcal{L}_{Ar}$;
- $\phi := Val(m_1, m_2)$ and $m_1 = \ulcorner \psi \urcorner$ and $m_2 = \ulcorner \chi \urcorner$ and $\langle \psi, \chi \rangle \in \Phi^-$;
- $\phi := \neg \psi$ and $v_{\Phi, M}(\psi) = 1$;
- $\phi := \psi \wedge \chi$ and $v_{\Phi, M}(\psi) = 0$ or $v_{\Phi, M}(\chi) = 0$; or
- $\phi := \forall x \psi(x)$ and $\exists m \in |M|$ such that $v_{\Phi, M}(\psi(m/x)) = 0$.

⁶ We could consider further restricting Mod such that the number domain is not permitted to be in the interpretation of relation symbols in \mathcal{L} : perhaps this would constitute some kind of category error.

⁷ R^N is the standard interpretation of R over ω .

Observe that if $\Phi = \langle \emptyset, \emptyset \rangle$, $v_{\Phi, \mathcal{M}}(Val(\ulcorner \phi^1, \urcorner \psi^1))$ is not defined. We shall then write $v_{\Phi, \mathcal{M}}(Val(\ulcorner \phi^1, \urcorner \psi^1)) = \infty$.

3.1.3. *A jump function that will not fly.* Our next task is to define the jump function, which will be the engine of our inductive construction. However, we first observe that the obvious generalisation of Kripke’s approach will not work. This gives some insight as to why a two-place *Val* predicate is more challenging to accommodate than its one-place cousin. Suppose we defined the jump as follows:

Definition 22. We let $j : \mathcal{P}(Sent_{\mathcal{L}_V})^2 \rightarrow \mathcal{P}(Sent_{\mathcal{L}_V})^2$ be such that for $\Phi = \langle \Phi^+, \Phi^- \rangle$

$$\begin{aligned} j(\Phi) &= \langle \{ \langle \phi, \psi \rangle \in Sent_{\mathcal{L}_V} \times Sent_{\mathcal{L}_V} \mid \forall \mathcal{M} \in Mod(v_{\Phi, \mathcal{M}}(\phi)) \\ &= 1 \rightarrow v_{\Phi, \mathcal{M}}(\psi) = 1 \} \rangle, \\ & \{ \langle \phi, \psi \rangle \in Sent_{\mathcal{L}_V} \times Sent_{\mathcal{L}_V} \mid \exists \mathcal{M} \in Mod(v_{\Phi, \mathcal{M}}(\phi)) \\ &= 1 \wedge v_{\Phi, \mathcal{M}}(\psi) \in \{0, \infty\} \} \rangle \end{aligned}$$

Suppose we start with the empty guess, $\langle \emptyset, \emptyset \rangle$ and then repeatedly apply the jump function j . If we want our construction to be non-decreasing, so that we can find a fixed point, then we want $j(\langle \emptyset, \emptyset \rangle)^+ \subseteq j \circ j(\langle \emptyset, \emptyset \rangle)^+$. Unfortunately, this does not occur. Consider the pair of sentences: $\langle Val(\ulcorner \phi^1, \urcorner \phi^1), \ulcorner \phi \wedge \neg \phi^1 \urcorner \rangle$. We see that $\langle \phi, \phi \rangle \in j(\langle \emptyset, \emptyset \rangle)^+$, since $\phi \models \phi$. Moreover since no model \mathcal{M} is such that $v_{\langle \emptyset, \emptyset \rangle, \mathcal{M}}(Val(\ulcorner \phi^1, \urcorner \phi^1)) = 1$, we also have $\langle Val(\ulcorner \phi^1, \urcorner \phi^1), \ulcorner \phi \wedge \neg \phi^1 \urcorner \rangle \in j(\langle \emptyset, \emptyset \rangle)^+$ (for trivial reasons).

But we can see when we consider the next iteration $j \circ j(\langle \emptyset, \emptyset \rangle)^+$, every model \mathcal{M} is such that

$$v_{\Gamma_1, \mathcal{M}}(Val(\ulcorner \phi^1, \urcorner \phi^1)) = 1$$

and

$$v_{\Gamma_1, \mathcal{M}}(Val(\ulcorner \phi \wedge \neg \phi^1 \urcorner)) = 0.$$

Thus $\langle Val(\ulcorner \phi^1, \urcorner \phi^1), \ulcorner \phi \wedge \neg \phi^1 \urcorner \rangle \notin j \circ j(\langle \emptyset, \emptyset \rangle)^+$. In fact, $\langle Val(\ulcorner \phi^1, \urcorner \phi^1), \ulcorner \phi \wedge \neg \phi^1 \urcorner \rangle \in j \circ j(\langle \emptyset, \emptyset \rangle)^-$.

This means that we won’t be able to get a fixed point with this jump function: its failure will be particularly apparent at limit levels of the construction. Moreover, it can be shown that the obvious generalisations of other traditional jump operations — like those of Kremer, Halbach and Horsten and Priest — will also suffer similar problems [Kremer, 1988, Halbach and Horsten, 2006, Priest, 1979]. At this point, we might think to adopt a revision theoretic strategy as in [Gupta and Belnap, 1993]. However, there is a way of revising the jump function definition which does work and is related to contemporary work in logics of truth [Ripley, Forthcoming].

3.1.4. *A jump function that does work.* Let's define the jump function somewhat differently.

Definition 22. We let $j_V : \mathcal{P}(\text{Sent}_{\mathcal{L}_V})^2 \rightarrow \mathcal{P}(\text{Sent}_{\mathcal{L}_V})^2$ be such that for $\Phi = \langle \Phi^+, \Phi^- \rangle$

$$\begin{aligned} j_V(\Phi) &= \langle \{ \langle \phi, \psi \rangle \in (\text{Sent}_{\mathcal{L}_V})^2 \mid \forall \mathcal{M} \in \text{Mod}(v_{\Phi, \mathcal{M}}(\phi)) \\ &\quad \in \{ \infty, 1 \} \rightarrow v_{\Phi, \mathcal{M}}(\psi) = 1 \} \}, \\ &\quad \{ \langle \phi, \psi \rangle \in (\text{Sent}_{\mathcal{L}_V})^2 \mid \exists \mathcal{M} \in \text{Mod}(v_{\Phi, \mathcal{M}}(\phi)) \\ &\quad = 1 \wedge v_{\Phi, \mathcal{M}}(\psi) = 0 \} \rangle \end{aligned}$$

This jump function gets around the problem above, because $v_{\emptyset, \mathcal{M}}(\text{Val}(\ulcorner \phi \urcorner, \ulcorner \phi \urcorner)) = \infty$ and thus the fact that $v_{\emptyset, \mathcal{M}}(\phi \wedge \neg \phi) = 0$ means that

$$\langle \text{Val}(\ulcorner \phi \urcorner, \ulcorner \phi \urcorner), \ulcorner \phi \wedge \neg \phi \urcorner \rangle \notin \Gamma_1^!$$

Remark 24. The idea for this is drawn from Ripley's notion of a tolerant-strict consequence [Ripley, Forthcoming]. However, as is often the case with these matters, there is a sense in which this consequence relation is particularly strict. We are only admitting those consequences where we are, so to speak, sure about the antecedent and consequent. The cost of this is reflexivity.

3.1.5. *The hierarchy.* We are now ready to provide our inductive definition. Recalling our goal to ensure that principles *VP* and *VD* are upheld, we set up the induction so that *VD* will be satisfied. We do this by simply putting it in at the beginning.

Definition 25. Let (i) $\coprod_{\alpha < \beta} \Phi_\alpha = \langle \bigcup_{\alpha < \beta} \Phi_\alpha^+, \bigcup_{\alpha < \beta} \Phi_\alpha^- \rangle$;
(ii) $\Psi \sqcup \Phi = \langle \Psi^+ \cup \Phi^+, \Psi^- \cup \Phi^- \rangle$;
(iii) $\Psi \sqsubseteq \Phi$ iff $\Psi^+ \subseteq \Phi^+$ and $\Psi^- \subseteq \Phi^-$; and
(iv) Ψ be called *consistent* if $\Psi^+ \cap \Psi^- = \emptyset$.

Definition 26. Let $\Gamma_\cdot : \text{On} \rightarrow \mathcal{P}(\text{Sent}_{\mathcal{L}_V})^2$ be defined by transfinite recursion as follows:

$$\begin{aligned} \Gamma_0 &= \langle \{ \langle \phi \wedge \text{Val}(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner), \psi \rangle \mid \phi, \psi \in \text{Sent}_{\mathcal{L}_V} \}, \emptyset \rangle \\ \Gamma_\alpha &= j_V \left(\coprod_{\beta < \alpha} \Gamma_\beta \right) \sqcup \coprod_{\beta < \alpha} \Gamma_\beta \end{aligned}$$

It is trivial to show that this construction is non-decreasing: it is built into the definition. However, it is possible that the jump may become undefined

if we ever get to a position where for some $\phi, \psi \in \text{Sent}_{\mathcal{L}_V}$ and ordinal α , $\langle \phi, \psi \rangle \in \Gamma_\alpha^+$ and $\langle \phi, \psi \rangle \in \Gamma_\alpha^-$. If this occurred, we could not use the valuation function, v , described above.

Proposition 27. *Suppose $\Phi \sqsubseteq \Psi$ where Φ and Ψ are consistent. Then*

- (1) if $v_{\Phi, \mathcal{M}}(\phi) = 1, v_{\Psi, \mathcal{M}}(\phi) = 1$; and
- (2) if $v_{\Phi, \mathcal{M}}(\phi) = 0, v_{\Psi, \mathcal{M}}(\phi) = 0$;

Proof. By induction on the complexity of formulae. □

The following claim is required for the main lemma which follows.

Claim 28. Let \mathcal{L} be an arbitrary countable language and let $\ulcorner \cdot \urcorner : \text{Sent}_{\mathcal{L}_V} \rightarrow \omega$ be an arbitrary coding system for \mathcal{L}_V . Then there is no sequence $(\chi_n)_{n \in \omega}$ of sentences of \mathcal{L}_V such that for all $n \in \omega$:

$$\chi_n \text{ is } \chi_{n+1} \wedge \text{Val}(\ulcorner \chi_{n+1} \urcorner, \ulcorner \psi \urcorner).$$

Proof. Suppose not and let $(\chi)_{n \in \omega}$ witness this. Then we have

$$\begin{aligned} \chi_0 &= \chi_1 \wedge \text{Val}(\ulcorner \chi_1 \urcorner, \ulcorner \psi \urcorner) \\ &= \chi_2 \wedge \text{Val}(\ulcorner \chi_2 \urcorner, \ulcorner \psi \urcorner) \wedge \text{Val}(\ulcorner \chi_1 \urcorner, \ulcorner \psi \urcorner) \\ &\vdots \\ &= \chi_n \wedge \text{Val}(\ulcorner \chi_n \urcorner, \ulcorner \psi \urcorner) \wedge \text{Val}(\ulcorner \chi_{n-1} \urcorner, \ulcorner \psi \urcorner) \wedge \bigcup \dots \wedge \text{Val}(\ulcorner \chi_1 \urcorner, \ulcorner \psi \urcorner) \\ &\vdots \end{aligned}$$

We then observe that any sentence of \mathcal{L}_V must contain a fixed and finite number of conjunctions. Suppose that χ_0 contains m many conjunctions. Then we see that

$$\chi_0 = \chi_{m+1} \wedge \text{Val}(\ulcorner \chi_{m+1} \urcorner, \ulcorner \psi \urcorner) \wedge \dots \wedge \text{Val}(\ulcorner \chi_0 \urcorner, \ulcorner \psi \urcorner)$$

and thus χ_0 contains at least $m + 1$ conjunctions: contradiction. □

Lemma 29. *For all α and $\phi, \psi \in \text{Sent}_{\mathcal{L}_V}$ it is not the case that $\langle \phi, \psi \rangle \in \Gamma_\alpha^+$ and $\langle \phi, \psi \rangle \in \Gamma_\alpha^-$.*

Proof. Suppose not and let α be the least ordinal such that for some $\phi, \psi \in \text{Sent}_{\mathcal{L}_V}$ we have $\langle \phi, \psi \rangle \in \Gamma_\alpha^+$ and $\langle \phi, \psi \rangle \in \Gamma_\alpha^-$.

Case 1: Suppose $0 < \zeta < \alpha$ is the least ordinal such that $\langle \phi, \psi \rangle \in \Gamma_\zeta^+$, and α is the least ordinal such that $\langle \phi, \psi \rangle \in \Gamma_\alpha^-$. Then $\langle \phi, \psi \rangle \notin \bigcup_{\beta < \alpha} \Gamma_\beta^-$, so the construction is well-defined below α and thus, for some \mathcal{M} we

have $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\phi) = 1$ while $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\psi) = 0$. We also see by assumption that either $v_{\sqcup_{\xi < \zeta} \Gamma_{\xi, \mathcal{M}}}(\phi) = 0$ or $v_{\sqcup_{\xi < \zeta} \Gamma_{\xi, \mathcal{M}}}(\psi) = 1$. Suppose $v_{\sqcup_{\xi < \zeta} \Gamma_{\xi, \mathcal{M}}}(\phi) = 0$. Then since $\sqcup_{\xi < \zeta} \Gamma_{\xi} \sqsubseteq \sqcup_{\beta < \alpha} \Gamma_{\beta}$, $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\phi) = 0$: contradiction. Suppose $v_{\sqcup_{\xi < \zeta} \Gamma_{\xi, \mathcal{M}}}(\psi) = 1$. Then since $\sqcup_{\xi < \zeta} \Gamma_{\xi} \sqsubseteq \sqcup_{\beta < \alpha} \Gamma_{\beta}$, $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\psi) = 1$: contradiction.

Case 2: Suppose $\zeta < \alpha$ is the least ordinal such that $\langle \phi, \psi \rangle \in \Gamma_{\zeta}^-$, and α is the least ordinal such that $\langle \phi, \psi \rangle \in \Gamma_{\alpha}^+$. Then by assumption $\langle \phi, \psi \rangle \notin \bigcup_{\beta < \alpha} \Gamma_{\beta}^+$ so the construction is well-defined below α . Thus we can see that $v_{\sqcup_{\xi < \zeta} \Gamma_{\xi, \mathcal{M}}}(\phi) = 1$ and $v_{\sqcup_{\xi < \zeta} \Gamma_{\xi, \mathcal{M}}}(\psi) = 0$ for some \mathcal{M} . Fix such an \mathcal{M} . Moreover, we also have $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\phi) = 0$ or $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\phi)\psi = 1$. Suppose $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\phi) = 0$. Then since $\sqcup_{\xi < \zeta} \Gamma_{\xi} \sqsubseteq \sqcup_{\beta < \alpha} \Gamma_{\beta}$, $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\phi) = 1$: contradiction. Suppose $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\psi) = 1$. Then since $\sqcup_{\xi < \zeta} \Gamma_{\xi} \sqsubseteq \sqcup_{\beta < \alpha} \Gamma_{\beta}$, $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\psi) = 0$: contradiction.

Case 3: Suppose $\langle \phi, \psi \rangle \in \Gamma_0^+$ and α is the least ordinal such that $\langle \phi, \psi \rangle \in \Gamma_{\alpha}^-$. We see that ϕ must be of the form $\chi_0 \wedge \text{Val}(\ulcorner \chi_0 \urcorner, \ulcorner \psi \urcorner)$. Moreover, for some \mathcal{M} we have $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\chi_0 \wedge \text{Val}(\ulcorner \chi_0 \urcorner, \ulcorner \psi \urcorner)) = 1$ while $v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\psi) = 0$. Fix such an \mathcal{M} . But then

$$v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\chi_0) = v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\text{Val}(\ulcorner \chi_0 \urcorner, \ulcorner \psi \urcorner)) = 1$$

and so $\langle \chi_0, \psi \rangle \in \bigcup_{\beta < \alpha} \Gamma_{\beta}^+$. Now either: there is some least $\xi > 0$ such that $\langle \chi_0, \psi \rangle \in \Gamma_{\xi}^+$ or $\langle \chi_0, \psi \rangle \in \Gamma_0^+$. In former situation, the argument from Case 1 or Case 2 may be applied, since it is clear that $\langle \chi_0, \psi \rangle \in \Gamma_{\alpha}^-$. In the latter situation, χ_0 must be a sentence of the form $\chi_1 \wedge \text{Val}(\ulcorner \chi_1 \urcorner, \ulcorner \psi \urcorner)$; and thus,

$$v_{\sqcup_{\beta < \alpha} \Gamma_{\beta, \mathcal{M}}}(\chi_1) = v_{\sqcup_{\beta < \alpha, \mathcal{M}}}(\text{Val}(\ulcorner \chi_1 \urcorner, \ulcorner \psi \urcorner)) = 1$$

and so $\langle \chi_1, \psi \rangle \in \bigcup_{\beta < \alpha} \Gamma_{\beta}^+$. Clearly the argument for χ_0 may be applied to χ_1 and we may end up with a sequence χ_0, \dots, χ_n . However by Claim 28, we see that this sequence must terminate after a finite number of steps. Thus, we shall eventually end up being able to apply the argument from Case 1 or Case 2. \square

Corollary 30. *The sequence $(\Gamma_{\alpha})_{\alpha \in \mathcal{O}_n}$ is well-defined.*

From here cardinality considerations imply the existence of a fixed point and we may define our intended extension for the validity predicate as follows:

Definition 31. Let Γ_{Val} be Γ_{α} for the least α such that $\Gamma_{\alpha} = \Gamma_{\alpha+1}$.

3.1.6. *Establishing the principles.* Let \top be an arbitrary sentence such that $v_{\emptyset, \mathcal{M}}(\top) = 1$ for all $\mathcal{M} \in \text{Mod}$; and let \perp be an arbitrary sentence such that $v_{\emptyset, \mathcal{M}}(\perp) = 0$ for all $\mathcal{M} \in \text{Mod}$.

- Proposition 32.** (i) $\langle \phi, \psi \rangle \in \Gamma_{Val}^+$ iff $\langle \ulcorner \top \urcorner, Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \rangle \in \Gamma_{Val}^+$;
(ii) For purely concrete ϕ, ψ , $\langle \phi, \psi \rangle \in \Gamma_{Val}^+$ iff $\phi \vdash \psi$ iff $\langle \phi, \psi \rangle \notin \Gamma_{Val}^-$.
(iii) For ϕ is purely arithmetic, $Val(\ulcorner \top \urcorner, \ulcorner \phi \urcorner) = 1$ iff $\mathbb{N} \models \phi$.

Proof. (i) (\rightarrow) Suppose $\langle \phi, \psi \rangle \in \Gamma_\beta^+ \subseteq \Gamma_{Val}^+$ for some β . Then every model $\mathcal{M} \in Mod$ is such that $v_{\Gamma_\beta, \mathcal{M}}(Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner)) = 1$, thus $\langle \ulcorner \top \urcorner, Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \rangle \in \Gamma_{\beta+1}^+ \subseteq \Gamma_{Val}^+$. (\leftarrow) Suppose β is the least ordinal such that $\langle \ulcorner \top \urcorner, Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \rangle \in \Gamma_\beta^+$. Then $\beta > 0$ and for all $\mathcal{M} \in Mod$, we have $v_{\sqcup_{\gamma < \beta} \Gamma_\gamma, \mathcal{M}}(Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner)) = 1$. But then $\langle \phi, \psi \rangle \in \bigcup_{\gamma < \beta} \Gamma_\gamma^+ \subseteq \Gamma_{Val}^+$.

(ii) Any putative consequence $\langle \phi, \psi \rangle$ where ϕ, ψ are purely concrete is in $\Gamma_1^+ \subseteq \Gamma_{Val}^+$ iff it is valid; and in $\Gamma_1^- \subseteq \Gamma_{Val}^-$ iff it is not valid.

(iii) If ϕ is arithmetic, then its interpretation is fixed across all models $\mathcal{M} \in Mod$. □

We may also establish weak versions of monotonicity, reflexivity and transitivity within Γ_{Val} .

- Proposition 33.** (i) If $\phi, \psi \in Sent_{(\mathcal{L} \cup \mathcal{L}_{Ar}) \setminus \mathcal{L}_V}$ and $\langle \phi, \psi \rangle \in \Gamma_{Val}^+$, then $\langle \delta \wedge \phi, \psi \rangle \in \Gamma_{Val}^+$;
(ii) $\langle \phi, \phi \rangle \in \Gamma_{Val}^+$ if $\phi \in Sent_{(\mathcal{L} \cup \mathcal{L}_{Ar}) \setminus \mathcal{L}_V}$; and
(iii) If $\phi, \psi, \chi \in Sent_{(\mathcal{L} \cup \mathcal{L}_{Ar}) \setminus \mathcal{L}_V}$, $\langle \phi, \psi \rangle \in \Gamma_{Val}^+$ and $\langle \psi, \chi \rangle \in \Gamma_{Val}^+$, then $\langle \phi, \chi \rangle \in \Gamma_{Val}^+$.

We are almost ready to define our target consequence relation, but there is a slight hitch that we need to address first. We would like our consequence relation to enjoy permutation of premises. Then we can write $\Delta \not\equiv \phi$ where Δ is a finite set of sentences from \mathcal{L}_V . Given that our Γ_{Val}^+ only takes pairs, it seems natural to use conjunction to weld together proxies for these finite sets. So given $\Delta = \{\gamma_1, \dots, \gamma_n\}$, we shall use $\gamma_1 \wedge \dots \wedge \gamma_n$ in its place. But it is critical that we could permute the order of these conjunction. So given $\sigma : n \rightarrow n$, a permutation, we should also be able to use $\gamma_{\sigma(1)} \wedge \dots \wedge \gamma_{\sigma(n)}$ to represent Δ and get the same result.

But our definition fails this at the first level, Γ_0^+ . Let $\phi, \psi \in Sent_{\mathcal{L}_V}$. Then $\langle \phi \wedge Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner), \psi \rangle \in \Gamma_0^+$, but $\langle Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \wedge \phi, \psi \rangle \notin \Gamma_0^+$. We added the first of these pairs to ensure that (VD) was satisfied, but given the syntactic nature of their addition, we do not get the second pair. This is easily remedied: we simply add them at the beginning too. Thus we amend the construction such that

$$\Gamma_0^+ = \{ \langle \phi \wedge Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner), \psi \rangle \mid \phi, \psi \in Sent_{\mathcal{L}_V} \} \cup \{ \langle Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \wedge \phi, \psi \rangle \mid \phi, \psi \in Sent_{\mathcal{L}_V} \}.$$

From now on, we shall take it that Γ_{Val} has been defined accordingly. It is easy to see that the proofs above go through after this change with at most minor revisions.

Definition 34. Let $\Delta = \{\gamma_1, \dots, \gamma_n\}$ be finite, then we let $\Delta \not\equiv_{Val} \phi$ iff $\langle \gamma_1 \wedge \dots \wedge \gamma_n, \phi \rangle \in \Gamma_{Val}^+$. We also write $\not\equiv_{Val} \phi$ iff $\langle \top, \phi \rangle \in \Gamma_{Val}^+$.

Theorem 35. (i) If $\phi \not\equiv_{Val} \psi$, then $\not\equiv_{Val} Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner)$;

(ii) $\phi, Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \not\equiv_{Val} \psi$.

Proof. (i) This follows from Proposition 32 (i).

(ii) By construction, $\langle \phi \wedge Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner), \psi \rangle \in \Gamma_0^+ \subseteq \Gamma_{Val}^+$. Thus $\phi, Val(\ulcorner \phi \urcorner, \ulcorner \psi \urcorner) \not\equiv_{Val} \psi$.

3.1.7. *What happens to the Beall-Murzi sentence?* To give a better idea what is going on in this construction, we illustrate what happens to the Beall-Murzi sentence. Recall that it is a sentence π such that we have established using the diagonal lemma that:

$$\vdash \pi \leftrightarrow Val(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner).$$

Given given the way we have set things up, we might be tempted to think that $\langle \pi, \perp \rangle$ actually gets into the extension of the validity predicate. This would be problematic since it would soon force $\langle \pi, \perp \rangle$ into the anti-extension causing the construction to jam and fail. This is not the case, but let us see how we might informally (and incorrectly) reason to this conclusion.

(1) We observe that, by definition:

$$(3.1) \quad \langle \pi \wedge Val(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner), \perp \rangle \in \Gamma_0^+$$

since it is an instance of *VD*.

(2) We then observe that π is equivalent to $Val(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner)$. There is no difference between saying π or $\pi \wedge Val(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner)$: they *mean the same thing*.

(3) Thus, (3.1) is actually telling us that $\langle \pi, \perp \rangle \in \Gamma_0^+ \subseteq \Gamma_{Val}^+$.

However, our construction is more restrictive than the reasoning used to get from step (2.) to step (3.). The fact that π and $Val(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner)$ mean the same thing is not sufficient. We are actually trying to make two moves here: first, on the basis of equivalence, we move to saying that $\langle \pi \wedge \pi, \perp \rangle \in \Gamma_0^+$; second, we argue that $\pi \wedge \pi$ means the same thing as π , we also have $\langle \pi, \perp \rangle \in \Gamma_0^+$. With regard to the first of these moves, we require syntactic equivalence between π and $Val(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner)$. However, our language has a paucity of terms and so this cannot occur [Richard G. Heck, 2007]. None-

theless, we may wish to move into richer languages, like *PRA*, where such terms are available. But even then, with regard to the second move, we are employing a form of contraction. While this may seem intuitively correct, the construction does not admit this rule and the argument is blocked at this point. The initial stage of the construction only contains consequences whose antecedents are conjunctions.

3.2. Remarks. We close with some remarks about the shortcomings of this definition and some suggestions for future investigation.

3.2.1. The weakness of strong Kleene systems. We should first note that in the broader church of the semantic paradox literature, our solution is essentially strong Kleene and as such has all the well known weaknesses of that approach. Most obviously, the conditional is so weak that we cannot always guarantee that it is reflexive. For example, we shall not get $\not\equiv_{\text{val}} \beta \rightarrow \beta$. Such problems are well known. We have two choices here: accept the weakness; or strengthen our conditional. Personally, I find the former option more palatable; however, there is good reason to think that the techniques developed by Field [2008] or Beall [2009] could be useful. In the former case, we might adopt a revision theoretic construction to strengthen the conditional. The cost of this is computational complexity and a less well-motivated story about how we have constructed our definition. In the latter case, we might take up a paraconsistent logic, which rejects and controls explosion. The cost here is heuristic complexity and a deep revision of our background logic. Both approaches are worth investigating. Moreover, the framework above could be a useful basis for such an investigation.

3.2.2. Not a genuine consequence relation. We also observe that $\not\equiv_{\text{val}}$ is not a genuine consequence relation in the sense that we do not have unrestricted reflexivity, monotonicity or transitivity. Monotonicity is lost at the first level when we let Γ_0^+ be all the instances of *VD* without the possibility of extra premises. This is not, however, difficult to recover. We simply add all extra cases to Γ_0^+ and keep the rest of the definition. The proof of Lemma 29 becomes slightly longer, but the adaptations are the obvious ones. Incorporating this change would have just made the exposition more difficult.

Reflexivity and transitivity, on the other hand, fail for deeper reasons. We lost reflexivity when we revised our jump function definition. We sacrificed it in order to get a simple non-decreasing sequence. We lost transitivity by forcing *VD* into the opening of the inductive construction.⁸ There are a couple of things to say here. First, we observe that the failure of reflexivity in the consequence relation is an exact parallel of the problems we had with

⁸ We have $\pi \wedge \text{Val}(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner) \not\equiv_{\text{val}} \perp$ and $\perp \not\equiv_{\text{val}} \phi$, but we do not have $\pi \wedge \text{Val}(\ulcorner \pi \urcorner, \ulcorner \perp \urcorner) \not\equiv_{\text{val}} \phi$.

the conditional. Thus we may accept it or try to strengthen the system. If we take the latter course, then revision theory or paraconsistency could be fertile options. On the other hand, we observe that generalised consequence relations that sacrifice reflexivity or transitivity have recently received greater attention in relation to semantic paradoxes [Ripley, Forthcoming]. We may be seeing further evidence for why such approaches are natural in the context of semantic paradox.

3.2.3. *The opening of the construction is very artificial.* The manner in which ρ is treated and the failure of monotonicity are evidence of a certain artificiality at the basis of the construction. It is extremely syntactic and as such, it ignores a lot of intuitive features that we may have expected the construction to pick up. The reason for this is that we essentially forced VD into the construction from the beginning and demanded it stay there. We did not consider how to make it fit naturally, we just made sure that we satisfied the goals for the construction set out at the beginning of Section 3. However, this seems like a natural place for more sophisticated constructions to start tinkering. There is a lot more that can be done here. Of course, there is a strict upper bound in the sense that if we closed Γ_0^+ under logical equivalence, then the argument in Section 3.1.7 would go through and the construction would fail.

On the other hand, we could consider jettisoning VD and just starting with $\Gamma_0^+ = \emptyset$. I think this is probably the most natural construction that can be formed in this way, in the sense that this is what you might build without the initial desiderata. However, VD is a plausible principle and without it, it might seem that we have just collapsed back into the McGee style option proposed in Section 2.2. This is, however, a little misleading: there is some gain. With the McGee style system we get Proposition 17, while with the fixed point we also get a converse as shown in 32(i).

3.3. Conclusion. In this paper, we have isolated the Beall-Murzi paradox among other semantic paradoxes, discussed canonical approaches to its solution; and provided a fixed point definition which satisfies both VP and VD .

References

- BEALL, J.C. (2009). *Spandrels of Truth*. Oxford University Press.
 BEALL, J.C. and MURZI, J. (2013). Two Flavors of Curry's Paradox. *Journal of Philosophy* 110 (3), p. 143-165.
 BOOLOS, G. (1979). *The Unprovability of Consistency*. CUP, Cambridge.
 FIELD, H. (2008). *Saving Truth from Paradox*. OUP, Oxford.

- GUPTA, A. and BELNAP, N. (1993). *The Revision Theory of Truth*. MIT Press, Cambridge.
- HÁJEK, P. and PUDLÁK, P. (1998). *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer.
- HALBACH, V. and HORSTEN, L. (2006). Axiomatizing kripke's theory of truth. *Journal of Symbolic Logic*, 71: 677-712.
- HODGES, W. (1997). *A Shorter Model Theory*. CUP, Cambridge.
- HECK, R.G. Jr. (2007). Self-reference and the languages of arithmetic. *Philosophia Mathematica*, pp. 1-29.
- KETLAND, J. (2012). Validity as a primitive. *Analysis*, 72(3):421-430.
- KREMER, M. (1988). Kripke and the logic of truth. *The Journal of Philosophical Logic*, 17(3):225-278.
- KRIPKE, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, 72:690-716.
- LEITGEB, H. (2008). Towards a logic of type-free modality and truth. In C. Dimitracopoulos et al., editor, *Logic Colloquium 05, Lecture Notes in Logic*. Cambridge University Press, Cambridge.
- MCGEE, V. (1991). *Truth, Vagueness, and Paradox*. Hackett, Indianapolis.
- MONTAGUE, R. (1963). Syntactical treatments of modality. *Acta Philosophica Fennica*, 16:153-167.
- PRIEST, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, 8: 219-241.
- PRIEST, G. (2008). *An Introduction to Non-Classical Logic: From If to Is*. Cambridge University Press, Melbourne.
- RIPLEY, D. Conservatively extending classical logic with transparent truth. *The Review of Symbolic Logic*, Forthcoming.
- SKYRMS, B. (1978). An immaculate conception of modality or how to confuse use and mention. *Journal of Philosophy*, 75:368-387.
- SMULLYAN, R.M. (1968). *First-Order Logic*. Dover, New York.

Toby MEADOWS
School of Mathematics,
University of Aberdeen