

Metaphysics as Assessment-Sensitive Explanation

John M. Bunke
University of Toronto

Düsseldorf Graduate Workshop
30 August 2018

1 Introduction

- Two recently popular points of view on the methodology and goals of metaphysics:
 - **Extraordinary realism** says that rival metaphysical theses are alternative *theories* of the relevant phenomena; e.g. Fine (2001), Schaffer (2009), and Sider (2011)
 - **Deflationism** says that rival metaphysical theses are more like alternative *languages*; e.g. Hirsch (2009, 2010), as well as Thomasson (2007, 2015) and Chalmers (2009, 2011)
- Example debate:
 - **Endurantism**: everyday material objects like books are three-dimensional things that persist through time by (wholly) existing at every moment in a given temporal interval
 - **Perdurantism**: everyday material objects like books are four-dimensional things that persist through time by having temporal parts located at every moment in a given temporal interval

2 Three Types of Discourse

- Let's use the term **disputant** for subjects with differing attitudes towards (seemingly) a single topic.
 - **Type I**: It *must* be the case that (at least) one disputant is wrong about the topic
 - A: "Sydney is the capital of Australia."

B: "Canberra is the capital of Australia."

- **Type II**: It *might* be that one (or more) disputant is wrong (and it might be that no disputant is wrong)

A: "Licorice is tasty."

B: "Licorice is not tasty."

- **Type III**: It *cannot* be the case that any disputant is wrong about the topic at issue, since there is no *single* topic at issue

A_[Toronto, spring 2014]: "Last winter was very cold."

B_[Toronto, spring 2018]: "Last winter was not particularly cold."

- *Extraordinary realism* treats metaphysical disputes as *type I* discourse: at most one of those two theories can be correct, so one of the disputants must be in error.
- *Deflationism* treats these as *type III* discourse: nobody is really *wrong*, the disputants just speak different languages, so each is as "correct" as any other.

3 Assessment Sensitivity

3.1 Kaplanian Apparatus

- We begin with some basic apparatus from Kaplan (1977).
- $\llbracket \cdot \rrbracket$ is the *semantic interpretation* function. With no superscripts, it takes an expression to its Kaplanian *character*, which is a function from context of utterance (or: use) to the *content* for an expression, which itself is a function from a world and a time (the *circumstance of evaluation*) to the *denotation* (or: extension) of the expression.
- ' $\lambda\alpha.\beta$ ' (where β is a non-sentential expression) is read: 'the (smallest) function that maps α to β '. ' $\lambda\alpha.\phi$ ' (where ϕ is a sentence) is read: 'the (smallest) function that maps α to 1 iff ϕ '.

- Contexts of utterance are centered worlds, i.e. ordered triples of a world, time, and agent.
- *Denotation*: $\llbracket \phi \rrbracket^{u,w,t} = 1$ iff ϕ is true in context of use u as evaluated with respect to circumstance of evaluation $\langle w, t \rangle$.
- *Content*: $\llbracket \phi \rrbracket^u = \lambda w \lambda t. \llbracket \phi \rrbracket^{u,w,t}$
- *Character*: $\llbracket \phi \rrbracket = \lambda u. \llbracket \phi \rrbracket^u$
- $S(x) = \lambda x$. the agent (subject) of context x ; $T(x) = \lambda x$. the time of x ; $W(x) = \lambda x$. the world of x ; $A(x) = \lambda x$. the addressee of x .
- A few example semantic entries:

$$\llbracket \mathbf{I} \rrbracket = \lambda u \lambda w \lambda t. S(u)$$

$$\llbracket \text{Now } \phi \rrbracket^u = \lambda w \lambda t. \llbracket \phi \rrbracket^{u,w,T(u)}$$

$$\llbracket \text{Actually } \phi \rrbracket^u = \lambda w \lambda t. \llbracket \phi \rrbracket^{u,W(u),t}$$

- All expressions have contents relative to a context of use; and most expressions will have denotations relative to a context of use and circumstance of evaluation.
- But, according to the assessment-sensitivity view, for some expressions a context of use and circumstance of evaluation are not enough to generate a determinate denotation (truth-value). For these expressions, denotation is also sensitive to the context of assessment, specifically to some feature of the subject of the context of assessment, such as her flavor-valuations, in the case of ‘tasty’.

3.2 Four Species of Context-Sensitive Semantics

- Four ways in which an expression, such as ‘tasty’, can be context-sensitive (cf. Lasnik 2017, 192):

	Assignment of <i>content</i> to expressions is sensitive	Assignment of <i>denotation</i> to content is sensitive
Sensitive to context of <i>use</i>	Indexical contextualism (IC)	Nonindexical contextualism (NIC)
Sensitive to context of <i>assessment</i>	Indexical assessment sensitivity (IAS)	Nonindexical assessment sensitivity (NIAS)

- **Indexical contextualism (IC)** is often assumed to be the correct semantics for ‘I’ and ‘you’ and it is the view of eternalists about tense, e.g. Richard (1981).
- **Nonindexical contextualism (NIC)** is the view of temporalists about tense, e.g. Prior (1968) or Kaplan (1977) (I follow Kaplan here for the purposes of this paper).
- Very quickly: It is common to relativize the truth-values of sentences in context to a possible world parameter to account for contingency. Thus the content of a (declarative) sentence (in a context of use) is a function from a world to the truth-value of the sentence evaluated at that world, or, equivalently, the set of worlds in which it is true.
- Temporalism relativizes truth-values to a world *and a time*. Temporalism says that many sentence-contents do not contain references to a time, even unpronounced ones.
- Eternalism, in contrast, says that all sentence-contents contain references to a time, but in some sentences these remain unpronounced; sentence-contents are functions only from worlds to truth-values.
- Consider the sentence-occurrence σ : ‘Turkey is in a recession’ said on 15 August 2018 by Robert in the actual world (@).¹ According to eternalism, the content of this occurrence is equivalent to the content of ‘Turkey is in a recession now’ in the same context; the ‘now’ is unpronounced in σ :

$$\text{Etern: } \llbracket \sigma \rrbracket^{u1} = \llbracket \text{Turkey is in a recession now} \rrbracket^{u1}$$

¹ A sentence-occurrence is a token sentence in a context of use—see: Kaplan (1977, 522).

$$= \lambda w. \llbracket \text{Turkey is in a recession at } T(u_1) \rrbracket^{u_1, w}$$

$$= \lambda w. \llbracket \text{Turkey is in a recession on 15/8/18} \rrbracket^{u_1, w}$$

- Compare this to the content of the same sentence-occurrence according to temporalism:

$$\text{Temp: } \llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t. \llbracket \text{Turkey is in a recession} \rrbracket^{u_1, w, t}$$

- Temporalism and eternalism differ on what σ 's content is. But they agree on its denotation in the context of utterance.
- **Indexical assessment sensitivity (IAS)** has rarely been endorsed so from now on I will ignore it.
- **Nonindexical assessment sensitivity (NIAS)** says that the denotation of an expression varies with some feature of the context of assessment. This is the type of view endorsed by MacFarlane (2014) and Lasersohn (2017) about 'tasty'.
- In general, IC and NIC (e.g. eternalism and temporalism) assign different contents to the same sentence-occurrence but the same denotation. In contrast, NIC and NIAS assign the same content, but different denotations.

3.3 Predicates of Personal Taste

- Absolutism (Ab) treats 'tasty' like 'weighs 8kg', that is, as expressing an objective property.
- We will use sentence occurrence σ : 'Licorice is tasty' as said in $u_1 = \langle @, 1/1/18, \text{Greta} \rangle$.

$$\text{Ab (content): } \llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t. \llbracket \text{Licorice is tasty} \rrbracket^{u_1, w, t}$$

$$\text{Ab (denotation): } \llbracket \sigma \rrbracket^{u_1, W(u_1), T(u_1)}$$

$$= [\lambda w \lambda t. \text{Licorice is tasty in } w \text{ at } t](@)(1/1/18)$$

$$= 1 \text{ iff licorice is tasty in } @ \text{ on } 1/1/18$$

- Like eternalism, the IC account of 'tasty' will posit an unpronounced element; we will use 'according to $S(x)$ ' (for simplicity):

$$\text{IC (content): } \llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t. \llbracket \text{Licorice is tasty, according to } S(u_1) \rrbracket^{u_1, w, t}$$

$$= \lambda w \lambda t. \llbracket \text{Licorice is tasty, according to Greta} \rrbracket^{u_1, w, t}$$

$$\text{IC (denotation): } \llbracket \sigma \rrbracket^{u_1, W(u_1), T(u_1)}$$

$$= [\lambda w \lambda t. \text{Licorice is tasty in } w \text{ at } t, \text{ according to Greta}](@)(1/1/18)$$

$$= 1 \text{ iff licorice is tasty in } @ \text{ on } 1/1/18, \text{ according to Greta}$$

- Ab and IC assign different contents and different denotations.
- NIC assigns a different content from IC, but the same denotation:

$$\text{NIC (content): } \llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t \lambda s. \llbracket \text{Licorice is tasty} \rrbracket^{u_1, w, t, s}$$

$$\text{NIC (denotation): } \llbracket \sigma \rrbracket^{u_1, W(u_1), T(u_1), S(u_1)}$$

$$= [\lambda w \lambda t. \text{Licorice is tasty in } w \text{ at } t, \text{ according to Greta}](@)(1/1/18)$$

$$= 1 \text{ iff licorice is tasty in } @ \text{ on } 1/1/18, \text{ according to Greta}$$

- According to an NIC account of 'tasty', the content is a function from world, time, *and subject* to the denotation. In the same way that temporalism says that contents are "neutral" with respect to times—i.e., that the same content can be evaluated with respect to different times—NIC about 'tasty' says that contents are neutral with respect to subjects (strictly speaking, with respect to flavor-valuations).
- Why does it matter that the content is different if the denotation is the same? Because content is "the level at which logical notions such as contradiction and entailment are fundamentally defined" (Lasersohn 2017, 8).²
- We can say that two contents Φ and Ψ are **contradictory** just in case the *intersection* of the set of worlds (or world-time pairs, or

² For Kaplan, content is supposed to capture "what is said" by an utterance in a context (Kaplan 1977, 500).

world–time–subject triples, etc.) in which Φ is true and the worlds (or world–time pairs, etc.) in which Ψ is true *is empty*.

- NIAS assigns the same content as NIC but the denotation is different:

NIAS (content): $\llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t \lambda s. \llbracket \text{Licorice is tasty} \rrbracket^{u_1, w, t, s}$

NIAS (denotation): $\llbracket \sigma \rrbracket^{u_1, W(u_1), T(u_1), S(a)}$

$= [\lambda w \lambda t. \text{Licorice is tasty in } w \text{ at } t, \text{ according to } S(a)](@)(1/1/18)$

$= 1$ iff licorice is tasty in @ on 1/1/18, according to $S(a)$

- $S(a)$ is the subject of the context of assessment.
- So when we evaluate σ for truth in its context of utterance, we don't yet get a truth value, since a is a free variable.

3.4 Disagreement

- NIC is an advance over IC in part because it has a better (though still unsatisfactory) account of disagreement.
- Let's consider two sentence-occurrences: σ is an utterance of 'Licorice is tasty' in $u_1 = \langle @, 1/2/13, \text{Jeff} \rangle$ and τ is an utterance of 'Licorice is not tasty' in $u_2 = \langle @, 30/8/17, \text{Jessica} \rangle$.
- According to IC, the contents of these sentence-occurrences aren't contradictory:

IC: $\llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t. \llbracket \text{Licorice is tasty, according to } S(u_1) \rrbracket^{u_1, w, t}$

$= \lambda w \lambda t. \llbracket \text{Licorice is tasty, according to Jeff} \rrbracket^{u_1, w, t}$

IC: $\llbracket \tau \rrbracket^{u_2} = \lambda w \lambda t. \llbracket \text{Licorice is not tasty, according to } S(u_2) \rrbracket^{u_2, w, t}$

$= \lambda w \lambda t. \llbracket \text{Licorice is not tasty, according to Jessica} \rrbracket^{u_2, w, t}$

- In contrast, the contents assigned to σ and τ by NIC *are* contradictory:

NIC: $\llbracket \sigma \rrbracket^{u_1} = \lambda w \lambda t \lambda s. \llbracket \text{Licorice is tasty} \rrbracket^{u_1, w, t, s}$

NIC: $\llbracket \tau \rrbracket^{u_2} = \lambda w \lambda t \lambda s. \llbracket \text{Licorice is not tasty} \rrbracket^{u_2, w, t, s}$

- Nevertheless, there is still a way in which Jeff and Jessica aren't really disagreeing on the NIC view. Although it is true that what Jeff said is that licorice is tasty, and what Jessica said is that licorice is not tasty, still, when we evaluate their utterances for truth value they are evaluated with respect to different subjects: they have jointly-satisfiable conditions under which their denotations are both 1. (Remember: IC and NIC assign the same denotations, just different contents.)
- NIAS is an advance over NIC because the conditions under which their denotations are both 1 will be incompatible. Suppose that Robert is the assessor; then the denotation of σ will be: *1 iff licorice is tasty in @ on 1/1/18, according to Robert* and the denotation of τ will be: *1 iff licorice is not tasty in @ on 1/1/18, according to Robert*. No matter who we substitute in for Robert as the assessor, including the interlocutors, e.g., Jeff, we get the same result (so long as all subjects have internally-consistent flavor-valuations).
- Still, though, it's not clear that arguing about whether one account or another *really* captures the sense of the disagreement or not; another sort of evidence for NIAS would be nice. The trouble is basically that it's not clear what "faultless disagreement" (Kölbel (2002)) is, or if it is, or if any of these theories capture that notion.

3.5 Retraction

- So here's a different sort of case: χ : 'I am standing up' as said in $u_3 = \langle @, 10\text{pm}, \text{Jeff} \rangle$ and ω : 'You are not standing up' as said (to Jeff) in $u_4 = \langle @, 11\text{pm}, \text{Jessica} \rangle$; and let's suppose that at 11pm Jeff is indeed not standing up, because he is lying down. And Jessica says (at 11pm):
 - (1) You lied to me—you said 'I am standing up', but you aren't standing up.
- The correct response for Jeff to make is *not*:
 - (2) #Yes, I see your point; I take back what I said.

- Instead, he is likely to say something along the lines of: “Well, look, I *was* standing up then, and *now* I am not.”
- Jessica’s utterance ω is correct: at 11pm Jeff is not standing up. But just because his utterance is *different* from hers, doesn’t mean he is or was *wrong*. This is why it would be absurd for him to “take back” or *retract* his previous assertion as in (2).³
- There is a difference between being *unwilling to make* an assertion that one previously made and *retracting* one’s previous assertion for being incorrect. In the context in which he uttered it, Jeff’s assertion was correct. He can no longer correctly make the same assertion, because the context has changed, but that is different than finding fault with his previous assertion.
- Notice that speakers are inclined to retract their assertions about which things are tasty, once their tastes change. Compare (3) and (4) to (5) and (6):

(3a) Q: Is orange soda tasty?

(3b) Jeff_[age 6]: Orange soda is tasty; in fact, it’s delicious!

Thirty years later:

(4a) Q: Is orange soda tasty? You said it was when I asked you before.

(4b) Jeff_[age 36]: Ugh, no: orange soda isn’t tasty, it’s disgusting. I don’t know how I could have been so wrong.

...

(5a) Q: Are you over 5 feet tall?

(5b) Jeff_[age 6]: No, I’m only 3 feet 8 inches tall.

Thirty years later:

(6a) Q: Are you over 5 feet tall? You said you weren’t when I asked you before.

(6b) Jeff_[age 36]: Yes, I am; I am 6 feet 2 inches tall. #I don’t know how I could have been so wrong.

- In (4), “I don’t know how I could have been so wrong”—which signifies a willingness to retract the earlier assertion—is felicitous, whereas in (6) it sounds quite strange.
- If you and I have differing attitudes about a topic, then we may be obliged to revise our attitudes, if one of us turns out to be incorrect (it may turn out that we are both correct—we do not yet know that one of us *must* be wrong, just that one *might* be). The felicity of retraction-statements indicates that this is a live possibility, since retraction is part of the process of engaging in such revision.
- If retraction is not felicitous, that suggests that the discourse is type III, because in this case the mere fact of differing attitudes has not thereby made incorrectness a live possibility. Why not? Because no one topic is (yet) at issue: our differing attitudes do not ipso facto jeopardize our attitudes.
- Like IC, NIC also incorrectly categorizes ‘tasty’-discourse as type III.
- Note, however, the felicity of retraction doesn’t decide between types I and II: e.g. if I assert that Sydney is the capital of Australia, then, on finding out that it is not, I should be willing to retract that assertion, because it is incorrect.

4 Metaphysics as Assessment-Sensitive Explanation

- In metaphysical discourse, speakers are willing to retract their previous assertions if they have changed their mind from being, e.g., an endurantist to a perdurantist. Suppose that Greta is an endurantist at noon, but, by 5pm she has been convinced by Robert that perdurantism is correct:

(7a) Robert_[noon]: Is this book (entirely) on the table?

³ See MacFarlane (2014, Ch. 5) for discussion of the notion of retraction.

(7b) Greta_[noon]: Yes, of course it is.

Five hours later:

(8a) Robert_[5pm]: Is this book (entirely) on the table? You said it was when I asked you before.

(8b) Greta_[5pm]: No, definitely not. I don't know how I could have been so wrong. I see now that I should have taken more care in thinking through the argument from general relativity.

- This sounds fine, including “I don't know how I could have been so wrong”, although it does reveal that Greta is being a bit hard on herself!
- This suggests that metaphysical discourse is not of type III; but why should we think it is type II rather than type I?
- One could point to the seeming *intractability* of these disagreements.
- But I think a better way is not to emphasize the lack of progress on metaphysical questions but instead to sketch an alternative positive picture of what metaphysics is, and to focus on the feeling that both endurantism and perdurantism might both somehow be correct.
- Instead of distinct *theories* of which at most one can be correct, or distinct *languages* of which none is more correct than any other, perhaps we should treat endurantism and perdurantism as *explanatory paradigms* (or: explanatory frameworks) which might both generate correct explanations.
- The model here are distinct styles or sorts of explanation, such as:
 - (i) explanations of behavior in neuro-physiological terms versus explanations in folk-psychological terms;
 - (ii) explanations of the “behavior” of objects in virtue of the stuff that makes them up and explanations of their behavior in virtue of their manner of organization;
 - (iii) explanations of a phenomenon in statistical terms versus in terms of causal mechanisms.

- In these cases the correctness of one (sort of) explanation does not preclude the correctness of others.
- This doesn't mean that there aren't better and worse explanations of the relevant phenomena. There can of course be correct and incorrect explanations, as well as better and worse (where both are correct).
- How would this work in metaphysics, specifically in the case of endurantism and perdurantism? There are “behaviors” of ordinary material objects (like books) that might be explained in either way and neither sort of explanation excludes the other from being correct. These provide ways of seeing how the same facts can “hang together” in different ways.
- Implementation of NIAS semantics: we would still have contents that are functions from worlds, times, and subjects to denotations. In evaluating this content for truth we would fill in the subject position with the agent of the context of assessment, not the agent of the context of use.
- The idea is that sentences containing, e.g., ‘book’, when used in the context of a metaphysical dispute, would differ in their denotations (truth-value) based on the explanatory paradigm of the assessor.
- Consider the sentence-occurrence $\sigma = \text{‘Books are three-dimensional’}$ as said in $u_5 = \langle @, 11\text{am}, \text{Suzie} \rangle$:

$$\llbracket \sigma \rrbracket^{u_5} = \lambda w \lambda t \lambda e. \llbracket \text{Books are three-dimensional} \rrbracket^{u_1, w, t, e}$$

$$\llbracket \sigma \rrbracket^{u_5, w, t, e} = \llbracket \text{Books are three-dimensional} \rrbracket^{u_1} (W(u_1))(T(u_1))(E(a))$$

$$= [\lambda w \lambda t \lambda e. \text{Books are three-dimensional in } w \text{ at } t, \text{ assuming } e](@)(11\text{am})(E(a))$$

$$= 1 \text{ iff books are three-dimensional in } @ \text{ at } 11\text{am}, \text{ assuming } E(a)$$
- $E(a)$ is the explanatory framework of the context of assessment.
- If the explanatory framework relevant at a context of assessment is endurantism, then the denotation of σ will be: *1 iff books are three-dimensional in @ at 11am, assuming endurantism* (so: the denotation will be 1).

- In contrast, if the relevant explanatory framework is perdurantism, the denotation will be 0.
- On a NIC treatment of such disputes, it would be the explanatory paradigm of Suzie (the agent of the context of *use*) that is relevant to determining the denotation of the sentence; on the NIAS treatment, it depends on what the relevant context of assessment is.
- Note that on the NIAS treatment I am describing, ‘Books are three-dimensional’ will be a *necessary* truth (in the sense of being *eternally* and *non-contingently* true), but it will not be a *logical truth*.⁴

5 Conclusion

- Perhaps this is a way of substantiating the idea that philosophy in general—and metaphysics in particular—seeks explanations (or *understanding*) of phenomena rather than theories, per se:

“[S]uccess in philosophy would be ... to ‘know one’s way around’ ... , a form of knowing *how* as contrasted with knowing *that*” (Sellars 1962, 1).

“[T]he philosopher searches for deeper explanatory principles To show that these principles, if true, would explain *p* involves deducing *p* from them Yet still, this is no attempt to prove *p*; and the explanatory hypotheses used in the explanation need not be known to be true, or be believed on grounds independent of *p* itself. To produce this possible explanation of *p* is, by seeing one way *p* is given rise to, to see how *p* can be true” (Nozick 1981, 11).

References

Chalmers, David J. 2009. “Ontological Anti-Realism”. In *Metametaphysics: New Essays on the Foundations of Ontology*, ed. by David J. Chalmers, David Manley, and Ryan Wasserman, 77–129. Oxford University Press.

- . 2011. “Verbal Disputes”. *Philosophical Review* 120 (4): 515–566.
- Fine, Kit. 2001. “The Question of Realism”. *Philosophers’ Imprint* 1 (1): 1–30.
- Hirsch, Eli. 2009. “Ontology and Alternative Languages”. In *Metametaphysics: New Essays on the Foundations of Ontology*, ed. by David John Chalmers, David Manley, and Ryan Wasserman, 231–258. Oxford University Press.
- . 2010. *Quantifier Variance and Realism: Essays in Metaontology*. Oxford University Press.
- Kaplan, David. 1977. “Demonstratives”. In *Themes From Kaplan*, ed. by Joseph Almog, John Perry, and Howard Wettstein, 481–563. Oxford University Press.
- Klbel, Max. 2002. *Truth Without Objectivity*. Routledge.
- Lasersohn, Peter. 2017. *Subjectivity and Perspective in Truth-Theoretic Semantics*. Oxford University Press.
- MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford University Press.
- Nozick, Robert. 1981. *Philosophical Explanations*. Harvard University Press.
- Prior, A.N. 1968. *Papers on Time and Tense*. Oxford University Press.
- Richard, Mark. 1981. “Temporalism and Eternalism”. In *Truth and Truth Bearers: Meaning in Context, Volume II*, 15–24. Oxford University Press.
- Schaffer, Jonathan. 2009. “On What Grounds What”. In *Metametaphysics: New Essays on the Foundations of Ontology*, ed. by David Manley, David J. Chalmers, and Ryan Wasserman, 347–383. Oxford University Press.
- Sellars, Wilfrid. 1962. “Philosophy and the Scientific Image of Man”. In *Science, Perception, and Reality*, 1–40. Routledge.
- Sider, Theodore. 2011. *Writing the Book of the World*. Oxford University Press.
- Thomasson, Amie L. 2007. *Ordinary Objects*. Oxford University Press.
- . 2015. *Ontology Made Easy*. Oxford University Press.

⁴ MacFarlane (2014, 69) distinguishes two ways of extending the Kaplanian account of logical truth but the sentence fails to meet both criteria.