

How (not) to be a modalist about essence

Draft of a chapter for a collection entitled 'Reality Making'

(eds. Mark Jago & Stephan Barker)

Please don't quote or cite

Abstract: Rather infamously, Kit Fine provided a series of counter-examples which purport to show that the modalist program of analysing essence in terms of metaphysical necessity is fundamentally misguided. Several would-be modalists have since responded, attempting to save the position from this Finean Challenge. This paper evaluates and rejects a trio of such responses, from Della Rocca (1996), Zalta (2006), and Gorman (2005). But I'm not here arguing for Fine's conclusion – ultimately, this is a fight amongst friends, with Della Rocca, Zalta, Gorman, and I all wanting to be modalists, but disagreeing on the details. As such, while my primary aim is to show what's *wrong* with this trio, the secondary aim is demonstrating how what's *right* about them in fact pushes one towards my own sparse modalist account. So while the primary conclusion of this paper is negative, the secondary, positive, conclusion is that modalists shouldn't give up hope – plausible responses to Fine are still out there.

Word count: 7,872 without references

§1. The Finean Challenge

Famously, Fine argued that essence 'is not to be understood in modal terms or even to be regarded as extensionally equivalent to a modal notion' (1994: 3). This is because the standard modal definition is vulnerable to counter-examples; specifically, the properties *being such that there are infinitely many prime numbers*, *being such that the Eiffel Tower essentially is a tower*, *being distinct from the Eiffel Tower*, and *being a member of {Socrates}* all satisfy the right-hand side of

M Φ is essential to x iff_{df} necessarily, if x exists, then x has Φ

yet are, according to Fine, non-essential to Socrates.

This objection has been well received, leading many metaphysicians to abandon the modalist account as fundamentally mistaken.¹ Yet, it has not gone *entirely* unchallenged – several would-be modalists have attempted to save the account from the Finean threat.

Two basic strategies are employed. The first strategy modifies **M**, adding further necessary conditions a property must satisfy to be essential. Obviously, such restrictions shouldn't be *ad hoc* or question-begging. Further, the resulting modalist definition should neither be vulnerable to Fine's initial set of, nor to any new, counter-examples. A second, seldom employed, strategy is to argue that Fine's properties aren't *counter-examples* – i.e. that such properties are, *contra*-Fine's intuitions, genuine essential properties of Socrates.²

The primary aim of this paper is evaluate a trio of responses – from Della Rocca (1996), Zalta (2006), and Gorman (2005), respectively – that, in the vein of the first strategy, offer ways to modify **M** so as to cut-off Fine's counter-examples. Ultimately, I conclude that this trio is unsuccessful; Della Rocca's account entails that no properties are essential to any object, Zalta's reply falls prey to new counter-examples, and Gorman's reply is based on a

¹ Further, the objection has served as the foundation for Fine's own fundamentalist account of essence, gradually spelled out in a series of papers over the past 30-odd years.

² Anecdotally, this seems to be the go-to move for modalists when discussing Fine's objection in conversation. However, the individuals who have advocated this 'bullet-biting' strategy in print are myself (Wildman forthcoming) and Zalta (2006: 684).

notion that is too vague to be load-bearing. Hence the primary conclusion of this paper is negative: if modalists want salvation from Fine, it won't be in these ways.

Yet, for all their flaws, these three replies aren't *entirely* off-base. Each can, after careful consideration, teach modalists valuable lessons about what a good response to Fine should look like – lessons which, I argue, push modalists towards modifying **M** by adding a sparse property restriction. Thus the secondary, positive, conclusion of this paper is that would-be modalists shouldn't give up hope – instead, they should embrace what I've elsewhere called sparse modalism. In effect, this paper is a fight amongst friends: Della Rocca, Zalta, Gorman, and I all want to be modalists, but disagree about the details. My aim is to show (1) what's wrong with their stories, and (2) how what's right about them in fact pushes one towards my preferred sparse modalist account.

Having here set-up the Finean challenge, the next section presents Della Rocca's response, which attempts to respond to Fine by excluding from essentiality 'trivial' necessary properties. After spelling the response out (§2.1), I raise a 'Pyhrric' objection (§2.2), showing how the suggested restriction entails that no properties are essential and, as such, the response must be rejected. I then turn to a reply from Zalta, based upon his unique metaphysical framework. Laying out both the framework and the response (§3.1-2) leads to the objection that the position is vulnerable to new, Fine-style counter-examples (§3.3). Following a lengthy discussion, I conclude that Zalta has no suitable way to dismiss these new problem cases, especially given the significant theoretical cost of his position. Moving on, I next (§4) look at Gorman's 'characteristic' response. While the basic idea here is a good one, the details are lacking; in effect, modalists can't trust Gorman's response because they simply don't know what they'd be committing themselves to if they did. The final section (§5) puts a positive spin on the above negative results, suggesting that, while the trio fail, they all shed light on what a successful modalist response ought to look like. After briefly describing the lessons the trio teach us, I conclude by offering modalists a potential ray of hope in the form of a sparse modalist account

§2. Trivial Properties: Della Rocca's Reply

The first response is motivated by the intuitive idea that properties that *all* objects necessarily possess are non-essential, because properties that *everything* necessarily has are too *trivial* to reveal an object's 'specific character' in any respect relevant to essentialism (Della Rocca 1996: 3). As such, the modalist definition should be restricted so as to exclude these trivial properties. A first stab at incorporating such a restriction results in:

TRIV Φ is essential to x iff_{df} (i) necessarily, if x exists, then x has Φ ; and (ii) it's not the case that, necessarily, every object has Φ if it exists

This modified definition captures the anti-trivial intuition but doesn't help against Fine: any 'degenerate' counter-example property that every object necessarily possesses can be conjoined to a non-degenerate, non-trivial property (Fine 1994: 7). For example, *being such that there are infinitely many prime numbers* is necessarily possessed by every object, so it is un-essential to Socrates, given **TRIV**. But take the property *being such that there are infinitely many prime numbers and human* – this isn't blocked by **TRIV**, yet it is no more essential to Socrates than its' universally necessary first conjunct.

§2.1 From Triv to Trivial

The way forward, Della Rocca suggests, is to note that there are *two* ways a property can be trivial: the first is to be a straightforward universally necessary property like *being self-identical*, *being male if a bachelor*, and *being human or not human*. For ease, call these

‘directly’ trivial properties. Meanwhile, a property Φ is *indirectly* trivial iff Φ is such that an object’s possession of it logically follows from the object’s possession of some property Ψ , where Ψ is a universally necessary property. For example,

...being identical with A is not universally necessary; in fact, this property is necessary only to A and, further, necessarily, only A has this property at all. However, since we can derive the fact that A has the property of being necessarily identical to A without having any information about A’s qualities other than the trivial fact that A is self-identical, the property of being identical with A is, though necessary, trivially so. (Della Rocca 1996: 3)

In other words, while *being Socrates* is not a universally necessary property, because Socrates’ necessary possession of it can be derived from the necessary truth of an instance of the universally necessary *being self-identical*, *being Socrates* is also trivial (albeit indirectly).

Acknowledging both ways that a property can be trivial, we can expand **TRIV** to get:

TRIVIAL Φ is essential to x iff_{df} (i) necessarily, if x exists, then x has Φ ; and (ii) it’s neither the case that, necessarily, every object has Φ if it exists, nor that x ’s having Φ logically follows from x ’s having Ψ , where Ψ is such that necessarily, every object has Ψ if it exists.

Unlike **TRIV**, **TRIVIAL** seems to offer the modalist a way to undercut Fine. Specifically, Socrates’ necessarily possessing *being a member of {Socrates}* logically follows from his possessing the universally necessary *being necessarily a member of {Socrates} if identical to Socrates*. And, as this latter property is directly trivial, the former is therefore indirectly trivial. Similar reasoning applies to Fine’s other examples – they are all trivial (either directly, since they are universally necessary, or indirectly, logically following from direct trivial properties). This is true even for the ‘non-degenerate’ modified counter-examples, like *being such that there are infinitely many prime numbers and human*: Socrates’ necessary possession of this property logically follows from his possessing the universally necessary *being such that there are infinitely many prime numbers and human if identical to Socrates*. As such, it seems modalism is saved by **TRIVIAL**!

§2.2 Rejecting **TRIVIAL**: a Pyrrhic argument

Sadly, any victory **TRIVIAL** provides modalism is entirely Pyrrhic. Grant that Socrates is necessarily human – i.e., that, necessarily, if Socrates exists, then he is human. From this, we can derive the property *being necessarily human if identical to Socrates*. This is a directly trivial property: every object necessarily possesses it! The problem is, from his having this property, we can derive the fact that Socrates necessarily possesses *being human* without any information about Socrates’s qualities other than the trivial fact that he is self-identical. In other words, from the conjunction of his possessing the directly trivial *being necessarily human if identical to Socrates* and the ‘trivial fact’ that Socrates is self-identical, it logically follows that Socrates is necessarily human. Then, because its necessary possession logically follows from the necessary possession of a property everything necessarily has, we must conclude that *being human* is an indirectly trivial property of Socrates’. So, by **TRIVIAL**, *being human* is non-essential to Socrates.

This generalizes: for any property Φ and any object x , if x necessarily has Φ , there is a property *being necessarily Φ if identical to x* that is directly trivial. Further, using only this property and the trivial fact that x is self-identical, it logically follows that x necessarily has Φ . And, as a property the necessary possession of which logically follows from the possession of a directly trivial property, Φ is therefore indirectly trivial. Then, given that no

trivial property is essential, it follows that Φ is not essential to x . In short: every property that satisfies **TRIVIAL**'s modal clause must fail to satisfy the non-triviality clause. The upshot is that **TRIVIAL** throws the baby out with the bathwater, entailing that there are *no* essential properties.

One might object to the above argument in the following way.³ To reach the conclusion that x necessarily has Φ , we appealed to two premises: first, the 'trivial fact' that x is self-identical, and, second, the universal fact that every object is such that, if it is x , then it is necessarily Φ . But, anyone who knows the second premise seems to know something very non-trivial about x – namely, *that it is necessarily Φ !* As a consequence, it seems that we appealed to more information about x 's qualities than the trivial fact that x is self-identical.

Note, however, that if this objection to my argument succeeds, it also undercuts Della Rocca's response to Fine. This is because the reasoning employed to generate my Pyrrhic conclusion exactly mirrors Della Rocca's reasoning in the quote above concerning the indirect triviality of *being Socrates*. More importantly, it is the exact same reasoning modalists must use to show the indirect triviality (and hence the non-essentiality) of properties like *being such that there are infinitely many prime numbers and human*. To clarify: the way a modalist can use **TRIVIAL** to undercut such a property is to show that it is indirectly trivial. This requires appealing to the direct triviality – i.e., the universally necessary possession – of properties *being such that there are infinitely many prime numbers and human if identical to Socrates*. But, if the objector is correct, then knowing that objects possess this property tells us something very non-trivial about Socrates – namely, that he is necessarily such that there are infinitely many prime numbers and is human! So the objection to my argument's success would leave the modalist unable to dismiss Fine's counter-examples.

However, I don't think the objection is successful, given that the following pair of sentences entail each other:

- (1) Necessarily, if Socrates exists, then Socrates is human
- (2) For all x , necessarily, if x is identical to Socrates, then x is human

In effect, all my Pyrrhic argument relies upon is this mutual entailment: (2) ascribes to every object a directly trivial, universally necessary property. Since (1) is entailed by (2), the property therein ascribed to Socrates must be indirectly trivial. That follows from the very definition of indirect triviality! The objector asserts that the property ascribed to Socrates in (1) must be non-trivial, which is just to assert the negation of the conclusion, not to point out a flaw in the reasoning. Of course, one man's *modus ponens* is another's *modus tollens* - perhaps this is where the argument stops.

Regardless, the end result is the same: despite its *prima facie* appeal, **TRIVIAL** does not help the modalist dismiss Fine's counter-examples. Modalists need indirect triviality to circumvent Fine, yet, it seems that either every would-be essential property is indirectly trivial (and hence not essential) or no property is indirectly trivial (and hence versions of Fine's counter-examples are mistakenly rendered essential). To that end, it is safe to say that no support for modalism can be found here.

§3. Bifurcation: Zaltarian Modalism

The second response, given by Zalta (2006), involves *bifurcating* the modalist definition, producing two distinct notions of essentiality applicable to abstract and concrete objects. After briefly detailing the requisite background for the response (§3.1), I then proceed to give

³ Thanks to an anonymous referee for pushing me on this objection.

the response itself (§3.2). As I go on to argue (§3.3), this bifurcation response is significantly flawed – so much so that modalists would do well to look elsewhere for a suitable reply to Fine.

§3.1 *The Foundations – Zalta’s four theses*

The Zaltarian response is built upon four theses that constitute part of Zalta’s idiosyncratic metaphysical picture. The first thesis concerns the distinction between abstract and concrete objects. According to the standard conception, abstract objects are necessarily abstract, concrete objects necessarily concrete, and the abstract/concrete divide is exhaustive and exclusive. Zalta rejects this traditionalist picture, in its place, suggesting an alternative *abstracta/concreta* demarcation,⁴ according to which abstracta are necessarily abstract and necessarily neither concrete nor non-concrete, concreta are necessarily either concrete or non-concrete, but necessarily not abstract, and *being non-concrete* is not the same as *being abstract*. In other words, abstracta are abstract, which concreta can never be, concreta are either concrete or non-concrete, which abstracta can never be, and being abstract or non-abstract is exhaustive and exclusive. This gives us the first thesis of Zalta’s picture:

AB-CON Objects are demarcated into the *abstracta* and the *concreta*

The second and third theses, meanwhile, concern the way objects can possess properties. Zalta calls the familiar mode of property possession *exemplification*. Exemplification is complete – necessarily, every object exemplifies either a property or its complement – and available to both abstracta and concreta. Yet along with exemplification, Zalta adds a second, special mode of property possession, exclusive to abstracta – that of *encoding* a property. Along with being abstracta exclusive, encoding is *gappy* – i.e., objects can fail to encode either a property or its complement. Thus we get our second thesis:

EXE-ENC Objects can either *exemplify* or *encode* a property

Directly building off the previous, the third thesis is about the *extension* of encoded properties:

EXTENSION Particular abstracta encode all and only those properties included in our conception of them

Take Sherlock Holmes, who is, according to Zalta, an abstractum. Sherlock encodes but does not exemplify *living in London*, though he does exemplify its complement. This is because, while his living in London is part of our conception of him, if you rounded-up all the Londoners, Holmes won’t be included. Similarly for *being a detective* – Holmes encodes this property (it’s part of our conception of him), fails to exemplify it (if you rounded up all the detectives, Holmes wouldn’t one of them), and exemplifies its complement (he is, in fact, a non-detective). Contrast this with *being fictional* and *being abstract*: Holmes doesn’t encode these properties, though he does exemplify them. Socrates, meanwhile, exemplifies *not living in London*, *being a philosopher*, and *being possibly non-concrete*, but doesn’t encode anything.

⁴ This is a deviation from (what I take to be) Zalta’s unfortunate terminology: Zalta uses ‘concrete object’ and ‘abstract object’, though I prefer the above so as to avoid straightforwardly paradoxical statements, e.g. ‘Concrete objects might be non-concrete’.

Finally, the fourth thesis concerns existence. On the standard conception, some ordinary objects (like Socrates) *contingently exist* – i.e., are such that there are some possible worlds that simply don't include them at all. Zalta rejects this contingentism in favour of

NECESSITISM All objects necessarily exist

An upshot of **NECESSITISM** is that we can use the simplest quantified modal logic, including the Barcan formula and its converse.⁵ Even better, **NECESSITISM** is consistent with the intuitive idea of there being 'contingent' ordinary objects, provided that we re-define contingency in terms of being *concrete* at some worlds and not at others. So, while it is true that Socrates necessarily *exists*, he remains a contingent object because he is contingently concrete (i.e., there are worlds where he exists and is concrete, as well as worlds where he exists but is non-concrete).

These four theses form the backbone of Zalta's metaphysical pictures and, more relevantly, play vital roles in structuring his reply to the Finean challenge. So, having now set the stage, we can proceed to the response itself.

§3.2 Zalta's Bifurcation Response

The first step in formulating the response is to grasp the essentialist consequences of **AB-CON** and **EXE-ENC**. Specifically, it is clear from the above that, 'in our metaphysical foundations, there are two fundamentally different kinds of objects... constituting mutually exclusive domains.' Further, while concreta

exemplify their properties in the classical way, [abstracta] are the kind of object which can both encode and exemplify properties. Such a basic distinction in kinds of objects merits a distinction in the notion of 'essential property' that applies to each kind. It is therefore natural to suppose that the notion of 'essential property' that is definable for [abstracta] differs from the notion definable for [concreta]. (Zalta 2006: 678)

In other words, the fundamental, metaphysical nature of the abstracta/concreta distinction implies a similar bifurcation concerning the notion of 'essential property': because abstracta and concreta are so radically different in what they are and how they have properties, the notion of *essential property* relevant to each will also be radically different. To that end, the best thing for the modalist to do is to offer *two* definitions of 'essential property', one appropriate for abstracta and another, distinct definition appropriate for concreta.

Regarding abstracta, Zalta suggests that, 'there is nothing more to the nature of an [abstractum] than the properties by which it is to be conceived' (2006: 662).⁶ Consequently, we get a straightforward modalist definition suitable for abstracta:

ABSTRACT Φ is *abstractly-essential* to x iff_{df} necessarily, x encodes Φ ⁷

Further, **EXTENSION** ensures that Fine's counter-examples are blocked for abstracta; as properties like *being such that there are infinitely many prime numbers* aren't part of our conception of, e.g. Sherlock Holmes, it follows that he fails to encode them, and they are therefore non-abstract-essential to him. So **ABSTRACT** seems a sensible modalist definition for abstracta. Of course, it is inappropriate for concreta, which cannot encode properties.

Zalta begins his treatment of concreta by noting that, setting aside Fine's objection, a definition along the lines of **M** is completely inappropriate given his metaphysical

⁵ See (Williamson 2013) for a recent defense of Necessitism in general, independent of Zalta's framework.

⁶ Note: I have altered terminology for consistency's sake.

⁷ See (Zalta 2006: 687)

framework: as Socrates necessarily exists but is only contingently concrete, it seems Zalta is committed to there being worlds where Socrates exists but fails to possess any *concreteness-entailing properties* (where a property Φ is concrete-entailing iff possessing Φ entails that one is concrete). This is pertinent because many of Socrates's essential properties (e.g. *being human, having a particular biological origin, etc.*) are concreteness-entailing. Consequently, **M**, within Zalta's framework, would leave Socrates' essence rather empty.

An alternative to an *existence-dependent* definition like **M** is a *concreteness-dependent* definition. This wouldn't suffer from the same flaw, allowing Socrates to remain essentially human, even though there are worlds where he fails to be so (i.e. because he is non-concrete). This leads Zalta to suggest:

WEAK Φ is *weakly-essential* to x iff_{df} necessarily, if x is concrete, then x has Φ ⁸

Of course, **WEAK** is susceptible to Fine's counter-examples: in every world where Socrates is concrete, he possesses *being such that the Eiffel Tower essentially is a tower* and all the rest of Fine's motley crew. Yet, note that Socrates has these properties not just in all the worlds where he is concrete – in fact, he has them in all the worlds where he exists (they are meant to be problems for **M** after all!). Further, given **NECESSITISM**, the set of worlds where Socrates exists is the set of all worlds. This means that Fine's counter-examples are properties that Socrates has in every world *simpliciter*. To that end, if we want to block them from being essential to Socrates, we only need add a further clause to **WEAK**:

STRONG Φ is *strongly-essential* to x iff_{df} (i) necessarily, if x is concrete, then x has Φ ; and (ii) it is not the case that, necessarily, x has Φ ⁹

Given that Fine's counter-example properties are possessed by Socrates in every world in which he exists and that Socrates necessarily exists, it follows that Fine's counter-examples fail to satisfy **STRONG**'s (ii). It therefore seems that strong-essentiality is a Fine-immune, modal definition of essential property, appropriate for concreta.¹⁰

Conjoining **STRONG** and **ABSTRACT**, we get:

BIFURCATION Φ is *essential* to x iff_{df} Φ is either abstractly- or strongly-essential to x

This is a universally applicable, modalist definition of essence, which, if the above is correct, is immune to Fine's counter-examples. So it seems Zalta's **BIFURCATION** offers modalists a reply to Fine's attack.

§3.3 Rejecting **BIFURCATION** – Costs, Counter-examples, & Epicycles

Unfortunately, for all its merits, **BIFURCATION** suffers from two flaws which, when taken together, show Zalta's response to be fundamentally unworkable.

The first problem concerns the theoretical 'cost' of the response. **BIFURCATION** relies upon Zalta's metaphysical picture, which, as noted above, involves commitment to the abstracta-concreta distinction, the exemplifying-encoding distinction, and to the necessary existence of every object. As such, it's fair to say that the response is theoretically costly,

⁸ See (Zalta 2006: 679)

⁹ *ibid*

¹⁰ It is worth noting that my presentation dramatically simplifies Zalta's argument against the essentiality of *being a member of {Socrates}*, which relies upon regarding mathematical theories and objects as unique kinds of abstract objects (2006: §5). As my objections to Zalta don't turn on these particulars, I feel it is safe to here white-wash over the matter.

requiring we take on a lot of (highly debatable) metaphysical baggage simply to get it going. Relatedly, the very heart of the response is a *bifurcation* – a split between two different definitions of essential property-hood. As theoretical unity is a virtue, this is yet another cost one must pay for embracing Zalta's response. Finally, **BIFURCATION** entails that, if a concretum x essentially has Φ , then necessarily, x doesn't necessarily possess Φ – e.g., if Socrates essentially is human, then necessarily, he isn't necessarily human. This violates what is arguably *the* general inferential connection between claims of essence and claims of necessity, namely, that from ' Φ is essential to x ' we can infer ' Φ is necessary to x '.¹¹ As intuitive as this connection is, rejecting it is yet another cost to be levied against Zalta. Of course, such 'costs' would be worth paying if the response was otherwise completely viable; then the baggage wouldn't be too heavy, the loss of theoretical unity so bad, the intuition connection acceptably forgone. But costs they remain, and costs add up – especially, when there are other troubles lurking.

The second problem is much more substantial: it seems that Fine's counter-examples can be revived as new problems for **STRONG** by conjoining them to any strongly-essential property.¹² For example, take Socrates and the property *being such that the Eiffel Tower essentially is a tower*. This property fails to satisfy the right-hand side of **STRONG** (since it is a property Socrates has in every world), so it is not strongly-essential to Socrates. However, the property *being human and such that the Eiffel Tower essentially is a tower* does satisfy **STRONG**'s right-hand side. Consequently, it is strongly-essential to Socrates. But such a property is clearly non-essential to Socrates – otherwise, discovering Socrates' nature would involve discovering 'the natures of all things' (Fine 1994: 6). Similar moves allow us to revive Fine's other counter-examples.

So **STRONG** is vulnerable to new counter-examples. Any apparent progress in dismissing Fine was illusory – Zalta's **BIFURCATION** has merely shifted the bubble in the carpet, not flattened it out. The only way forward for Zalta seems to be mimicking the initial response strategies: either (i) add further restrictions (to **STRONG**, in this case) thereby blocking the new counter-examples, or (ii) argue that, *contra* Finean intuitions, if it is of the nature of Socrates to be human, then it is equally of his nature to be human *and* such that the Eiffel Tower essentially is a tower.

Taking the first line, the challenge becomes specifying the appropriate restriction. One initial option is to restrict **STRONG** by excluding any property which necessarily implies a property possessed in every world (Zalta 2006: 684); e.g., since *being human and such that the Eiffel Tower essentially is a tower* necessarily implies a property Socrates possesses in every world (namely, *being such that the Eiffel Tower essentially is a tower*), this property is excluded from essentiality. Such a restriction is far too strict, however. An object x 's exemplifying the property Φ necessarily implies that x exemplifies the disjunctive property *being Φ or $\neg\Phi$* , which x possesses in every world where it exists (remember that exemplification is complete). And, as every object exists in every world, this entails that Φ necessarily implies a property possessed by x in every world. So *no* properties would, given this restriction, satisfy **STRONG** and hence no properties would be essential to any concretum.

A second option is to exclude *vacuously strongly-essential* properties, where a property Φ is vacuously strongly-essential to an object x iff Φ is a conjunctive property with at least one strongly-essential conjunct and another necessary conjunct (Zalta 2006: 685, fn17).¹³ Incorporating this into **STRONG** results in

¹¹ See e.g. Fine (2005: 332) for more concerning this connection.

¹² The one exception might be *being a member of {Socrates}*; see note 9 above.

¹³ This is, of course, a bit fast and loose since *properties* aren't the sort of thing that contain operators or connectives; however, following Zalta's lead, I will speak of properties in this manner.

VERYSTRONG Φ is *strongly-essential* to x iff_{df} (i) necessarily, if x is concrete, then x has Φ ; and (ii) it is not the case that, necessarily, x has Φ ; and (iii) Φ is not vacuously strongly-essential to x

While this blocks the counter-examples to **STRONG**, it is vulnerable to further counter-examples: e.g., *being such that if the Eiffel Tower is essentially a tower, then concrete and distinct from the Eiffel Tower* satisfies **VERYSTRONG** yet is, for the same reasons, obviously non-essential to Socrates. To block this new counter-example, Zalta must add even further restrictions, as in the following:

VERY²STRONG Φ is *strongly-essential* to x iff_{df} (i) necessarily, if x is concrete, then x has Φ ; and (ii) it is not the case that, necessarily, x has Φ ; and (iii) Φ is not vacuously strongly-essential to x ; and (iv) Φ is not a conditional property with a vacuously strongly-essential property in the consequent

But this too is vulnerable: *being human or such that if the Eiffel Tower essentially is a tower, then concrete and distinct from the Eiffel Tower* satisfies **VERY²STRONG**, but is, for familiar reasons, non-essential to Socrates. So Zalta must pile on more restrictions, e.g.:

VERY³STRONG Φ is *strongly-essential* to x iff_{df} (i) necessarily, if x is concrete, then x has Φ ; and (ii) it is not the case that, necessarily, x has Φ ; and (iii) Φ is not vacuously strongly-essential to x ; and (iv) Φ is not a conditional property with a vacuously strongly-essential property in the consequent and (v) Φ is not a disjunctive property, one of whose disjuncts satisfies (iv)

Paraphrasing Williamson (1988: 487), this is the sort of definition for which only lack of ingenuity prevents discovery of further intuitive counter-examples. More importantly though, even if it *isn't* susceptible to new counter-examples, **VERY³STRONG** is a terribly complicated definition, which manages to avoid counter-examples only via the accumulation of gerrymandered, *ad hoc* restrictions.

Where does this leave Zalta? Perhaps we can add epicycle upon epicycle until the resulting definition is, *contra* Fine, an extensionally adequate modal definition of essence. But whatever is gained through crafting such a definition, any explanatory potency it might have had is long since lost. In other words, even if **VERY³STRONG** is extensionally adequate, it is hardly an informative, intuitive, and useful account – it starts to sound like a ridiculous Ptolemaic system: a complex, contorted set of exceptions and arbitrary calculations that do a really good job of 'fitting the facts', but are explanatorily unimportant. If this mess is the best modalism can offer, it would be better to just give up on the project.

But perhaps there is another option.¹⁴ The new counter-examples all involve properties generated by what Almog (2003: 221) calls 'logical tricks'; i.e., conjunction, disjunction, or the material conditional are employed to produce, from a property that satisfies **STRONG** (or a variant), a property that (i) satisfies the definition, but (ii) intuitively isn't essential.¹⁵ Noting this, Zalta could insist that such 'trick' properties are somehow

¹⁴ Thanks to an anonymous referee for pushing for more on this point.

¹⁵ More precisely, the process seems to be something like the following: Step 1: start with a true property ascription, eg. 'Socrates has the property of *being human*'. Step 2: replace the singular term with a variable to get an open property ascription, e.g. ' x has the property of *being human*'. Step 3: Via a process of denormalization, go from the open property ascription to an open predication (' x is human'). Step 4: Employing 'logical tricks', show that the open predication is logically equivalent to some other open predication that includes some further, undesirable elements (' x is human and necessarily, if the Eiffel Tower exists, then it is a

illegitimate or inappropriate – *genuine* counter-examples to **STRONG** can't include trick properties. This wouldn't be entirely out of the blue, as one can see the above restriction epicycles as (ineffective) gestures at the broader idea that trick properties are somehow cheating. Obviously, this would get Zalta out of the frying pan.

The problem is that the distinction between inappropriate, trick and essentiality-acceptable, non-trick properties doesn't seem compatible with Zalta's metaphysical picture. In particular, Zalta is committed to the claim that, 'for every expressible set of properties, there is an abstract object which encodes just the members of the set.' (1983: 34) Take the set {*being round, being square, being round and square*}. Some abstractum – Mr. Round-&-Square, say – encodes exactly the properties in this set. Given **ABSTRACT**, it follows that these properties are essential to Mr. Round-&-Square. The same applies to all the above complex counter-example properties: necessarily, such properties will be encoded by, and hence be essential to, *some* objects. As such, these properties can't be irrelevant or inappropriate when it comes to essentiality – quite the opposite in fact.

Further, *why* should we then think trick properties are inappropriate? What exactly makes them so? The only plausible justification is that there is some underlying *metaphysical* significance to the distinction – e.g., the trick properties don't exist or don't genuinely characterize objects. But this seems even more problematic for Zalta. First, without any non-circular motivation for this claim, the 'metaphysical' distinction cited here sounds arbitrary. Second, even if there was an underlying metaphysical element in play, this would amount to yet another theoretical commitment required to get Zalta's account up and running.

Third and finally though, if the metaphysical distinction worked, such that it provided a means for excluding trick properties as inappropriate to essentialism, then it would work in the previous cases as well – i.e., the exclusion could be used to dismiss Fine's *initial* counter-examples and hence defend good old **M**. And, while imposing this distinction then turns out to be a good strategy for the modalist, it isn't a good one *for Zalta*: if modalists are going to use some metaphysical distinction between the trick and non-trick properties to do the requisite heavy lifting, why bother doing so after buying into Zalta's framework, with all of its prohibitive theoretical costs and counter-intuitive consequences? Better to simply stick with **M** and rely on the metaphysical exclusion to block Fine's initial counter-examples (indeed, the next section looks at a response that does exactly this!).

A similar point applies to the second strategy of accepting that properties like *being human and such that the Eiffel Tower is essentially a tower* are just as much a part of Socrates' nature as *being human*. If bullet-biting works for the counter-examples to **STRONG**, then it will work for the counter-examples to **M**. And, as before, a simple analysis of the theoretical costs shows Zalta's picture to be a worse deal than **M**. So Zalta is trapped. If this strategy fails, then the new counter-examples undercut **STRONG**. Meanwhile, if the strategy succeeds, then modalists can use it from the get-go, and Zalta's reply (with all of its costly metaphysical baggage) turns out to be fundamentally unnecessary. Regardless, bullet-biting – like appealing to an underlying metaphysical distinction concerning the trick versus non-trick properties – isn't conducive to the Zaltarian.

Overall then, it isn't clear how Zalta's response is at all appealing. His account is costly, counter-intuitive, and, most worryingly, faces new counter-examples. The only responses to these new counter-examples either push Zalta towards postulating a convoluted, complex, caveat-filled monstrosity of a definition or require appeal to further resources that,

tower'). Step 5: Using lambda abstraction and expansion, generate a new open property ascription from Step 4's open predication (*x has the property of being human and necessarily, if the Eiffel Tower exists, then it is a tower*). Step 6: Re-instantiate Socrates, to get a new, 'trick' property ascription ('Socrates has the property of *being human and necessarily, if the Eiffel Tower exists, then it is a tower*').

if they work, render the rest of Zalta's account redundant. Taken together, these issues make Zalta's response fundamentally unattractive – modalists should look elsewhere for help.

§4. Characteristics & Mere Features: Gorman's Response

The third and final reply, suggested by Gorman [2005], relies upon a metaphysical distinction between properties. The basic idea is that modalists demarcate appropriate from (in some sense) 'inappropriate' properties, then use this split to block Fine's counter-examples. To spoil the ending: while I think this general strategy is extremely promising, Gorman's particular version – discussed in this section – relies upon a distinction that is simply too obscure.

As Gorman points out, modalists can object to the Finean Challenge, arguing that,

...Fine's arguments fail because the counterexamples he appeals to are too artificial. Whatever kind of fact it is that there are infinitely many primes, it is not a fact about Socrates. Although one can say, 'Socrates is such that there are infinitely many primes', doing so provides no information about what he is like, and therefore *a fortiori* it provides no information about what is essential to him. (2005: 279)

Such a retort to Fine does seem justified, and we've already seen something like this in the previous section, in the discussion of rejecting as inappropriate those properties generated by 'logical tricks'.

Now, it seems natural to think that some properties of Socrates really tell us something about him, while others don't – i.e., while *being a man* does provide some genuine information, *being such that there are infinitely many primes* provides no information about what Socrates is, let alone *essentially* is, like. Building on this, Gorman distinguishes between an object's *mere features*, which are all the properties an object has, regardless of how (un)informative they might be, from its *characteristics*, the subset of features that provide genuine information about the object.

The essentialist upshot of the distinction seems straightforward: obviously, only an object's characteristics are appropriate in essentialist contexts – as the above indicates, mere features provide *no* information about an object, let alone about their essences. So, incorporating this into **M**, Gorman offers:

CHARACTER Φ is essential to x iff_{df} (i) necessarily, if x exists, then x has Φ ; and (ii) Φ is a characteristic of x ¹⁶

The dismissal of Fine's counter-examples follows quite naturally: because they tell us nothing about what Socrates is – as Fine notes when he uses them against **M** – Fine's counter-examples must be mere features of Socrates. Consequently, they fail to satisfy **CHARACTER**'s clause (ii). Thus we have our third response: employing the characteristic/mere feature distinction, Gorman blocks Fine's counter-examples and seems to save modalism.

The primary problem with Gorman's response isn't that it doesn't succeed. Rather, it is that the distinction that lies at its heart succeeds all too well. In short: Gorman owes us an account of what makes a property a characteristic rather than a mere feature, which must tell us *why* the distinction has the essentialist upshot that his response to Fine requires; without such an explanation, the whole account looks incredibly *ad hoc*. This is particularly worrisome because one rough definition of 'essential property' is those properties which 'bear, in the metaphysically significant sense of the phrase, on what an object is' (Fine 1994: 2) and this rough definition sounds an awful lot like Gorman's loose characterization of

¹⁶ See (Gorman 2005: 279)

characteristics – it is natural to think that a property bears, in the metaphysically significant sense of the phrase, on what an object is if and only if the property provides genuine ‘information’ about what the object is like. As such, one might be tempted to *equate* characteristics with essential properties. Yet, if ‘characteristic’ is simply a synonym for ‘essential property’, then Gorman’s attempt to demarcate appropriate from inappropriate properties relies upon the very distinction it’s meant to support. Thus it seems that defining characteristics – and thereby distinguishing them from mere features – isn’t an irrelevant dispute; rather, whether the response is viciously circular depends upon elucidating it.

Gorman would likely retort that,

That there is a distinction [between features and characteristics] is important, but precisely where the line is to be drawn is not, so as to avoid irrelevant disputes I will leave the matter vague. Doubtless there will be unclear cases, but that is no argument against the distinction in general, and the cases that Fine brings forward are hardly borderline one. ... whether a given feature is or is not a characteristic of a thing is clearer than what the thing’s essence is: one can, for instance, be unsure about what Socrates is while being sure that his color is one of his characteristics and that his membership in {Socrates} is not. [2005: 279]

Yet this is cold comfort. Why is *being a member of {Socrates}* not a genuine characteristic of Socrates? It does, after all, provide some rather significant, genuine information about Socrates. Similarly for *being distinct from the Eiffel Tower* – what makes this a mere feature? Without a clear, independent grip on what makes a characteristic a characteristic, Gorman’s distinction seems too good to be true. Further, it leaves modalists unable to buy into his response – as it stands, one simply doesn’t know what one is going in for by embracing it.¹⁷ This is enough to warrant a return to the drawing board for the modalist seeking a reply to Fine.

§5. Learning our lessons?

The above trio of responses are all, in their various ways, unsuccessful. We might conclude that this shows Fine was right all along – modalism is dead, and any attempts to revive it are wasted efforts. This certainly would be a useful, albeit bleak, lesson for would-be modalists to take away.

But, I contend that this conclusion is too hasty. For while the trio doesn’t do the job, there is an undercurrent running through all three that indicates a plausible way forward. With this in mind, in this final, positive section, I’ll highlight what lessons I think these three (flawed) responses can teach us, show how these lessons push modalists towards adding to **M** a restriction to what Lewis calls ‘sparse’ properties, and gesture at how I think the resulting *sparse modalism* offers would-be modalists a ray of hope in the face of the Finean Challenge.

Let’s begin with Della Rocca’s **TRIVIAL**. The basic idea underlying this response was a good one: properties that *all* objects necessarily have are too generic, too *trivial* to be

¹⁷ In fact, Gorman goes on to reject **CHARACTER** in favour of an alternative definition formulated in terms of *explanation*, understood as an ontic relation one entity bears to another when the former is the way it is *because* the latter is the way it is [2005: 282-3]. Gorman suggests that an object’s essential properties are those characteristics the object has which explain, but are not explained by, the object’s other properties; i.e.:

EXPLANATION Φ is essential to x iff_{df} (i) Φ is a characteristic of x ; and (ii) Φ is not explained by some other characteristic of x

Setting aside the fact that **EXPLANATION** relies upon the unclear characteristic/mere feature distinction, it faces two problems. First, as Gorman’s *explanation* relation is rather spooky, **EXPLANATION** simply doubles-down, rather than reduces, the mystery. Second, as Gorman points out [2005: 289], accepting **EXPLANATION** entails abandoning modalism. Hence, *for modalists*, **EXPLANATION** is fundamentally unacceptable, equivalent to abandoning ship.

relevant to essentialism. The trouble came once we tried to apply this notion to block all of Fine's counter-examples – it quickly turned out that every necessary property was indirectly trivial, and hence non-essential. Yet, setting the details aside, we can note that distinguishing kinds of properties – the generic, trivial ones, from the specific, non-trivial – definitely seems to line-up with the aims of essentialists.

A related point came up in the discussion of the counter-examples to Zalta's **STRONG**; one potential rebuttal to these new counter-examples is that the properties involved are somehow illegitimate or inappropriate, the products of 'logical tricks'. And while this didn't help the Zaltarian, such a restriction could be employed by a modalist to assist in undercutting Fine's original counter-examples. And such a line was directly advanced in Gorman's **CHARACTER**, which tried to use the characteristic/mere feature distinction to demarcate something like the appropriate/inappropriate property distinction that both Della Rocca and the Zaltarian were gesturing at (though Gorman's particular effort was unacceptable because what exactly a 'characteristic' is remains fundamentally unclear).

In all three cases, what we had were attempts to demarcate something about the properties involved – be it their triviality, their 'tricky' nature, or their uninformativeness – which rendered some ineligible for essentiality. So, if there were some plausible, widely-accepted picture that distinguishes properties into two divergent types, only one of which includes informative, non-tricky, and non-trivial properties, there would be strong intuitive motivation for incorporating this restriction into the modalist definition. I take this to be the first lesson the trio teaches modalists: many properties are, by their nature, going to be irrelevant to essentialist concerns, and a good modalist definition will exclude such 'inappropriate' properties.

A second lesson, most explicit in Zalta but lurking in the others as well, is also worth flagging: modalists can and should use their background metaphysical picture in constructing a response to the Finean Challenge. Zalta freely employed his (rather peculiar) metaphysics in order to generate **BIFURCATION**, and while this response (and the picture it relies upon) isn't very attractive, the thought here again is a good one.

What is suggested by these lessons is that a suitable modalist response (i) like Gorman, distinguishes those properties which are genuinely informative from those which aren't; (ii) like Della Rocca, rules the trivial properties out; (iii) like Zalta, excludes properties generated by 'logical tricks'; and (iv) naturally emerges from our background metaphysical picture. The resulting definition must also be simple, non-gerrymandered (unlike, say, **VERY³STRONG**), and of course, immune to Fine's counter-examples.

Interestingly, a suitable distinction seems readily available, waiting for would-be modalists to employ it: namely, Lewis' distinction between the abundant and the sparse properties. According to Lewis,

abundant properties may be as extrinsic, as gruesomely gerrymandered, as miscellaneously disjunctive, as you please. They pay no heed to the qualitative joints, but carve things up every which way. Sharing of them has nothing to do with similarity. Perfect duplicates share countless properties and fail to share countless others; things as different as can be imagined do exactly the same... The sparse properties are another story. Sharing of them makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are ipso facto not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy. (1986: 59-60)

Lewis' distinction seems to do the job Gorman, Zalta, & Della Rocca's restrictions were gesturing towards; sparse properties genuinely characterise things (completely and without redundancy!), are not generated by 'logical tricks' (for such tricky properties wouldn't carve nature at the joints), and are not going to be trivial. Further, there are strong independent reasons for believing that there is such a distinction to be drawn. As Lewis pointed out, sparse

properties are necessary to fix qualitative similarity, track causal powers, help in analysing laws of nature, causation, intrinsicity, and supervenience, to account for Moorean facts of common sense, provide a minimal notion of physicalism, handle Kripke's rule-following worries, and respond to Putman's objections to metaphysical realism – a laundry list of applications that makes believing in them an offer you can't refuse.

What we get then is Della Rocca, Zalta, and Gorman are all indirectly pushing the modalist towards incorporating into **M** a sparse property restriction, such as in:

SPM A property Φ is essential to x iff_{df} (i) necessarily, if x exists, then x has Φ ; and
 (ii) Φ is a sparse property

This sparse modalism has a lot going for it. First, it links back with the previous responses, incorporating the intuitive motivations without inheriting any of the particular defects. For example, unlike Della Rocca, all the work isn't being done by trying to stretch the notion of triviality, and unlike Gorman, many have a grip on what these sparse properties are meant to be. Second, given that the preponderance of metaphysicians already accept the sparse/abundant distinction (especially since it's so useful), sparse modalism seems to be rather low 'cost' (unlike Zalta's bifurcation account). Third, and most importantly, sparse modalism is immune to Fine's counter-examples.¹⁸

Thus the final, take-home lesson is this: the three responses discussed were all on the right track – they get the heart of the solution right, in that we need to find some way to exclude certain properties as inappropriate to essentialist investigations. What they get wrong was the details. But we shouldn't take the difficulties facing the particular accounts as signs that modalism is dead. Rather, building off the lessons Della Rocca, Zalta, and Gorman teach us, we should explore the potential a new, sparse modalist response – *that* may be how to be a modalist about essence.¹⁹

References

- Almog, J. 2003. 'The Structure-in-things: existence, essence and logic', *Proceedings of the Aristotelian Society*, 103: 197–225
- Della Rocca, M. 1996. 'Essentialism,' *Philosophical Books* 37: 1-20; 81-9
- Fine, K. 1994. 'Essence and modality', *Philosophical Perspectives* 8: 1-16
- Fine, K. 2005. 'Necessity and non-existence', in his *Modality and Tense*, Oxford University Press
- Gorman, M. 2005. 'The essential and the accidental,' *Ratio* 18: 276-89
- Lewis, D. 1986. *On the Plurality of Worlds*. Blackwell
- Wildman, N. *forthcoming*. 'Modality, sparsity, and essence,' in *The Philosophical Quarterly*
- Williamson, T. 1988. 'Review of R. Hale, *Abstract Objects*', *Mind* 97: 487-490
- Williamson, T. 2013. *Modal Logic as Metaphysics*. Oxford University Press

¹⁸ See my *forthcoming* for an extensive discussion of how sparse modalism handles the Finean challenge, in particular *being a member of {Socrates}*.

¹⁹ Acknowledgements go here.

Zalta, E. 1983. *Abstract Objects: an introduction to Axiomatic Metaphysics*. D. Reidel

Zalta, E. 2006. 'Essence and modality', *Mind* 115: 659-693