THE REVIEW OF SYMBOLIC LOGIC
Volume 7, Number 4, December 2014

ON PATHOLOGICAL TRUTHS

LUCAS ROSENBLATT
CONICET and University of Buenos Aires
DAMIÁN E. SZMUC
University of Buenos Aires

Abstract. In Kripke's classic paper on truth it is argued that by adding a new semantic category different from truth and falsity it is possible to have a language with its own truth predicate. A substantial problem with this approach is that it lacks the expressive resources to characterize those sentences which fall under the new category. The main goal of this paper is to offer a refinement of Kripke's approach in which this difficulty does not arise. We tackle this characterization problem by letting certain sentences belong to more than one semantic category. We also consider the prospect of generalizing this framework to deal with languages containing vague predicates.

§1. Introduction. A defect usually attributed to Kripke (1975)'s approach to truth is that his theory is expressed in a language that cannot appropriately specify the status of some sentences expressible in that same language. This is sometimes called the 'semantic characterization problem'. A notable example of this is a Liar sentence λ which is its own negation.[1] In Kripke's theory it is not possible to truly express that the Liar sentence is neither true nor false. There is no such "sentence-classifier" capable of performing this task in the language of the theory.

The difficulty cannot be straightforwardly overcome by introducing as a primitive symbol a specific operator designed for the task. Whether such an operator is consistently definable depends on the sort of valuation schema employed. Kripke's theory of truth can be formulated using different valuation schemata. If the Weak Kleene schema is used, it is in fact possible to introduce such an operator (see Gupta & Martin, 1984 and Gupta & Belnap, 1993, ch. 2 for the details). However, there might be good reasons to prefer other valuation schemata. For example, if one thinks that the Liar does not fail to express a proposition, and hence is not meaningless, there is a sense in which the Strong Kleene schema provides a conceptually better motivated logic.[2] Unfortunately, it turns out that there is no simple way of adding a sentence-classifying operator over this schema without generating an inconsistency. Assume that we are in a three-valued setting and that we characterize an operator N by stipulating that for every valuation v the following holds:

$$v(N\phi) = \begin{cases} 1 & \text{if } v(\phi) = \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

Received: February 14, 2014.

[1] We will not concern ourselves for now over the way in which self-referentiality is achieved. If the reader prefers, she can take λ to be *equivalent* to the sentence stating that λ is not true.

[2] We will assume that the reader is familiar with the details of these schemata and we will ignore here other valuation schemata such as Supervaluationism.

© Association for Symbolic Logic, 2014

doi:10.1017/S1755020314000239

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

Once this operator is available it is possible to truly say that any Liar sentence λ is neither true nor false: $N\lambda$. However, this comes at a high cost. For let δ be the sentence $NTr\langle\delta\rangle \vee \neg Tr\langle\delta\rangle$. It is easy to see that there is no consistent assignment of truth values to δ if $v(NTr\langle\delta\rangle \vee \neg Tr\langle\delta\rangle) = max\{v(NTr\langle\delta\rangle), v(\neg Tr\langle\delta\rangle)\}$, which is how $\vee$ is defined in the Strong Kleene schema.

This does not mean that there is no way of introducing such an operator, even if the Strong Kleene schema is employed. An idea that has been recently explored by Beall (2006) and Beall (2009) is to separate the notion of being-neither-true-nor-false from the notion of being pathological or ungrounded.[3] The idea is as follows. Assume that we have a certain operator π available in the language. Since the language already contains the truth predicate, we can form the following sentences:

$$\delta_1 := \pi Tr\langle\delta_1\rangle$$
$$\delta_2 := \neg\pi Tr\langle\delta_2\rangle.^4$$

If we interpret π as 'is-neither-true-nor-false', there is no major problem with these sentences: δ_1 is plainly false and δ_2 is plainly true. For let δ_1 be neither-true-nor-false, then what it says is the case, and hence it is true. But if it is true, then what it says is the case again, and so it is neither-true-nor-false. So δ_1 can only be false. For δ_2 assume that it is neither-true-nor-false, then what it says is not the case, and so it is false. But if it is false, then what it says is not the case, and so it is neither-true-nor-false. So δ_2 can only be true.

However, if we think of π as 'pathological', a different diagnosis is available. According to any reasonable theory of pathologicality, both δ_1 and δ_2 are pathological. So we reason as follows. First, given that δ_1 is pathological, what it says is indeed the case. Thus, it is true. Hence, δ_1 is both pathological *and* true. Second, given that δ_2 is pathological, what it says is not the case, and therefore it is false. Hence, δ_2 is both pathological *and* false.

Beall thinks that we can deal with these sentences in a non-paraconsistent setting by allowing certain values (other than truth and falsity) to overlap. More specifically, assuming *pathological* is a third semantic category different from truth and falsity, some sentences (such as δ_1) will be categorized both as true and pathological, while others (such as δ_2) as false and pathological.

We think that this is indeed a very nice idea and that in fact it has a wider scope of application: vague languages. Vague languages contain sentences that are prima facie neither clearly true nor clearly false. In other words, we could say that some vague sentences are unclearly true and that other vague sentences are unclearly false. So just as there are pathologically true (false) sentences in the case of semantic languages, there are unclearly true (false) sentences in languages containing vague predicates. Moreover, just as this sort of overlap between the semantic values allows for the introduction of a pathologicality operator, it also allows, in the case of vagueness, for the introduction of an unclarity operator.

The rest of the paper is devoted to argue for these claims more rigorously. In Section 2, we try to fill in the details in the sketchy picture given in Beall (2006). More specifically,

[3] Beall prefers the more neutral 'paranormal'. We think that 'pathological' and 'ungrounded', although less neutral, are certainly more interesting. Moreover, in what follows we will use 'pathological' and 'ungrounded' equivalently.

[4] $\pi Tr\langle\delta_1\rangle$ is not intended to be read as 'δ_1 is pathological *and* true' but only as 'δ_1 is pathological'. The truth predicate is needed for the purpose of self-reference. Of course, we could have done things differently. For example, it would do to represent the concept of pathologicality by a predicate. However, nothing important hangs on this.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

we provide a formal construction that gives a reasonable interpretation for the pathologicality operator π. We will see that the construction enjoys some nice properties such as a form of monotonicity and the fixed point property for the truth predicate. In Section 3, we sketch how the formal framework might be extended to deal with a language containing vague predicates. In Section 4 we discuss some problems the present account faces and Section 5 contains some concluding remarks.

§2. Interpreting the pathologicality operator.

2.1. Some definitions. Since we want to semantically categorize certain self-referential formulas, we need to start with a language capable of constructing such formulas. The obvious choice is to use the language of arithmetic. However, for reasons that will become clear below, this complicates things quite a bit. Hence, here we use a different kind of naming system, one that is not purely syntactic.

Let $\mathcal{L}$ be a first-order language and let $\mathcal{L}^{Tr}$ be $\mathcal{L}$ together with the truth predicate 'Tr(x)' and a set of distinguished constants $\mathcal{A} = \{a_1, a_2, a_3, ...\}$. The denotation of each of these constants is given by a function f_1 such that $f_1 : \mathcal{A} \longrightarrow Form_{\mathcal{L}^{Tr}}$, where as usual $Form_{\mathcal{L}^{Tr}}$ stands for the set of $\mathcal{L}^{Tr}$-formulas. Let $\mathcal{L}^{+}$ be $\mathcal{L}^{Tr}$ plus the pathologicality operator π and another set of distinguished constants $\mathcal{B} = \{b_1, b_2, b_3, ...\}$. The denotation of each of these constants is given by a function f_2 such that $f_2 : \mathcal{A} \cup \mathcal{B} \longrightarrow Form_{\mathcal{L}^{+}}$, where $Form_{\mathcal{L}^{+}}$ stands for the set of $\mathcal{L}^{+}$-formulas, and for each i, $f_2(a_i) = f_1(a_i)$. In this way, we make sure that $\mathcal{L}^{Tr}$ contains no names for formulas of $\mathcal{L}^{+}$.[5] To ease the notation, we'll write $\langle\phi\rangle$ for the distinguished name $c \in \mathcal{A} \cup \mathcal{B}$ such that $f_2(c) = \phi$.[6]

The set of semantic values V we will work with is $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$. The order between the values is given by the following diagram:

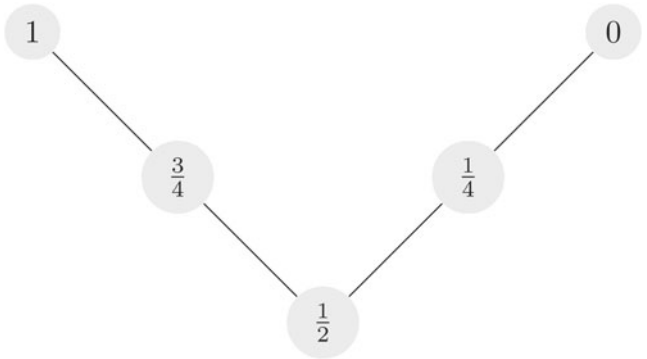


Fig. 1. The ordering of the semantic values.

[5] If we were using arithmetic as our naming system, it would be no easy task to obtain something like this. The reason is that under the standard coding of the formulas of, say, Peano arithmetic (plus a truth predicate and a pathologicality operator) there is no easy way to separate those numbers that code formulas that contain an explicit occurrence of the pathologicality operator from those in which there is an implicit occurrence of it. Later on, in definition 2.3, we'll make use of the fact that the names in $\mathcal{L}^{Tr}$ cannot denote formulas with explicit *or implicit* occurrences of the pathologicality operator.

[6] We borrow this way of getting self-reference from Ripley (2012).

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

Why five semantic values? The following definition gives an answer:

DEFINITION 2.1 (The set of true, false, and pathological formulas).
The set of True formulas is $\{\phi \in \mathcal{L}^+ | v(\phi) = \frac{3}{4}$ *or* $v(\phi) = 1\}$
The set of False formulas is $\{\phi \in \mathcal{L}^+ | v(\phi) = \frac{1}{4}$ *or* $v(\phi) = 0\}$
The set of Pathological formulas is $\{\phi \in \mathcal{L}^+ | v(\phi) \neq 1$ *and* $v(\phi) \neq 0\}$

This means that the semantic categories overlap. If a formula receives the value $\frac{3}{4}$, it is true and pathological, and if it receives the value $\frac{1}{4}$, it is false and pathological. A formula with value $\frac{1}{2}$ is pathological but neither true nor false.[7] As stated before, we want to use Strong Kleene as our valuation schema. But since we want to have the truth predicate around, we will use Strong Kleene-Kripke valuations (sometimes we'll simply speak of SKK-valuations or of fixed points, for short):

DEFINITION 2.2 (SKK-valuations for $\mathcal{L}^+$). *A valuation* v *is SKK if and only if*

1. $v(\neg\phi) = 1 - v(\phi)$
2. $v(\phi \vee \psi) = max\{v(\phi), v(\psi)\}$
   relative to the order $0 < \frac{1}{4} < \frac{1}{2} < \frac{3}{4} < 1$
3. $v(\exists x\phi) = sup\{v'(\phi) : v'$ is an x-variant of $v\}$
4. $v(Tr\langle\phi\rangle) = v(\phi)$

We need one more basic definition:

DEFINITION 2.3 (Extension). *A valuation* v_2 *extends a valuation* v_1 *iff* $\forall\phi \in \mathcal{L}^+$,

1. if $\phi \in Form_{\mathcal{L}^+} - Form_{\mathcal{L}^{Tr}}$, then
   if $v_1(\phi) \in \{\frac{3}{4}, 1\}$, then $v_2(\phi) \in \{\frac{3}{4}, 1\}$
   if $v_1(\phi) \in \{\frac{1}{4}, 0\}$, then $v_2(\phi) \in \{\frac{1}{4}, 0\}$
2. if $\phi \in Form_{\mathcal{L}^{Tr}}$, then $v_1(\phi) = v_2(\phi)$

We say that a formula ϕ is in $Form_{\mathcal{L}^{Tr}}$ if and only ϕ contains no occurrences of π and no occurrences of b_i for any $b_i \in \mathcal{B}$. Notice that Definition 2.3 is a bit restrictive, because clause 2. requires that the value of an $\mathcal{L}^{Tr}$-formula remains invariant under extensions. A consequence of this is that if we consider a truth teller sentence τ such that τ is $Tr\langle\tau\rangle$, and a valuation v_1 such that $v_1(\tau) = \frac{1}{2}$, all extensions v_2 will be such that $v_2(\tau) = \frac{1}{2}$. So in this respect the truth teller sentence τ will behave just like the liar sentence λ. This way of doing things will allow us to claim that they are both pathological in the same sense.

2.2. The construction. It is still unclear how to give a semantic interpretation for the pathologicality operator. The more straightforward way of doing it, once we have the extra semantic categories around, is by characterizing it in the following way:

$$v(\pi\phi) = \begin{cases} 0 & \text{if } v(\phi) \in \{0, 1\} \\ \frac{3}{4} & \text{otherwise} \end{cases}$$

[7] If the reader dislikes penta-valued semantics, she should note that a three-valued *relational* semantics can be used in its place. Instead of having the values $\frac{1}{4}$ and $\frac{3}{4}$, we simply assign certain formulas both 0 and $\frac{1}{2}$ and other formulas both 1 and $\frac{1}{2}$.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

This is how Beall (2006) does it. Unfortunately, the operator lacks the following property:

DEFINITION 2.4 (Monotonicity). *Let O be a monadic operator. We say that O is monotonic if and only if for all formulae ϕ, ψ, if $v(\phi) \leq v(\psi)$, then $v(O\phi) \leq v(O\psi)$ (where $\leq$ is the order relation of the space of values V as defined in Figure 1).*

Clearly, π is not a monotonic operator if defined as before. Say that $v(\phi) = \frac{1}{2}$ and $v(\psi) = 1$. Then we have $v(\pi\phi) = \frac{3}{4}$ and $v(\pi\psi) = 0$. So it is not obvious how to give a Kripke-style fixed point construction for the truth predicate if this operator is around.[8] In addition, both $\pi\lambda$ and δ_1 will have value $\frac{3}{4}$. However, this seems like a rather strange diagnosis. While there is a clear argument for the claim that δ_1 is pathological and true, no similar argument is available for $\pi\lambda$, unless every (non-vacuous) occurrence of the Liar is considered to be enough to infer that the corresponding sentence is pathological.

We take a different route here. The construction we give below uses a technique introduced in Yablo (2003). Yablo's original construction was intended to give a suitable semantics for a conditional that behaves nicely in the presence of a transparent truth predicate.[9] Roughly, Yablo's idea is that the truth value of a conditional statement at a certain valuation depends on how the antecedent and the consequent of that statement behave at the different fixed points extending that valuation. The account is similar to a possible world semantics, where the fixed points play the role of possible worlds. As Field (2005) observes, this account faces a number of problems. Among others, there is no (obvious) way to define a pathologicality operator using Yablo's conditional, and it gives rather bad results for formulas containing nested conditionals. On top of that we would add that it is unclear why we have to look at different fixed points to determine the semantic value of a conditional claim.

Our hope is that all these difficulties can be put aside if instead of using this technique to define a conditional connective, we use it to define a primitive pathologicality operator. Conceptually speaking, when we say that a sentence is pathological or non-pathological, we are claiming that it behaves in a certain way across different fixed points, so it makes sense to evaluate attributions of pathologicality by considering different fixed points.

This is how the construction "intuitively" works. First, we start with Kripke's minimal fixed point,[10] which we call P^0. We stipulate that in P^0 every formula of the form $\pi\phi$ is neither true nor false. More formally, $P^0(\pi\phi) = \frac{1}{2}$, for every formula ϕ. Then we consider the set of SKK-valuations (i.e., the set of fixed points) extending P^0. We call this set $\mathcal{R}^0$. At successor stages $P^{\alpha+1}$ we semantically evaluate the formulas involving the pathologicality operator by looking at the set of SKK-valuations $\mathcal{R}^\alpha$ extending P^α. For instance, if we want to find out what $P^1(\pi\phi)$ is, we need to look at $R(\phi)$ for all $R \in \mathcal{R}^0$. At limit stages P^λ we look at the intersection of the set of SKK-valuations extending each P^β for $\beta < \lambda$. Later on we will show that this construction has the fixed point property. Hence, there is an ordinal Y (after Yablo), such that $P^Y = P^{Y+1}$. However, to obtain the intended value of some formulas involving π it will be necessary to construct a new fixed point P^* which will contain some pathologically true (false)

[8] Of course, this is not to say that it cannot be done. In fact, there are schemata with non-monotonic operators that have been shown to enjoy the fixed point property.

[9] Actually, Yablo's paper uses a four-valued semantics and is a critique of Field's approach to paradoxes. Field's theory can be found, for example, in Field (2005), Field (2007), and Field (2008)

[10] Naturally, we could start with a different fixed point, so nothing crucial depends on this.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

formulas. This fixed point gives us the extension (and the anti-extension) of the truth predicate.

Figure 2 below should be useful to depict the way in which we obtain new valuations from old ones:

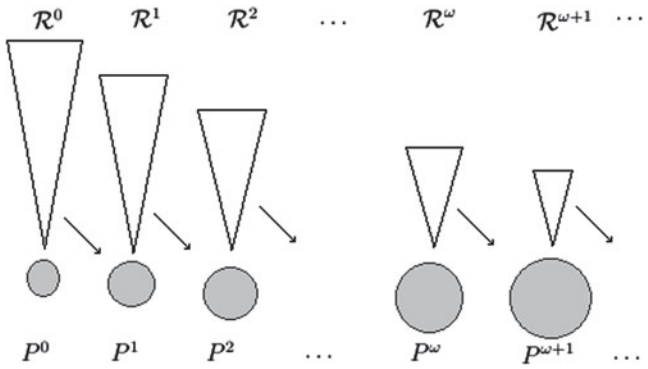


Fig. 2. The Yablo sequence.

More rigorously, we can provide the following definition for the valuations we use:

DEFINITION 2.5 (The Yablo Sequence). *We define P^0 as the minimal Kripke fixed point (where all formulas of the form $\pi\phi$ have value $\frac{1}{2}$).*

$\mathcal{R}^0 = \{R | R$ *is an SKK-valuation extending* $P^0\}$

For successor ordinals $\alpha + 1$, let $P^{\alpha+1}$ be the valuation obtained by letting the formulas of the form $\pi\phi$ behave as specified below and then applying the SKK operations:

$$P^{\alpha+1}(\pi\phi) = \begin{cases} 0 & \text{if } \forall R \in \mathcal{R}^\alpha R(\phi) = 1 \text{ or } \forall R \in \mathcal{R}^\alpha R(\phi) = 0 \\ 1 & \text{if } \forall R \in \mathcal{R}^\alpha R(\phi) = \frac{1}{2} \\ P^\alpha(\pi\phi) & \text{otherwise} \end{cases}$$

$\mathcal{R}^{\alpha+1} = \{R | R \in \mathcal{R}^\alpha$ *is an SKK-valuation extending* $P^{\alpha+1}\}$

For limit ordinals λ, let P^λ be the valuation obtained by letting the formulas of the form $\pi\phi$ behave as specified below and then applying the SKK operations:

$$P^\lambda(\pi\phi) = \begin{cases} 0 & \text{if } \forall R \in \bigcap\{\mathcal{R}^\beta | \beta < \lambda\} R(\phi) = 1 \text{ or} \\ & \forall R \in \bigcap\{\mathcal{R}^\beta | \beta < \lambda\} R(\phi) = 0 \\ 1 & \text{if } \forall R \in \bigcap\{\mathcal{R}^\beta | \beta < \lambda\} R(\phi) = \frac{1}{2} \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

$\mathcal{R}^\lambda = \{R | R \in \bigcap\{\mathcal{R}^\beta | \beta < \lambda\}$ *is an SKK-valuation extending* $P^\lambda\}$

Spelling this out a bit more, we obtain each P^α by a two step process. We first assign new values to some formulas of the form $\pi\phi$. This gives us a valuation that is *not* a fixed point, but that can be built up into a fixed point by closing under the SKK operations. This fixed point is our P^α.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

That each of the resulting valuations P^α are indeed fixed points can be proved using the standard argument involving the monotonicity of the Kripke jump. More specifically, let's call P^α_0 the valuation obtained by assigning new values to some formulas of the form $\pi\phi$ and leaving everything else the same. By applying the jump operation together with the Strong Kleene operations, we can construct a sequence of valuations P^α_1, P^α_2, P^α_3,...., P^α_ω, $P^\alpha_{\omega+1}$,... . It is straightforward to prove (by an induction on the complexity of ϕ) that this sequence is monotonic, in the sense that if $\zeta \leq \eta$, then for any formula $\phi \in \mathcal{L}^+$, $P^\alpha_\zeta(\phi) > \frac{1}{2}$ only if $P^\alpha_\eta(\phi) > \frac{1}{2}$ and $P^\alpha_\zeta(\phi) < \frac{1}{2}$ only if $P^\alpha_\eta(\phi) < \frac{1}{2}$. This means that after a while we reach a fixed point, which we call P^α.[11]

We shall prove that, in addition to this, the entire construction enjoys the fixed point property. First we need a straightforward lemma:

LEMMA 2.6 (The set of sets of extensions forms a decreasing sequence). *If $\alpha \leq \beta$, then $\mathcal{R}^\beta \subseteq \mathcal{R}^\alpha$*

Proof. We prove this by transfinite induction on β. If β is 0 or a limit ordinal λ then the result follows trivially from the definitions of $\mathcal{R}^0$ and $\mathcal{R}^\lambda$, respectively. If β is a successor ordinal η+1, then by the inductive hypothesis we know that $\mathcal{R}^\eta \subseteq \mathcal{R}^\alpha$. Since by the definition of $\mathcal{R}^{\eta+1}$, $\mathcal{R}^{\eta+1} \subseteq \mathcal{R}^\eta$, we can infer that $\mathcal{R}^{\eta+1} \subseteq \mathcal{R}^\alpha$. So we have $\mathcal{R}^\beta \subseteq \mathcal{R}^\alpha$. □

LEMMA 2.7 (Monotonicity of the sequence of fixed points). *If $\alpha \leq \beta$, then P^β extends P^α*

Proof. Since no formula obtains a value in $\{\frac{1}{4}, \frac{3}{4}\}$ in any of the P valuations, it is enough to prove that if $\alpha \leq \beta$, then for every formula $\phi \in \mathcal{L}^+$:

1. If $P^\alpha(\phi) = 1$, then $P^\beta(\phi) = 1$.
2. If $P^\alpha(\phi) = 0$, then $P^\beta(\phi) = 0$.

This is proved by induction on the complexity of ϕ. The only interesting case is where ϕ is of the form $\pi\psi$ (a case for which, curiously, we do not need the inductive hypothesis). To complete the proof of the theorem it is enough to prove the following four facts:

1. if $P^\alpha(\pi\psi) = 1$, then $P^{\alpha+1}(\pi\psi) = 1$;
2. if $P^\alpha(\pi\psi) = 0$, then $P^{\alpha+1}(\pi\psi) = 0$;
3. if $\forall\eta{<}\lambda\ P^\eta(\pi\psi) = 1$, then $P^\lambda(\pi\psi) = 1$; and
4. if $\forall\eta{<}\lambda\ P^\eta(\pi\psi) = 0$, then $P^\lambda(\pi\psi) = 0$.

We just prove 1. and 3., since item 2. is similar to 1. and item 4. is similar to 3.

1. Assume that $P^{\alpha+1}(\pi\psi) \neq 1$. Then $P^{\alpha+1}(\pi\psi) = 0$ or $P^{\alpha+1}(\pi\psi) = \frac{1}{2}$ (remember that $P^{\alpha+1}(\pi\psi)$ is never $\frac{1}{4}$ or $\frac{3}{4}$ by the definition of π). If $P^{\alpha+1}(\pi\psi) = \frac{1}{2}$, then $\frac{1}{2} = P^\alpha(\pi\psi) = P^{\alpha+1}(\pi\psi) \neq 1$. If $P^{\alpha+1}(\pi\psi) = 0$, we can use Lemma 2.6 to infer that $P^{\alpha+1} \in \mathcal{R}^{\alpha+1} \subseteq \mathcal{R}^\alpha$

[11] In fact, there are many fixed points P^α. Naturally, we pick the one that is "less informative" because we think that it is the best motivated, but there are others. To see why, consider a sentence γ such that γ is $Tr\langle\gamma\rangle \vee \pi\top$ (we assume that there is a truth constant $\top$ in $\mathcal{L}^+$). Since already at P^1 we have $P^1(\pi\top) = 0$, the value of the disjunction at P^1 completely depends on the value of $Tr\langle\gamma\rangle$, which in our construction will be $\frac{1}{2}$, but could in principle be different. In this sense, the sentence γ is similar to a truth teller.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

and so $\exists R \in \mathcal{R}^{\alpha} R(\pi\psi) = 0$. And from this it follows that $P^{\alpha}(\pi\psi) \neq 1$, for otherwise there would be an $R \in \mathcal{R}^{\alpha}$ which is not an extension of P^{α}.

3. Assume that $P^{\lambda}(\pi\psi) \neq 1$. Then $P^{\lambda}(\pi\psi) = 0$ or $P^{\lambda}(\pi\psi) = \frac{1}{2}$. In either case it follows that $\exists R \in \mathcal{R}^{\lambda} R(\pi\psi) = 0$. By Lemma 2.6 it holds that $\mathcal{R}^{\lambda} \subseteq \bigcap\{\mathcal{R}^{\eta} : \eta < \lambda\}$. Hence, $\exists R \in \bigcap\{\mathcal{R}^{\eta} : \eta<\lambda\} R(\pi\psi) = 0$. Therefore it is not the case that $\forall R \in \bigcap\{\mathcal{R}^{\eta} : \eta<\lambda\} R(\pi\psi) = 1$. So, we can infer that it is not the case that $\forall\eta<\lambda\ P^{\eta}(\pi\psi) = 1$. □

As the sequence of sets of valuations $\mathcal{R}$ decrease, the sequence of valuations P increase. This means that the extension (and the antiextension) of the truth predicate gets larger as more formulas obtain a value different from $\frac{1}{2}$. As a corollary we can infer that if $\alpha < \beta$, then the extension (antiextension) of the the truth predicate at α is a subset of the extension (antiextension) of the truth predicate at β. Furthermore, the following can be proved.

THEOREM 2.8 (The Fixed Point P^{Y}). *There is an ordinal α such that $P^{\alpha} = P^{\alpha+1}$*

Proof. As usual, a cardinality argument can be used (see Kripke, 1975 for example). □

It is not hard to see that all the $\mathcal{L}^{Tr}$-formulas have in the Yablo fixed point P^{Y} the same value they receive in P^{0}. But in addition, in P^{Y} some formulas of the extended language $\mathcal{L}^{+}$ obtain the value they are expected to obtain. Some examples will help to see why this is so.

EXAMPLE 2.9 (Sentences that are just true (false) are not pathological). *Let ϕ be $\top$ or any tautology. Our starting policy is that $P^{0}(\pi\phi) = \frac{1}{2}$. Since P^{0} is Kripke's minimal fixed point, $P^{0}(\phi) = 1$. This means that $\forall R \in \mathcal{R}^{0}, R(\phi) = 1$, by the definition of extension. From this, it follows that $P^{1}(\pi\phi) = 0$. It is also clear that this will not change, so for every $\beta \geq 1$, it holds that $P^{\beta}(\pi\phi) = 0$. Since the fixed point P^{Y} is such that $Y \geq 1$, it follows that $P^{Y}(\pi\phi) = 0$.*

EXAMPLE 2.10 (The Liar sentence is pathological). *Let λ be $\neg Tr\langle\lambda\rangle$. Clearly, $P^{0}(\lambda) = \frac{1}{2}$ and since $\lambda \in Form_{\mathcal{L}^{Tr}}$, it holds that $\forall R \in \mathcal{R}^{0} R(\lambda) = \frac{1}{2}$. Hence, $P^{1}(\pi\lambda) = 1$, and it is not hard to see that this will not change, so for every $\beta \geq 1$, it holds that $P^{\beta}(\pi\lambda) = 1$. Therefore, $P^{Y}(\pi\lambda) = 1$, as expected. (A similar argument shows that $P^{Y}(\pi\tau) = 1$).*

This is an important difference and -we believe- an advantage over the approach in Beall (2006), where $\pi\lambda$ and similar sentences are classified as true and pathological. In our approach, it is simply true that the Liar is a pathological sentence.

We remarked above that one of the problems in Yablo (2003) was the lack of valid embedded conditionals. Let's see how our theory does with iterations of the pathologicality operator.

EXAMPLE 2.11 (Iterations of the pathologicality operator). *Let ϕ be $\top$ again, and $\pi^{n}\phi$ be*

$$\underbrace{\pi \ldots \pi}_{n-times}\phi$$

It is not hard to check that for each $j<n$, $P^{j}(\pi^{n}\phi) = \frac{1}{2}$, but that $P^{n}(\pi^{n}\phi) = 0$. Hence, for any n, $P^{Y}(\pi^{n}\phi) = 0$, as expected.

So far we have seen how to apply the pathologicality operator to the usual non-pathological sentences, and also to pathological sentences like the Liar. But what about our target sentences δ_1 and δ_2? Recall that until now we have not made use of the extra truth values $\frac{1}{4}$ and $\frac{3}{4}$. Hence, for no formula ϕ it holds that $P^{Y}(\phi) = \frac{1}{4}$ or that $P^{Y}(\phi) = \frac{3}{4}$. In particular, we have $P^{Y}(\delta_1) = P^{Y}(\delta_2) = \frac{1}{2}$.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

Nevertheless, there is an easy way to fix this:

DEFINITION 2.12 (The new fixed point P*). *P^* is obtained from P^Y by letting $P^*(\pi\phi) = \frac{3}{4}$ whenever $P^Y(\phi) = \frac{1}{2}$ and $\exists R \in \mathcal{R}^Y R(\phi) \neq \frac{1}{2}$, and then applying the SKK operations.*

Spelling this out a bit more, Definition 2.12 ensures that:

$$P^*(\pi\phi) = \begin{cases} 0 & \text{if } P^Y(\phi) \in \{0, 1\} \\ 1 & \text{if } P^Y(\phi) = \frac{1}{2} \text{ and } \forall R \in \mathcal{R}^Y, R(\phi) = \frac{1}{2} \\ \frac{3}{4} & \text{otherwise} \end{cases}$$

It is not hard to show that P^* is well-defined and that it monotonically extends P^Y.[12] In fact, something stronger can be proved: that the only way in which P^* differs from P^Y involves (some of) the formulas with value $\frac{1}{2}$ in P^Y.

LEMMA 2.13. *For every formula $\phi \in Form_{\mathcal{L}^+}$,*

1. $P^Y(\phi) = 1$ if and only if $P^*(\phi) = 1$.
2. $P^Y(\phi) = 0$ if and only if $P^*(\phi) = 0$.

Proof. By induction on the complexity of ϕ. Naturally, the only interesting case is that where ϕ is of the form $\pi\psi$ (a case for which we don't have to use the inductive hypothesis). For (2) we reason as follows. $P^*(\pi\psi) = 0$ if and only if $P^Y(\psi) \in \{0, 1\}$ if and only if $P^{Y+1}(\pi\psi) = 0$ if and only if $P^Y(\pi\psi) = 0$. The reasoning for (1) is similar. □

So some formulas with value $\frac{1}{2}$ in P^Y will obtain the value $\frac{3}{4}$ in P^*, and since P^* is an SKK-valuation, some formulas with value $\frac{1}{2}$ in P^Y, will obtain the value $\frac{1}{4}$ in P^*. This is as it should be, for we want some formulas to be categorized as pathologically true and others to be categorized as pathologically false.

Furthermore, with the previous Lemma we can prove some nice facts about the behavior of the operator π in P^*:

THEOREM 2.14 (The behavior of π in P^*). *For every formula $\phi \in Form_{\mathcal{L}^+}$:*

1. $P^*(\pi\phi) = 0$ if and only if $P^*(\phi) \in \{0, 1\}$.
2. If $P^*(\pi\phi) = 1$, then $P^*(\phi) = \frac{1}{2}$.
3. If $P^*(\phi) \in \{\frac{1}{4}, \frac{3}{4}\}$, then $P^*(\pi\phi) = \frac{3}{4}$.
4. If $P^*(\pi\phi) = \frac{3}{4}$, then $P^*(\phi) \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$.

Proof. The proofs of all these facts depend on Lemma 2.13, the definition of P^* and the fact that P^Y is a fixed point. For 1., assume that $P^*(\pi\phi) = 0$. By the definition of P^*, this holds if and only if $P^Y(\phi) \in \{0, 1\}$, which in turn is true if and only if $P^*(\phi) \in \{0, 1\}$ by Lemma 2.13. 4. clearly follows from 1., and we leave 2. and 3. to the reader. □

[12] Just as with each P^α in the Yablo sequence (see footnote 11), there is more than one SKK valuation P^* obtainable from P^Y. To see this consider a sentence γ_1 such that γ_1 is $\pi Tr\langle\gamma_1\rangle \vee Tr\langle\gamma_1\rangle$. Given that $P^*(\pi Tr\langle\gamma_1\rangle) = \frac{3}{4}$, the value of γ_1 at P^* partially depends on the value of $Tr\langle\gamma_1\rangle$. In this sense γ_1 is similar to a truth teller, and although our construction is such that $P^*(\gamma_1) = \frac{3}{4}$, things could in principle have been done differently, since $P^*(\gamma_1)$ could have been 1.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

Naturally, we can now define validity in the following way:

DEFINITION 2.15 (Validity). *An argument from* Γ *to* ϕ *is valid (in notation,* $\Gamma \vDash \phi$*) if and only if* $P^*(\phi) \geq \frac{3}{4}$ *whenever* $P^*(\gamma) \geq \frac{3}{4}$ *for every* $\gamma \in \Gamma$. *And a formula* ϕ *is valid (*$\vDash \phi$*) if and only if* $P^*(\phi) \geq \frac{3}{4}$.

Validity is still preservation of truth, although some valid arguments are not truth-only or truth-and-not-pathological preserving. With this definition, we can prove that the pathologicality operator π interacts with the logical constants (and with itself) as follows:

THEOREM 2.16 (Facts about π and the logical expressions).

$$\pi\phi \wedge \pi\psi \vDash \pi(\phi \wedge \psi) \text{ but } \pi(\phi \wedge \psi) \nvDash \pi\phi \wedge \phi\psi$$
$$\pi(\phi \vee \psi) \vDash \pi\phi \vee \pi\psi \text{ but } \pi\phi \vee \pi\psi \nvDash \pi(\phi \vee \psi)$$
$$\pi\phi \vDash \pi\neg\phi \text{ but } \neg\pi\phi \nvDash \pi\neg\phi$$
$$\pi\neg\phi \vDash \pi\phi \text{ but } \pi\neg\phi \nvDash \neg\pi\phi$$
$$\pi\pi\phi \vDash \pi\phi \text{ but } \pi\phi \nvDash \pi\pi\phi$$
$$\forall x\pi\phi(x) \vDash \pi\forall x\phi(x) \text{ but } \pi\forall x\phi(x) \nvDash \forall x\pi\phi(x).$$
$$\pi\exists x\phi(x) \vDash \exists x\pi\phi(x) \text{ but } \exists x\pi\phi(x) \nvDash \pi\exists x\phi(x)$$

All these facts are predictable except perhaps for the failure of the inference from $\pi\phi$ to $\pi\pi\phi$. There are pathological formulas such that we can apply the pathologicality operator once to them, but not twice (nor thrice, and so on). For instance, if λ is a Liar sentence, $\pi\lambda$ will receive the value 1, but $\pi\pi\lambda$ or, more generally, $\pi^n\lambda$ for $n \geq 2$ will receive the value 0.

It is straightforward to check that the sentences considered in the previous examples ($\pi\top$, $\pi\lambda$, and $\pi^n\top$) and sentences of the same kind maintain their value in P^*. Next we show how the construction works with our target sentences δ_1 and δ_2.

EXAMPLE 2.17 (Sentences that are true and pathological). *Consider again a sentence* δ_1 *such that* δ_1 *is* $\pi Tr\langle\delta_1\rangle$. *Since our starting policy is that* $P^0(\pi\phi) = \frac{1}{2}$ *for every formula of the form* $\pi\phi$, $P^0(\pi Tr\langle\delta_1\rangle) = \frac{1}{2}$. *To obtain the value of* $P^1(\pi Tr\langle\delta_1\rangle)$, *we reason in the following way. By the definition of* extension, *there are valuations* $R^a, R^b, R^c, R^d, R^e \in \mathcal{R}^0$ *such that* $R^a(Tr\langle\delta_1\rangle) = 0$, $R^b(Tr\langle\delta_1\rangle) = \frac{1}{4}$, $R^c(Tr\langle\delta_1\rangle) = \frac{1}{2}$, $R^d(Tr\langle\delta_1\rangle) = \frac{3}{4}$ *and* $R^e(Tr\langle\delta_1\rangle) = 1$. *This means that we are in the* otherwise case, *which implies that* $P^1(\pi Tr\langle\delta_1\rangle) = P^0(\pi Tr\langle\delta_1\rangle) = \frac{1}{2}$. *Moreover, since nothing changes at limit ordinals, it is not hard to see that for each ordinal* β, $P^\beta(\pi Tr\langle\delta_1\rangle) = \frac{1}{2}$ *and there are valuations* $R \in \mathcal{R}^\beta$ *such that* $R(Tr\langle\delta_1\rangle) \neq \frac{1}{2}$. *Hence,* $P^Y(\pi Tr\langle\delta_1\rangle) = \frac{1}{2}$ *and* $\exists R \in \mathcal{R}^Y R(Tr\langle\delta_1\rangle) \neq \frac{1}{2}$. *By Definition 2.12, we can obtain* $P^*(\pi Tr\langle\delta_1\rangle) = \frac{3}{4}$, *as desired.*

EXAMPLE 2.18 (Sentences that are false and pathological). *For* δ_2 *such that* δ_2 *is* $\neg\pi Tr\langle\delta_2\rangle$, *the reasoning is similar. But it should be noted that every* P^β *in the sequence is an SKK-valuation. Therefore,* $P^\beta(\neg\pi Tr\langle\delta_2\rangle) = 1 - P^\beta(\pi Tr\langle\delta_2\rangle)$. *So for every* β, $P^\beta(\delta_2) = \frac{1}{2}$. *Hence,* $P^Y(\neg\pi Tr\langle\delta_2\rangle) = P^Y(Tr\langle\delta_2\rangle) = \frac{1}{2}$. *As before, by Definition 2.12, this implies that* $P^*(\pi Tr\langle\delta_2\rangle) = \frac{3}{4}$. *But since* P^* *is an SKK-valuation too,* $P^*(\neg\pi Tr\langle\delta_2\rangle) = 1 - P^*(\pi Tr\langle\delta_2\rangle) = \frac{1}{4}$.

This is enough to establish that δ_1 is a true pathological sentence and that δ_2 is a false pathological sentence.

Quantified sentences involving π also behave appropriately. An easy case is the following:

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

EXAMPLE 2.19 (Quantifiers and pathologicality I). *Let us consider the pair of sentences* $\pi\forall x(x = x)$ *and* $\forall x(\pi x = x)$. *Clearly,* $P^0(\forall x(x = x)) = 1$. *Hence,* $\forall R \in \mathcal{R}^0$, $R(\forall x(x = x)) = 1$, *by the definition of extension. From this it follows that* $P^1(\pi\forall x(x = x)) = 0$. *As for* $\forall x(\pi x = x)$, *we know that* $P^0(x = x) = 1$, *regardless of what object is assigned to* x. *Hence,* $\forall R \in \mathcal{R}^0$, $R(x = x) = 1$, *and so for each* x*-variant of* P^1, *the* x*-variant assigns* $\pi x = x$ *the value* 0. *Therefore,* $P^1(\forall x(\pi x = x)) = 0$. *Since nothing will change from there on, we have* $P^*(\forall x(\pi x = x)) = P^*(\pi\forall x(x = x)) = 0$.

For a more complicated example, consider this one:

EXAMPLE 2.20 (Quantifiers and pathologicality II). *Let* $\pi^n\phi$ *be as in Example 2.11 Now consider the following sentence:*

$$\pi(\pi\top \vee \pi^2\top \vee \pi^3\top \vee ....)$$

It seems clear that this can be expressed in our language by means of a quantification of the form:

$$\pi\exists n(\pi^n\top)$$

This sentence should receive the value 0, since it says that a non-pathological sentence is pathological. Observe that for each n, *the sentence* $\pi^n\top$ *acquires the value 0 only at stage* P^n. *Hence, at each finite stage* P^n *of the construction, the sentence* $\exists n(\pi^n\top)$ *has value* $\frac{1}{2}$. *And so does the sentence* $\pi\exists n(\pi^n\top)$. *At stage* P^ω *every* $\pi^n\top$ *acquires the value 0, so now* $\exists n(\pi^n\top)$ *has the value 0. So although at* P^ω *the sentence* $\pi\exists n(\pi^n\top)$ *has value* $\frac{1}{2}$, *at* $P^{\omega+1}$ *we can infer that it has value 0.*

Although our construction for π gives, in our opinion, the right diagnosis for a vast number of pathological sentences, it has been pointed out to us by an anonymous referee that it is rather *ad hoc*, because it does not conform closely to intuitions about the concept of pathologicality. In this sense, it could be argued that there is a stark contrast with Kripke's fixed point construction for truth, which is very well motivated.

First of all, we should note that while truth is an intuitive concept, at least to the extent that competent speakers have more or less clear intuitions about how to use the truth predicate, the same cannot be said of pathologicality (or ungroundedness) which is, at best, a semi-intuitive concept. Hence, perhaps it is too demanding to expect our construction to conform closely to "intuitions" about pathologicality. In fact, it is arguably a controversial matter whether there are such intuitions.

Nevertheless, this is not to say that the construction needs no conceptual justification, and in fact, we do think that there is a nice one available. To see how it works, an analogy might be helpful. Just as being necessarily true is a form of being true, that is, true at every possible world, being non-pathologically true is also a form of being true, in this case, true at every fixed point. The pathologicality operator works as a kind of modal operator that tracks down how a sentence behaves at different fixed points. The pathological sentences are, roughly, those that are neither true nor false at every fixed point and those that behave differently at different fixed points. For instance, to determine the truth value of $\pi\lambda$ or $\pi 0 = 0$ we have to look at how λ and $0 = 0$ behave at all fixed points. This is what the construction does.

Naturally, we also want to see what the truth values of sentences like $\pi\pi\lambda$ and $\pi\pi 0 = 0$ are. These sentences resemble things of the form *It is necessary that it is necessary that* ϕ.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

For them, we need to see how $\pi\lambda$ and $\pi 0 = 0$ behave at different fixed points. That's why the construction need to be iterated.[13]

An additional complication has to do with the fact that we need to tweak the Yablo fixed point P^Y into a more suitable fixed point P^* to correctly interpret sentences like δ_1 and δ_2. The rationale for this new fixed point is that the circular character of δ_1 and δ_2 has, as a consequence, that no amount of iterations in the Yablo construction is going to be enough to settle their intended truth value. In this respect, the construction resembles the revision theory of truth, that categorizes as pathological those sentences which do not stabilize across the revision sequence.

§3. Vagueness.

3.1. Preliminary remarks. In this section we consider the possibility of extending the framework to languages containing at least one vague predicate. There is a reason for this. Typical non-classical solutions to the semantic paradoxes and to the paradoxes of vagueness have certain things in common. For example, in both cases it is standard to introduce semantic categories different from truth and falsity, and in both cases these new semantic categories cause problems. We have already seen that for semantic predicates the introduction of a third category usually produces new inconsistencies. For vague predicates a similar problem appears: the new category introduces new sharp boundaries. In both cases these problems show up most clearly when a sentence-classifying device is added to the language. For semantic predicates, this role is played by the pathologicality (or some similar) operator, whereas for vague predicates this role is played by an unclarity (or some similar) operator.

So it seems interesting to see whether this characterization project -that aims to correctly classify each sentence according to its semantic category- can also be carried out for languages including vague predicates by assigning overlapping semantic categories to some sentences. The idea is that just as we managed to introduce an operator in the case of the semantic predicates by letting certain sentences be true (false) and pathological, in the case of vague predicates, a similar strategy is available. We can introduce an unclarity operator by letting certain sentences be true (false) and unclear, or equivalently, unclearly true (false).

3.2. How should the unclarity operator behave? To deal with vague predicates it is useful to have infinitely many truth values around. This can be accomplished by the use of a fuzzy framework.[14] In such a framework it seems that there are some natural constraints

[13] It might be replied that the problem lies in the fact that we take the intersections of the fixed points at limit ordinals and that that's where the artificiality of the construction is. However, we take the move of using the intersection of the fixed points to be no more problematic that the move of taking the intersection of all the extensions of the truth predicate at limit ordinals, something that is sometimes done to interpret the truth predicate in some paraconsistent logics, such as LP. The only difference is that we are intersecting on "more complicated things", which is only to be expected, since, as we've already remarked, the pathologicality operator has a sort of modal flavor to it.

[14] We are not trying to claim here that a fuzzy account is the best possible solution to the paradoxes of vagueness. Hence, we won't be dealing with all the problems that have been attributed in the literature to such an account. As we have stressed before, our goal is to see whether there is an interesting characterization problem for vague sentences as there is for semantic sentences. Since we think that the notion of unclarity is crucial to this sort of project and an infinitely-valued semantics seems suitable to deal with such a notion, the fuzzy framework seems appropriate.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

on how an unclarity operator $\mathbb{U}$ should behave. If the language has vague predicates, a sentence can be clearly true, clearly false, or something in between. If there are degrees of truth, there are also degrees of unclarity, that is, degrees concerning how clearly true (false) a sentence is. The fuzzy framework allows us to have a fine-grained characterization of the status of those intermediate sentences which are neither clearly true nor clearly false. Since we assign clearly true sentences the value 1 and clearly false sentences the value 0, we will say that the more a sentences approaches those values the more clear it is. As a consequence, we will stipulate that a clearly unclear sentence has the value $\frac{1}{2}$. This already gives a hint as to how the unclarity operator should work.

If a formula ϕ has value 0 or 1, then it is completely clear, so $\mathbb{U}\phi$ should have value 0. On the other hand, if ϕ has value $\frac{1}{2}$, $\mathbb{U}\phi$ should have value 1. What about all formulas that have a value other than 0, $\frac{1}{2}$, or 1? Since we are interpreting the value $\frac{1}{2}$ as being neither true nor false, it seem plausible to say that if a formula has a value strictly greater than $\frac{1}{2}$ (and different from 1), then the formula is true to some degree, or equivalently, unclearly true. Symmetrically, if a formula has a value strictly less than $\frac{1}{2}$ (and different from 0), then the formula is false to some degree, or equivalently, unclearly false. Given this reading of the space of values, it also seems plausible to ask that for any two formulas ϕ and ψ, if the the value of ϕ is "closer" to $\frac{1}{2}$ than the value of ψ, then the value of $\mathbb{U}\phi$ should be "closer" to 1 than the value of $\mathbb{U}\psi$ (and both should be greater than $\frac{1}{2}$). Also, if the value of ϕ and the value of ψ are "at the same distance" from $\frac{1}{2}$, then the value of $\mathbb{U}\phi$ and that of $\mathbb{U}\psi$ should be the same (and greater than $\frac{1}{2}$).

These constraints can be rigorously characterized in a straightforward way. Consider three functions:

$$g_1 : \left(0, \frac{1}{2}\right] \longrightarrow \left(\frac{1}{2}, 1\right]$$
$$g_2 : \left(\frac{1}{2}, 1\right) \longrightarrow \left(\frac{1}{2}, 1\right]$$
$$g_3 = \{< 0, 0 >, < 1, 0 >\}.$$

Before stating how g_1 and g_2 work we need to introduce the following definition (where $d(x, y)$ is the absolute value of the subtraction $x - y$):

DEFINITION 3.1 (Increasing, Decreasing, Symmetric).
A function f *is* increasing *if* $f(x) \leq f(y)$ *whenever* $x \leq y$.
A function f *is* decreasing *if* $f(x) \geq f(y)$ *whenever* $x \leq y$.
Two functions f *and* g *are* symmetric *with respect to* [0, 1] *if:* $f(x) = g(y)$ *if and only if* $d(x, \frac{1}{2}) = d(y, \frac{1}{2})$.

Let h be $g_1 \cup g_2 \cup g_3$. Assume that g_1 is increasing, g_2 is decreasing, and that g_1 and g_2 are symmetric w.r.t. [0, 1]. What we are claiming is that any adequate fuzzy unclarity operator for a vague language should be represented by a function satisfying the conditions imposed on h.[15]

This time the operator will be added to a language $\mathcal{L}$ which contains no self-referential expressions, but that has at least one vague predicate. We obtain our target language $\mathcal{L}^+$ by

[15] It also seems adequate to demand that the unclarity operator behaves like a continuous function and that it satisfies uniformity, where a function f is *uniform* if $d(d(x, \frac{1}{2}), d(y, \frac{1}{2})) = d(d(f(x), \frac{1}{2}), d(f(y), \frac{1}{2}))$. As we will see soon, our unclarity operator satisfies these extra requirement. However, we do not have strong reasons to dismiss operators that do not.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

adding this unclarity operator, so $\mathcal{L}^+$ is $\mathcal{L}+\mathbb{U}$. The set of values is $V = \{x \in \mathbb{R} | 0 \leq x \leq 1\}$, where $\mathbb{R}$ is the set of real numbers, and the order for the set of values is a generalization of the one appearing in Figure 1.

The crucial definitions of Section 2.1 work in this case too, but now generalized to the infinite space of values.

DEFINITION 3.2 (The set of true, false, and unclear formulas).
The set of True formulas is $\{\phi \in \mathcal{L}^+ | v(\phi) \in (\frac{1}{2}, 1]\}$
The set of False formulas is $\{\phi \in \mathcal{L}^+ | v(\phi) \in [0, \frac{1}{2})\}$
The set of Unclear formulas is $\{\phi \in \mathcal{L}^+ | v(\phi) \in (0, 1)\}$

Moreover, we can still use SK-valuations:

DEFINITION 3.3 (Strong-Kleene valuations for $\mathcal{L}^+$). *A valuation v is SK if and only if*

1. $v(\neg\phi) = 1 - v(\phi)$
2. $v(\phi \vee \psi) = max\{v(\phi), v(\psi)\}$
   now relative to the usual order of the real numbers in [0, 1]
3. $v(\exists x\phi) = sup\{v'(\phi) : v' \text{ is an } x\text{-variant of } v\}$

DEFINITION 3.4 (The unclarity operator). *The unclarity operator $\mathbb{U}$ is defined in the following way:*

$$v(\mathbb{U}\phi) = \begin{cases} \frac{1}{2} + v(\phi) & \text{if } 0 < v(\phi) \leq \frac{1}{2} \\ \frac{1}{2} + (1 - v(\phi)) & \text{if } \frac{1}{2} < v(\phi) < 1 \\ 0 & \text{otherwise} \end{cases}$$

Next we give a couple of examples to show how $\mathbb{U}$ works.

EXAMPLE 3.5 (An unclearly true sentence). *Let ϕ and ψ be sentences containing some vague predicate such that $v(\phi) = \frac{1}{8}$ and $v(\psi) = \frac{1}{4}$. Now we compute the value of the sentence $\mathbb{U}\mathbb{U}\phi \vee \mathbb{U}\psi$. Since $v(\phi) = \frac{1}{8}$, $v(\mathbb{U}\phi) = \frac{5}{8}$, and $v(\mathbb{U}\mathbb{U}\phi) = \frac{7}{8}$. Also, since $v(\psi) = \frac{1}{4}$, $v(\mathbb{U}\psi) = \frac{3}{4}$. Hence, $v(\mathbb{U}\mathbb{U}\phi \vee \mathbb{U}\psi) = \frac{7}{8}$, given that $\vee$ is defined as the maximum. So the sentence is unclearly true.*

EXAMPLE 3.6 (A clearly false sentence). *Let ϕ be as before and consider the sentence $\mathbb{U}\mathbb{U}(\phi \vee \top)$. Given that $v(\top)$ is always 1, $v(\phi \vee \top) = 1$, so $v(\mathbb{U}(\phi \vee \top)) = 0$. Therefore, $v(\mathbb{U}\mathbb{U}(\phi \vee \top)) = 0$ too. So this sentence is clearly false.*

More generally, it is not hard to see that for any formula ϕ with a value other than 0 or 1, $\mathbb{U}\phi$ can be represented using the function depicted in Figure 3 below.

Furthermore, the next two theorems show that this function respects the constraints imposed above on any function h adequately representing the unclarity operator:

THEOREM 3.7 (Symmetry I). *Let $d(x, y)$ be the distance between x and y, i.e. $d(x, y)$ is the absolute value of $x - y$. For all pairs of formulas ϕ and ψ in $\mathcal{L}^+$, if $d(v(\phi), \frac{1}{2}) = d(v(\psi), \frac{1}{2})$, then $v(\mathbb{U}\phi) = v(\mathbb{U}\psi)$.*

Proof. If $v(\phi), v(\psi) \in \{0, 1\}$, then the proof is straightforward. So assume that $d(v(\phi), \frac{1}{2}) = d(v(\psi), \frac{1}{2})$ for any two formulas ϕ and ψ such that $v(\phi), v(\psi) \in (0, 1)$. To reach a contradiction assume also that $v(\mathbb{U}\phi) \neq v(\mathbb{U}\psi)$. If $v(\phi) = v(\psi)$, then the result holds trivially. If $v(\phi) \neq v(\psi)$, we can suppose without loss of generality that $v(\phi) > v(\psi)$. More specifically, $v(\phi) \in [\frac{1}{2}, 1)$ and $v(\psi) \in (0, \frac{1}{2}]$. Since $v(\mathbb{U}\phi) \neq v(\mathbb{U}\psi)$,

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

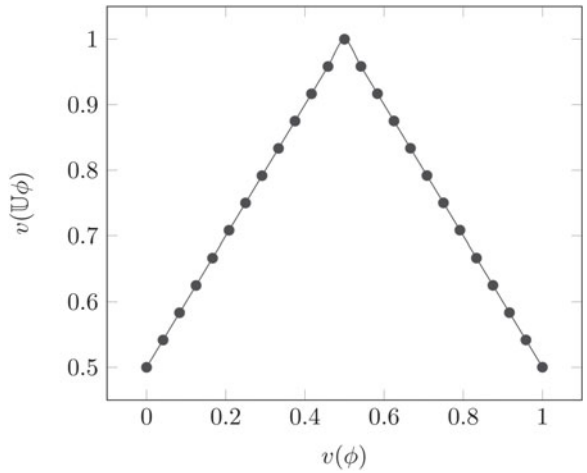


Fig. 3. The unclarity operator $\mathbb{U}$.

by the definition of $\mathbb{U}$, we obtain:
$\frac{1}{2} + (1 - v(\phi)) \neq \frac{1}{2} + v(\psi)$.
Given that $d(v(\phi), \frac{1}{2}) = d(v(\psi), \frac{1}{2})$, we know that $v(\phi) - \frac{1}{2} = \frac{1}{2} - v(\psi)$. Therefore, we can infer that:
$\frac{1}{2} + (1 - v(\phi)) + (v(\phi) - \frac{1}{2}) \neq (\frac{1}{2} + v(\psi)) + (\frac{1}{2} - v(\psi))$.
Simplifying, we obtain $1 \neq 1$, which is a contradiction. □

THEOREM 3.8 (Symmetry II). *Let $d(x, y)$ be as before, and let ϕ and ψ be such that $v(\phi), v(\psi) \in (0, 1)$. If $d(v(\phi), \frac{1}{2}) < d(v(\psi), \frac{1}{2})$, then $v(\mathbb{U}\phi) > v(\mathbb{U}\psi)$.*

Proof. We have four cases:

1. $v(\phi)$ and $v(\psi)$ are both in $(0, \frac{1}{2}]$.
2. $v(\phi)$ and $v(\psi)$ are both in $[\frac{1}{2}, 1)$.
3. $v(\phi)$ is in $(0, \frac{1}{2}]$ and $v(\psi)$ is in $[\frac{1}{2}, 1)$.
4. $v(\phi)$ is in $[\frac{1}{2}, 1)$ and $v(\psi)$ is in $(0, \frac{1}{2}]$.

For all of them assume that $v(\mathbb{U}\phi) \leq v(\mathbb{U}\psi)$.

1. By the definition of $\mathbb{U}$ and the assumption we have $\frac{1}{2} + v(\phi) \leq \frac{1}{2} + v(\psi)$. And from this it follows that $d(v(\phi), \frac{1}{2}) \geq d(v(\psi), \frac{1}{2})$.

2. This time the definition of $\mathbb{U}$ together with the assumption give us $\frac{1}{2}$ + (1 - $v(\phi)$) $\leq$ $\frac{1}{2} + (1 - v(\psi))$. But this implies that $v(\phi) \geq v(\psi)$. Hence, $d(v(\phi), \frac{1}{2}) \geq d(v(\psi), \frac{1}{2})$.

3. Using the definition of $\mathbb{U}$ we obtain $\frac{1}{2} + v(\phi) \leq \frac{1}{2} + (1 - v(\psi))$. Hence $v(\phi) \leq 1 - v(\psi)$. But clearly, for any formula ϕ, $d(v(\phi), \frac{1}{2}) = d(1 - v(\phi), \frac{1}{2})$. So we can infer again that $d(v(\phi), \frac{1}{2}) \geq d(v(\psi), \frac{1}{2})$.

4. Similar to the previous case. □

Notice that under this framework the unclarity operator can be seen as a generalization of the pathologicality operator of the previous section. More specifically, $\mathbb{U}$ behaves exactly as π if we restrict ourselves to formulas with a semantic value in $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$.

Furthermore, although we will not get into the details here, it is possible to combine both approaches. Roughly, the idea is to have a richer language having self-referential

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

expressions, a truth predicate, at least one vague predicate, and an operator that can be interpreted as the unclarity operator or as the pathologicality operator, depending on the case. Technically, this can be done by "plugging in" the definition of the unclarity operator into the Yablo construction provided above, but we leave this for another occasion.

§4. Further Issues. There are a couple of issues that we haven't dealt with in this paper. First of all, the lack of a suitable conditional in the Strong Kleene schema is usually considered as an important problem for theories of transparent truth and for theories of vagueness. On the one hand, even though truth is transparent in the sense that any formula ϕ is intersubstitutable *salva veritate* with $Tr\langle\phi\rangle$, the absence of a proper conditional connective implies that the theory does not prove the unrestricted T-Schema, something that is usually expected of theories of truth employing a non-classical logic. On the other hand, the lack of a suitable conditional also implies that in this framework there can be no proof of the Tolerance Principle, if it is expressed as a sentence rather than as an inference.

However, this is not as serious as it sounds. We have only tried to solve *one* of the problems usually attributed to Kripke's account of truth, but naturally there are many others (the same applies in the case of vagueness). In any case, nothing of what we have said is incompatible with the possibility of adding a nice conditional on top of the theories we have presented.

Secondly, in Theorem 2.16 we showed how the pathologicality operator interacts with the propositional connectives. Some of those facts, however, might be regarded as controversial for the unclarity operator. In particular, there seem to be no obvious reason to believe that $\mathbb{U}\phi$ should be implied by $\mathbb{U}\mathbb{U}\phi$. If it is unclear whether ϕ is unclear, how can we be sure that ϕ is unclear? In the case of pathologicality, if a formula of the form $\pi\pi\phi$ holds, this is because ϕ is pathological, so $\pi\phi$ should hold too. In the case of unclarity there is a similar line of reasoning available, although perhaps it is not as compelling. The only way for a formula of the form $\mathbb{U}\mathbb{U}\phi$ to hold is for ϕ to be unclear, so $\mathbb{U}\phi$ should hold as well.

Thirdly, a sentence like $\pi(\lambda \wedge \pi\delta_1)$ will have value $\frac{3}{4}$, but one would perhaps expect it to be equivalent to $\pi\lambda$. This can be fixed by tweaking the definition of π. One way of doing this is by not letting any formula of the form $\pi\phi$ to have value 1. This is similar to what Beall (2006) does, where $\pi\lambda$, $\pi(\lambda \wedge \pi\delta_1)$ and every other pathological sentence to which the pathologicality operator is applied receives the value $\frac{3}{4}$. However, as we have stressed before, this seems a high price to pay. Although there is a straightforward argument for the claim that $\pi\delta_1$ is pathological and true, any argument establishing that $\pi\lambda$ is pathological and true would have to be different. A different line of response is that it is not straightforward that $\lambda \wedge \pi\delta_1$ should be equivalent to λ. After all there is a sense in which the first sentence does seem to be pathological in addition to being true. If that is so, then there is no reason to expect $\pi(\lambda \wedge \pi\delta_1)$ to be equivalent to $\pi\lambda$.

We do not consider these problems to be fatal for the present account, but they are indeed problems, and at the moment we are unaware of any elegant way of solving them without substantively changing the frameworks.

§5. Closing remarks. The main goal of this paper was to show how the semantic characterization problem can be solved by letting certain sentences belong to more than one semantic category. In doing so we provided a formal construction that gives a reasonable interpretation for the pathologicality operator π. We have also sketched a way of extending this approach to languages that include vague predicates.

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press

§6. Acknowledgments. Earlier versions of the material in this paper were presented at conferences in Buenos Aires and Paris. We would like to thank the audiences of those conferences for their valuable comments and suggestions. We are particularly grateful to Eduardo Barrio, Alexandre Billon, Natalia Buacar, Catrin Campbell-Moore, Roy Cook, Eleonora Cresto, Paul Egré, Graham Leigh, Carlo Nicolai, David Nicolas, Federico Pailos, Lavinia Picollo, Thomas Schindler, Johannes Stern, James Studd, Diego Tajer, Paula Teijeiro, Jérémy Zehr, and an anonymous referee of this journal. We would like to add that the paper could not have been written without the financial aid of CONICET and CIN.

BIBLIOGRAPHY

Beall, J. (2006). True, false and paranormal. *Analysis*, **66**(2), 2–14.

Beall, J. (2009). *Spandrels of Truth*. Oxford: Oxford University Press.

Field, H. (2005). Variations on a theme by yablo. In Beall, J., & Armour-Garb, B. editors. *Deflationism and Paradox*. Oxford: Oxford University Press, pp. 53–74.

Field, H. (2007). Solving the paradoxes, escaping revenge. In Beall, J. editor. *Revenge of the Liar: New Essays on the Paradox*. Oxford: Oxford University Press, pp. 78–144.

Field, H. (2008). *Saving Truth from Paradox*. New York: Oxford University Press.

Gupta, A., & Belnap, N. (1993). *The Revision Theory of Truth*. Cambridge, Mass: MIT Press.

Gupta, A., & Martin, R. (1984). A fixed point theorem for the weak Kleene valuation scheme. *Journal of Philosophical Logic*, **13**, 131–35.

Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, **72**(19), 690–716.

Ripley, D. (2012). Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, **5**(2), 354–378.

Yablo, S. (2003). New grounds for naive truth. In Beall, J. editor. *Liars and Heaps*. Oxford: Oxford University Press, pp. 312–330.

PUAN 480, INSTITUTO DE FILOSOFIA
FACULTAD DE FILOSOFIA Y LETRAS, UNIVERSIDAD DE BUENOS AIRES
C.P. 1406, CIUDAD DE BUENOS AIRES, ARGENTINA
E-mail: l_rosenblatt@hotmail.com

PUAN 480, INSTITUTO DE FILOSOFIA
FACULTAD DE FILOSOFIA Y LETRAS, UNIVERSIDAD DE BUENOS AIRES
C.P. 1406, CIUDAD DE BUENOS AIRES, ARGENTINA
E-mail: dami.tih@gmail.com

https://doi.org/10.1017/S1755020314000239 Published online by Cambridge University Press