Intrinsic Properties and Relations

Forthcoming in *Inquiry*

Jan Plate
Università della Svizzera italiana

March 2018

Contents

Abstract

This paper provides an analysis of the intrinsic/extrinsic distinction, as applied both to properties and to relations. In contrast to other accounts, the approach taken here locates the source of a property's intrinsicality or extrinsicality in the manner in which that property is 'logically constituted', and thus—plausibly—in its nature or essence, rather than in, e.g., its modal profile. Another respect in which the present proposal differs from many extant analyses lies in the fact that it does not seek to analyse the 'global' distinction between intrinsic and extrinsic properties on the basis of the 'local' distinction between having a property intrinsically and having it extrinsically. Instead, the latter distinction is explicated on the basis of the former.

Keywords: Properties; relations; intrinsicality

1 Introduction

The distinction between intrinsic and extrinsic properties has proven difficult to analyse, but it is undeniably useful in analysing other notions. For instance, in explicating the concept of an intrinsically valuable object, it is quite natural to appeal to that of an intrinsic property, and to say (to a first approximation) that an object is intrinsically valuable iff it is valuable 'in virtue of' its intrinsic properties.[1] More generally, the same could be said about the notion of an intrinsically F thing, where 'F' might be replaced with any of a wide variety of adjectives, such as 'beautiful', 'virtuous', 'funny', etc.: for it will be an at least initially promising hypothesis that a thing is intrinsically F iff it is F in virtue of—or at least *largely* in virtue of—its intrinsic properties.[2] Even if it were denied that the intrinsic/extrinsic distinction can be applied in this way, one could still list numerous other applications.[3]

[1]Though see Kagan (1998) for a contrasting view. For an overview of the issues involved in the distinction between intrinsic and extrinsic value, see e.g. Zimmerman (2015) and Rønnow-Rasmussen (2015). Throughout this paper, I shall use 'extrinsic' as synonymous with 'non-intrinsic'.

[2]For a more detailed analysis, see Section 12. Several authors, especially in recent years, have taken the 'local' notion of an entity's being intrinsically such-and-such to be more basic than the 'global' notion of an intrinsic property. (Cf., e.g., Francescotti [1999; 2014b], Parsons [2001: 10], Witmer, Butchard & Trogdon [2005: 333], Figdor [2008; 2014], Williams [2013: 435], Bader [2013: 554], and Marshall [2009; 2016a].) If the analysis provided in Section 12 is successful, it would suggest that the opposite view is at least equally defensible.

[3]For illustration, here is the third paragraph of Robert Francescotti's introduction to his recent anthology:

Given the philosophical importance of the distinction, the question of how, if at all, one might formulate an informative analysis of the distinction is of corresponding interest. This interest is impressively reflected in the number of proposals that have been published in the decades since David Lewis's (1983a) seminal paper on this topic.[4] The main purpose of the present paper is to add to this number of proposals, though the target of our analysis will be the more general concept of an intrinsic property *or relation*.[5] As a shorthand for 'property or relation', I shall use the term 'attribute'.

Why another analysis of intrinsicality? One motivating factor lies in the fact that there exists so far no generally accepted account. (A majority of the analyses that have been proposed in the literature have eventually come under criticism, most notably in a series of articles by Dan Marshall.[6]) But more pertinent to the present paper is the fact that no general account of the intrinsic/extrinsic distinction has so far tried to locate the 'source' of an attribute's intrinsicality or extrinsicality in the attribute's *logical constitution*, i.e. (roughly) in

> In the philosophy of physics, there is the question of whether the causal powers of objects supervene on their intrinsic features, whether the relational properties of physics require an intrinsic ground, and whether all or even any fundamental physical properties are intrinsic. And assuming that fundamental physical items have an intrinsic nature, how can we ever have knowledge of that intrinsic nature given that we can only be aware of the effects? In metaphysics, there is the question of what constitutes genuine change (which seems to require a change in intrinsic properties), and whether change in the intrinsic properties of an individual is compatible with its *enduring* through time (being numerically identical at different times). Philosophers of mind wonder whether the content of our mental states supervenes on our intrinsic features or whether mental content is partly a function of the external items toward which our thoughts are directed. There is also the issue of whether consciousness extends beyond the intrinsic features of one's brain or even the rest of one's body. For the philosopher of art, there is the issue of whether aesthetic value is an intrinsic feature of an object, and if not, what relations to which external items ground aesthetic properties; and in epistemology, there is the long-standing question of whether the justification of one's beliefs is solely a function of one's intrinsic features. (Francescotti 2014a: 1; emphasis in the original)

A further potential application from metaphysics is the task of explicating the concept of a *duplicate*, as two entities may be considered duplicates iff they share the same purely qualitative intrinsic properties. In addition, Cian Dorr (2016a: 242) has recently proposed a definition of 'determinism' that makes use of the notion of an intrinsic property. Still other applications can be found in Weatherson & Marshall (2017: §1.1).

[4]For an overview, see Weatherson & Marshall (2017: §3).

[5]The idea that the concept of intrinsicality is applicable not only to properties but also to relations can be made plausible both by way of example and by generalising typical characterisations of what it is for a property to be intrinsic. See, e.g., Lewis (1983b: 356n.), Langton & Lewis (1998: §8), Parsons (2001: §2.1.9), Weatherson & Marshall (2017: §1.3), and Hoffmann-Kolss (2010: ch. 8). The characterisation relevant for the present paper is given in Section 4 below.

[6]See Marshall (2012; 2013; 2014; 2015). Also see Weatherson & Marshall (2017) for (among other things) a critique of a proposal by Francescotti (1999).

the way the attribute is 'constructed' from other entities by way of logical operations.[7] This *logical-constitution approach* will be briefly outlined in Section 3. As for the rest of the paper, Section 4 takes some first steps towards constructing an account of intrinsicality that implements this approach, and Sections 5 and 6 provide the formal framework that the account will rely on. (In Section 6 we will also engage in some methodological considerations.) The account itself is then developed in Sections 7–10, and a summary of the account is given in Section 11. Finally, Section 12 offers a brief analysis of the 'local distinction' between having a property intrinsically and having it extrinsically, and Section 13 concludes the paper. To begin with, however, we have to take a look at two analyses of intrinsicality that have recently been put forward by Dan Marshall.

2 Two Marshallian Proposals

2.1 Aboutness intrinsicality

In his 'Varieties of Intrinsicality' (2016b), Marshall has developed a taxonomy of several forms of intrinsicality, one of which he refers to as 'absolute aboutness intrinsicality'.[8] He finds this latter concept suggested in the following passage by Lewis:

> A sentence or statement or proposition that ascribes intrinsic properties to something is entirely about that thing; whereas an ascription of extrinsic properties to something is not entirely about that thing, though it may well be about some larger whole which includes that thing as part. (1983a: 197)

In Marshall's reading, this passage suggests that a property P is intrinsic iff, by metaphysical necessity, for any entity x, the "ascription" of P to x is "intrinsically about x". In his terminology, an 'ascription' of a property P to an entity x is the state of affairs that might be paraphrased as 'x's being F' (where 'F' is replaced by an adjective signifying P). So, e.g., the ascription of paleness to Socrates is Socrates' being pale, or the state of affairs that

[7] A near-exception is Brad Skow's (2007: 115) analysis of the intrinsic/extrinsic distinction for shape properties, according to which "a shape property Px is intrinsic just in case every quantifier in its analysis is restricted to x and x's parts". (The "analysis" of the respective property P—or Px in Skow's notation—is here supposed to be in terms of "fundamental spatial relations".)

[8] The 'absolute' is somewhat optional: Marshall adds it only for the purpose of distinguishing the notion in question from what he calls 'possession aboutness intrinsicality'.

Socrates is pale. Marshall takes a state of affairs s to be *intrinsically about* an entity x iff (roughly):

> s (either truly or falsely) describes how x and its parts are and how they are related to each other, as opposed to how x and its parts are related to other things and how other things are. (2016b: 240)

What Marshall refers to as an ascription of a property P to an entity x, I shall refer to as P's *instantiation* by x, or alternatively, as x's instantiation *of* P. By way of a more concise expression, if 'P' names a property P and 'x' names some entity x, I will further write '$P(x)$' to refer to x's instantiation of P.[9] Marshall's definition of aboutness intrinsicality can then be rephrased as follows:

(AI) A property P is *intrinsic* iff it holds by metaphysical necessity that, for any entity x, the state of affairs $P(x)$ is intrinsically about x.

This analysis yields in many cases intuitively correct results. For instance, on the (intuitively plausible) assumption that, for at least one entity x, the state of affairs that x is a father fails to be 'intrinsically about' x, the property of *being a father* is duly classified as extrinsic. However, the analysis fares less well when we consider non-obtaining instantiations of the parthood relation. For example, while the non-obtaining state of affairs that Italy has Paris as a part may plausibly be identified with Italy's instantiation of the property of *having Paris as a part*, that state of affairs is not intrinsically about Italy, since it concerns how Italy is related to something—viz., Paris—that is not a part of it. (This reasoning is borne out by Marshall's own account of intrinsic aboutness, as given in his [2016a]; cf. footnote 15 below.) Consequently, *having Paris as a part* is under (AI) classified as extrinsic, whereas intuitively that property seems to be *in*trinsic.[10] (AI) is thus intuitively inadequate.

[9] Note that, on the present conception of states of affairs, a state of affairs can exist without obtaining. Even a false statement may thus express a state of affairs.

[10] At least as long as the relevant parthood relation is itself intrinsic, which I take to be plausible. A similar example, viz., the property of *being identical with Obama*, has been used by Marshall (2015: 11n.) to argue against an account of intrinsicality proposed by Rosen (2010: 112). (Also cf. Marshall [2016b: 246f.].)

To see why it is necessary to assume here that the relevant parthood relation is intrinsic, suppose, e.g., that parthood is the relation whose instantiation by x and y (in this order) is the state of affairs that God cannot think of y without thinking of x. I submit that in this case it would no longer be intuitively plausible to think that *having Paris as a part* is intrinsic.

2.2 Possession intrinsicality

Fortunately, the sort of counter-example just described can be avoided if one explicates the concept of intrinsicality along the lines of another of Marshall's notions, namely—most plausibly—that of 'possession aboutness intrinsicality'. This is the concept that Marshall, in another recent paper, effectively identifies with the concept of intrinsicality *tout court*, suggesting that it is "arguably the notion philosophers typically use 'intrinsic' to express" (2016a: 704n.). Using the formalism adopted above, we can state the proposal as follows:

(PI) A property P is *intrinsic* iff it holds by metaphysical necessity that, for any entity x: if x instantiates P, then the state of affairs $P(x)$ is intrinsically about x.

The *analysans* of this proposal is evidently much weaker than that of (AI), and in this way it manages to avoid the counter-example that renders the latter so implausible. There are, however, several worries.

First, there is the worry that (PI) trivially classifies any *uninstantiable* property as intrinsic: that is, any property P that is such that, necessarily, no entity instantiates P. On a suitably abundant ontology, coupled with a moderately fine-grained conception of attributes, there will plausibly exist properties that are both extrinsic and uninstantiable; an example would be that of *being such that Socrates is non-self-identical*. For any such property, however, the right-hand side of (PI) will be vacuously true (since the antecedent of the embedded conditional will always be false), and the property will consequently be classified as intrinsic. By contrast, the same problem does not—or at least not obviously—arise for (AI): e.g., Plato's instantiation of the mentioned property does *not* appear to be intrinsically about Plato.

An adherent of a starkly coarse-grained conception of properties might try to defend (PI) by arguing that the property in question is really nothing else than that of *being non-self-identical* (on the grounds, say, that the 'two' properties are necessarily coextensive). Since, furthermore, *being non-self-identical* is plausibly intrinsic, we have here a situation where common sense returns two contradictory verdicts on one and the same property, and so it might be said that the classification of that property can be left as 'spoils to the victor'.[11] However, at least in the present case this sort of response cuts no ice, since it relies on a highly controversial assumption concerning the coarse-grainedness of properties.[12] And it would also be something of a shortcoming if, in order for (PI) to work, one would have to assume the non-existence of uninstantiable properties.

A second and related worry has to do with the fact that (PI), like all other accounts discussed by Marshall (2016b), makes the intrinsicality of a property a *modal* issue. To be sure, there is ample precedent for this.[13] But as can perhaps be seen from the previous paragraph's examples, we appear to have a tolerably good grasp of what makes a property intrinsic or extrinsic even independently of whether we regard that property as instantiable. This observation in turn supports the claim (which strikes me as intuitively plausible) that the question of whether a given property is intrinsic or not is determined by the property's 'essence', rather than by its modal profile. We may refer to this as the *essentiality intuition*. To borrow Yablo's (1999) suggestive phrase, we want an account of intrinsicality that is primarily sensitive to the *de jure* features of the respective properties, rather than to their *de facto* features, even if these happen to be modal in character. Further, owing mainly to the influence of Kit Fine (1994; 1995a; 1995b), it is now commonly assumed that an entity's essence does not simply boil down to modal issues. It would accordingly seem natural to expect of an account of intrinsicality that the essence of the respective attribute should play in it a prominent role.[14] (PI) disappoints this expectation.

[11]Cf. Lewis (1986b). For some critical discussion of this manoeuvre, see Eddon (2011). For more sympathetic (albeit brief) discussion, see Section 6.1 below. The classification of *being non-self-identical* as intrinsic is not entirely uncontroversial. For example, Josh Parsons (2001: 15) and Ralf Bader (2013: 555) both classify uninstantiable properties as "neither intrinsic nor extrinsic".

[12]It is worth noting here that Marshall himself (2016a: 713) finds it desirable to have an account of intrinsicality that is "independent of which theory of the identity conditions of properties [is] correct". Although this statement would be grist for my mill in the present context, I would not unreservedly endorse it. If a *prima facie* sensible account of some category of attributes produces counter-intuitive results when combined with an unintuitive conception of attributes, the thing to blame might not be the account itself but rather that conception of attributes. I will expand on this issue in Section 6.1 below.

[13]Cf., e.g., Moore (1922: 268f.) and Lewis (1983a). An early dissenting voice is Dunn (1990: 184).

[14]In this way intrinsicality differs markedly from *essentiality*. Cf. Figdor (2008: 694): "We think the definition of a property determines whether the property is intrinsic or extrinsic, whereas the definition plays no such role in whether a property is essential, accidental, innate, or acquired". (By an 'essential property', Figdor here means a property that is "had essentially by some individual".)

As far as extensional adequacy is concerned, our discussion so far has shown that (AI) and (PI) have somewhat complementary shortcomings: intrinsic properties such as *having Paris as a part* are classified as intrinsic by (PI) but not by (AI), and extrinsic uninstantiable properties such as *being such that Socrates is non-self-identical* are classified as extrinsic by (AI) but not by (PI). It would be desirable to have an account of intrinsicality that correctly classifies both of these sorts of properties.

Before we move on, it will be worth flagging five further concerns that one might have about the extensional adequacy of Marshall's account of intrinsicality. In contrast to the preceding worries, these additional concerns arise not from (PI) alone but rather only from the *conjunction* of (PI) and Marshall's account of intrinsic aboutness. Suppose we call a property a *haecceitistic inclusion property* iff, for some entity x, it is the property of *having x as a part*. Very briefly, these concerns may then be stated as follows (to save space, I have delegated some of the justifications to footnotes):

1. Negations of haecceitistic inclusion properties, such as the property of *not having Paris as a part*, may be considered intuitively intrinsic; but under Marshall's account they are classified as *ex*trinsic.[15]

[15] According to Marshall, a state of affairs s is *intrinsically about* an entity x iff s is "a non-qualitative state of affairs that is expressed by a sentence that is pre-intrinsically about x". Here a state of affairs is 'non-qualitative' iff it "haecceitistically concerns" some entity (p. 707), in the sense in which, e.g., "the state of affairs expressed by 'Obama is an uncle' haecceistically concerns Obama" (p. 706). Further, an expression e is *pre-intrinsically about* an entity x iff it is such that:

for any Xs, IF

(i) for any expression f, if f contains at most brackets, commas, variables, names referring to parts of x, predicates expressing perfectly natural relations, and operator expressions expressing perfectly natural operators, then f is one of the Xs;

(ii) for any expression f, if f is of the form $\ulcorner[Qv|\phi]\psi\urcorner$, where Q is either a perfectly natural two-place quantifier or a two-place correlate of a one-place perfectly natural quantifier, ϕ and ψ are each one of the Xs, v is a variable, and for some set Π of names of all the parts of x, under any assignment g to v and the free variables of ϕ, ϕ necessitates $\ulcorner\bigvee\{\ulcorner v \leq b\urcorner | b \in \Pi\} \vee \bigvee\{\ulcorner v \leq u\urcorner | u \text{ is a free variable in } \phi \text{ other than } v\}\urcorner$, then f is one of the Xs; and

(iii) for any expression f, if f is a concatenation of expressions that are among the Xs, then f is one of the Xs;

THEN e is one of the Xs. (p. 727)

This is the first full version of Marshall's account of intrinsic aboutness; he later on (p. 729f.) provides another version that is purged of all references to 'linguistic' entities but is otherwise meant to be equivalent.

To see that *not having Paris as a part* is under Marshall's account classified as extrinsic, it is enough to notice that, under Marshall's definition, a sentence cannot be pre-intrinsically about an entity x if it contains any name of something that is not a part of x. (Predicates, of course, are not names in Marshall's framework.) For let P be the property of *not having Paris as a part*, and let x be Rome. Then x instantiates P, and hence, if P is

2. Disjunctions of haecceitistic inclusion properties, such as the property of *having either Paris or Rome as a part*, are (fairly clearly) intuitively intrinsic, but are classified as extrinsic under Marshall's account.[16]

3. The same goes for properties like *having a part that is distinct from x* (for some entity x). These are classified as extrinsic under Marshall's account, but intuitively they seem to be intrinsic (although this may be controversial).[17]

4. His account also classifies as extrinsic any properties that—like *being composed of cells* or *having as many red parts as green parts*—would have to be expressed, at a suitable level of analysis, by quantifying over (possibly deeply nested) sets or pluralities that are ultimately composed of parts of the respective property's bearer.[18]

5. Finally, suppose that P is a perfectly natural property and that there exists a 'perfectly natural operator' O that, when applied to some entity x, yields another entity (such as x's singleton) that is not a part of x.[19] Let x be some entity, let n be a name of x, let $\mathcal{O}$ be an 'operator expression' expressing O (such that $\ulcorner \mathcal{O}n \urcorner$ is a composite singular term denoting the result of applying O to x), and let F be some predicate that signifies P. Then the sentence $\ulcorner F(\mathcal{O}n) \urcorner$ will express a state of affairs that, by Marshall's lights, is intrinsically about x, and the property that can be abstracted

---

to be classified as intrinsic, the state of affairs $P(x)$ will have to be intrinsically about x, i.e. Rome. So there has to exist some sentence that expresses $P(x)$ and is pre-intrinsically about Rome. But now, unless we adopt an absurdly coarse-grained conception of states of affairs, any sentence that expresses $P(x)$ at the rather deep level of analysis required by Marshall's definition of pre-intrinsicality will have to contain a name of Paris; and Paris is not a part of Rome. Hence there seems to be no sentence of the required sort, which would mean that P is not classified as intrinsic under Marshall's account.

[16]The justification for this is similar to the argument of the previous footnote. For instance, Italy's instantiation of the mentioned property will (arguably) not be intrinsically about Italy.

[17]For example, consider the property of *having a part that is distinct from Dublin*. Arguably, Italy's instantiation of this property is not intrinsically about Italy.

[18]For instance, if we opt for a set-theoretic analysis, the property of *being composed of cells* will have to be expressed by something along the lines of, '$\lambda x\, \exists S$ (each member of S is a cell as well as a part of $x \wedge$ each part of x has a part in common with some member of S)'. If P is this property and x any entity that has P, then x's instantiation of P will under Marshall's analysis not be intrinsically about x. The basic reason for this lies in the fact that the variable 'S' in the λ-expression just sketched is not restricted to the parts of the respective bearer of P.

[19]For the sake of argument, I here follow Marshall in assuming that the predicate 'perfectly natural' is not semantically defective and can be applied to both properties and 'operators' (where the latter are the semantic values of what Marshall calls 'operator expressions').

from this (and which might be denoted by $\ulcorner \lambda x\, F(\mathcal{O}x) \urcorner$) will accordingly be classified as intrinsic. But intuitively, it is far from obvious that this classification is correct, since the result of applying O to x is by hypothesis not a part of x.

Not all of these points are uncontroversial. I expect that there will be some disagreement over the first and third items on this list, since properties like *not having Paris as a part* or *having a part that is distinct from Dublin* may conceivably be regarded as extrinsic.[20] Nor are the other three concerns supposed to be fatal. Still, to address them all would require a considerable amount of work, and would likely lead to a significantly more complicated account of intrinsic aboutness than the one offered by Marshall. Apart from their relevance for his analysis of intrinsicality, the first four worries are worth raising here also because the underlying intuitions about the in- or extrinsicality of the mentioned properties (and others like them) will inform the construction of our account in Sections 7–10. But unlike the concerns about uninstantiable attributes and about the modal character of (PI), they will not help to motivate our general approach as set out in the next two Sections.

3 The Logical-Constitution Approach

The defining characteristic of the logical-constitution approach lies in the fact that it locates the 'source' of an attribute's intrinsicality or extrinsicality in the attribute's logical constitution. By an attribute's 'logical constitution', I roughly mean the way the attribute is 'built up' from other entities, provided that the relevant operations are at least broadly speaking logical. For instance, if a given property P is the conjunction of two other properties Q_1 and Q_2, then its being so will be part of P's logical constitution, and hence (plausibly) of its essence. And similarly, if P is, for some relation R and entity x, the property of *bearing R to x* (e.g. if P is the relation of *being a friend of Aristotle*), then it will have R and x among its

[20]Fortunately, it would not be very difficult to accommodate conflicting intuitions (as far as these two points are concerned) by suitable adjustments to the account developed below. Thus, to accommodate a possible intuition to the effect that properties like *not having Paris as a part* are extrinsic, one could easily assimilate our treatment of this sort of property to that of properties like *being distinct from Paris*. (Cf. Section 10.1. This move would however have the probably undesirable effect that properties like *having Paris or Rome as a part* would also be classified as extrinsic.) And to do justice to a possible intuition to the effect that properties like *having a part that is distinct from Dublin* are extrinsic, one would only have to remove a certain complication that is (with the aim of accommodating the contrary intuition) introduced in Section 10.3.

logical constituents. A convenient way of expressing an attribute's logical constitution (at a lesser or greater level of granularity) is by means of λ-expressions. For example, the conjunction of two properties Q_1 and Q_2 could be represented by '$\lambda x\,(Q_1(x) \wedge Q_2(x))$', and the property of *being a friend of Aristotle* could be represented by 'λx friend-of$(x,$ Aristotle$)$'.[21]

The general idea of an attribute's being 'constructed' from other entities will already be familiar from algebraic treatments of attributes, of the sort found in Bealer (1982; 1994), Zalta (1983), Menzel (1993), and Swoyer (1998). A caveat, however: the present talk of logical constitution should *not* be understood as implying that for any given attribute there is only one way in which it is constituted from other entities. For example, if P is the conjunction of two other properties Q and R, where Q is in turn the conjunction of two properties Q_1 and Q_2, then P may at the same time be the conjunction of Q_1 and a further property S, namely if S is the conjunction of Q_2 and R. Whether there are such cases depends on the coarse-grainedness of the underlying conception of attributes, and our talk of logical constitution should not be understood as committing us to any particular stance on that issue.

The logical constitution of an attribute can plausibly be regarded as part of the attribute's essence. Thus it is essential to a conjunction of two properties P and Q that it should be the conjunction of these two properties; and likewise it is essential to the property of

[21]Readers of Marshall (2009) might suspect that, without the appeal to some extra-logical notions, the pursuit of a logical-constitution approach will be a fool's errand: for doesn't his argument (which goes back to Parsons [2001: 22f.] and is reminiscent of an earlier argument by Sider [1996: 23]) show that intrinsicality cannot be "defined using only broadly logical notions"? The argument runs roughly as follows. Let E be the property of *being an electron*, and let E^* be that of *being either a lonely positron or an accompanied electron*, where something is 'lonely' iff it is the only thing there is, and 'accompanied' iff it is not lonely. Plausibly, E is intrinsic while E^* is extrinsic. But E and E^* satisfy the same "rigid broadly logical" formulas, where (i) a formula A with one free variable is *rigid* iff we have $\forall x\,(\Box A(x) \vee \Box\neg A(x))$ and (ii) a formula is *broadly logical* iff it is a formula of a certain restricted language L. (I have reproduced Marshall's description of this language at the beginning of Appendix A.) Moreover, if 'x is intrinsic' were to abbreviate a broadly logical formula, the latter would have to be rigid. Hence 'x is intrinsic' cannot abbreviate a broadly logical formula, or else E and E^* would either both count as intrinsic or both count as extrinsic. But that 'x is intrinsic' abbreviates a broadly logical formula is precisely what it would mean for 'intrinsic' to be definable using only broadly logical notions; and so 'intrinsic' cannot be defined in this way.

A partial response to this argument on behalf of the present paper's project would be to point out that the language in which we will be formulating our account of intrinsicality differs in some crucial respects from Marshall's language L, but it takes some effort to work out what those crucial respects are. (See Appendix A. Many thanks to an anonymous referee for pressing me on this issue.)

It is interesting to note that Marshall himself goes some way in the direction of a logical-constitution approach, not only in his own account of intrinsic aboutness but also already in his (2014) discussion of Yablo's (1999) account of intrinsicality. However, he there deals with the logical constitution not of attributes but rather of states of affairs, and he goes beyond the broadly logical in employing a concept of perfect naturalness.

being a friend of Aristotle that it should be in a suitable way composed from Aristotle and the friendship relation. In this way, an attribute's logical constitution can be said to reflect its essence, and this in turn makes it plausible to say that the logical-constitution approach preserves what I have in the previous Section called the 'essentiality intuition'.[22]

4 Intrinsicality and Logical Complexity

Before we try to implement the logical-constitution approach and construct a formal account of intrinsicality, it will be useful to have an informal orienting characterisation to serve as a guideline. According to a typical such characterisation, used by Marshall (2012: 531), a property is intrinsic iff it is necessarily the case that any entity x that has that property has it "in virtue of how it [i.e. x] is, as opposed to how it is related to things wholly distinct from it or how things wholly distinct from it are". This characterisation largely conforms to the way in which I shall understand the term 'intrinsic' in the present paper, though with a few qualifications. First, Marshall's characterisation is concerned only with the intrinsicality of properties, rather than with that of properties *and relations*, though this is only a minor point.[23] A more important point has to do with the fact that, by employing the notion of metaphysical necessity, Marshall's characterisation stacks the deck in favour of a modal analysis of intrinsicality, as opposed to one that is framed in terms of logical constitution. Moreover, in focusing on "how" a given entity x "is", his characterisation appears to set aside the question of *which* entities x has as a part, and thereby risks giving the impression that a property like that of *having Paris as a part* is not intrinsic. For these reasons, the orienting characterisation that I shall take as a guideline in this paper deviates slightly from Marshall's. It runs as follows:

[22]It might be objected that an attribute's logical constitution need not always reflect its *whole* essence. For example, given that it is essential to Socrates' singleton to have Socrates as a member, it will plausibly also be essential to the property of *being identical with Socrates' singleton* to have a constituent that has Socrates as a member. So the essence of an attribute can sometimes be said to 'outstrip' its logical constitution. However, I am not aware of any case where this would uncontroversially affect the question of whether the respective attribute is intrinsic. (Thanks here to Alex Skiles.)

[23]For a formulation that also applies to relations, see Weatherson & Marshall (2017: §1.3).

(In_0) A κ-adic attribute A is *intrinsic* iff to instantiate A is, for any entities $\underbrace{x_1, x_2, \ldots}_{\kappa\text{-many}}$, purely a matter of which parts the x_i have, what the x_i and their parts are like, and how the x_i and their parts are related among each other, as opposed to what other entities there are, or how any of the x_i and their parts are related to any other entities.

To forestall misunderstanding, it should be noted that the use of the 'any' in "for any entities $x_1, x_2, \ldots$" is here best understood as a concession to English grammar. Ideally, the variables 'x_1', 'x_2', etc.—or more precisely their occurrences—would in (In_0) be bound not by some quantifier but rather by the 'is purely a matter of' construction. More specifically, the right-hand side of (In_0) should be read as *trying to articulate something about the logical constitution* of the respective attribute A.[24]

The interpretational problem that arises from the 'any' in (In_0) appears to be somewhat less acute if we limit the scope of the characterisation to properties and use a combination of 'something' and 'it':

(In_0') A property P is *intrinsic* iff for something to instantiate P is purely a matter of which parts it has, what it and its parts are like, and how it and its parts are related among each other, as opposed to what other entities there are or how it or any of its parts are related to any other entities.

For here the 'something' seems to allow for a reading that is not simply quantificational.

Given that we have now set up an informal characterisation, it might be thought that in order to construct an appropriately precise account, we only have to formalise the various components of (In_0). This might well be an instructive exercise, but there is also reason to believe that it would lead us into some unnecessary detours.[25] For a more direct approach,

[24]Compare: 'To instantiate the sister-of relation is, for any x and y, to be such that x is a female sibling of y.' On a strict reading, the 'any' here has to be interpreted as a quantifier, but it would not be uncharitable to understand the statement instead as trying to tell us *what it is* to instantiate the sister-of relation. Read in this way, it could be regimented (as it would be in Rayo's [2013] formalism) with the help of an operator '$\equiv$' that takes the variables 'x' and 'y' as subscripts and thereby binds their occurrences, as in, 'sister-of$(x, y) \equiv_{x,y}$ female$(x) \wedge$ sibling-of(x, y)'. The 'is purely a matter of' construction might be regimented in a similar fashion, as far as variable-binding is concerned. (Thanks to Fabrice Correia and an anonymous referee for pressing me on this issue.)

[25]For example, the point of the 'as opposed to' clauses in (In_0) and (In_0') is to give an indication as to what kinds of relation are supposed to be referred by the 'how ... are related among each other'. Namely, they are

I propose to start instead with the question of whether intrinsicality or extrinsicality (or neither) should be considered a 'complex-making' feature of attributes, where the relevant sense of 'complex' is one of *logical* rather than, say, *mereo*logical complexity.[26] In other words: Must an attribute be (logically) complex in order to be intrinsic? Or must it be complex in order to be *ex*trinsic? At least *prima facie*, the correct answer to the first question is 'no', for it appears plausible to say that all logically simple properties are intrinsic; and one would thus expect that the same will be true for logically simple relations.[27] But if this is correct, then the answer to the *second* question will clearly be 'yes'. That is, in order to be extrinsic, an attribute has to be complex.

If an attribute has to be logically complex in order to be extrinsic, it stands to reason that in order to be extrinsic it has to be logically complex *in a certain way*, since not all complex attributes are extrinsic. (*Having Paris as a part*, for instance, is complex but *in*trinsic.) But how are we to specify a "way" in which a given attribute is logically complex? Here λ-expressions promise to be useful, since they allow us to express how a given attribute is 'logically constituted' from various other entities. What is less clear is whether, in order for a given attribute to be extrinsic, there has to exist *some* λ-expression that (as we shall say) denotes the attribute while expressing the requisite sort of complexity; or whether, instead, it has to be the case that *all* λ-expressions by which the attribute is denoted express that sort of complexity.[28] A moment's reflection suggests that the more plausible option is

---

supposed to be *intrinsic* relations. Accordingly, one of the challenges in constructing a formalisation of (In_0) (if one were to go this route) would be to formalise this point without running into circularity. A natural way of doing so would be to give the formalisation a recursive character, which would immediately lead to the question of what sort of attribute should be regarded as constituting the 'base case' of an intrinsic attribute, and why. Here a promising answer would be to say that the base case is constituted by the *logically simple* attributes (see below). However, if the intrinsicality of the logically simple attributes were to be hard-wired into our account via an explicit declaration, it would fail to provide an answer as to *what it is* about the logically simple attributes that makes them intrinsic.

[26] On a very straightforward approach, an attribute may be said to be logically simple iff it "is not itself a negation, conjunction, disjunction, quantification, modalization, etc. of any other properties or relations" (Menzel 2016). Unfortunately, under any but the most fine-grained conceptions of attributes, this analysis has the consequence that no attribute at all is logically simple. In my (2016: §3.3) I have proposed an alternative account that does not suffer from this defect. (Cf. footnote 96 below.)

[27] An argument for the thesis that all simple properties are intrinsic can be found in Marshall (2012: 535).

[28] To say that a λ-expression 'denotes' a given attribute will sound odd to someone who thinks of a λ-expression as a kind of abbreviation of a natural-language predicate (since such predicates are not usually taken to denote anything). In the present paper, however, λ-expressions are expressly used as *names* of attributes.

the first. After all, if a given attribute is exhibited as extrinsic by even only a single λ-expression, why should that not be enough? A potential problem arises (as we will see in Section 8) in connection with making sure that the λ-expression in question does not merely *seem* to exhibit the attribute as having the requisite sort of complexity. But for now, there is little reason not to stick with the first option. We will therefore say that, to a first approximation, an attribute is extrinsic iff it is denoted by some λ-expression that satisfies certain conditions. However, there is a somewhat technical problem that should be addressed right away.

The formulation just given (i.e. "an attribute is extrinsic iff it is denoted by . . . ") would be acceptable if it did not falsely suggest that the λ-expression in question belongs to some language whose semantics fully settles which λ-expressions denote which attributes. While it is true that we will be relying on a certain language (to be described in the next Section), the semantics of that language will be extremely 'gappy', insofar as the only atomic expressions to have a fixed meaning will be a handful of operators and a single constant denoting the identity relation. As a result, any λ-expression that contains occurrences of one or more other constants will have a denotation *only* relative to a function (called 'interpretation') that maps each of the respective constants to some entity. Similarly, if a given λ-expression contains free occurrences of some variables, it will have a denotation only relative to a function that maps each one of those variables to some entity. In the general case, we will thus have to speak of λ-expressions as having—or failing to have—a denotation only relative to some interpretation and variable-assignment.

As it is often tedious to keep track of relativisations to interpretations and variable-assignments, it might be thought convenient to fix on a particular interpretation $\mathbf{I}$ and variable-assignment $\mathbf{g}$ and to declare that all further talk of denotation is to be understood

---

Some philosophers, wishing to avoid what Gilmore (2013) has in a different context called a 'priority problem', may prefer to give an account of intrinsicality not in terms of λ-expressions (or any other sort of 'linguistic' entity), but rather exclusively in terms of 'ways in which an attribute may be constituted from other entities', or *modes of constitution* for short. In principle this may be possible, but it is likely to run into some practical difficulties. In particular, the linear structure of λ-expressions makes it easy to single out individual occurrences of expressions, such as variables or quantifiers. Modes of constitution, by contrast, are not naturally thought of as having a linear structure. (The importance of being able to single out term- and operator-occurrences—or whatever the appropriate analogue would be if we were instead talking about modes of constitution—will become apparent in Sections 7–10.)

as relativised to **I** and **g**. But this proposal suffers from a serious shortcoming. Since there are too many things to form a set (if one counts the ordinals, for instance), *there simply does not exist* any interpretation or variable-assignment that assigns a constant or variable to every entity—at least not if interpretations and variable-assignments are conceived of as functions and functions as sets of ordered pairs. To avoid the limitations that would result from relativising to just a single interpretation and variable-assignment, we can instead quantify over *all* interpretations I and variable-assignments g and then (in the scope of such a quantification) relativise to the respective I and g. We will thus say that an attribute A is extrinsic iff for some interpretation I, variable-assignment g, and λ-expression L: (i) L denotes A relative to I and g, and (ii) L, together with I and g, satisfies such-and-such further conditions. Conversely, since an attribute is intrinsic iff it fails to be extrinsic, our account of *in*trinsicality will take the following form:

(In_1) An attribute A is *intrinsic* iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes A relative to I and g, then . . . .

This is of course only a start. We still need to find a sensible way of filling in the ellipsis, and in fact most of the rest of this paper will be taken up by this task. But first we have to outline the formal framework that we will be operating in.

5 A Formal Language

We have to outline a system of expressions—a formal language—by which attributes and states of affairs can be denoted. The language in question is essentially a version of the familiar language of first-order logic, except for a few modifications.[29]

[29]More specifically, the system sketched below is for the most part identical to the one described in my (2016: §2). The only substantial difference is that the present language admits infinitely 'deep' formulas, i.e. formulas with an infinite number of levels of parentheses. (The existence of such formulas is a consequence of the fact that the language allows for infinitary conjunctions. For suppose that $\varphi_1, \varphi_2, \ldots$ are countably many formulas such that, for each n with $0 < n < \omega$, φ_n is of depth n. The infinitely long conjunction of the φ_i, i.e. $\ulcorner \varphi_1 \wedge \varphi_2 \wedge \ldots \urcorner$, will then be of infinite depth. The language of my [2016] also allows for infinitary conjunctions, but one of the constraints that are there imposed on the syntax of formulas prohibits infinitely deep formulas, thereby leading to inconsistency.)

5.1 Syntax

The operators of the language are '$\neg$', '$\wedge$', '$\exists$', and 'λ', to which the parentheses '(' and ')' and the comma are added as auxiliary symbols. Italicised letters (in upper or lower case, and with or without sub- or superscripts) are used as *variables*, and unitalicised English words or hyphenated phrases (again with or without sub- or superscripts) are used as *constants*. We will be assuming that we have available proper-class many variables and constants. The symbol '$\mathtt{I}$' will be employed as a special 'logical constant' to denote the identity relation. We will count as a *term* anything that is a variable, constant, formula, or λ-expression. This inclusiveness is motivated by the fact that formulas and λ-expressions are intended to serve as names (of states of affairs and attributes, respectively). A term will be said to be *atomic* iff it is either a variable or a constant.

The syntax of non-atomic terms (i.e. of formulas and λ-expressions) can be specified in the usual recursive manner. To begin with, a *formula* is any expression that conforms to one of the following four patterns:

(i) $\ulcorner P(t_1, t_2, \ldots)\urcorner$, where P is a constant, variable, or λ-expression, and where $t_1, t_2, \ldots$ are one or more terms.

(ii) $\ulcorner \neg t\urcorner$, where t is a constant, variable, or formula.

(iii) $\ulcorner (t_1 \wedge t_2 \wedge \ldots)\urcorner$, where the t_i are two or more constants, variables, or formulas.

(iv) $\ulcorner \exists v_1, v_2, \ldots\ \varphi\urcorner$, where the v_i are one or more pairwise distinct variables and φ is a formula such that: each v_i has at least one free occurrence in φ, and each free occurrence of each v_i in φ stands at subject-position.

To say that a term-occurrence stands at *subject-position* means here that it occurs as an element of an argument-list, which is a list of one or more term-occurrences enclosed in parentheses, delimited by commas, and immediately preceded by an occurrence of a constant, variable, or λ-expression. Two examples would be the occurrences of 'x' and 'y' in '$P(x, y)$'. By contrast, we will say that a term-occurrence stands at *predicate-position* iff it immediately precedes an argument-list. An example is the occurrence of 'P' in '$P(x, y)$'.

Further, a *λ-expression* is any expression $\ulcorner \lambda v_1, v_2, \ldots\ \varphi \urcorner$, where the v_i are one or more pairwise distinct variables and φ is a formula such that: each of the v_i has in φ at least one free occurrence, and each free occurrence of each v_i in φ stands at subject-position. These restrictions, as also the parallel restrictions in (iv) above, are motivated by the need to avoid semantic paradox.[30] The distinction between free and bound occurrences can be drawn in the usual way. (For formal definitions, see Appendix B.) In connection with λ-expressions, we will also need to speak of the *λ-variables* and the *matrix* of such an expression. In particular, for any λ-expression $\ulcorner \lambda v_1, v_2, \ldots\ \varphi \urcorner$, the v_i will be said to be its λ-variables and φ to be its matrix. We will say that a term-occurrence stands *at sentence-position* iff it is either (a) the operand of some occurrence of '$\neg$' or '$\wedge$' or (b) a formula-occurrence that, for some variables $v_1, v_2, \ldots$, is immediately preceded by an occurrence of $\ulcorner \lambda v_1, v_2, \ldots \urcorner$ or $\ulcorner \exists v_1, v_2, \ldots \urcorner$.

By way of abbreviation, we will write $\ulcorner (t = t') \urcorner$ instead of $\ulcorner \mathtt{I}(t, t') \urcorner$ and $\ulcorner (t \neq t') \urcorner$ instead of $\ulcorner \neg \mathtt{I}(t, t') \urcorner$, where t and t' may be any terms. In addition, $\ulcorner (\varphi_1 \vee \varphi_2 \vee \ldots) \urcorner$ will abbreviate $\ulcorner \neg(\neg\varphi_1 \wedge \neg\varphi_2 \wedge \ldots) \urcorner$, $\ulcorner (\varphi \rightarrow \psi) \urcorner$ and $\ulcorner (\varphi \leftrightarrow \psi) \urcorner$ will respectively abbreviate $\ulcorner \neg(\varphi \wedge \neg\psi) \urcorner$ and $\ulcorner \neg(\varphi \wedge \neg\psi) \wedge \neg(\psi \wedge \neg\varphi) \urcorner$, and we will write $\ulcorner \forall v_1, v_2, \ldots\ \varphi \urcorner$ as shorthand for $\ulcorner \neg\exists v_1, v_2, \ldots\ \neg\varphi \urcorner$. Occasionally, we will also write $\ulcorner (t_1 \,\&\, t_2 \,\&\, \ldots) \urcorner$ (where $t_1, t_2, \ldots$ are atomic terms distinct from 'x') to abbreviate $\ulcorner \lambda x\ (t_1(x) \wedge t_2(x) \wedge \ldots) \urcorner$. As usual, outermost parentheses may be omitted.

5.2 Semantics

In general, a term has or lacks a denotation only relative to some interpretation and variable-assignment. An *interpretation* is here a partial function from constants to entities that maps the constant '$\mathtt{I}$' to the identity relation, and a *variable-assignment* is simply a partial function from variables to entities. As usual, these functions will be treated as sets of ordered pairs. (It will thus, e.g., make sense to speak of a superset of an interpretation.) We will often write 'a denotes$_{I,g}$ b' as an abbreviation of 'a denotes b relative to I and g', and we will say that two terms are *coreferential* relative to I and g—or coreferential$_{I,g}$ for

[30]Cf. Plate (2016: 7n., 12n.).

short—iff they both denote$_{I,g}$ the same entity. (If two terms fail to denote$_{I,g}$ anything, they will thus *not* be coreferential relative to I and g.) We further say that a cardinality κ is an *adicity of* an attribute A iff A has an instantiation by some κ-sequence of entities.[31] An attribute with an adicity κ will also be called κ-adic.

The semantics of formulas and λ-expressions can be recursively specified by the following six stipulations, where the relevant notions of *instantiation*, *negation*, *conjunction*, and *existential quantification* are taken as primitive and where I and g may be any interpretation and variable-assignment:

(S1) A constant or variable t has a denotation relative to I and g iff either I or g maps t to some entity. In this case t denotes$_{I,g}$ that entity.

(S2) A formula $\ulcorner P(t_1, t_2, \ldots)\urcorner$ has a denotation relative to I and g iff (i) each t_i has a denotation relative to I and g and (ii) P denotes$_{I,g}$ some attribute whose adicity matches the length of the argument-list. In this case $\ulcorner P(t_1, t_2, \ldots)\urcorner$ denotes$_{I,g}$ the instantiation of the attribute denoted$_{I,g}$ by P by the entities that are respectively denoted$_{I,g}$ by $t_1, t_2, \ldots$ (in this order).

(S3) A formula $\ulcorner \neg t \urcorner$ has a denotation relative to I and g iff t denotes$_{I,g}$ some state of affairs s. In this case $\ulcorner \neg t \urcorner$ denotes$_{I,g}$ the negation of s.

(S4) A formula $\ulcorner t_1 \wedge t_2 \wedge \ldots \urcorner$ has a denotation relative to I and g iff each t_i denotes$_{I,g}$ a state of affairs. In this case $\ulcorner t_1 \wedge t_2 \wedge \ldots \urcorner$ denotes$_{I,g}$ the conjunction of the states of affairs respectively denoted$_{I,g}$ by the t_i.

(S5) A formula $\ulcorner \exists v_1, v_2, \ldots\ \varphi \urcorner$ has a denotation relative to I and g iff $\ulcorner \lambda v_1, v_2, \ldots\ \varphi \urcorner$ denotes$_{I,g}$ an attribute A. In this case $\ulcorner \exists v_1, v_2, \ldots\ \varphi \urcorner$ denotes$_{I,g}$ the existential quantification of A.

(S6) A λ-expression $\ulcorner \lambda \underbrace{v_1, v_2, \ldots}_{\kappa\text{-many}}\ \varphi \urcorner$ has a denotation relative to I and g iff φ denotes a

[31]This talk of instantiation "by some κ-sequence of entities" should be understood as shorthand for talk of instantiation by some entities $\underbrace{x_1, x_2, \ldots}_{\kappa\text{-many}}$. To say that an attribute 'has' such an instantiation means that there *exists* such an instantiation, not that it obtains.

state of affairs relative to I and some variable-assignment that differs from g at most in what it assigns to the v_i. In this case $\ulcorner\lambda v_1, v_2, \ldots\ \varphi\urcorner$ denotes$_{I,g}$ an attribute A such that, for any κ-sequence of entities $x_1, x_2, \ldots$:[32] the instantiation of A by $x_1, x_2, \ldots$ (in this order) is the state of affairs that is denoted by φ relative to I and a variable-assignment that differs from g at most insofar as, for each $i \in \{1, 2, \ldots\}$, it assigns x_i to v_i.

As can be seen from these stipulations, we are making use of a rather abundant ontology of states of affairs (and relatedly of attributes), since we are assuming that every state of affairs has a negation, that every set of states of affairs has a conjunction, and so on.

6 On the Individuation of Attributes

While it may seem desirable to have an account of intrinsicality that yields intuitively correct results more or less independently of how attributes are individuated (e.g., whether in a coarse- or fine-grained fashion), this desideratum should arguably not be regarded as an absolute requirement.[33] In this Section, I lay out some considerations that favour a certain *moderately* coarse-grained way of individuating attributes. For brevity's sake, ways of individuating attributes will here be referred to as 'conceptions' of attributes. The moderately coarse-grained conception at which we will arrive at the end of this Section will form the background for the account of intrinsicality to be developed in Sections 7–10.

6.1 Two auxiliary principles

In order to eliminate a certain ambiguity in the semantics of λ-expressions, I propose to adopt the following two auxiliary principles:[34]

---

[32] More precisely: for any sequence of entities $x_1, x_2, \ldots$ whose length exactly matches that of $v_1, v_2, \ldots$ (i.e. the sequence of the λ-variables of $\ulcorner\lambda v_1, v_2, \ldots\ \varphi\urcorner$), so that for each v_i in the latter sequence there is an x_i in the former, and *vice versa*. (This precisification is needed for the case of infinitely long sequences.)

[33] See the second half of Section 6.1 below. (Also cf. Appendix A, footnote 12.)

[34] Cf. Plate (2016: 10). The first of these principles will be controversial, as it is often held that some attributes can be instantiated by variable numbers of entities. (E.g., see Yi [1999], MacBride [2005: §§2.2ff.], McKay [2006: 13].) If there are such attributes, I will here have to ignore them to keep the exposition manageable.

(P1) Every attribute has exactly one adicity.

(P2) If A and B are two κ-adic attributes, then there exists some κ-sequence of entities $x_1, x_2, \ldots$ such that A's instantiation by $x_1, x_2, \ldots$ (in this order) is distinct from B's instantiation by $x_1, x_2, \ldots$ (in this order).

Taken together, (P1) and (P2) entail that no attribute shares all of its instantiations with any other. As already hinted at, these principles are motivated by considerations about the semantics of λ-expressions. For what the above stipulation (S6) tells us about the denotation of such expressions is only (i) under what conditions a given λ-expression $L = \ulcorner \lambda \underbrace{v_1, v_2, \ldots}_{\kappa\text{-many}} \varphi \urcorner$ has a denotation and (ii) that, if those conditions are satisfied, L denotes *an* attribute whose instantiation by any given κ-sequence $x_1, x_2, \ldots$ is the state of affairs denoted by φ relative to a certain interpretation and variable-assignment.[35] With (P1) and (P2) in place, that indefinite article can be replaced with a definite one, since there will then be only one attribute that satisfies the description.

Given those two principles, and in particular (P2), attributes may in the present framework be said to be individuated by their instantiations. It follows that the coarse-grainedness of our conception of attributes will closely correspond to the coarse-grainedness of our conception of states of affairs.[36]

Prima facie it is desirable to have an account of intrinsicality that correctly classifies attributes as intrinsic or extrinsic regardless of how finely or coarsely they may be individuated. On reflection, however, this is not so clear. For suppose that a certain conception of attributes is so coarse-grained as to identify a certain intuitively intrinsic property P with a certain intuitively *ex*trinsic property Q. (An example of such a conception is the familiar view on which properties are functions from possible worlds to extensions. Assuming that Socrates could not have failed to be a son of Phaenarete, the intuitively intrinsic property

[35]To reduce clutter, I am here largely suppressing references to interpretations and variable-assignments.

[36]This will become clearer in the next subsection. By and large, to say that a given conception of states of affairs is more coarse-grained than another means, in the terms of our framework, that under the first conception a state of affairs can typically be denoted by a broader range of formulas than under the second. Correspondingly, to say that a given conception of *attributes* is more coarse-grained than another means that, under the first conception, an attribute can typically be denoted by a broader range of λ-expressions than under the second.

of *being Socrates* is on this conception identical with the intuitively extrinsic property of *being Socrates and a son of Phaenarete*.) Should we say that any account of intrinsicality, even when combined with such a conception, ought still to yield the verdict that P is intrinsic and Q extrinsic? On one possible view the answer is 'yes': the account has to classify P and Q in this way in order to be extensionally adequate—even if this means that, under the conception in question, *one and the same* property is classified as both intrinsic and extrinsic. On another possible view, the answer is 'no': if the conception in question is correct, then P and Q really are one and the same property, and on pain of contradiction we shouldn't classify one and the same property as both intrinsic and extrinsic. Rather we should classify it as *either* intrinsic or extrinsic, and we are free to do so in accordance with our account ("spoils to the victor").

Which of these two views is correct? That plausibly depends on the circumstances. If one has better reasons to stick with that coarse-grained conception of attributes than to accept the intuitive assessment according to which P is intrinsic and Q extrinsic, then the second view may very well be the appropriate one to take.[37] But what if one has better reason to accept the intuitive assessment than to stick with the coarse-grained conception? It is not obvious that in this case the first view would be the correct one to hold. For if P is intrinsic and Q extrinsic, it immediately follows that any conception of attributes that identifies the two is false; and it is far from clear that we should want our account of intrinsicality to yield, when combined with such a conception, the "correct" but self-contradictory verdict that one and the same property is both intrinsic and non-intrinsic. After all, the overall theory that results from combining our account with such a conception would in that case be not only false but inconsistent. Accordingly, if our aim is a correct overall theory, it appears that we should simply reject that conception.

These remarks straightforwardly generalise to a broad range of cases. For presumably it does not matter very much *why* a given conception of attributes is rejected, as long as it is rejected with good reason; and so we may draw the (tentative) conclusion that, in constructing our account of intrinsicality, we should feel free to ignore conceptions of

[37] Another option might be to say that some, or all, properties are intrinsic or extrinsic *only relative to* a certain description. But this would go against the way the intrinsic/extrinsic distinction is commonly applied.

attributes that we have good reason to reject. In the following two subsections, I shall identify (and then reject) two sorts of such conceptions: first, those that are in a certain sense *irregular*, and second, those that on closer inspection turn out to be implausibly coarse-grained.

6.2 A regularity assumption

Suppose that, on a certain view about properties, the property of *being Socrates* is identical with *being Socrates and someone's son*, on the grounds that Socrates is essentially a son of Phaenarete. Apart from identifying an intuitively intrinsic property with one that is intuitively extrinsic, the view under consideration appears problematic also because it renders the individuation of properties an 'irregular' affair: it takes the individuation of properties to depend on questions that are strictly extraneous to the theory of properties *per se*, such as the question of what properties Socrates has essentially. To be sure, questions of this nature will be quite relevant and not at all extraneous if we think of properties as functions from possible worlds to extensions, or as sets of *possibilia*. But as we have seen above, this sort of conception should be rejected for being overly coarse-grained (if not already for its commitment to *possibilia*).[38]

Similar remarks hold for the individuation of states of affairs: if we reject the possibilist identification of states of affairs with sets of possible worlds (or with functions from possible worlds to truth-values), there will be little reason to think that Socrates' being self-identical (in symbols: Socrates = Socrates) is the same as the state of affairs that Socrates is self-identical *and* a son of Phaenarete. Instead, we will be free to individuate states of affairs in a manner that depends only on their respective logical constitution, as expressed by the formulas by which they can be denoted.

This is of course quite general and abstract. A natural way of formulating a fully specific thesis about the individuation of states of affairs in accordance with this general idea would be to take some concept of *equivalence* (where, under any such concept, the equivalence of two formulas is solely a matter of their logical form), and to use it in framing a thesis along

[38]Also cf. the next subsection.

the following lines:[39]

(CS) For any interpretation I and variable-assignment g, no two equivalent formulas denote$_{I,g}$ distinct states of affairs.

The weaker the concept of equivalence that is here 'plugged in' as the relevant sense of 'equivalent', the more coarse-grained the individuation of states of affairs will have to be in order for (CS) to be true. (Thus, weaker concepts of equivalence correspond to more demanding readings of (CS).) Assuming that (CS) is true under a given concept of equivalence, (P1) and (P2), together with (S6), can be seen to give rise to the following thesis about the coarse-grainedness of the individuation of attributes (in which the term 'equivalent' is to be read in the same sense as in (CS)):

(CA) For any interpretation I and variable-assignment g, no two λ-expressions denote$_{I,g}$ distinct attributes if they have equivalent matrices and the same λ-variables in the same order.

Now, for any given concept of equivalence, (CS) and (CA) only articulate a *lower* bound on the coarse-grainedness of the individuation of (respectively) states of affairs and attributes. Thus they do not rule out such irregularities as the identity of *being Socrates* and *being Socrates and a son of Phaenarete*. To enforce a 'regular' individuation of attributes and states of affairs, one would in addition have to adopt a thesis that imposes an *upper* bound on the coarse-grainedness of states of affairs. This might be done, to a first and very rough approximation, by saying that *only* equivalent formulas can denote the same state of affairs. But this would be a serious mistake, for if two formulas contain distinct sets of constants denoting the same entities, they may easily fail to be equivalent and yet denote the same state of affairs; and this need not have anything to do with the coarse-grainedness of states of affairs.[40] In order to account for such cases, it will be helpful to draw on the notion of a 'reduction' of a term. Roughly, a reduction of a term t can be understood as being simply

[39]The 'C' in the label stands for 'coarse-grained' and the 'S' for 'state of affairs'. In the following we will also use 'F' for 'fine-grained' and 'A' for 'attribute'.

[40]For example, relative to an interpretation that maps the constant 'bald' to some property P and maps both 'Cicero' and 'Tully' to the same entity x, the two inequivalent formulas 'bald(Cicero)' and 'bald(Tully)' will denote one and the same state of affairs, viz., P's instantiation by x.

the result of replacing in t zero or more occurrences of atomic terms with coreferential other terms. More precisely:[41]

Definition 6.1. A term t' is a *reduction* of a term t relative to an interpretation I and a variable-assignment g iff t' is the result of replacing in t zero or more occurrences of atomic terms by occurrences of other terms, in such a way that the following three conditions are satisfied:[42]

(i) Each replaced occurrence o is replaced with a term that is coreferential$_{I,g}$ with the term of which o is an occurrence.

(ii) No replaced occurrence is a bound variable-occurrence.

(iii) No replacing occurrence o contains a variable-occurrence that is bound by an operator-occurrence not in o.

It is easy to check that any reduction of a formula is again a formula, and that any reduction of a λ-expression is again a λ-expression. Moreover, any term is a reduction of itself (relative to any given interpretation and variable-assignment). For the sake of brevity, we will often write 'reduction$_{I,g}$' instead of 'reduction relative to I and g'.

With the help of the concept of reduction, we can impose an upper bound on the coarse-grainedness of states of affairs (and consequently attributes) by adopting a thesis of the following form:

(FS) For any interpretation I, variable-assignment g, and formulas φ and ψ: if φ and ψ denote$_{I,g}$ the same state of affairs, then there exist an interpretation $I' \supseteq I$ and a variable-assignment $g' \supseteq g$ such that some reduction$_{I',g'}$ of φ is equivalent to some reduction$_{I',g'}$ of ψ.[43]

[41] Cf. Plate (2016: 24).

[42] Formal definitions relating to the notions of occurrence and containment may be found in Appendix B.

[43] To see the need for quantification over supersets of the respective interpretation I and/or the respective variable-assignment g, suppose that $\varphi =$ '$a \wedge b$', $\psi =$ '$c \wedge d$', and that three states of affairs s_1, s_2, and s_3 are such that (i) 'a' denotes$_{I,g}$ $s_1 \wedge s_2$, (ii) 'b' denotes$_{I,g}$ s_3, (iii) 'c' denotes$_{I,g}$ s_1, (iv) 'd' denotes$_{I,g}$ $s_2 \wedge s_3$, and (v) no term denotes$_{I,g}$ s_2. Finally, suppose that, for any atomic terms or formulas t_1, t_2, and t_3, the formula $\ulcorner(t_1 \wedge t_2) \wedge t_3\urcorner$ is in the relevant sense equivalent to $\ulcorner t_1 \wedge (t_2 \wedge t_3)\urcorner$. In this case, it may still turn out that no reduction$_{I,g}$ of φ is equivalent to any reduction$_{I,g}$ of ψ. However, if g' is a variable-assignment that maps each variable in g's

In the opposite way to (CS), weaker concepts of equivalence correspond to less demanding versions of (FS). But analogously to (CS), under any particular reading, the truth of (FS) will give rise to a corresponding thesis about the coarse-grainedness of attributes:

(FA) For any interpretation I, variable-assignment g, and λ-expressions L_1 and L_2: if L_1 and L_2 have the same λ-variables in the same order and denote$_{I,g}$ the same attribute, then there exist an interpretation $I' \supseteq I$ and a variable-assignment $g' \supseteq g$ such that the matrix of some reduction$_{I',g'}$ of L_1 is equivalent to the matrix of some reduction$_{I',g'}$ of L_2.

This thesis follows from (FS) by a relatively straightforward proof.[44]

Piggybacking (as it were) on (CS) and (FS), our regularity assumption can now be stated as follows:

(R) There exists a concept of equivalence under which (CS) and (FS) are both true.

By the above observations, any concept of equivalence that renders both (CS) and (FS) true will do the same for (CA) and (FA). Informally, we will say of a concept of equivalence that it 'captures' the coarse-grainedness of a given conception of attributes iff, according to that conception, the concept in question renders both (CA) and (FA) true.

---

domain to the same entity as g (which is to say that $g' \supseteq g$) and in addition maps the variable 'e' to s_2, then '$(c \wedge e) \wedge b$' will be a reduction$_{I,g'}$ of φ, and the formula '$c \wedge (e \wedge b)$', which is equivalent to '$(c \wedge e) \wedge b$', will be a reduction$_{I,g'}$ of ψ. By the corresponding version of (CS) and the semantics of formulas, it then follows that φ and ψ denote$_{I,g}$ the same state of affairs; for, by the semantics of formulas, φ and ψ are coreferential$_{I,g'}$ with their respective reductions$_{I,g'}$ '$(c \wedge e) \wedge b$' and '$c \wedge (e \wedge b)$', and by (CS), '$(c \wedge e) \wedge b$' is coreferential$_{I,g'}$ with '$c \wedge (e \wedge b)$'. (Analogous considerations apply to any interpretation $I' \supseteq I$ that maps some constant to s_2.) Hence, if (FS) did not quantify over supersets of either I or g, it would risk being false in a way that has nothing to do with the coarse-grainedness of states of affairs.

[44] Suppose (FS) holds, and assume for *reductio* that (FA) doesn't. The falsity of (FA) would mean that there exist an interpretation I, a variable-assignment g, and λ-expressions L_1 and L_2 such that L_1 and L_2 have the same λ-variable(s) in the same order and denote$_{I,g}$ the same κ-adic attribute A, but do not have reductions$_{I',g'}$ (for any $I' \supseteq I$ and $g' \supseteq g$) with equivalent matrices. From this it follows that the matrices of L_1 and L_2 themselves do not have equivalent reductions$_{I',g'}$ (again, for any $I' \supseteq I$ and $g' \supseteq g$). Using (FS), one can then infer that, for any κ-sequence of entities $x_1, x_2, \ldots$ and any variable-assignment g': if g' differs from g at most insofar as it maps the λ-variables of L_1 (and thus of L_2) to $x_1, x_2, \ldots$, respectively, then those matrices denote$_{I,g'}$ distinct states of affairs. But by the semantics of λ-expressions—i.e. (S6) of Section 5.2—those two states of affairs would *both* have to be "the" instantiation of A by $x_1, x_2, \ldots$ (in this order), and that cannot be.

6.3 Fineness of grain

How fine-grained would a correct conception of attributes have to be? Very generally speaking, we have a defeasible reason to favour coarser-grained conceptions of attributes over finer-grained ones, at least insofar as properties and relations are supposed to be 'ways for things to *be*', rather than ways in which things might be represented. A natural way to spell this out is in modal terms: if P and Q are distinct properties, then it should be possible for something to have P without having Q, or *vice versa*. But this is only a rough guide. The properties of *being non-self-identical* and *being such that Socrates is non-self-identical*, for instance, are both uninstantiable, and so it is trivially impossible for anything to have the one without the other. Nonetheless, it is intuitively plausible to distinguish them, on the grounds that they differ with respect to what it takes for them to exist. In order for *being non-self-identical* to exist, we seem to need (under certain assumptions about the combinatorial abundance of attributes) only the existence of the identity relation (and it may well be doubted whether this is a substantial requirement, as opposed to one that is trivially satisfied), while, for *being such that Socrates is non-self-identical* to exist, we also need the existence of Socrates. A friend of a starkly coarse-grained conception of attributes, on which these two properties are one and the same, may find ways of dismissing this line of thought, but its intuitive pull seems rather strong.[45] What it suggests is that the individuation of attributes is considerably more fine-grained than it would be under standard possibilist construals (under which properties are either sets of *possibilia* or functions from possible worlds to extensions). Of course the same is suggested, as we have seen above, by such pairs of properties as *being Socrates* and *being Socrates and a son of Phaenarete*, of which one is intuitively intrinsic and the other intuitively extrinsic.

Intuitions of this sort have to be treated with some suspicion, however, or else they might quickly lead us to some overly fine-grained conception of attributes. It may for instance be tempting to think that, for any property P, the property $\lambda x\,(P(x) \land \exists y\,P(y))$ *must* be extrinsic, given that it can be denoted by this particular λ-expression (with that

[45]The issue is loosely related to the debate over what Plantinga (1983) calls 'existentialism', as well as to the 'being constraint' discussed in Williamson (2013: §4.1).

unrestricted quantifier in the second conjunct!), and that it must therefore be distinct from P if P is intrinsic. By this line of thought, a theorist may be swiftly led to an unusually fine-grained conception of attributes that completely overturns the original conception of properties as 'ways for things to be'. The mistake here lies in a tendency to think that a complex λ-expression can only denote an attribute that is at least equally complex. To guard against this mistake, it will be well to keep in mind that λ-expressions are representations, and representations, like appearances, can be misleading.[46] More particularly, unless we have already embraced an extremely (and implausibly) fine-grained conception of attributes, we should expect that some λ-expressions may exhibit a certain amount of what might be called 'redundant complexity'.[47]

What exactly counts as 'redundant complexity' will depend on the coarse-grainedness of the conception in question. Somewhat more precisely, it will depend on what the weakest concept of equivalence is under which (CS) still comes out true. (Given the regularity assumption, this will also be the strongest concept of equivalence under which (FS) comes out true.[48]) Under a suitable such concept, the second conjunct of '$\lambda x\,(P(x) \wedge \exists y\, P(y))$', for instance, will constitute an instance of redundant complexity, in particular if the formula '$P(x) \wedge \exists y\, P(y)$' is in the relevant sense equivalent to '$P(x)$'.

To return to the main topic of this subsection, the answer to the question of how fine-grained a correct conception of attributes would have to be would, in keeping with the regularity assumption, have to be given by some concept of equivalence. To define such a concept, we may in a first step introduce the notion of a 'semantically well-formed term':

Definition 6.2. A term t is *semantically well-formed* iff t has a denotation relative to at least one interpretation and variable-assignment.

For example, under the assumption that no attribute is a state of affairs, the formula '$P(x) \wedge$

[46]For a possibly helpful analogy, consider the notion of a *disjunctive* property. Presumably one would not think that any property that can be denoted by '$\lambda x\,(P(x) \vee P(x))$' must be disjunctive (despite that conspicuous disjunction sign). Rather, the second disjunct in '$\lambda x\,(P(x) \vee P(x))$' will be most reasonably treated as a case of redundant complexity and any property P as identical with $\lambda x\,(P(x) \vee P(x))$.

[47]Cf. Plate (2016: 21).

[48]At least if equivalent concepts are treated as identical. To see this, note first that the strongest concept under which (FS) is true cannot be stronger than the weakest concept under which (CS) is true. By (R), the former must then be equivalent to the latter.

P' will not count as semantically well-formed. This concept is useful to delineate the class of pairs of formulas that may properly be said to entail each other. (Of a contradictory formula, such as '$x \neq x$', it may be appropriate to say that it trivially entails every formula whatsoever, but it would be very odd to say the same thing of a semantically ill-formed formula such as '$P(x) \wedge P$'.) In the next step, we can now use this notion to define the following concept of entailment:[49]

Definition 6.3. A formula φ *entails* a formula ψ iff φ is semantically well-formed and, for any interpretation I and variable-assignment g, the following two conditions are satisfied:

(i) If φ has a denotation relative to I and g, then so does ψ.

(ii) If φ denotes$_{I,g}$ an obtaining state of affairs, then so does ψ.

On this basis, the corresponding concept of equivalence can be introduced in the usual way, viz., by stipulating that two formulas are *equivalent* iff they entail each other.

An important characteristic of the present notion of entailment lies in the fact that a formula will entail another *only* if the second contains no free occurrence of any variable or non-logical constant that does not also have a free occurrence in the first. For instance, the formulas '$a = a$' and '$b = b$' do not entail each other, despite being classically equivalent. Hence, if the present concept of equivalence is such as to render (FS) true, it very plausibly follows (given that Socrates is distinct from Plato) that the state of affairs (Socrates = Socrates) is distinct from (Plato = Plato), and similarly, the property of *being non-self-identical* will be distinct from that of *being such that Socrates is non-self-identical*. By what has been said at the beginning of this subsection, these would be welcome consequences.

It can be argued, however, that the present concept of equivalence is not yet strong enough. For example, let E be the property of *being an electron*, let S be the property of *being self-identical*, and consider the property of *being either an electron distinct from S or a non-electron identical with S*.[50] Relative to a variable-assignment that maps 'E' to E, this

[49]Cf. Plate (2016: 13).

[50]This might be more concisely expressed as: *being an electron iff distinct from S*. The example is adapted from Parsons' (2001: 23) example of *being either a lonely positron or an accompanied electron*. (For a discussion of the latter, see Appendix A.)

latter property—call it 'E^*'—is denoted by the following λ-expression:

$$\lambda x\,\big((E(x) \wedge (x \neq \lambda x(x = x))) \vee (\neg E(x) \wedge (x = \lambda x(x = x)))\big). \tag{1}$$

Further, if we individuate attributes in accordance with the present concept of equivalence, then E itself can quite analogously be denoted (relative to a variable-assignment that maps 'E^*' to E^*) by the following:

$$\lambda x\,\big((E^*(x) \wedge (x \neq \lambda x(x = x))) \vee (\neg E^*(x) \wedge (x = \lambda x(x = x)))\big). \tag{2}$$

For if we replace the two occurrences of 'E^*' in (2) with (1), we obtain a λ-expression whose matrix is equivalent to '$E(x)$', whence it follows by (CA) that, relative to an assignment that maps 'E' to E, (2) denotes $\lambda x\, E(x)$, which is to say, E.

Given the perfect analogy between (1) and (2), we can now see that the logical constitution of E^* and E must also be perfectly alike; and hence, if the logical constitution of E is such as to render it intrinsic, the same must be true of E^*. Yet intuitively, whereas E (i.e. the property of *being an electron*) may for all we know be logically simple and intrinsic, E^* fairly clearly seems to be logically complex and *ex*trinsic.

Assuming that it won't do to accept E^* as intrinsic, it might be concluded that attributes are not rendered intrinsic or extrinsic by their respective logical constitutions alone (if at all). In other words, it might be thought that it is time to give up on the logical-constitution approach and to start enlisting the help of genuinely metaphysical notions, such as Lewis's concept of a perfectly natural attribute.[51] But this sort of proposal, however tempting,

[51] This concept has notably been employed in Lewis's own (1983b; 1986a) accounts of intrinsicality as well as in those of Langton & Lewis (1998) and Marshall (2016a). What makes this concept a natural choice is perhaps the fact that it is in some suitable sense a concept of *fundamentality* and thus gives rise to the hope that, once E^* is represented in perfectly natural terms, its true extrinsic character will be evident from the representation in question. If this is the rough idea, one might implement it by replacing the 'top level' of our account (as developed so far), i.e. (In_1), with the following:

(In_1') An attribute A is *intrinsic* iff there exist an interpretation I, a variable-assignment g, and a λ-expression L such that the following three conditions are satisfied:

(i) L denotes$_{I,g}$ A.

(ii) Each atomic term that occurs in L at predicate-position denotes$_{I,g}$ a perfectly natural attribute.

(iii) . . . .

is in the present context clearly too narrow. For, just as it is intuitively hard to swallow that E^* will be intrinsic if E is, it is likewise difficult to accept that E and E^* should be perfectly alike in their logical constitution. Moreover, if we were to accept it in this case, we would, more generally, have to say that every logically simple property P is as it were surrounded by a halo of other simple properties that are related to P just like E^* is to E. This point obviously generalises to logically simple *relations*. And as if this were not enough, those additional attributes, being themselves logically simple, would all have their own halos. While we might learn to live with this consequence, it is certainly counter-intuitive. Possibly it might be suggested that we have to revise our notions of logical constitution and logical simplicity, but how to do so? Enlisting 'genuinely metaphysical' notions for this purpose would appear both *ad hoc* and inappropriate.

This apparently leaves us with only one remaining option, viz., to move towards a less coarse-grained conception of attributes. In other words, we have to assume that the weakest concept of equivalence under which (CS) and (CA) still hold true is stronger than the one defined above. To solve our problem, this new concept has to be chosen in such a way that the matrix of the reduction of (2) that results from replacing the two occurrences of 'E' with (1) will no longer be equivalent to '$E(x)$'; for only then will it no longer follow from some true reading of (CA) that E is denoted by (2).[52]

This manoeuvre would spare us some of the complications that we will get into in Section 8.1 below. However, the appeal to perfect naturalness is itself not unproblematic. For example, in Lewis's own explications of the concept, he makes use of some descriptions that are not obviously all satisfied by a single kind of attribute (cf. Schaffer [2004] and Dorr & Hawthorne [2013]), which lends some credence to the view that the notion is simply too obscure to be useful. (Cf., e.g., Taylor [1993: 88], Bealer & Mönnich [2003: 195], Witmer et al. [2005: 329].) As a way out, it might be suggested that we reconceive the perfectly natural attributes as those that are "invoked in the scientific understanding of the world" (to borrow Schaffer's [2004: 92] phrase). However, while this promises to reduce the obscurity, it would also burden our analysis with questions as to what attributes are or are not invoked in the scientific understanding of the world, and one would not normally have thought that such questions have anything to do with intrinsicality. (Cf. also Yablo [1999: 480].) Another alternative, which I have suggested in my (2016: 37), would be to identify the perfectly natural properties with "those that are both logically simple and 'positive'", where the 'positive' properties are those that are "more restrictive than their respective negations as to what other attributes they can—within the bounds of nomological necessity—be coinstantiated with" (p. 30). But in the present context, this proposal would arguably not be helpful, either, since E^* appears to be no less 'positive' than E.

[52]At this point it would be desirable to have a thesis that is able to express the fine-grainedness of a given conception of attributes more fully than (FA) does. For in the absence of an assumption to the effect that E is denoted by such-and-such a λ-expression *other* than '$\lambda x\, E(x)$', (FA) will always allow E to be denoted by (2) (or any other λ-expression, even '$\lambda x\, \neg E(x)$'), as long as the 'equivalent' in (FA) is understood in the sense of a reflexive relation. The task of devising a more complete alternative to (FA) will be best left for another occasion, however.

I can here offer only a very tentative and provisional solution to the problem. In a first step we may introduce the notion of 'S-entailment', as follows:[53]

Definition 6.4. A formula φ *S-entails* a formula ψ iff the following three conditions are satisfied:

(i) φ entails ψ.

(ii) For any atomic term t: if ψ contains a free occurrence of t at subject-position, then so does φ.

(iii) If ψ contains an occurrence of a complex term at subject-position, then so does φ.

The 'entails' in condition (i) should be read in the sense of Definition 6.3 above. Let us further say that two formulas are *S-equivalent* just in case they S-entail each other. Under a conception of attributes whose coarse-grainedness is captured by the concept of S-equivalence thus defined, we are free to hold that E fails to be denoted by (2), since no reduction of (2) will have a matrix that is S-equivalent to '$E(x)$'.[54] Individuating attributes in accordance with the concept of S-equivalence thus solves at least part of our problem. Friends of a *very* fine-grained approach, e.g. along the lines of Menzel (1993), will of course wish to go even further, but for the purposes of this paper there will be no need to determine just how much more fine-grained our conception of attributes ought to be.

7 A First Condition

7.1 Quantifier-restriction

Returning now to the concept of intrinsicality, let us start by recalling the rough characterisation given in Section 4:

[53] The 'S' may be read as mnemonic for 'strict'.

[54] To produce this result, the second clause of Definition 6.4 is in fact not needed. However, without the second clause, an analogous problem would arise for properties like $\lambda x\,(E(x) \leftrightarrow (x \neq E))$ and $\lambda x\,(E(x) \leftrightarrow E(E))$. Even with all three clauses in place, there are still going to be 'halos' around logically simple *relations*. For example, if R is a simple dyadic relation, then the relations $\lambda x, y\,(R(x,y) \leftrightarrow R(x,x))$ and $\lambda x, y\,(R(x,y) \leftrightarrow R(y,x))$ will also count as simple. This may not be particularly palatable, but at least it does not appear to lead to the result that intuitively extrinsic attributes will on the present approach be misclassified as intrinsic.

(In$_0$) A κ-adic attribute A is *intrinsic* iff to instantiate A is, for any entities $\underbrace{x_1, x_2, \ldots}_{\kappa\text{-many}}$, purely a matter of which parts the x_i have, what the x_i and their parts are like, and how the x_i and their parts are related among each other, as opposed to what other entities there are, or how any of the x_i and their parts are related to any other entities.

In addition, recall the preliminary sketch of a formal analysis at which we arrived at the end of that Section:

(In$_1$) An attribute A is *intrinsic* iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes A relative to I and g, then . . . .

Given (In$_0$), it may at first be natural to think that at least *one* of the conditions that have to take the place of the ellipsis in (In$_1$) will require that any quantification in L be restricted in its range to parts of the respective relata of A.[55] However, this would go a little too far. Arguably it would be a mistake to require that *all* quantification in L be restricted in the mentioned way, for this would also extend to quantifier-occurrences that are contained in term-occurrences at subject-position, such as the '$\exists$' in '$\lambda x\, \mathrm{part}(\exists y\, \mathrm{horse}(y), x)$'. Relative to a suitable interpretation and variable-assignment, this expression denotes the (admittedly unusual) property of *having the state of affairs that there are horses as a part*. I submit that this property is intuitively intrinsic, despite the fact that the '$\exists$' in '$\lambda x\, \mathrm{part}(\exists y\, \mathrm{horse}(y), x)$' is *not* restricted in its range to parts of the property's respective bearer. If this assessment is correct, then our condition should be concerned with only those occurrences of quantified formulas that stand in L at *primary sentence-position*, i.e. at sentence-position but not contained in any term-occurrence at subject-position.[56] Taking this restriction into account, we can provisionally formalise the desired condition as follows:

(C1) For any formula φ and any variables $v_1, v_2, \ldots$: if L contains at primary sentence-position an occurrence o of $\ulcorner \exists v_1, v_2, \ldots\ \varphi \urcorner$, then, for any variable $v \in \{v_1, v_2, \ldots\}$,

[55]Cf., e.g., Skow's (2007: 151) analysis of intrinsicality for shape properties or Marshall's (2016a) definition of 'intrinsically about', reproduced in footnote 15 above. (In the case in which A is a property, the respective 'relatum' of A would more appropriately be called a 'bearer' of A.)

[56]For definitions of the terms 'subject-position' and 'sentence-position', see Section 5.1 above.

there exists a non-empty index set J together with a set of variables $\{u_i \mid i \in J\}$ and a set of terms $\{P_i \mid i \in J\}$ such that the following five conditions are satisfied:

(i) Each u_i is a λ-variable of L.

(ii) Each P_i denotes$_{I,g}$ a parthood relation.

(iii) φ entails $\ulcorner \bigvee_{i \in J} P_i(v, u_i) \urcorner$.

(iv) Each occurrence of each u_i that is free in o is bound by the initial occurrence of 'λ' in L.[57]

(v) Each occurrence of each P_i that is free in o is free *simpliciter*.

One of the things that are achieved by the combination of (In_1) and this condition—i.e. by the account of intrinsicality that results from letting (C1) take the place of the ellipsis in (In_1)—is that, whenever a given property P is denoted by '$\lambda x\, \exists y\, R(y, x)$', relative to some interpretation I and variable-assignment g, then P will not be classified as intrinsic, unless 'R' denotes$_{I,g}$ a parthood relation. So, e.g., the property of *being a father* will be correctly classified as extrinsic, because it will be denoted$_{I,g}$ by '$\lambda x\, \exists y\, R(y, x)$' if '$R$' denotes$_{I,g}$ the converse of the father-of relation, which is plausibly not a parthood relation.

In Section 9 we will see that (C1) has to be considerably strengthened. First, however, we will need to add two clarifications, and in Section 8 turn to the task of revising (In_1).

7.2 Parthood relations

At least two aspects of (C1) call for clarification. First, in talking in condition (ii) of "a parthood relation" rather than 'the' parthood relation, we are deliberately leaving room for mereological pluralism, different versions of which have been endorsed by Armstrong (1997), McDaniel (2004; 2009; 2014), and Fine (2010). For our purposes there will be no

[57]To see the need for this condition, it may be helpful to reflect, e.g., on the λ-expression '$\lambda x\, ((x = x) \wedge \exists x, y\, (P(y, x) \wedge Q(x) \wedge Q(y)))$', where '$P$' denotes$_{I,g}$ a parthood relation. (Note that the occurrences of 'x' within that of '$P(y, x) \wedge Q(x) \wedge Q(y)$' are bound by the occurrence of '$\exists$' rather than by the initial 'λ'.)

Similarly, to see the need for condition (v), consider the expression $L :=$ '$\lambda x\, \exists y, z\, ((\lambda u, w\, R(u, w, z))\, (y, x) \wedge Q(y))$', supposing that '$\lambda u, w\, R(u, w, z)$' denotes$_{I,g}$ a parthood relation. The problem in this latter case is that the second occurrence of 'z' in L is not free but rather bound by the occurrence of '$\exists$'.

need to take a stance on exactly what a parthood relation is, though we will impose the following two (relatively weak) constraints:[58]

(PR1) Any parthood relation is a dyadic relation whose instantiation by any entities x and y, in this order, is the state of affairs that x is (in the relevant sense) a part of y.

(PR2) The relation of identity is a parthood relation.

The first constraint is only intended to establish the adicity and 'direction' of parthood relations. The purpose of the second constraint is similarly straightforward: its aim is to allow our analysis to accommodate the intuition that 'identity properties' such as $\lambda x\,(x = \text{Socrates})$ are intrinsic, without having to treat identity separately from parthood.[59]

7.3 Entailment

The second aspect of (C1) that stands in need of clarification is the notion of *entailment* that is supposed to be operative in condition (iii). A natural choice may be the concept introduced in Definition 6.3 (p. 28 above). However, that concept is arguably not the one to be used in interpreting (C1), for its definition has the consequence that, e.g., the formula '$\text{part}(y, x)$' is entailed by any contradictory formula that contains free occurrences of 'part', 'x', and 'y'.[60,61] This would afford an easy way for a λ-expression like

[58]For some discussion of what it is to be a parthood relation, or for one thing to be a part of another, see, e.g., Skiles (2014: §2.2) and Yablo (2016). Skiles suggests that intrinsicality cannot be analysed in terms of parthood (at least not in such a way that the resulting analysis satisfies his criteria for being a *reductive* analysis), and explores ways of analysing parthood in terms of various notions in the vicinity of intrinsicality, such as that of a property's being 'intrinsic to' a given entity. Suppose that such an analysis is forthcoming and that parthood can be adequately analysed in terms of some intrinsicality-related notion X. If, in addition, it is possible to analyse X in terms of intrinsicality, then this would mean that parthood can indeed be analysed in terms of intrinsicality. But as far as I can see, this would not give us a reason to think that intrinsicality cannot *also* be analysed in terms of parthood.

[59]In taking $\lambda x\,(x = \text{Socrates})$ to be intuitively intrinsic, I am assuming that the identity relation is itself intrinsic. In the following, this assumption will be left implicit. The identity relation would arguably *not* be intrinsic if it were, e.g., analysable in terms of quantification over properties (along the lines of a proposal to the effect that for an entity x to be identical with an entity y is for x to have each one of its properties in common with y and *vice versa*), given that this quantification would not be restricted to the parts of the respective relata.

On a related note, I will be tacitly assuming that no unexpected 'sources of extrinsicality' are lurking in our primitive notions of instantiation, negation, conjunction, and existential quantification. For example, the negation of a state of affairs s should *not* be thought of (roughly *à la* Church [1951: 9]) as some state of affairs to the effect that, if s obtains, then *every* state of affairs obtains.

[60]For a formal definition of 'contradictory', see Definition 7.1 below.

[61]The constant 'part' will in the following serve as our default way of denoting parthood relations within formulas or λ-expressions. More particularly, I shall take it to be understood that—relative to whatever in-

'$\lambda x\, \exists y\, (\text{part}(x, x) \wedge (y \neq y))$' to pass the requirement expressed by (C1). But this is undesirable, since the λ-expression just mentioned denotes a property (namely, that of *having oneself as a part and being such that something is non-self-identical*) that, on a suitably fine-grained conception of attributes, will plausibly be regarded as extrinsic. We therefore need a concept of entailment under which it is not the case that any contradictory formula will (more or less indiscriminately) entail any other formula. More specifically, we need a concept of *relevant* entailment.

In contrast to the usual practice of relevance logic, we will here take a purely semantic approach, insofar as we will not be primarily concerned with proof theory and inference rules. At the same time, our approach will also differ from ordinary treatments of the *semantics* of relevance logic, since we will not be operating with model-theoretic structures or possible worlds. Instead, the general idea will be as follows: What separates relevant entailment in our sense from 'normal' entailment is a constraint requiring that one of the two relata (which are formulas) should not contain any subformulas that *are not needed to produce the entailment*; and according as this requirement (or some version of it) is imposed on the entailing or on the entailed formula, we obtain a corresponding concept of 'L-relevant' or of 'R-relevant' entailment.[62]

To formalise the notion of a subformula's being "not needed to produce the entailment", we will rely on the familiar concept of substitution, together with the following notions of tautology and contradiction:

Definition 7.1. A formula φ is *tautologous* (*contradictory*) iff φ is semantically well-formed and, for any interpretation I and variable-assignment g: if φ has a denotation relative to I and g, then φ denotes$_{I,g}$ an obtaining (non-obtaining) state of affairs.

As usual, tautologous formulas will also be referred to as 'tautologies' and contradictory formulas as 'contradictions'. On this basis we next define a concept of 'pruning':

terpretation is relevant in the respective context—'part' denotes an *intrinsic* parthood relation that is correctly paraphrasable by the English predicate 'is a part of'.

[62]The prefixes 'L' and 'R' stand for 'left' and 'right', respectively.

Definition 7.2. A formula φ' is a *pruning* of a formula φ iff either φ' itself is a tautology or contradiction or there exists a non-empty set S of term-occurrences in φ that satisfies the following three conditions:

(i) Each term-occurrence $o \in S$ stands at primary sentence-position.

(ii) φ' results from φ by replacing each member of S with some formula.[63]

(iii) Each term-occurrence $o \in S$ is in φ' replaced with a tautology or contradiction.

In other words, a formula φ' is a pruning of a formula φ iff φ' is a tautology or a contradiction or results from φ by replacing one or more subformulas (i.e. term-occurrences at primary sentence-position) with tautologies or contradictions.

On the basis of this concept, together with the concept of entailment introduced in Definition 6.3, we can now define two notions of 'strictly relevant' entailment:

Definition 7.3. A formula φ *strictly L-relevantly* entails (for short: *strictly L-entails*) a formula ψ iff φ entails ψ, and no non-contradictory pruning of φ entails ψ.

Definition 7.4. A formula φ *strictly R-relevantly* entails (for short: *strictly R-entails*) a formula ψ iff φ entails ψ but entails no non-tautologous pruning of ψ.

These concepts are rather narrow. For instance, a disjunction φ will not strictly L-entail any formula at all *unless* φ is a contradiction; and similarly, no formula strictly R-entails any non-tautologous conjunction.[64] The concept of strict L-entailment will turn out to be useful in Section 12 below, and a welcome feature of strict R-entailment, for our present purposes, lies in the fatct that no formula is strictly R-entailed by a contradiction. However, as a result of the fact that no formula strictly R-entails any non-tautologous conjunction, this notion is unfortunately too restrictive: as we will see in Section 10.2, we need a concept

[63] N.B.: The replacing formulas need not be distinct from the terms whose occurrences they are used to replace.

[64] To prove the first claim, let φ be any non-contradictory disjunction. It is clear that φ has then at least one non-contradictory disjunct. By replacing one or more of the other disjuncts with contradictions, one can thus obtain a non-contradictory pruning of φ that entails whatever φ entails. Hence, φ entails no formula that is not also entailed by some non-contradictory pruning of φ. So, as claimed, φ does not strictly L-entail any formula at all. The second claim can be proved along similar lines.

of entailment under which there is no general obstacle to a non-tautologous conjunction's being entailed by some formula.

To define such a concept, we will first need two more specific notions of pruning:

Definition 7.5. A formula φ' is a $\top$-*pruning* ($\bot$-*pruning*) of a formula φ iff either φ' itself is a tautology (contradiction) or there exists a non-empty set S of term-occurrences in φ that satisfies the following three conditions:

(i) Each term-occurrence $o \in S$ stands at primary sentence-position.

(ii) φ' results from φ by replacing each member of S with some formula.

(iii) For each term-occurrence $o \in S$: if o stands in the scope of an even (odd) number of occurrences of '$\neg$', then o is in φ' replaced with a tautology, and otherwise with a contradiction.[65]

For example, given that '$x = x$' a tautology, '$(x = x) \wedge q$' and '$(x = x) \vee q$' are $\top$-prunings of '$p \wedge q$' and '$p \vee q$', respectively. Care has to be taken with conditionals: e.g., '$(x = x) \to q$' is *not* a $\top$-pruning of '$p \to q$', since the latter is an abbreviation of '$\neg(p \wedge \neg q)$', so that the replaced occurrence of 'p' stands in an odd (rather than the required even) number of occurrences of '$\neg$'. In general, $\top$-pruning can be thought of as a weakening and $\bot$-pruning as a strengthening of formulas: the former effectively shortens conjunctions (by 'eliminating' conjuncts) and turns disjunctions into tautologies, while the latter effectively shortens disjunctions and turns conjunctions into contradictions.

Using the concepts of $\top$-pruning and $\bot$-pruning, we can now define two corresponding notions of *non*-strictly relevant entailment:

Definition 7.6. A formula φ *L-relevantly* entails (for short: *L-entails*) a formula ψ iff φ entails ψ, and no $\top$-pruning of φ entails ψ.

Definition 7.7. A formula φ *R-relevantly* entails (for short: *R-entails*) a formula ψ iff φ entails ψ but does not entail any $\bot$-pruning of ψ.

[65] We say that a term-occurrence stands in the scope of an even or odd number of occurrences of '$\neg$' according as the number of *all* occurrences of '$\neg$' in whose scope that term-occurrence stands is even or odd. If the term-occurrence stands in the scope of no occurrences of '$\neg$' at all, then that number is zero, and hence even.

The concept that we need for interpreting the 'entails' in condition (iii) of (C1) is that of *R-relevant* entailment, given that a contradictory formula does not R-entail any formula at all, and *a fortiori* does not R-entail any formula like 'part(y, x)'. Let this, then, be the sense of the 'entails' in (C1).[66]

8 Revising the Top Level

Our account of intrinsicality, as developed so far, consists of two parts: the 'top level', given by (In_1) (p. 32), and what might be called its 'core condition', given by (C1). Both parts still have to be extensively revised in order to yield an adequate account. The present Section will be devoted to revising (In_1), and in Sections 9 and 10 we will take up the task of improving (C1).

8.1 Trivialisation troubles

Suppose we adopt a conception of attributes whose coarse-grainedness is 'captured' by the concept of S-equivalence defined at the end of Section 6.3 above. Under such a conception, any property P is identical with $\lambda x\, (P(x) \wedge \exists y\, P(y))$. The property of *being a carbon atom*, for instance, will be identical with the property of *being a carbon atom and such that there exists at least one carbon atom*. This identification is hardly absurd, since for an entity *to be a carbon atom and such that there exists at least one carbon atom* imposes for all intents and purposes exactly the same requirements on the world (and on that entity) as for it to be a carbon atom. Meanwhile, however, the combination of (In_1) and (C1) classifies any property P as extrinsic if there exists some interpretation I and variable-assignment g such that P is denoted$_{I,g}$ by '$\lambda x\, (P(x) \wedge \exists y\, P(y))$'. The present conception of attributes thus leads to the result that any property whatsoever—and, by parallel considerations, any relation whatsoever—is classified as extrinsic. This consequence is obviously undesirable. One

[66]For some illustration of the two concepts just defined, let A and B be two non-empty sets of variables and/or non-logical constants, such that B is a proper subset of A. It can then be seen that the conjunction of the members of A will R-entail but not L-entail the conjunction of the members of B, whereas the disjunction of the members of B will L-entail but not R-entail the disjunction of the members of A. (In the special case in which B has only one member, 'the conjunction of the members of B' should be read as referring to that single member, and likewise for 'the disjunction of the members of B'.)

possible way to avoid it is to move towards a finer-grained conception of attributes under which it is no longer the case that any property P is identical with $\lambda x\,(P(x) \wedge \exists y\, P(y))$. But this would arguably be an overreaction, for it is possible to modify our account in a way that avoids the trivialising consequence just noted without requiring a finer-grained conception of attributes. All that is needed is that we enable the account to ignore any instances of what we have in Section 6.3 called 'redundant complexity'.

As a first step in this direction, we may insert into (In_1) an existential quantification over λ-expressions M whose matrices are *equivalent* to the matrix of the respective λ-expression L, where the relevant notion of equivalence is the one that captures the coarse-grainedness of our conception of attributes:[67]

(In_2) An attribute A is *intrinsic* iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes$_{I,g}$ A, then there exists some λ-expression M that has the same λ-variables as L and whose matrix is equivalent to that of L, such that ....

Having replaced (In_1) with (In_2), corresponding changes have to be made to our 'core condition' (C1): in particular, each occurrence of 'L' has to be replaced with 'M'. To form an intuitive gloss of (In_2), one may think of L as 'trying' to exhibit A as extrinsic, whereas M tries to reveal that L at best only *seems* to exhibit A as extrinsic.[68] Here the particular way in which L would seem to exhibit A as extrinsic is by denoting A while failing to satisfy the core condition, where this failure would be due to some complex-making feature of L (such as the quantification in '$\lambda x\,(P(x) \wedge \exists y\, P(y))$'); and the particular way in which M would reveal that A *merely* seems to exhibit A as extrinsic is by being equivalent to L while nonetheless satisfying the core condition. For in this way M would reveal that, if L does in fact fail to satisfy the core condition, it does so only due to some redundant complexity. What (In_2) accordingly tells us is, in a nutshell, that an attribute is intrinsic iff any given

[67]The steps that we will be taking in this subsection in order to deal with the trivialisation problem are essentially analogous to those I have taken in my (2016: §3.3) to solve a similar problem that arises in connection with the analysis of the concept of logical simplicity (as applied to attributes).

[68]In a typical game-semantic approach, one would here speak of two players A and B that respectively choose L and M with the goal of producing an assignment of entities to variables under which the core condition is either true or false. (In particular, A wants it to come out false, while B wants it to come out true. Along with L, A also chooses I and g.) For the sake of conciseness, I am in the main text talking of L and M as if they were themselves those players. (For more on game semantics, see Hodges [2013].)

λ-expression at best only *seems* to exhibit it as extrinsic.

We can now immediately see, however, that (In_2) does not yet address the full breadth of our problem. For suppose that I and g are an interpretation and a variable-assignment such that the variables 'P' and 'Q' both denote$_{I,g}$ a certain property P. Under the present conception of attributes (whose coarse-grainedness is captured by the concept of S-equivalence), the expression '$\lambda x\, P(x)$' will then of course be coreferential$_{I,g}$ not only with '$\lambda x\, (P(x) \wedge \exists y\, P(y))$' but also with '$\lambda x\, (P(x) \wedge \exists y\, Q(y))$'; yet clearly the matrix of this latter expression is not in the relevant sense equivalent to either '$P(x)$' or '$Q(x)$'. And so we are still faced with the undesirable consequence that any property P is under our account classified as extrinsic, simply because it can be denoted by '$\lambda x\, (P(x) \wedge \exists y\, Q(y))$'.

Fortunately, there is a natural way to address this issue: rather than to quantify over λ-expressions whose matrices are equivalent to the matrix of L itself, we should quantify, more broadly, over λ-expressions whose matrices are equivalent to the matrix of some *reduction* of L:[69]

(In_3) An attribute A is *intrinsic* iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes$_{I,g}$ A, then there exist an interpretation $I' \supseteq I$ and a variable-assignment $g' \supseteq g$ such that, for some λ-expression M that has the same λ-variables as L and whose matrix is equivalent to that of some reduction$_{I',g'}$ of L: . . . .

This revision again calls for some changes in (C1). As before, all occurrences of 'L' in (C1) have to be replaced with 'M', and in addition all occurrences of 'I' and 'g' have to be replaced with, respectively, 'I'' and 'g''. Let (C1$'$) be the result of these modifications. The combination of (In_3) and (C1$'$) avoids the threat of trivialisation, since it no longer has the consequence that any property that can be denoted by '$\lambda x\, (P(x) \wedge \exists y\, Q(y))$' is classified as extrinsic.[70] We now only have to add one last complication.

[69] The relevant concept of reduction has been introduced in Definition 6.1 (p. 24 above).

[70] The reason why (In_3) has to quantify over supersets of I and/or g is similar to the reason for the analogous feature of (FS) and (FA) in Section 6.2 above (cf. in particular footnote 43). Thus, suppose that P_1, P_2, and P_3 are three properties whose conjunction $P_1 \,\&\, P_2 \,\&\, P_3$ is intrinsic, and let Q_1, Q_2, and Q_3 be, respectively, the properties $P_1 \,\&\, P_2$, $P_2 \,\&\, P_3$, and $P_1 \,\&\, P_3$. Further, let I be some interpretation, let g be some variable-assignment

8.2 Redundant predications of parthood

Let us assume that the property of *being an electron* is intrinsic. The property of *having an electron as a part* will then be intrinsic as well,[71] and so will its negation, i.e. the property of *not having an electron as a part*. The disjunction of the latter two properties—i.e. *either having an electron as a part or not having an electron as a part*—can be denoted by the following:

$$\lambda x \left(\exists y \left(E(y) \wedge \mathrm{part}(y, x) \right) \vee \neg \exists y \left(E(y) \wedge \mathrm{part}(y, x) \right) \right), \tag{1}$$

where 'E' denotes *being an electron*. Notice that (1) satisfies the current 'core condition' of our account, viz., (C1′). But now, under the moderately coarse-grained conception of attributes that we have been working with so far, the property denoted by (1) turns out to be identical with that of *being such that every electron is an electron and, for any entity z, every part of z is a part of z*, or in symbols:

$$\lambda x \left((x = x) \wedge \forall y \left(E(y) \rightarrow E(y) \right) \wedge \forall y, z \left(\mathrm{part}(y, z) \rightarrow \mathrm{part}(y, z) \right) \right). \tag{2}$$

Plausibly, this latter property is extrinsic, for to instantiate it does not seem to be "purely a matter of which parts [something] has, what it and its parts are like, and how it and its parts are related among each other".[72] Yet, due to the fact that the matrix of (2) is S-equivalent to that of (1), our account threatens to classify this property as intrinsic.

We could accommodate the intuition that the property in question is extrinsic by moving once more to a finer-grained conception of attributes, whose coarse-grainedness is captured by a concept of equivalence that is strong enough to render the matrix of (2) inequivalent to that of (1). However, while the coreferentiality of (1) and (2) may perhaps be

that maps the variables 'Q_1', 'Q_2', and 'Q_3' to, respectively, Q_1, Q_2, and Q_3, and suppose that (as may be the case) no term at all denotes$_{I,g}$ P_1, P_2, or P_3. Under our moderately coarse-grained conception of attributes, the *ex hypothesi* intrinsic property $P_1 \,\&\, P_2 \,\&\, P_3$ will then be denoted$_{I,g}$ not only by '$\lambda x \left(Q_1(x) \wedge Q_2(x) \right)$' but also by '$\lambda x \left(Q_1(x) \wedge Q_2(x) \wedge \exists y\, Q_3(y) \right)$'. Let L be this latter expression. Since no term denotes$_{I,g}$ the properties P_1 and P_3 of which Q_3 is the conjunction, it can now easily turn out that there exists no λ-expression M that has the same λ-variables as L and whose matrix is equivalent to that of some reduction$_{I,g}$ of L such that (C1) (with 'L' replaced by 'M') is satisfied. To ensure that $P_1 \,\&\, P_2 \,\&\, P_3$ is classified as intrinsic, we thus have to consider reductions of L relative to *supersets* of either I or g.

[71] Cf. Sider (2007) on the "inheritance of intrinsicality".

[72] Cf. (In_0'), p. 12 above.

unexpected, I think it is far from absurd. And if we do accept that (1) and (2) denote one and the same property P, then the intuitively correct classification of this property seems to be the one that classifies it as extrinsic. For there is a tolerably good sense in which the two disjuncts in the matrix of (1) may be said to 'cancel each other out', leaving behind only a tautology not especially concerned with the parts of x. Accordingly, the better course of action seems to be to revise instead our account of intrinsicality, so as to ensure that properties like P are classified as extrinsic.[73]

The "tolerably good sense" in which the two disjuncts in (1)'s matrix cancel each other out may be captured by saying that those two disjuncts can be replaced by tautologies or contradictions in such a way as to yield a formula that is equivalent to the original matrix of (1), where the relevant notion of equivalence is the one that captures the coarse-grainedness of our conception of attributes (that is to say, the concept of S-equivalence).[74] So I propose to modify our account in such a way that it effectively ignores any λ-expression M whose matrix is 'simplifiable' in this sense. More formally, we can define this notion of simplifiability as follows:

Definition 8.1. A formula φ is *simplifiable* iff it is equivalent to some formula φ' that, for some non-empty set S of term-occurrences, satisfies the following three conditions:

(i) Each term-occurrence $o \in S$ stands in φ at primary sentence-position.

(ii) φ' results from φ by replacing each member of S with a tautology or contradiction.

(iii) No term-occurrence $o \in S$ is an occurrence of a tautology or contradiction.

Let us further say that a formula ψ is *unsimplifiably equivalent* to a formula φ just in case ψ is equivalent to φ and fails to be simplifiable. To put this concept to use, we may now replace the 'equivalent' in (In_3) with 'unsimplifiably equivalent':

---

[73] It is sometimes assumed (as in Sider [1996: 15] or Weatherson [2001: 370]) that all disjunctions of intrinsic properties must again be intrinsic. The case considered here would be a counter-example to that assumption. (Also cf. Parsons [2001: 14f.] and Marshall [2016b: 258].)

[74] In particular, the two occurrences of '$\exists y\,(E(y) \wedge \mathrm{part}(y,x))$' in (1) can be replaced with the tautologous matrix of (2) to produce a formula that is S-equivalent to the matrix of (1).

(In$_4$) An attribute A is *intrinsic* iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes$_{I,g}$ A, then there exist an interpretation $I' \supseteq I$ and a variable-assignment $g' \supseteq g$ such that, for some λ-expression M that has the same λ-variables as L and whose matrix is unsimplifiably equivalent to that of some reduction$_{I',g'}$ of L: . . . .

This solves our problem, since it renders irrelevant the fact that the matrix of (2) is equivalent to that of (1), given that the latter is simplifiable.

In the following, (In$_4$) will serve as the 'top level' of our account. The main task ahead is now the revision of (C1′).

9 Strengthening the Core Condition

9.1 Free term-occurrences

In constructing (C1′), we have been ignoring all those term-occurrences in M that are not bound by some occurrence of '$\exists$', and so the combination of (In$_4$) and (C1′) still allows such properties as *being a friend of Aristotle* to be misclassified as intrinsic (assuming no surprises in the metaphysics of the 'friend-of' relation). This is clearly a major defect. Hence, we have to either strengthen (C1′) or conjoin it with a further condition. In any case, we will have to impose a requirement roughly to the effect that any term that has in M a free occurrence at subject-position is in the matrix of M 'specified' to denote a part of one or more of the respective relata of A (i.e. of the attribute in question). This talk of specification may at present be somewhat opaque, but the basic idea is not much different from that of quantifier-restriction as formalised in (C1′). For, roughly in the same way in which M can contain restrictions on its quantified variables, it can also contain restrictions on the referents of the terms that have in it free occurrences at subject-position. More particularly, we can say that a term t is in M *specified to denote* a part of at least one of "the respective relata of A" just in case the matrix of M R-entails a predication of a parthood relation, or a disjunction of such predications, analogous to the formula mentioned in clause (iii) of

(C1′).[75] The relevant requirement, to be conjoined with (C1′), can thus be stated as follows:

(C2) For any term t and any free occurrence o of t in M: if o stands at subject-position, then there exists an index set J together with a set of variables $\{u_i \mid i \in J\}$ and a set of terms $\{P_i \mid i \in J\}$ such that the following four conditions are satisfied:

(i) Each u_i is a λ-variable of M.

(ii) Each P_i denotes$_{I',g'}$ a parthood relation.

(iii) The matrix of M R-entails $\ulcorner \bigvee_{i \in J} P_i(t, u_i) \urcorner$.

(iv) Each occurrence of each P_i that is free in the (first) occurrence of M's matrix in M is free *simpliciter*.[76]

But even the conjunction of (C1′) and (C2) leaves several kinds of term-occurrence out of account. Consider, e.g., the case of a term-occurrence that does not stand at subject-position. Certainly any free occurrence that stands in M at *primary predicate-position*—i.e. at predicate-position, but not contained in any occurrence that stands at subject-position—should be given a 'free pass', in the sense of not being required to satisfy the consequent of the main conditional of (C2). If M is, e.g., the λ-expression '$\lambda x\, E(x)$', it would make little sense to require that the matrix of M should R-entail $\ulcorner P(E, x) \urcorner$, where P is a term denoting$_{I',g'}$ a parthood relation. It is also arguable that we should give a free pass to any free formula-occurrence that stands in M at *primary sentence-position*, for otherwise we would be excluding such innocuous λ-expressions as '$\lambda x\, (\text{part}(e, x) \wedge E(e))$', in which the formula '$E(e)$' has a free occurrence at primary sentence-position.[77]

[75] The appeal to R-entailment, as opposed to entailment *simpliciter*, is motivated by considerations similar to those given in Section 7.3 above.

[76] This clause is meant to rule out that any P_i has in M an occurrence that contains a variable-occurrence bound by the initial 'λ' in M. (Cf. footnote 57 above.) The parenthetical qualification ("first") is needed for the case that M is infinitely deeply nested.

[77] If 'e' and 'E' respectively denote$_{I',g'}$ a certain entity x and the property of *being an electron*, then '$\lambda x\, (\text{part}(e, x) \wedge E(e))$' will denote$_{I',g'}$ a property that can be appropriately described as that of *having x as a part and being such that x is an electron*. Since we are assuming that the parthood relation denoted$_{I',g'}$ by 'part' is itself intrinsic (cf. footnote 61 above), this property will also plausibly be intrinsic (and it will be so regardless of whether x is in fact an electron). For a similar example, the property of *having Paris as a part and being such that Paris has the Kremlin as a part* likewise seems intrinsic, regardless of the fact that Paris does not actually have the Kremlin as a part.

What if the free occurrence in question stands at *non*-primary predicate-position, like the occurrence of 'horse' in '$\lambda x\,\text{part}(\exists x\,\text{horse}(x), x)$' or the occurrence of 'human' in '$\lambda x\,(x = \text{human}(\text{Socrates}))$'? A defender of the conjunction of (C1′) and (C2) might argue that this will be of no concern, given that any entity that instantiates $\lambda x\,\text{part}(\exists x\,\text{horse}(x), x)$—and thus has the state of affairs $\exists x\,\text{horse}(x)$ as a part—thereby also has any *constituent* of that state of affairs as a part, including the property of *being a horse*. Analogously for any entity that instantiates $\lambda x\,(x = \text{human}(\text{Socrates}))$. An opponent might object, however, that this solution requires a somewhat broad conception of parthood and that, for the purposes of analysing the concept of intrinsicality, it will be preferable not to take sides in the debate over what counts as a part of what. If the opponent is right, then free term-occurrences at non-primary predicate-position *should* be required to satisfy the consequent of the main conditional of (C2); and the same will hold (by similar considerations) for free term-occurrences at non-primary subject- or sentence-position. This line of reasoning cannot be easily dismissed. But on the other hand, a state of affairs like $\exists x\,\text{horse}(x)$ may well have constituents other than the property of *being a horse*, even though these constituents are not named by any terms occurring in '$\exists x\,\text{horse}(x)$' or any other formula denoting the state of affairs that there are horses; and if we follow the opponent's reasoning, then some analogous requirement should be put in place for those unnamed constituents as well, on pain of creating a double standard for named and unnamed constituents. Such a requirement could certainly be formulated, but it would make the overall account of intrinsicality even more ungainly (and by a considerable margin), whereas the theoretical benefit seems rather slim by comparison. For this reason, I propose that we continue to 'give a free pass' to term-occurrences that are properly contained in term-occurrences standing at subject-position. The only free term-occurrences to which we are so far *not* giving a free pass are accordingly those that stand at *primary subject-position*, i.e. at subject-position while not being properly contained in any term-occurrences standing at subject-position, like the occurrence of 'Socrates' in '$\lambda x\,R(x, \text{Socrates})$'.

Let us return to free term-occurrences at primary *sentence*-position. Two paragraphs back we said that free formula-occurrences at this kind of position should be given a free pass, but we have thereby not yet said anything about free occurrences at primary sentence-position of *constants and variables*, such as the occurrence of 's' in '$\lambda x\,((x = x) \wedge s)$'. Arguably, such occurrences should *not* be given a free pass. To see this, we have to begin by noting that, typically, a property like that of *being such that s obtains* (for some state of affairs s) is intuitively not intrinsic. After all, for something x to instantiate such a property will not be "purely a matter of which parts it [i.e. x] has, what it and its parts are like, and how it and its parts are related among each other", but also of whether s obtains; and s may very well have nothing to do with x or the parts of x. In most familiar cases, this can be readily enough brought out by choosing a λ-expression that exhibits, at least to some extent, the logical structure of s. For example, if s is the state of affairs that there are horses, then the property of *being such that there are horses* will, relative to a suitable interpretation I and variable-assignment g, be denoted by $L :=$ '$\lambda x\,((x = x) \wedge \exists y\,\text{horse}(y))$'. And unless the property of *being a horse* has a very surprising metaphysics (or we adopt an extremely coarse-grained conception of attributes and work with a correspondingly liberal concept of equivalence), there will then exist no interpretation $I' \supseteq I$ and variable-assignment $g' \supseteq g$ such that, "for some λ-expression M that has the same λ-variables as L and whose matrix is unsimplifiably equivalent to that of some reduction$_{I',g'}$ of L", (C1′) is satisfied. So the property of *being such that there are horses* will be duly classified as extrinsic, at least if (C1′) figures as one of the conditions that replace the ellipsis in (In$_4$). However, if s does not *have* a 'logical structure', i.e. if s is logically simple, the problem persists.[78] Let I and g be

[78] I have proposed an analysis of logical simplicity as applied to states of affairs in my (2016: 34f.):

> A state of affairs s is *logically simple* iff, for any interpretation I, variable-assignment g, and formula φ: if φ denotes$_{I,g}$ s, then there exist an interpretation $I' \supseteq I$, a variable-assignment $g' \supseteq g$, and an atomic term t such that φ has relative to I' and g' a reduction that is equivalent to $\ulcorner\neg\neg t\urcorner$.

This analysis presupposes that the underlying conception of states of affairs is coarse-grained enough to treat each state of affairs as identical with its own 'double negation'. If this is not granted, the '$\neg\neg t$' on the last line of the analysis may, e.g., be replaced with '$[t]$', where the bracket notation would be introduced as follows (cf. *op. cit.*, p. 35n.):

> For any term t, $\ulcorner[t]\urcorner$ is a formula, and for any interpretation I, variable-assignment g, and any term t, $\ulcorner[t]\urcorner$ has a denotation relative to I and g iff t denotes$_{I,g}$ a state of affairs; in which case $\ulcorner[t]\urcorner$ denotes$_{I,g}$ the same state of affairs.

This notation is also useful insofar as it fills a lacuna left by the thesis (CS) in Section 6.2 above. For instance, (CS) now has the consequence—provided that $\ulcorner[t]\urcorner$ is in the relevant sense equivalent to $\ulcorner t \vee t\urcorner$—that, for any interpretation I and variable-assignment g: if t denotes$_{I,g}$ a given state of affairs s, then so does $\ulcorner t \vee t\urcorner$.

any interpretation and variable-assignment, and let L be any λ-expression that denotes$_{I,g}$ the property in question, i.e. $\lambda x\,((x = x) \wedge s)$: given the logical simplicity of s, there will then always exist some interpretation $I' \supseteq I$, variable-assignment $g' \supseteq g$, and λ-expression M such that (i) M's matrix is equivalent to that of some reduction$_{I',g'}$ of L and (ii) M is identical with $\ulcorner \lambda v\,((v = v) \wedge t) \urcorner$, where v is the λ-variable of L and t some atomic term denoting$_{I',g'}$ s. Such a λ-expression neither (C1′) nor (C2) is able to rule out: not the former, since it is only concerned with bound variable-occurrences, and nor the latter, since it is only concerned with term-occurrences in subject-position. Hence, as long as s is logically simple, we are currently not classifying $\lambda x\,((x = x) \wedge s)$ as extrinsic. But intuitively, such a property is no less extrinsic than that of *being such that there are horses*.

The upshot of these considerations is that free occurrences of atomic terms at primary sentence-position should not be given a free pass, meaning that they *should* be required to satisfy the consequent of the main conditional of (C2). To summarise, we should give a free pass to all and only those free term-occurrences that (i) do not stand at primary subject-position and (ii) are not occurrences of atomic terms at primary sentence-position. Modifying (C2) accordingly, we obtain:

(C2′) For any term t and any free occurrence o of t in M: if o either stands at primary subject-position or is an occurrence of an atomic term at primary sentence-position, then there exists an index set J together with a set of variables $\{u_i \mid i \in J\}$ and a set of terms $\{P_i \mid i \in J\}$ such that the following four conditions are satisfied:

(i) Each u_i is a λ-variable of M.

(ii) Each P_i denotes$_{I',g'}$ a parthood relation.

(iii) The matrix of M R-entails $\ulcorner \bigvee_{i \in J} P_i(t, u_i) \urcorner$.

(iv) Each occurrence of each P_i that is free in the (first) occurrence of M's matrix in M is free *simpliciter*.

9.2 Bound occurrences of non-atomic terms

Another class of term-occurrences that we have so far been leaving out of account is that of *bound* occurrences of non-atomic terms, i.e. of formulas and λ-expressions. We have said above (p. 44) that free term-occurrences at primary predicate-position should be given a 'free pass' in the sense that they should not be required to satisfy the consequent of the main conditional of (C2). In the same paragraph, it was said that free formula-occurrences at primary sentence-position should similarly be given a free pass. Now, arguably the same goes for *bound* term-occurrences at primary predicate-position and for bound formula-occurrences at primary sentence-position. For instance, while the λ-expression '$\lambda x\, \exists y\, \text{part}(y, x)$' contains at primary sentence-position a bound occurrence of '$\text{part}(y, x)$', the denoted property will certainly be intrinsic as long as 'part' denotes an intrinsic parthood relation. By contrast, it can be argued that a bound occurrence of a non-atomic term should *not* be given a free pass if it stands at primary *subject*-position.

For example, consider the occurrences of '$P(x)$' and '$Q(x)$' in '$\lambda x\, (P(x) = Q(x))$'. Assuming that '$P$' and '$Q$' respectively denote two properties P and Q, the property denoted by this expression is that of *being such that one's instantiation of P is identical with one's instantiation of Q*, which is intuitively extrinsic. From this we may infer that bound term-occurrences at primary subject-position should not receive a free pass. On the other hand, when it comes to bound term-occurrences at non-primary subject-position, or at non-primary predicate- or sentence-position, we can apply the same considerations that we have above applied to *free* term-occurrences in these same sorts of position. As a result, it seems reasonable to conclude that a bound occurrence of a non-atomic term should receive a free pass if and only if it does not stand at primary subject-position.

To modify (C2′) accordingly, a natural first step would be to delete the 'free' on the first line of (C2′). The variable 'o' of the resulting version of (C2′) will then range also over bound *variable*-occurrences, which have so far been the responsibility of (C1′) (with the exception of those that are bound by an occurrence of 'λ'). But more importantly, if the revision is to work as intended, the consequent of the main conditional of (C2′) will have to be modified as well. For if the respective occurrence o of the respective term t

contains a variable that is bound not by the initial 'λ' but rather by (say) some quantifier-occurrence within M, then the requirement that the *matrix* of M should R-entail the disjunction $\ulcorner \bigvee_{i \in J} P_i(t, u_i) \urcorner$ will typically not yield the desired result. This becomes clear if one considers, e.g., the case in which M is the λ-expression '$\lambda x\, \exists y\, \mathrm{part}(P(y), x)$', where '$P$' denotes$_{I',g'}$ some property. The matrix of M—i.e. the formula '$\exists y\, \mathrm{part}(P(y), x)$'—will in this case not R-entail anything like '$\mathrm{part}(P(y), x)$'. But nor should it be required to.

In order to formulate the appropriate requirement, it will be useful to introduce first the notion of a term-occurrence's being 'governed' by an occurrence of a formula, which can be provisionally defined as follows:[79]

Definition 9.1. A term-occurrence o is *governed* by a formula-occurrence o' iff the following three conditions are satisfied:

(i) o' stands at primary sentence-position.

(ii) o is free in o'.

(iii) o is not free in any formula-occurrence that properly contains o'.

For example, consider the λ-expression '$\lambda x\, \neg\exists y\, \neg\mathrm{part}(y, x)$'—which, relative to a suitable interpretation and variable-assignment, denotes the property of (what Lewis has called) *loneliness*—and let o be the second occurrence of 'y' in this expression. We then have that o is governed by the embedding occurrence of '$\neg\mathrm{part}(y, x)$'. By contrast, due to clause (iii) of the present definition, o is *not* governed by the occurrence of '$\mathrm{part}(y, x)$', since the latter is properly contained in the occurrence of '$\neg\mathrm{part}(y, x)$'. As will become clearer shortly, this helps us ensure that properties such as loneliness are not classified as intrinsic.

The notion of governing allows us to formulate the needed revision of (C2′)'s main conditional. The basic idea can be put as follows: if the respective occurrence o satisfies certain conditions, then o has to be either (a) a variable-occurrence bound by the initial occurrence of 'λ' in M or (b) *governed* by some occurrence o' of a formula that R-entails a suitable formula $\ulcorner \bigvee_{i \in J} P_i(t, u_i) \urcorner$. To specify the "certain conditions", we can for the most part use

[79]For the final definition, see Section 10.1 below.

the antecedent of the main conditional of (C2′), except for one small modification: since the variable 'o' now ranges also over bound variable-occurrences, its range now includes the elements of lists of variable-occurrences that are immediately preceded by occurrences of '$\exists$' or 'λ' (like the first occurrences of 'y' and 'z' in '$\lambda x\, \exists y, z\, R(x, y, z)$'); and it patently makes no sense to require of *these* that they be governed by formula-occurrences that R-entail certain disjunctions—for one thing because any such variable-occurrence is bound in any formula-occurrence that contains it, and can therefore not be governed by any formula-occurrence. Variable-occurrences of this sort should thus be given a 'free pass'. To do so, we will let the variable 'o' in our revision of (C2′) range only over *referential* term-occurrences, in the following sense:

Definition 9.2. A term-occurrence is *referential* iff it is not an element of a list of one or more variable-occurrences that is immediately preceded by an occurrence of 'λ' or '$\exists$'.

In addition, it will be convenient to have at our disposal the notion of an 'argument-occurrence':

Definition 9.3. A term-occurrence is an *argument-occurrence* iff, for some λ-expression L, it is a referential variable-occurrence bound by the initial occurrence of 'λ' in L.

For example, of the four occurrences of 'x' in '$\lambda x\, (P(x) \wedge \exists x\, Q(x))$', only the second is an argument-occurrence: the first and third are not referential, and the third and fourth are not bound by the initial occurrence of 'λ'.[80]

With the help of the three concepts just introduced, we can now formulate a revised version of (C2′) that will supersede the conjunction of (C1′) and (C2′) as our core condition:

(CC_0) For any term t and any referential occurrence o of t in M: if o either stands at primary subject-position or is an occurrence of an atomic term at primary sentence-position,

[80]On a somewhat pedantic note (though this will help to prevent misunderstanding later on), consider the second occurrence of 'x' in the formula '$y = \lambda x\, R(x, z)$': this occurrence is *also* not an argument-occurrence, since it is not bound by the initial occurrence of 'λ' *in any λ-expression*. Instead it is only bound by the initial occurrence of 'λ' in an *occurrence* of a λ-expression. The only *expression* that the occurrences just mentioned are contained in is the formula '$y = \lambda x\, R(x, z)$'. More generally, any occurrence of an expression is contained in exactly one—typically larger—expression; and if a variable-occurrence is bound by some occurrence of 'λ', then both will be contained in the same expression. (See Appendix B for the relevant definitions.) It follows from this that, for any expression E and any occurrence o: if o is an argument-occurrence in E, then E is a λ-expression, and o is bound by the initial occurrence of 'λ' in E.

then o is either an argument-occurrence or governed by an occurrence o' of a formula φ such that, for some index set J together with a set of variables $\{u_i \mid i \in J\}$ and a set of terms $\{P_i \mid i \in J\}$, the following five conditions are satisfied:

(i) Each u_i is a λ-variable of M.

(ii) Each P_i denotes$_{I',g'}$ a parthood relation.

(iii) φ R-entails $\ulcorner \bigvee_{i \in J} P_i(t, u_i) \urcorner$.

(iv) Each occurrence of each u_i that is free in o' is an argument-occurrence.[81]

(v) Each occurrence of each P_i that is free in o' is free *simpliciter*.

For many applications, this already constitutes an adequate core condition, but there are still three ways in which it is too restrictive. First, it does not allow negations or disjunctions of what we may call 'inclusion attributes', such as the property of *having Paris as a part*, to be classified as intrinsic.[82] Second, it does not allow an attribute to be classified as intrinsic if an analysis of the latter will make reference to pluralities or sets, as in the case of *being composed of cells*. And third, it does not properly deal with 'chains' of parthood relations, as in *having an atom as a part that has a neutron as a part*. In the next Section, we will deal with each of these issues in turn.[83]

[81] From this it follows that each occurrence of each u_i free in o' is bound by the initial occurrence of 'λ' in M. To see this, note that o' governs o, which means that o is free in o', hence contained in o', and therefore contained in the same expression as o'. But the expression that o is contained in is M. So o' is also contained in M, and the same goes for any occurrence of any u_i that is free in o' (and hence contained in o'). To say that such an occurrence is an *argument*-occurrence then means that it is bound by the initial occurrence of 'λ' in M. (Cf. the previous footnote.)

[82] In Section 2.2 above, properties of this kind have been referred to as 'haecceitistic inclusion properties', but for the sake of brevity I will here simply call them 'inclusion properties'. An example of an inclusion *relation* would be $\lambda x, y\,(\text{part}(\text{Paris}, x) \wedge \text{part}(\text{Moscow}, y))$.

[83] Another limitation of (CC$_0$), though only a minor one, lies in the fact that it effectively assumes that *β-reductions* do not affect the denotation of a λ-expression. (A β-reduction is the syntactic transformation by which, e.g., '$\lambda x\,(\lambda y\, R(y, a))(x)$' can be transformed into '$\lambda x\, R(x, a)$'. See Dorr [2016b: §§5f.] for relevant discussion.) If this assumption were false, it might happen that the weakest concept of equivalence that renders (CA) true is still strong enough that, e.g., the matrix of

$$\lambda x\, \exists y\, (\text{part}(y, x) \wedge (\lambda z\,(z = z))(y)) \tag{$*$}$$

is not equivalent to '$\exists y\,(\text{part}(y, x) \wedge (y = y))$'. In this way it would turn out that the property denoted by $(*)$ is by our account classified as extrinsic, since the second and third occurrences of 'z' in $(*)$ are not governed by any formula that R-entails $\ulcorner \text{part}(z, x) \urcorner$. Yet the denoted property is plausibly *in*trinsic. The assumption in question could be dispensed with if we adopted a more complicated definition of governing, but it does not seem to me that much would be gained by this.

10 Three Refinements

10.1 Negations and disjunctions of inclusion attributes

Relative to a suitable interpretation I and variable-assignment g, the property of *not having Paris as a part* is denoted by the λ-expression '$\lambda x\, \neg\text{part}(\text{Paris}, x)$'. Let L be this expression. Intuitively, the property it denotes is intrinsic, for to instantiate it appears to be purely a matter of "which parts [something] has".[84] Under our present account of intrinsicality, which is obtained by replacing the ellipsis in (In_4) with (CC_0), there should then exist an interpretation $I' \supseteq I$, a variable-assignment $g' \supseteq g$, and a λ-expression M such that: (i) M has the same λ-variables as L; (ii) the matrix of M is unsimplifiably equivalent to that of some reduction$_{I',g'}$ of L; and (iii) the 'core condition' (CC_0) is satisfied. As above, let us suppose that the relevant concept of equivalence is that of S-equivalence.

Suppose now further that $M = L$, and let o be the occurrence of 'Paris' in M: clearly o then satisfies the antecedent of the main conditional of (CC_0). But the consequent is *not* satisfied, since the only formula-occurrence governing o is the occurrence of '$\neg\text{part}(\text{Paris}, x)$'. And this outcome can apparently not be avoided by an alternative choice of M. It thus turns out that our account misclassifies *not having Paris as a part* as extrinsic. By analogous considerations, it can be seen that our account also misclassifies the property of *having either Paris or Rome as a part*. For suppose that M is the λ-expression '$\lambda x\, \neg(\neg\text{part}(\text{Paris}, x) \wedge \neg\text{part}(\text{Rome}, x))$', and let o again be the occurrence of 'Paris' in M. Then, as in the previous example, we find that o satisfies the antecedent of the main conditional of (CC_0) but not its consequent. Yet intuitively, *having either Paris or Rome as a part* should come out intrinsic, given that for something to instantiate it is "purely a matter of which parts it has".

A straightforward way to repair this defect is to revise our above definition of 'is governed by' (p. 49) by exempting from its third condition all those cases in which o is free *simpliciter*:

Definition 10.1. A term-occurrence o is *governed* by a formula-occurrence o' iff the following three conditions are satisfied:

[84] Cf. (In_0'), p. 12 above.

(i) o' stands at primary sentence-position.

(ii) o is free in o'.

(iii) If o is not free *simpliciter*, then o is not free in any formula-occurrence that properly contains o'.

Thus, if o is free *simpliciter*, then it will be governed by any formula-occurrence o' that satisfies conditions (i) and (ii). The occurrence of 'Paris' in '$\lambda x\, \neg\text{part}(\text{Paris}, x)$', for instance, is now governed by the occurrence of '$\text{part}(\text{Paris}, x)$' *as well as* by that of '$\neg\text{part}(\text{Paris}, x)$'.

Unfortunately, this weakening of the concept of governing has the unwelcome side effect that 'distinctness properties', such as *being distinct from Paris*, are also classified as intrinsic. To see this, suppose that M is the λ-expression '$\lambda x\, (\text{Paris} \neq x)$'—or, without use of abbreviatory devices: '$\lambda x\, \neg\mathtt{I}(\text{Paris}, x)$'—and let o once more be the occurrence of 'Paris' in M. This occurrence will then be governed by that of '$\mathtt{I}(\text{Paris}, x)$', and, by what has been said in Section 7.2 above, the identity relation denoted$_{I',g'}$ by '$\mathtt{I}$' is a parthood relation. Hence, o satisfies the consequent of the main conditional of (CC_0), and in this way our account threatens to classify distinctness properties as intrinsic. But intuitively, such properties are *ex*trinsic.[85] In order to accommodate this intuition, we will need to strengthen (CC_0) by adding a sixth clause:

(vi) If o is free *simpliciter* and at least one P_i denotes$_{I',g'}$ the identity relation, then o' does not stand in the scope of an odd number of occurrences of '$\neg$'.

This modification ensures that distinctness properties are classified as extrinsic. Let us refer to the new core condition that results from adding (vi) to (CC_0) as '(CC_1)'.

[85]At first sight this assessment may seem to conform with our orienting characterisations (In_0) and (In_0'): for something to be distinct from Paris is plausibly *not* purely a matter of "which parts it has, what it and its parts are like, and how it and its parts are related among each other". But on reflection the matter is not so clear. Since parthood is reflexive, it might be argued that for something x to be distinct from Paris is very much "purely a matter of which parts it has [etc.]", since it is purely a matter of whether that particular part of *x that is x itself* is identical with Paris. And after all, isn't this just how we have to argue in order to get the result that 'identity properties', such as *being Socrates*, should count as intrinsic? It seems, then, that the differential classification of distinctness properties as extrinsic and of identity properties as *in*trinsic is not borne out by our orienting characterisations. Nonetheless I think that it is nearly non-negotiable.

10.2 Sets of parts

The first five numbered clauses of (CC_1) roughly amount to the requirement that φ should R-entail a formula to the effect that the respective referent of t (in the context of its occurrence o) is a part of at least one of the respective relata of A. But this requirement—and in particular condition (iii)—is a good deal too strong.

To see why, let P be the plausibly intrinsic property of *being composed of cells*. One approach towards analysing P would be to use the machinery of plural logic. Another approach, which I shall follow here, is to employ standard set-theoretic machinery and to make use of sets as 'auxiliary entities'. On this approach, we will first of all identify P with the property of *being a fusion of some set of cells*. A fusion of a set S may here be understood as the unique entity that, first, has each member of S as a part and, second, is such that each of its parts has a part in common with at least one member of S.[86] Relative to an interpretation that maps the constants 'set', 'el', and 'cell' to, respectively, the property of *being a set*, the relation of set-membership, and the property of *being a cell*, our property P can then be denoted by

$$\begin{aligned}\lambda x\, \exists S \Big(&\mathrm{set}(S) \wedge \forall y \left(\mathrm{el}(y,S) \rightarrow (\mathrm{cell}(y) \wedge \mathrm{part}(y,x))\right) \wedge \\ &\forall y \left(\mathrm{part}(y,x) \rightarrow \exists z, w \left(\mathrm{el}(z,S) \wedge \mathrm{part}(w,z) \wedge \mathrm{part}(w,y)\right)\right)\Big).\end{aligned} \tag{1}$$

Suppose that M is this λ-expression, I' an interpretation of the sort just described, and g' some variable-assignment. The thing to note, then, is that the referential occurrences of 'S', the referential occurrences of 'y' on the first line, and those of 'z' and 'w', all satisfy the antecedent of (CC_1)'s main conditional (since they all stand at primary subject-position), but fail to satisfy the consequent. This suggests that, in order to ensure that our account classi-

[86]Cf. Hovda (2009). I deliberately restrict myself to talk of fusions of *sets* because, given our other theoretical commitments, it would quickly lead to paradox if we allowed that any class of entities, no matter how large, has a fusion. (Cf. McCarthy [2015].)

As an alternative to the set-theoretic approach pursued here, it might be suggested that we make use of infinitely long disjunctions. Thus P might be identified with the property of *being an entity x such that (a) for some cell c_1, x has c_1 as a part and each part of x has a part in common with c_1, or (b) for some cells c_1 and c_2, x has c_1 and c_2 as a part and each part of x has a part in common with c_1 or c_2, or …* . At least one problem with this approach lies in the fact that, in order to cover every set-sized cardinality, the disjunction would have to be of more than set-sized length.

fies properties like P as intrinsic, (CC_1) has to be weakened rather substantially, namely in two respects.

First, we have to allow referential occurrences o of terms t in M to count as 'redeemed' if they are governed by an occurrence of a formula to the effect that t is a *set of parts* of one or more of the respective relata of A. (To say that a given occurrence 'counts as redeemed' means here that it does not prevent the classification of A as intrinsic.) But this is still too narrow: we should also allow such occurrences to count as redeemed if they are governed by occurrences of formulas to the effect that their respective referents are sets of *sets of* parts of one or more of A's relata, and so on for even more deeply nested sets. One way to see this is by considering the property—call it 'Q'—of *having as many green parts as red parts*. Assuming for the sake of example that the properties of *being red* and *being green* are intrinsic, Q should be classified as intrinsic, too; but on a standard set-theoretic analysis it will be denotable by a λ-expression '$\lambda x\, \exists B\, (\ldots)$' in whose matrix the referential occurrences of 'B' (for 'bijection') are governed by an occurrence of a formula to the effect that B is a set of *ordered pairs* of parts of x (where x is the respective bearer of the property).[87] Moreover, under the standard conception of ordered pairs (viz., Kuratowski's), each such pair is a nested set $\{\{a\}, \{a, b\}\}$, where a and b are the two coordinates of the pair in question. So we are here dealing with quantification over sets of sets *of sets* of parts of x. The analysis of other intrinsic properties may require quantification over even more deeply nested sets.[88]

---

[87]More specifically, relative to an interpretation that maps the constant 'pair' to the triadic relation that holds among entities x, y, and z (in this order) iff x is the ordered pair $\langle y, z\rangle$, Q will be denotable by the following λ-expression:

$$\begin{aligned}\lambda x\, \exists B\, \Big(&\mathrm{set}(B) \wedge \forall y\, (\mathrm{el}(y, B) \rightarrow \exists z, w\, (\mathrm{pair}(y, z, w) \wedge \mathrm{part}(z, x) \wedge \mathrm{part}(w, x) \wedge \mathrm{red}(z) \wedge \mathrm{green}(w))) \wedge \\ &\forall y\, ((\mathrm{part}(y, x) \wedge (\mathrm{red}(y) \vee \mathrm{green}(y))) \rightarrow \exists z, w\, (\mathrm{el}(z, B) \wedge (\mathrm{pair}(z, y, w) \vee \mathrm{pair}(z, w, y)))) \wedge \\ &\forall y, z, w, u, r, s\, ((\mathrm{el}(y, B) \wedge \mathrm{el}(z, B) \wedge \mathrm{pair}(y, w, u) \wedge \mathrm{pair}(z, r, s)) \rightarrow ((w = r) \leftrightarrow (u = s)))\Big).\end{aligned}$$

Here the first line 'says' that B is a set and each member of B is a pair of (in this order) some red and some green part of x; the second line says that each red or green part of x is either the first or the second coordinate of some ordered pair that is a member of B; and the third line says that B is a bijection.

[88]An objector might at this point raise a Benacerraf-style worry about there being no uniquely adequate conception of ordered pair. (Cf. Benacerraf [1965].) But it is not clear to me what such an objection would aim to show. If the point were that there exists no uniquely adequate set-theoretic analysis of properties like Q, it would be innocuous. And if the point were that, under the set-theoretic approach, we would have to distinguish between multiple properties that all equally deserve to be described as a 'property of *having as many green parts as red parts*', there would still not (as far as I can see) be a serious problem.

It is interesting to ask whether a term-occurrence that is governed by a suitable predication of 'set-of-parts-hood' may also have to count as redeemed if it is free *simpliciter*. An example of such an occurrence is that of the constant 'Paris' in the λ-expression

$$\lambda x\,\big(\mathrm{set}(\mathrm{Paris}) \wedge \forall y\,(\mathrm{el}(y,\mathrm{Paris}) \to \mathrm{part}(y,x))\big), \tag{2}$$

which denotes the somewhat bizarre property of *having Paris as a set of one's parts*. Fairly clearly, for something to instantiate this property is not "purely a matter of which parts it has [etc.]", and so we would seem to have good reason to regard it as extrinsic. But appearances may be deceiving. If Paris *were* simply a set of some entities—say, a set of three entities a, b, and c—then it would presumably be less strange to think of *having Paris as a set of one's parts* as intrinsic: to instantiate it would be purely a matter of having a, b, and c as one's parts. But now, what sort of thing Paris actually is (e.g. whether it is a set or a non-set) should make no difference with respect to whether *having Paris as a set of one's parts* is intrinsic. For consider: for any entity x, the property of *having x as an atomic part* is plausibly intrinsic, even if x is not actually atomic. For another example, just as *having Mars as a part and being such that Mars is made of rock* is intrinsic, the property of *having Paris as a part and being such that Paris is made of rock* should likewise count as intrinsic (never mind whether Paris is actually made of rock). Analogously, then, if *having* $\{a,b,c\}$ *as a set of one's parts* counts as intrinsic, then so should the property of *having Paris as a set of one's parts*. I would accordingly propose to treat this property as intrinsic.

The second respect in which (CC_1) has to be weakened concerns the 'chaining' of specifications. For example, the referential occurrences of 'y' on the first line of (1) ought to count as redeemed (in the sense of not preventing the classification of P as intrinsic) in virtue of the fact that they are governed by an occurrence o of a formula to the effect that y is a member of S, where the occurrence of 'S' in o is in turn governed by an occurrence of a formula to the effect that S is a set of parts of x. In contrast to the case of *having an atom as a part that has a neutron as a part* (briefly mentioned on p. 51 above), the present chain consists not of two predications of parthood, but rather of a predication of set-of-parts-hood followed (in 'descending' order) by a predication of set-membership. Let us say that a chain of predications is *redeeming* iff, as a result of its presence, the term-occurrence at the chain's 'lower' end should count as redeemed (in the sense specified above).

All redeeming predication chains have in common that the term-occurrences at their respective 'upper' ends are *argument-occurrences*, i.e. referential variable-occurrences bound by the initial 'λ' of the containing λ-expression.[89] So, e.g., a chain whose upper end is a variable-occurrence bound by a quantifier will (for obvious reasons) not count as redeeming.[90] As for the types of predication that a redeeming chain can consist of, we have so far encountered (i) predications of parthood, (ii) predications of set-membership, and (iii) predications of 'set-of-parts-hood'. By what has been said above, this last term should here be understood as applying not only to set-of-parts-hood properly so called, but also to all its nested variants, such as set-of-*sets-of*-parts-hood and even set-of-(parts-*or-sets-of-parts*)-hood, and so on.

It is somewhat controversial whether a redeeming chain can also consist of *nothing but* predications of set-membership. We can here remain neutral on the question of whether set-membership is a parthood relation. Independently of this question I propose to treat set-membership as 'on a par' with parthood as far as the intrinsic/extrinsic distinction is concerned. One reason for this move lies in the fact that, from a purely intuitive point of view, it seems in fact quite plausible (at least to my mind) to say that properties like *having Paris as a member* are intrinsic.[91] The second reason lies in the fact that this move will lead to a simpler account of intrinsicality (without any obvious loss in extensional adequacy). To be sure, on a strict reading of our orienting characterisation (In_0'), it is not the case that for something to instantiate *having Paris as a member* is "purely a matter of which parts it has [etc.]", unless set-membership is a parthood relation. At the same time, however, I think that it would not be wildly counter-intuitive to broaden that characterisation so as to treat set-membership in a way that is largely analogous to the way in which it treats parthood

[89] See Definition 9.3 above.

[90] An example of such a chain would be the single predication of parthood that, in '$\lambda x\, \exists y\, \mathrm{part}(\mathrm{Paris}, y)$', connects the occurrence of 'Paris' to the subsequent occurrence of 'y'.

[91] This is *pace* Marshall (2015: 14), who classifies the property of *having a member* as extrinsic.

relations.

The above considerations have led us to a view on which single predications of parthood, set-membership, and set-of-parts-hood all make for redeeming (albeit short) chains. If this is correct, it is hard to see why longer chains made up of any combination of such predications should not also count as redeeming. This means in particular that, in the case of a predication of set-of-parts-hood near the top of a given chain, we need not require that the depth of the nesting should be matched by the number of subsequent (i.e. lower) predications of set-membership. For example, the property of *having S as a set of one's parts and being such that Paris is a member of a member of a member of S*, for some entity S, will still count as an (admittedly unusual) intrinsic property. Furthermore, given that set-membership is for the purposes of the intrinsic/extrinsic distinction treated as on a par with parthood, it stands to reason that predications of 'set-of-parts-*or-members*-hood' (where this label again also applies to nested variants of what it applies to *stricto sensu*) should be treated in the same way as predications of set-of-parts-hood.

So much for the two respects in which (CC_1) has to be weakened. In order to implement the necessary changes, we will first address the problem of formally characterising the class of all predications that are to the effect that the referent of a given term t is a part, or a member, or a set of parts or members (etc.) of one or more other entities. In the next subsection, we will then turn to chains consisting of multiple such predications.

The natural approach to a formal characterisation of the mentioned class of predications is to proceed in a recursive manner, starting with simple predications of parthood and set-membership, as well as disjunctions of such predications. Predications of set-of-parts-or-members-hood will then be added in the recursion step. As parameters, we will use an interpretation I, a variable-assignment g, a term t, and two sets of terms T and U. (By way of orientation, it may help to think of t as the term of which o is an occurrence, of T as the set of M's λ-variables, and of U as a set of terms of which each member denotes the property of *being a set*, the relation of set-membership, or a parthood relation.) The definition runs as follows:

Definition 10.2. For any interpretation I, variable-assignment g, term t, and any sets of terms T and U, the following holds:

(i) $\mathscr{D}_0(I, g, t, T, U)$ is the class of all formulas φ such that there exists a non-empty set of formulas Δ satisfying the following two conditions:

1. φ is equivalent to $\ulcorner \bigvee_{\delta \in \Delta} \delta \urcorner$.[92]
2. For each $\delta \in \Delta$, there exist terms $P \in U$ and $u \in T$ such that P denotes$_{I,g}$ either a parthood relation or set-membership, and $\delta = \ulcorner P(t, u) \urcorner$.

(ii) For any ordinal $\alpha > 0$, $\mathscr{D}_\alpha(I, g, t, T, U)$ is the class of all formulas φ such that, for some terms $S, E \in U$, some variable v, and some non-empty set of formulas Δ, the following four conditions are satisfied:

1. φ is equivalent to $\ulcorner S(t) \wedge \forall v \left(E(v, t) \rightarrow \bigvee_{\delta \in \Delta} \delta \right) \urcorner$.
2. S and E respectively denote$_{I,g}$ the property of *being a set* and set-membership.
3. For any $\delta \in \Delta$, there exists an ordinal $\beta < \alpha$ such that $\delta \in \mathscr{D}_\beta(I, g, v, T, U)$.
4. The variable v does not occur free in t or any member of $T \cup U$.

The class of all formulas that express (relative to a given interpretation I and variable-assignment g) the requirement that the referent of a given term t be a part, or a member, or a set of parts or members (etc.) of one or more entities $x_1, x_2, \ldots$—where the x_i are the respective referents of some of the members of a given set of terms T—can now be formally characterised as the class of all formulas φ such that, for some ordinal α and some set of terms U, φ is a member of $\mathscr{D}_\alpha(I, g, t, T, U)$. Equipped with this notation, we can take a first step towards the desired weakening of (CC$_1$):

(CC$_2$) For any term t and any referential occurrence o of t in M: if o either stands at primary subject-position or is an occurrence of an atomic term at primary sentence-position, then o is either an argument-occurrence or governed by an occurrence o' of a formula φ such that, for some set V of λ-variables of M, some set of terms U, and some set of formulas Δ, the following five conditions are satisfied:

[92]The 'equivalent' should here be understood as mutual entailment in the sense of Definition 6.3.

(i) φ R-entails $\ulcorner \bigvee_{\delta \in \Delta} \delta \urcorner$.

(ii) For any $\delta \in \Delta$, there exists an ordinal α such that δ is a member of $\mathscr{D}_\alpha(I', g', t, V, U)$.

(iii) For any variable $v \in V$: any occurrence of v free in o' is an argument-occurrence.

(iv) For any term $u \in U$: any occurrence of u free in o' is free *simpliciter*.

(v) If o is free *simpliciter* and there exist a formula $\delta \in \Delta$ and terms P and u such that δ is equivalent to $\ulcorner P(t, u) \urcorner$ and P denotes$_{I',g'}$ the identity relation, then o' does not stand in the scope of an odd number of occurrences of '$\neg$'.

Here the clauses (i), (iii), (iv), and (v) respectively correspond—with some adjustments due to the switch from talk of an index set J to talk about a set of formulas Δ—to the clauses (iii), (iv), (v), and (vi) of (CC_1).

10.3 Chains of Predications

(CC_2) still suffers from the defect that it ignores redeeming chains consisting of more than one predication, as a result of which it fails to accommodate not only the intrinsicality of *having an atom as a part that has a neutron as a part*, but also that of *being composed of cells*. To remedy this, we will have to replace the entire consequent of (CC_2)'s main conditional with a requirement to the effect that the respective term-occurrence o sits at the bottom of a redeeming chain. More specifically, the requirement will be to the effect that o should be directly or indirectly linked to one or more argument-occurrences via a possibly branching graph (technically, a *rooted directed graph*), which I shall call an '(I, g)-graph'. Before we set about defining this concept, I should note two complications that arise in connection with redeeming chains of more than one predication.

First, while negated parthood predications, as in the case of *not having Paris as a part*, do not in general impugn an attribute's intrinsicality,[93] a chain of *two or more* negated parthood predications arguably does, since a property such as *neither having Paris as a part nor being such that Paris has the Kremlin as a part* (or in symbols: $\lambda x\,(\neg\text{part}(P, x) \wedge \neg\text{part}(K, P))$) seems clearly extrinsic. Further, the property of *not having Paris as a part but being such that Paris*

[93] Cf. Section 10.1 above.

has the Kremlin as a part seems likewise extrinsic. To accommodate these observations, I propose that a chain containing a negated predication of parthood (where the respective non-part is the referent of a free term-occurrence) should count as redeeming *only* if that predication stands at the very bottom of the chain. Negated predications of set-membership and set-of-parts-or-members-hood should arguably be treated in the same way.

Second, we should take into account that, while *being distinct from Paris* is extrinsic, the slightly more complex property of *having a part that is distinct from Paris* appears to be intuitively *in*trinsic, since for something to instantiate it is (as may be argued) "purely a matter of which parts it has".[94] If this assessment is correct, then a chain containing a negated predication of identity (where one of the two relata is the referent of a free term-occurrence) should count as redeeming if and only if that predication satisfies two conditions: it has to stand at the bottom of the chain (by what has been said in the previous paragraph, given that identity is a parthood relation), and the chain has to contain a least two links. The requirements for negated predications of identity are thus somewhat stronger than those for negated predications of other parthood relations. As far as I can see, no similar exceptions hold for any other parthood relation, nor for the relations of set-membership and set-of-parts-or-members-hood.

In the following definition of '(I, g)-graph', the conclusions of the previous two paragraphs are incorporated in the form of condition (ii.5). The main idea is that each node of an (I, g)-graph has to be either an argument-occurrence or to satisfy what is essentially an analogue of the consequent of (CC_2)'s main conditional. In addition, some machinery is needed to create the links, or 'edges', that lead from the graph's 'root', possibly via intermediate nodes, to one or more argument-occurrences.

Definition 10.3. For any interpretation I and variable-assignment g, an (I, g)*-graph* is an ordered triple $\langle N, E, r \rangle$, where N and E are sets (of the graph's 'nodes' and 'edges', respectively) and r (the graph's 'root') is a term-occurrence such that, for some λ-expression L, some function Γ from N to the set of formula-occurrences in L, and some function Φ from

[94] We could not rightly be said to know which parts a given entity x has if we did not know, for any given entity y, whether x has y as a part. So, if we know which parts x has, and x does not, e.g., have Paris as a part, we will know this. But then we will also know that x has at least one part that is distinct from Paris.

N to the class of sets of terms, the following three conditions are satisfied:

(i) Each member of N is a term-occurrence in L.

(ii) For any term t: any occurrence o of t with $o \in N$ is either an argument-occurrence or governed by an occurrence o' of a formula φ such that, for some set of terms U, some set of formulas Δ, and some function F from Δ to the class of sets of terms, the following five conditions are satisfied:

1. φ R-entails $\ulcorner \bigvee_{\delta \in \Delta} \delta \urcorner$.
2. For any $\delta \in \Delta$, $F(\delta)$ is minimal under the condition that, for some ordinal α, δ is a member of $\mathscr{D}_\alpha(I, g, t, F(\delta), U)$.[95]
3. $\Gamma(o) = o'$ and $\Phi(o) = \bigcup_{\delta \in \Delta} F(\delta)$.
4. For any term $u \in U$: any occurrence of u free in o' is free *simpliciter*.
5. If o is free *simpliciter* and o' stands in the scope of an odd number of occurrences of '$\neg$', then $o = r$ and, for any formula $\delta \in \Delta$ and any terms τ and u: if δ is equivalent to $\ulcorner \tau(t, u) \urcorner$ and τ denotes$_{I,g}$ the identity relation, then u is *not* a variable of which each occurrence free in o' is an argument-occurrence.

(iii) N and E are minimal under the following two conditions:

1. $r \in N$.
2. For any $o \in N$, any term $t \in \Phi(o)$, and any occurrence o' of t free in $\Gamma(o)$, the following holds: $o' \in N$ and $\langle o, o' \rangle \in E$.

In the simplest case, an (I, g)-graph has only a single node, viz., the graph's root (which then has to be an argument-occurrence). In a more complex case, the graph's nodes form a single chain without branches. For example, if g is some variable-assignment and I an interpretation that maps the constants 'set' and 'el' to, respectively, sethood and set-membership, then the second occurrence of 'y' in the λ-expression (1) on p. 54 above is the

[95] $F(\delta)$ is here minimised because, for each ordinal α, set of terms T, and formula $\delta \in \Delta \cap \mathscr{D}_\alpha(I, g, t, T, U)$, we want only those terms $t \in T$ that are actually mentioned in δ. (This restriction is needed for the proper functioning of condition (iii.2).)

root of an (I, g)-graph with exactly two edges, leading from the root to the third occurrence of 'S' and from there to the second occurrence of 'x'. A simple example of a *branching* (I, g)-graph can be found in '$\lambda x, y\, \exists z\, (\text{part}(z, x) \vee \text{part}(z, y))$': each of the two referential occurrences of 'z' is the root of a V-shaped (I, g)-graph, in which the respective occurrence of 'z' is directly linked to the referential occurrences of 'x' and 'y'.

Let us say that an (I, g)-graph $\langle N, E, r\rangle$ is *well-founded* iff there exists no infinite sequence of term-occurrences $o_1, o_2, \ldots$ such that E contains $\langle o_1, o_2\rangle$, $\langle o_2, o_3\rangle$, and so on. The final version of our account's core condition can then be stated as follows:

(CC) For any term t and any referential occurrence o of t in M: if o either stands at primary subject-position or is an occurrence of an atomic term at primary sentence-position, then o is the root of a well-founded (I', g')-graph.

To give a somewhat informal gloss of this condition, one might say that it requires every referential term-occurrence that stands in M at primary subject-position (as well as every referential occurrence of an atomic term that stands in M at primary sentence-position) to be either identical with or 'properly linked up' to one or more referential variable-occurrences that are bound by the initial occurrence of 'λ' in M. This does not exactly roll off the tongue; but if we are prepared to cut some corners, we might instead perhaps say that (CC) "requires the term-occurrences in M to be properly linked up to M's λ-variables".

11 Summary

Using (CC) to replace the ellipsis in (In_4), we arrive (after a few cosmetic changes) at the following account:

(In) An attribute A is *intrinsic* iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes$_{I,g}$ A, then there exist an interpretation $I' \supseteq I$, a variable-assignment $g' \supseteq g$, and a λ-expression M that satisfy the following three conditions:

(i) M has the same λ-variables as L.

(ii) M's matrix is unsimplifiably equivalent to that of some reduction$_{I',g'}$ of L.

(iii) For any term t and any referential occurrence o of t in M: if o either stands at primary subject-position or is an occurrence of an atomic term at primary sentence-position, then o is the root of a well-founded (I', g')-graph.

Combining the informal paraphrases offered in Section 8.1 and at the end of the previous Section, we can say that an attribute A is intrinsic, according to (In), iff every λ-expression L that denotes A at best only *seems* to exhibit A as extrinsic, by containing some term-occurrence that does not 'properly link up' to one or more of L's λ-variables.

A selection of 'data points' that have helped us arrive at this account is given in Table 1. Concerning its formal characteristics, I should begin by noting that the account has exactly two parameters: first, the notion of a *parthood relation*, which is operative in the definition of the concept of an (I, g)-graph, and second, the concept of *equivalence* that is both explicitly and implicitly at play in condition (ii) (implicitly, because it is also used in the definition of the relevant concept of simplifiability). The notion of a parthood relation has been left unspecified except for the two constraints (PR1) and (PR2) listed in Section 7.2. Naturally, the broader the notion of parthood that is 'plugged' into (In), the broader the corresponding concept of intrinsicality. As for the relevant concept of equivalence, we have seen in Section 8.1 that this concept should be chosen in such a way as to capture the coarse-grainedness of our conception of attributes.

In Section 4 it was suggested that all logically simple attributes are intrinsic. This is borne out by the present account, at least if 'logically simple' is understood in the sense of the analysis proposed in my (2016: §3.3).[96] Some other consequences worth mentioning have to do with closure under Boolean operations. Thus, if (In) is correct, then neither the

[96] According to that analysis, an attribute A is logically simple

> iff, for any interpretation I, variable-assignment g, and λ-expression L: if L denotes$_{I,g}$ A, then there exist an interpretation $I' \supseteq I$, a variable-assignment $g' \supseteq g$, and an atomic term F such that L has relative to I' and g' a reduction whose matrix is equivalent to $\ulcorner F(v_1, v_2, \ldots)\urcorner$, where $v_1, v_2, \ldots$ are (in this order) the λ-variables of L.

The relevant concept of reduction is here the same that has been introduced in Section 6.2 above, and the relevant concept of equivalence depends, as in the case of (In), on the coarse-grainedness of our conception of attributes. (Cf. *op. cit.*, p. 23n.) Given this understanding of 'logically simple', the proof of the thesis that every logically simple attribute is intrinsic essentially boils down to showing that, for any interpretation I', variable-assignment g', atomic term F, and variables $v_1, v_2, \ldots$, the λ-expression $M = \ulcorner \lambda v_1, v_2, \ldots\ F(v_1, v_2, \ldots)\urcorner$ has a non-simplifiable matrix and satisfies condition (iii) of (In).

§	Intrinsic	Extrinsic
2.1	Having Paris as a part	Being a father
2.2	$\lambda x\,(x \neq x)$	Being such that Socrates is non-self-identical[a]
6.2		Being Socrates and someone's son
6.3		Being an electron iff distinct from $\lambda x\,(x = x)$
7.1	$\lambda x\,\text{part}(\exists y\,\text{horse}(y), x)$	
7.2	$\lambda x\,(x = \text{Socrates})$	
7.3		$\lambda x\,\exists y\,(\text{part}(x, x) \wedge (y \neq y))$[a]
8.2		Either having an electron as a part or not having an electron as a part[a,b]
9.1	Having Paris as a part and being such that Paris has the Kremlin as a part	Being a friend of Aristotle
		Being such that s obtains
		Being such that there are horses
9.2	$\lambda x\,\exists y\,\text{part}(y, x)$	$\lambda x\,(P(x) = Q(x))$
		$\lambda x\,\neg\exists y\,\neg\text{part}(y, x)$
10.1	Not having Paris as a part	Being distinct from Paris
	Having either Paris or Rome as a part	
10.2	Being composed of cells[c]	
	Having as many red parts as green parts[d]	
	Having Paris as a set of one's parts	
	Having Paris as a member	
10.3	Having an atom as a part that has a neutron as a part[e]	$\lambda x\,(\neg\text{part}(\text{Paris}, x) \wedge \neg\text{part}(\text{Kremlin}, \text{Paris}))$
	Having a part that is distinct from Paris	$\lambda x\,(\neg\text{part}(\text{Paris}, x) \wedge \text{part}(\text{Kremlin}, \text{Paris}))$

[a] On a suitably fine-grained conception of attributes.
[b] On a suitably coarse-grained conception of attributes.
[c] Assuming that *being a cell* is intrinsic.
[d] Assuming that *being red* and *being green* are intrinsic.
[e] Assuming that the properties of *being an atom* and *being a neutron* are intrinsic.

Table 1: Examples of intrinsic and extrinsic attributes.

class of intrinsic nor the class of extrinsic attributes is closed under negation. *Being Socrates*, e.g., is classified as intrinsic, but its negation, *being distinct from Socrates*, is classified as *ex*trinsic, and the negation of *that* property is again classified as intrinsic.[97] The question of whether the classes of intrinsic and extrinsic attributes are closed under conjunction and disjunction depends on the coarse-grainedness of the underlying conception of attributes. For example, on a conception of attributes that is sufficiently (but not *too*) coarse-grained, the disjunction of the intrinsic properties of *having an electron as a part* and *not having an electron as a part* turns out extrinsic, and so does their conjunction.[98] Likewise, the conjunction of the two extrinsic properties of *being such that there are electrons* and *being either an electron or such that there are no electrons* is on a sufficiently coarse-grained conception of attributes identical with the (plausibly) intrinsic property of *being an electron*, and the same goes for the *dis*junction of the two extrinsic properties of *being an electron such that something is not an electron* and *being such that everything is an electron*.

12 The Local Distinction

Unlike the concepts of intrinsicality and extrinsicality that we have been concerned with above, the 'local' notions of having a property *intrinsically* and having it *extrinsically* are not contradictories: for example, a given electron might have the extrinsic property of *being such that there exists at least one electron* not only intrinsically but also (at the same time) *ex*trinsically, viz., if there exists some other electron.[99] Nonetheless, I think that the two notions can be adequately understood on the basis of the distinction between intrinsic and extrinsic attributes. We only have to help ourselves, in addition, to two concepts of necessitation: a 'basic' one that corresponds to the (first) notion of entailment defined in Section 6.3 above and another that is instead based on the notion of 'strictly L-relevant entailment' introduced in Section 7.3. The basic concept can be defined as follows:

[97] See condition (ii.5) of the definition of '(I, g)-graph'.

[98] Cf. Section 8.2 above. We have there only considered the case of the *dis*junction of the two properties just mentioned, but exactly parallel considerations apply to their conjunction.

[99] For other examples, see, e.g., Dunn (1990: 183) and Humberstone (1996: 228).

Definition 12.1. A state of affairs s_1 *necessitates* a state of affairs s_2 iff there exist an interpretation I, a variable-assignment g, and formulas φ and ψ such that: (i) φ denotes$_{I,g}$ s_1, (ii) ψ denotes$_{I,g}$ s_2, and (iii) φ entails ψ.

Next, a concept of *L-relevant* necessitation can be defined by replacing the 'entails' in condition (iii) with 'L-entails'. Similarly, by replacing the 'entails' with 'strictly L-entails', one obtains a definition of *strictly* L-relevant necessitation:

Definition 12.2. A state of affairs s_1 *strictly L-relevantly necessitates* (for short: strictly L-necessitates) a state of affairs s_2 iff there exist an interpretation I, a variable-assignment g, and formulas φ and ψ such that (i) φ denotes$_{I,g}$ s_1, (ii) ψ denotes$_{I,g}$ s_2, and (iii) φ strictly L-entails ψ.

An at least approximate account of the local distinction might now be formulated along the following lines:[100]

($\mathrm{In}_{\mathrm{loc}}$) An entity x has a property P *intrinsically* iff there exists some intrinsic property Q such that $Q(x)$ obtains and necessitates $P(x)$.

($\mathrm{Ex}_{\mathrm{loc}}$) An entity x has a property P *extrinsically* iff there exists a state of affairs s that satisfies the following two conditions:

(i) s obtains and strictly L-relevantly necessitates $P(x)$.

(ii) No intrinsic property Q is such that $Q(x)$ obtains and necessitates s.

For the sake of illustration, let P be the extrinsic property of *being such that there exists at least one electron*, or in symbols: $\lambda x\,((x = x) \wedge \exists y\, E(y))$, and let e be some electron. It is then easy to see that e has P 'intrinsically' in the sense of ($\mathrm{In}_{\mathrm{loc}}$), at least provided that *being an electron* is intrinsic. At the same time, e has P *extrinsically*. For let e' be some electron distinct from e, and let s be the state of affairs that e is such that e' is an electron, or in symbols: $(e = e) \wedge E(e')$. It then follows that s strictly L-necessitates $P(e)$, as can be seen from the fact that the formula '$(e = e) \wedge E(e')$' strictly L-entails '$(e = e) \wedge \exists y\, E(y)$'; and plausibly there is no intrinsic property Q such that $Q(e)$ obtains and necessitates s.

[100] Both of the following two definitions can be readily generalised to attributes of higher adicities.

The reason why ($\mathrm{Ex}_{\mathrm{loc}}$) makes use of *strictly* L-relevant necessitation lies in the need to avoid spurious instances of a property's being had extrinsically. For example, suppose that all the pens in my drawer are ballpoint pens, let p be one such pen, and let D and B be, respectively, the properties of *being a pen in my drawer* and *being a ballpoint pen*. Then the conjunction of the states of affairs $\forall x\,(D(x) \to B(x))$ and $D(p)$ obtains and necessitates—even L-relevantly necessitates—the state of affairs $B(p)$. At the same time, that conjunction is not necessitated by any state of affairs $Q(p)$, where Q is an intrinsic property. Hence, without the 'strictly L-relevantly' in condition (i), p would be classified as having B extrinsically, which would be an unwelcome result. By contrast, the conjunction of $\forall x\,(D(x) \to B(x))$ and $D(p)$ does not *strictly* L-relevantly necessitate $B(p)$, as can be seen from the fact that '$\forall x\,((x = x) \to B(x)) \wedge D(p)$' is a non-contradictory pruning of '$\forall x\,(D(x) \to B(x)) \wedge D(p)$' that, like the latter, entails '$B(p)$'.

If ($\mathrm{In}_{\mathrm{loc}}$) and ($\mathrm{Ex}_{\mathrm{loc}}$) are by and large correct, we can apparently conclude (contrary to what has sometimes been suggested) that the distinction between having a property intrinsically and having it extrinsically need not be regarded as more basic than the 'global' distinction between intrinsic and extrinsic attributes.[101]

13 Conclusion

There is no denying that the account of intrinsicality proposed in this paper is unusually complex. In itself, this need not be a bad thing: that seemingly simple notions can turn out to have unexpectedly complex analyses is a familiar phenomenon. (Take, for example, the concept of a smooth curve, of a random number, or of a computable function.) But sometimes, at least in philosophy, analyses can become too complex for comfort, leaving one puzzled as to why one should ever have cared about a concept as complex as the proposed analysis would suggest.[102] As far as I can see, there are three possible ways to proceed in such a situation. First, one might try to find an analysis that is equivalent to

[101] An interestingly different conception of the intrinsically/extrinsically distinction has been explored by Carrie Figdor (2008), who, following Ellis (2002), lays heavy emphasis on the consideration of "relevant counterfactual circumstances" (p. 696).

[102] Timothy Williamson (2000) has famously felt this way about certain analyses of the concept of knowledge.

the original, but simpler. Second, one might try to see if one can 'prune' the analysis in some reasonable way, so as to end up with a concept that is no longer *a priori* coextensive with the original but more 'natural' and easier to work with. A third option would be to abandon the analytic project and to take the concept as primitive.[103]

Suppose that the complexity of the present account of intrinsicality is indeed too much for comfort: which of the three options should we choose? Concerning the first, I have to confess that I do not know how to simplify the account without sacrificing extensional adequacy. Concerning the third, I am afraid that it would leave it a mystery what the intrinsicality of an attribute might consist in—e.g., what it is about *having Paris as a part* that makes this property intrinsic. Best, then, to take the second option. For example, we might decide to settle for an account that abandons some or all of the refinements introduced in Section 10. But it is not clear to me that the gains in simplicity would be worth the resulting cost in extensional adequacy.

When it comes to the analysis of relatively familiar notions, our preference for simplicity might in part be rooted in the hope that a simpler analysis would provide a better explanation of *why* such-and-such properties happen to fall, or not to fall, under a given concept. This is certainly true if one compares any finitely stateable analysis with one that merely enumerates infinitely many instances (as in, 'An attribute A is *intrinsic* iff A is the property of *being self-identical* or A is the property of *having a proper part* or ...'). By being less 'disjunctive', the simpler analysis will articulate non-disjunctive features that several intrinsic attributes have in common, and may thereby claim to provide a better explanation as to 'what makes' those attributes intrinsic. However, this would seem to be less an argument for pruning away complexities than a reason to look for a simpler but *equivalent*

[103] An adherent of a particularly demanding view as to what counts as an 'analysis' or as a 'reductive account' may argue that the account given in this paper qualifies neither as an analysis nor as reductive. On such a view, I have effectively been treating 'intrinsic' as a primitive all along and have unwittingly provided an account that is no analysis at all. Since this may be a possible view to hold, I should perhaps clarify in which sense I take myself to have given an 'analysis of intrinsicality'. As I understand the term, what makes something an analysis is—besides its logical form—a matter of its aims, viz., (i) that the coextensiveness of *analysans* and *analysandum* is meant to follow in large part from what may be called *semantic intuitions* (cf., e.g., Eklund [2015], who talks in a similar vein of 'competence intuitions'), and (ii) that the analysis itself is meant to be informative. (This latter aim requires that the analysis should not be circular.) A successful analysis may thus be said to be 'justified' by semantic intuitions. This justification need not, however, extend to the *framework* within which the analysis is constructed. Different frameworks can be expected to give rise to correspondingly different analyses; but this is not in itself objectionable.

analysis. And as for the question of why we group such-and-such attributes together under a common label, an *analysis* may in any case be the wrong place to look for an answer. An analysis will usually not tell us what biases and desiderata have shaped our conceptual landscape and have led us to apply the label 'intrinsic' to some properties as opposed to others. But such biases and desiderata may better reflect why we "care" about intrinsicality than any analysis could.[104]

References

Armstrong, D. M. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.

Bader, R. M. (2013). Towards a Hyperintensional Theory of Intrinsicality. *The Journal of Philosophy*, *110*, 525–63.

Bealer, G. (1982). *Quality and Concept*. Oxford: Clarendon.

Bealer, G. (1994). Property Theory: The Type-Free Approach v. the Church Approach. *Journal of Philosophical Logic*, *23*, 139–71.

Bealer, G. & Mönnich, U. (2003). Property Theories. In D. M. Gabbay & F. Guenthner (eds.), *Handbook of Philosophical Logic*, volume 10 (pp. 143–248). Dordrecht: Kluwer.

Benacerraf, P. (1965). What Numbers Could Not Be. *The Philosophical Review*, *74*, 47–73.

Church, A. (1951). A Formulation of the Logic of Sense and Denotation. In P. Henle, H. M. Kallen, & S. K. Langer (eds.), *Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer* (pp. 3–24). New York, NY: The Liberal Arts Press.

Dorr, C. (2016a). Against Counterfactual Miracles. *The Philosophical Review*, *125*, 241–86.

Dorr, C. (2016b). To Be F Is to Be G. *Philosophical Perspectives*, *30*, 39–134.

[104]I would like to thank Claudio Calosi, Fabrice Correia, Richard Glauser, Ghislain Guigon, Benedikt Löwe, Dan Marshall, Robert Michels, Kevin Mulligan, Benjamin Neeser, Pierre Saint-Germier, Alexander Skiles, and an anonymous referee for valuable comments and discussion. I am also grateful for financial support provided by the Swiss National Science Foundation (projects CRSII1_147685 and 100012_173040).

Dorr, C. & Hawthorne, J. (2013). Naturalness. In K. Bennett & D. W. Zimmerman (eds.), *Oxford Studies in Metaphysics*, volume 8 (pp. 3–77). Oxford: Oxford University Press.

Dunn, J. M. (1990). Relevant Predication 2: Intrinsic Properties and Internal Relations. *Philosophical Studies*, *60*, 177–206.

Eddon, M. (2011). Intrinsicality and Hyperintensionality. *Philosophy and Phenomenological Research*, *82*, 314–36.

Eklund, M. (2015). Intuitions, Conceptual Engineering, and Conceptual Fixed Points. In C. Daly (ed.), *The Palgrave Handbook of Philosophical Methods* (pp. 363–85). New York: Palgrave Macmillan.

Ellis, B. (2002). *The Philosophy of Nature: A Guide to the New Essentialism*. Chesham: Acumen.

Figdor, C. (2008). Intrinsically/Extrinsically. *The Journal of Philosophy*, *105*, 691–718.

Figdor, C. (2014). What's the Use of an Intrinsic Property? In R. M. Francescotti (ed.), *Companion to Intrinsic Properties* (pp. 139–155). Berlin: De Gruyter.

Fine, K. (1994). Essence and Modality. *Philosophical Perspectives*, *8*, 1–16.

Fine, K. (1995a). Senses of Essence. In W. Sinnott-Armstrong (ed.), *Modality, Morality, and Belief* (pp. 53–73). Cambridge: Cambridge University Press.

Fine, K. (1995b). The Logic of Essence. *Journal of Philosophical Logic*, *24*, 241–73.

Fine, K. (2010). Towards a Theory of Part. *The Journal of Philosophy*, *107*, 559–89.

Francescotti, R. M. (1999). How to Define Intrinsic Properties. *Noûs*, *33*, 590–609.

Francescotti, R. M. (2014a). *Companion to Intrinsic Properties*. Berlin: De Gruyter.

Francescotti, R. M. (2014b). Intrinsic/Extrinsic: A Relational Account Defended. In R. M. Francescotti (ed.), *Companion to Intrinsic Properties* (pp. 175–198). Berlin: De Gruyter.

Gilmore, C. S. (2013). Slots in Universals. In K. Bennett & D. W. Zimmerman (eds.), *Oxford Studies in Metaphysics*, volume 8 (pp. 187–233). Oxford: Oxford University Press.

Hodges, W. (2013). Logic and Games. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2013 ed.). Metaphysics Research Lab, Stanford University.

Hoffmann-Kolss, V. (2010). *The Metaphysics of Extrinsic Properties*. Frankfurt am Main: Ontos Verlag.

Hovda, P. (2009). What Is Classical Mereology? *Journal of Philosophical Logic*, *38*, 55–82.

Humberstone, I. L. (1996). Intrinsic/Extrinsic. *Synthese*, *108*, 205–67.

Kagan, S. (1998). Rethinking Intrinsic Value. *The Journal of Ethics*, *2*, 277–97.

Langton, R. & Lewis, D. K. (1998). Defining 'Intrinsic'. *Philosophy and Phenomenological Research*, *58*, 333–45.

Lewis, D. K. (1983a). Extrinsic Properties. *Philosophical Studies*, *44*, 197–200.

Lewis, D. K. (1983b). New Work for a Theory of Universals. *Australasian Journal of Philosophy*, *61*, 343–77.

Lewis, D. K. (1986a). *On the Plurality of Worlds*. Oxford: Blackwell.

Lewis, D. K. (1986b). Postscripts to "Causation". In Lewis (ed.), *Philosophical Papers*, volume 2 (pp. 172–213). Oxford: Oxford University Press.

MacBride, F. (2005). The Particular–Universal Distinction: A Dogma of Metaphysics? *Mind*, *114*, 565–614.

Marshall, D. (2009). Can 'Intrinsic' Be Defined Using Only Broadly Logical Notions? *Philosophy and Phenomenological Research*, *78*, 646–72.

Marshall, D. (2012). Analyses of Intrinsicality in Terms of Naturalness. *Philosophy Compass*, *7*, 531–42.

Marshall, D. (2013). Analyses of Intrinsicality without Naturalness. *Philosophy Compass*, *8*, 186–97.

Marshall, D. (2014). Yablo's Account of Intrinsicality. In R. M. Francescotti (ed.), *Companion to Intrinsic Properties* (pp. 199–220). Berlin: De Gruyter.

Marshall, D. (2015). Intrinsicality and Grounding. *Philosophy and Phenomenological Research*, *90*, 1–19.

Marshall, D. (2016a). An Analysis of Intrinsicality. *Noûs*, *50*, 704–39.

Marshall, D. (2016b). The Varieties of Intrinsicality. *Philosophy and Phenomenological Research*, *92*, 237–63.

McCarthy, T. (2015). A Note on Unrestricted Composition. *Thought*, *4*, 202–11.

McDaniel, K. (2004). Modal Realism with Overlap. *Australasian Journal of Philosophy*, *82*, 137–52.

McDaniel, K. (2009). Ways of Being. In D. J. Chalmers, D. Manley, & R. Wasserman (eds.), *Metametaphysics* (pp. 290–319). Oxford: Oxford University Press.

McDaniel, K. (2014). Compositional Pluralism and Composition as Identity. In A. J. Cotnoir & D. L. M. Baxter (eds.), *Composition as Identity* (pp. 130–42). Oxford: Oxford University Press.

McKay, T. J. (2006). *Plural Predication*. Oxford: Oxford University Press.

Menzel, C. (1993). The Proper Treatment of Predication in Fine-Grained Intensional Logic. *Philosophical Perspectives*, *7*, 61–87.

Menzel, C. (2016). Problems with the Actualist Accounts (supplement to the entry on actualism). In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.). Metaphysics Research Lab, Stanford University.

Moore, G. E. (1922). The Conception of Intrinsic Value. In Moore (ed.), *Philosophical Studies* (pp. 253–275). London: Kegan Paul.

Parsons, J. (2001). *Theories of Persistence*. PhD thesis, Australian National University.

Plantinga, A. (1983). On Existentialism. *Philosophical Studies*, *44*, 1–20.

Plate, J. (2016). Logically Simple Properties and Relations. *Philosophers' Imprint*, *16*(1), 1–40.

Rayo, A. (2013). *The Construction of Logical Space*. Oxford: Oxford University Press.

Rønnow-Rasmussen, T. (2015). Intrinsic and Extrinsic Value. In I. Hirose & J. Olson (eds.), *The Oxford Handbook of Value Theory* (pp. 29–43). Oxford: Oxford University Press.

Rosen, G. (2010). Metaphysical Dependence: Grounding and Reduction. In B. Hale & A. Hoffmann (eds.), *Modality* (pp. 109–35). Oxford: Oxford University Press.

Schaffer, J. (2004). Two Conceptions of Sparse Properties. *Pacific Philosophical Quarterly*, *85*, 92–102.

Sider, T. (1996). Intrinsic Properties. *Philosophical Studies*, *83*, 1–27.

Sider, T. (2007). Parthood. *The Philosophical Review*, *116*, 51–91.

Skiles, A. (2014). Primitivism about Intrinsicality. In R. M. Francescotti (ed.), *Companion to Intrinsic Properties* (pp. 221–52). Berlin: De Gruyter.

Skow, B. (2007). Are Shapes Intrinsic? *Philosophical Studies*, *133*, 111–30.

Swoyer, C. (1998). Complex Predicates and Logics for Properties and Relations. *Journal of Philosophical Logic*, *27*, 295–325.

Taylor, B. M. (1993). On Natural Properties in Metaphysics. *Mind*, *102*, 81–100.

Weatherson, B. (2001). Intrinsic Properties and Combinatorial Principles. *Philosophy and Phenomenological Research*, *63*, 365–80.

Weatherson, B. & Marshall, D. (2017). Intrinsic vs. Extrinsic Properties. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2017 ed.). Metaphysics Research Lab, Stanford University.

Wetzel, L. (1993). What Are Occurrences of Expressions? *Journal of Philosophical Logic*, *22*, 215–20.

Williams, J. R. G. (2013). Part-Intrinsicality. *Noûs*, *47*, 431–52.

Williamson, T. (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

Witmer, D. G., Butchard, W., & Trogdon, K. (2005). Intrinsicality without Naturalness. *Philosophy and Phenomenological Research*, *70*, 326–50.

Yablo, S. (1999). Intrinsicness. *Philosophical Topics*, *26*, 479–505.

Yablo, S. (2016). Parts and Differences. *Philosophical Studies*, *173*, 141–57.

Yi, B. (1999). Is Two a Property? *The Journal of Philosophy*, *96*, 163–90.

Zalta, E. N. (1983). *Abstract Objects: An Introduction to Axiomatic Metaphysics*. Dordrecht: Reidel.

Zimmerman, M. J. (2015). Intrinsic vs. Extrinsic Value. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2015 ed.). Metaphysics Research Lab, Stanford University.

Appendix A: The Marshall–Parsons Argument

Greatly expanding on a brief argument by Josh Parsons (2001: 22f.), Dan Marshall (2009) has argued that the intrinsic/extrinsic distinction cannot be analysed "using only broadly logical notions". This argument may at first sight seem to block the project pursued in the present paper, which is to develop an account under which the intrinsicality or extrinsicality of a given attribute is a matter of the attribute's logical constitution. The purpose of this Appendix is to discuss whether Marshall's argument does in fact pose a threat to this project, and if so, how it can be answered from within the present paper's framework.

Marshall's argument revolves around two properties, viz., (i) *being an electron* and (ii) *being either a lonely positron or an accompanied electron*, of which the first may plausibly be regarded as intrinsic and the second as extrinsic. (To say that a thing is 'lonely' means here that there exists nothing else besides it, and to say that it is 'accompanied' means that it is not lonely.) Let us refer to the first property as 'E' and to the second as 'E^*'. In addition, let us refer to the property of *being a positron* as 'P'. A central part of Marshall's argument is devoted to establishing the claim that

(S) E and E^* satisfy the same rigid broadly logical formulas.[1]

Here the notion of a *rigid* formula is defined as follows: "A one-place formula A with a free variable x is *rigid* iff $\ulcorner\forall x[\text{Necessarily}(A(x)) \vee \text{Necessarily}(\neg A(x))]\urcorner$ is true" (p. 654). As for 'broadly logical', we can infer from what is said at several places in Marshall's paper (e.g. on p. 672) that a formula is *broadly logical* iff it is a formula in the language L that he describes in his footnote 2:

> Let L be a language containing the variables 'x', 'y', 'z', 'x_1' ... together with the operators and predicates: '$\neg$' (meaning 'it is not the case that'), '$\wedge$' (meaning 'and'), '$\exists$' (meaning 'for some'), '$\Box$' (meaning 'necessarily'), 'at', 'is a possible world', 'is a set', 'exists', '$=$', '$\in$' (meaning 'is a member of'), 'is a proper part of', 'instantiates', and 'is a property'. If X and Y are variables in L then $\ulcorner(X \text{ is a possible world})\urcorner$, $\ulcorner(X \text{ is a set})\urcorner$, $\ulcorner(X \text{ exists})\urcorner$, $\ulcorner(X = Y)\urcorner$, $\ulcorner(X \in Y)\urcorner$, $\ulcorner(X \text{ is a proper part of } Y)\urcorner$, $\ulcorner(X \text{ instantiates } Y)\urcorner$, and $\ulcorner(X \text{ is a property})\urcorner$ are atomic formulas in L. If A and B are formulas in L, and X is a variable in L, then $\ulcorner\neg A\urcorner$, $\ulcorner(A \wedge B)\urcorner$, $\ulcorner(\exists X)A\urcorner$, $\ulcorner\Box A\urcorner$, and $\ulcorner(\text{at } X)A\urcorner$ [read: at world X, it is the case that A] are formulas in L. All formulas in L are specified by the previous two sentences. (p. 648n.)

Apart from clarifying the notion of a broadly logical formula, this passage helps explain what it would mean for 'intrinsic' to be definable "using only broadly logical notions" (as Marshall understands this phrase); for as he explains in the same footnote, "'intrinsic' can be defined using only broadly logical notions iff 'x is intrinsic' is an abbreviation of a formula in L". He goes on to allow that "[w]e might wish to broaden what counts as a broadly logical definition either by i) turning L into an infinitary language by allowing infinite conjunctions and infinite blocks of quantifiers, or ii) by adding vocabulary such as the plural quantifier 'there are', the operator 'actually', or the term 'the actual world' to the list of broadly logical vocabulary", but he suspects that his argument can be "modified so that it applies to various such expanded conceptions of what counts as a broadly logical definition".

From the above thesis (S), together with three *prima facie* reasonable assumptions—viz., (i) that E is intrinsic, (ii) that E^* is extrinsic, and (iii) that any property is either necessarily intrinsic or necessarily extrinsic—it is only a few steps to the conclusion that 'intrinsic' cannot be defined (or analysed) using only broadly logical notions. (I give a very brief sketch of the reasoning in footnote 21 above.[2]) I have no objection to these assumptions, nor to the inferences that lead from

[1]The label 'S' (which is not Marshall's) is intended as mnemonic for 'same'.

[2]In that footnote I essentially follow the "informal argument" that Marshall offers in §2 of his paper, in which the central claim (S) occurs for the first time on p. 654. This claim is then put on a more rigorous footing in the "formal argument" of §4. While (S) does not *explicitly* reappear in that later Section, one can see that it is implicitly established when Marshall concludes that E and E^* would "either both satisfy 'x is intrinsic' or both fail to satisfy 'x is intrinsic'" if 'x is intrinsic' abbreviated a rigid broadly logical formula (p. 665). Another difference lies in the fact that in §4 Marshall gives an argument for the extrinsicality of E^*, based on the assumption that positronhood is "intrinsically qualitatively complete" (p. 664f.). These differences are not relevant for our purposes, however.

(S) to Marshall's conclusion, nor to the way in which he establishes (S). Instead, in order to assess whether his argument poses an obstacle to the present paper's project, I will focus on the question of whether it is possible to establish an analogue of (S) in which the phrase 'broadly logical' is understood in a sufficiently broad sense that it applies to the right-hand side of the present paper's account of intrinsicality.

Answering this question turns out to be more complicated than one might at first expect, for not only is the language of our account quite different from Marshall's language L, but there are also important differences with respect to the larger framework. Thus, while Marshall works with an ontology of *possibilia*, notions of metaphysical modality play no role in our framework, and accordingly, attributes are in it not individuated in terms of necessary coextensiveness. Instead, they are individuated on the basis of a suitable concept of equivalence, applied to matrices of λ-expressions; and formulas are used as names of states of affairs.

To see what sort of difference all this makes *vis-à-vis* Marshall's argument, suppose we wish to adopt a moderately coarse-grained conception of attributes: coarse-grained enough that any property Q is identical with $\lambda x\,(Q(x) \wedge (x = x))$, but also *fine*-grained enough to distinguish the property of *being self-identical* from that of *being self-identical and such that Socrates is self-identical*. To formulate such a conception, we would first introduce a concept of equivalence, understood as mutual entailment in, e.g., the following sense:[3]

Definition A.1. A formula φ *entails* a formula ψ iff φ has a denotation relative to some interpretation and variable-assignment and, for any interpretation I and variable-assignment g, the following two conditions are satisfied:

(i) If φ has a denotation relative to I and g, then so does ψ.

(ii) If φ denotes relative to I and g an obtaining state of affairs, then so does ψ.

Equipped with the corresponding concept of equivalence (on which two formulas are equivalent just in case they entail each other in the sense just defined), we would in the next step adopt the following principle:[4]

(P) If a λ-expression L_1 has the same λ-variables in the same order as another λ-expression L_2, then they satisfy the following two conditions:

[3]The following definition is essentially the same as Definition 6.3 (p. 28).

[4]For details, see Section 6.2. (P) is in effect a simplified version of the conjunction of the two theses (CA) and (FA) that are formulated in that same Section. To reduce clutter, references to interpretations and variable-assignments are here suppressed as far as possible.

(i) If the matrices of L_1 and L_2 are equivalent and both L_1 and L_2 have a denotation, then they both denote the same attribute.

(ii) If L_1 and L_2 denote the same attribute, then there exist a reduction L_1' of L_1 and a reduction L_2' of L_2 such that the matrix of L_1' is equivalent to that of L_2'.

A 'reduction' of a given λ-expression L is here, roughly, a λ-expression (not necessarily distinct from L) that results from L by replacing zero or more occurrences of atomic terms with other terms that denote the same entities as those whose occurrences they replace.[5]

With these preliminaries out of the way, let us now consider E^*. Relative to a variable-assignment that maps 'E' and 'P' respectively to E and P (i.e. to the properties of *being an electron* and *being a positron*), E^* will be denoted by

$$\lambda x \left(\left(P(x) \wedge \neg\exists y\, (y \neq x) \right) \vee \left(E(x) \wedge \exists y\, (y \neq x) \right) \right). \qquad (*)$$

Under the above concept of equivalence, the matrix of $(*)$ turns out to be equivalent to '$(P = P) \wedge E(x)$', i.e. (written without abbreviatory devices) to '$\mathtt{I}(P, P) \wedge E(x)$'.[6] This is due to the fact that the formula '$\exists y\, (y \neq x)$' denotes an obtaining state of affairs relative to any interpretation and variable-assignment relative to which it denotes anything at all. And the reason for *this* is simply that everything is 'accompanied', since there are at least two entities. Indeed, analogous considerations apply to any formula according to which x is accompanied by at least κ-many entities (for any set-sized cardinality κ), due to the 'abundance' of our ontology of attributes and states of affairs (as well as sets). Now, given that the matrix of $(*)$ is equivalent to '$\mathtt{I}(P, P) \wedge E(x)$', it follows from (P) that E^* is in fact the same property as that of *being an electron such that* $P = P$. Intuitively, this property is still extrinsic. It may be extrinsic in a different way than might occur to one if the property is instead described as that of *being either a lonely positron or an accompanied electron*, but that does not matter for a discussion of Marshall's argument. The challenge is still the same, viz., to see whether we can classify E^* as extrinsic without at the same time having to classify E as extrinsic.

As it turns out, our account *does* classify E^* as extrinsic, at least under two assumptions. The first of these simply states that

(A1) Instantiating E^* does not require having P as a part.

This thesis could be made more precise, but for present purposes there will be little need to do so.

---

[5]Cf. Definition 6.1 (p. 24).

[6]'$\mathtt{I}$' functions here as a logical constant denoting the identity relation. (Cf. Section 5.1.)

I take it to be eminently plausible on any reasonable precisification. For the second thesis, we first have to introduce two concepts of analysability:[7]

Definition A.2. An entity x is *analysable* in terms of an entity y iff there exists a term t (which could be a formula or λ-expression), as well as some interpretation I and variable-assignment g, such that t denotes$_{I,g}$ x and contains a free occurrence of a term that denotes$_{I,g}$ y.[8]

Definition A.3. An entity x is *fully analysable* in terms of a given set S of entities iff there exists a term t, as well as an interpretation I and a variable-assignment g, such that:

(i) t denotes$_{I,g}$ x, and

(ii) Every variable or constant that occurs free in t denotes$_{I,g}$ some member of S.

With the help of these two concepts, the second assumption can be formulated as follows:

(A2) The state of affairs $\mathtt{I}(P, P)$ is not fully analysable in terms of any set that contains only entities in terms of which E is analysable.

Whether this assumption holds depends in part on the metaphysics of E and P (e.g. on whether P is fully analysable in terms of $\{E\}$). For the sake of the example, let us suppose that it is true.

To illustrate how, under these two assumptions, the account developed in this paper classifies E^* as intrinsic, I will for reasons of space rely on a very rudimentary description of that account. Thus, very roughly: the account classifies an attribute A as extrinsic iff, *first*, A is denoted by some λ-expression L that at least *seems* to exhibit A as extrinsic by virtue of certain 'complex-making' features of L and, *second*, there exists no reduction of L that shows that those features are misleading. To apply this to our example, let L be the expression '$\lambda x\,(\mathtt{I}(P, P) \wedge E(x))$', which, by what has been said above, denotes E^*.[9] As may be seen from the details of our account (though admittedly not from the rudimentary description that has just been given), the complex-making feature of L by virtue of which it "seems to exhibit" E^* as extrinsic are the two occurrences of 'P'.

Let now M be some reduction of L. M will then denote the same property as is denoted by L, viz., E^*. Further, M's matrix will contain a first conjunct γ_1 that results from '$\mathtt{I}(P, P)$' by replacing zero or more occurrences of either '$\mathtt{I}$' or 'P' with some coreferential other term(s), and a second

[7]The following definitions are essentially taken from my (2016: §3.5), though with one important change in the second definition: in order to be 'fully analysable' in terms of a given set S, an entity need not be analysable in terms of each member of S.

[8]Here 'denotes$_{I,g}$' is short for 'denotes relative to I and g'. (Cf. Section 5.2.)

[9]As already above, I am here suppressing references to interpretations and variable-assignments in order to simplify the discussion.

conjunct γ_2 that results from '$E(x)$' by replacing the occurrence of 'E'. There are two ways in which M might 'save' E^* from being classified as extrinsic.[10] The first way would be for M's matrix (in which 'x' occurs free) to entail that P is a part of x; but this is ruled out by (A1).[11] The second way would be for γ_1 to be an instance of 'redundant complexity', which would require that M's matrix is equivalent to γ_2 (in the sense defined above). But M's matrix cannot be equivalent to γ_2 if γ_1 contains some free occurrence of a variable or constant (other than '$\mathtt{I}$') that does not occur free in γ_2. And that will precisely be the case if (A2) is true. So M does not, after all, 'save' E^* from being classified as extrinsic; but M was *any* reduction of L. Hence, given (A1) and (A2), no reduction of '$\lambda x\, (\mathtt{I}(P, P) \wedge E(x))$' will show that what this λ-expression suggests—viz., that E^* is extrinsic—is not in fact the case. And so E^* will be classified as extrinsic. By contrast, E may (for all that has been said here) still be classified as *in*trinsic.

As I hope can be seen from this example, the use of λ-expressions, together with some auxiliary notions, such as denotation and occurrence, and combined with a moderately fine-grained conception of attributes, allows us to articulate differences between E and E^* that could not be expressed using only the set of notions that Marshall considers broadly logical for the purposes of his argument. Even though E and E^* satisfy exactly the same formulas of Marshall's language L, it need not be the case, for all we know, that both of them fail to satisfy the right-hand side of the present paper's account of intrinsicality. Further, the description given by the right-hand side of our account can be fairly regarded as 'rigid', since it does not appeal to any contingent features of the to-be-classified attribute (such as the number of things—or sequences of things—that happen to instantiate it). Hence, if we interpret the 'broadly logical' in (S) in a sufficiently broad sense that it applies to the language of our account, that thesis no longer seems to hold.[12]

[10] Here I am drawing on further details of the account, which the above description leaves implicit.

[11] To keep the example reasonably simple, I am ignoring various even more outlandish possibilities, such as that M's matrix might entail that P is a *set* of parts of x.

[12] An objector might at this point reply that a good account of intrinsicality should be extensionally adequate even when combined with a highly coarse-grained conception of attributes, rather than the merely *moderately* coarse-grained conception that we have appealed to here. This may at first blush seem a plausible maxim, but I do not think that it makes for an effective objection. One reason for this is given in the discussion at the end of Section 6.1: some conceptions of attributes are simply too coarse-grained to be plausible. Another reason is the following. If, working with a very coarse-grained conception, we still regard E^* as extrinsic, on the basis of its satisfying the following description: *the property of being either a lonely positron or an accompanied electron*, then presumably some of the complexities of ($*$) are not in fact redundant under the conception in question; for ($*$) is just a formalisation of that description. (It would be a dialectical mistake to go on insisting on E's intrinsicality and E^*'s extrinsicality regardless of how redundant those features become as we move to more and more coarse-grained conceptions of attributes, *unless* we want to maintain that one and the same attribute can be both intrinsic and extrinsic. Some conceptions are after all coarse-grained enough that E^* collapses not just into $\lambda x\, ((P = P) \wedge E(x))$ but into E itself.) Moreover, at least one of those non-redundant complexities of ($*$) will have to be of a suitably 'extrinsicality-indicating' character; and finally it will have to be the case that not all of these non-redundant, extrinsicality-indicating complexities turn out to be redundant, or to lose their

Appendix B: Occurrences and Containment

Concepts of occurrence, containment, and variable-binding are familiar from basic mathematical logic and computer science. However, since the discussion in this paper sometimes relies on some finer points relating to these notions, it will be useful to provide suitably precise definitions that apply, in particular, to the formal language described in Section 5.

To begin with, an *expression* of that language can be thought of as a function mapping ordinal numbers to symbols (where each atomic term—i.e. each constant or variable—counts as one symbol), such that the function's domain is an initial segment of the ordinals. Symbols need for our purposes not be thought of as 'genuinely linguistic' entities (such as types of inscriptions): we can remain entirely noncommittal about their nature, and even allow them to be pure sets. As for *occurrences* of expressions, these will here be taken to be ordered triples $\langle E, e, \alpha \rangle$, where the first and second coordinates are expressions and the third coordinate is an ordinal. Intuitively, the first coordinate is the expression that contains the occurrence in question, the second is the expression that the occurrence in question is an occurrence *of*, and the third indicates the occurrence's starting position within the containing expression, counting from zero.[1] A triple $\langle E, e, \alpha \rangle$ may thus be said to be "the occurrence of e in E that begins with the $(\alpha + 1)$th symbol of E", *provided* that E and e are expressions and that α is an ordinal such that, for each ordinal β in the domain of e, the symbol $E(\alpha + \beta)$—i.e. the $(\alpha + \beta + 1)$th symbol of E—is identical with $e(\beta)$. Unless these conditions are met, $\langle E, e, \alpha \rangle$ will not be said to be an 'occurrence' at all.

Based on this conception of occurrences, we can now go on to define three useful concepts of containment:

Definition B.1. An expression E *contains* an expression e iff, for some ordinal α, there exists an occurrence $\langle E, e, \alpha \rangle$.

Definition B.2. An expression E *contains* an occurrence o iff, for some expression e and ordinal α, o is identical with $\langle E, e, \alpha \rangle$.

Definition B.3. An occurrence $\langle E, e, \alpha \rangle$ *contains* an occurrence $\langle E', e', \alpha' \rangle$ iff $E = E'$, and for each ordinal β in the domain of e' there exists an ordinal γ in the domain of e such that $\alpha' + \beta = \alpha + \gamma$.

The following are some easy consequences of the last two definitions:

---

extrinsicality-indicating character, once we consider some reduction of $(*)$. If so, however, there will again be no reduction of $(*)$ that saves E^* from being classified as extrinsic.

[1]Cf. Wetzel (1993), who construes an occurrence of e in E as an ordered triple $\langle n, e, E \rangle$, where n indicates the occurrence's position in the sequence of all occurrences of e in E, ranked by their starting points.

- Every occurrence is contained in exactly one expression.
- Every occurrence contains itself.
- An occurrence contains another *only* if they are both contained in the same expression.
- For any occurrences o and o': if o contains an occurrence that contains o', then o contains o'.

For the sake of brevity, we sometimes speak of an expression or occurrence "in" an expression or occurrence ε, meaning an expression or occurrence that is (in the relevant sense) *contained* in ε. If an occurrence is contained in an occurrence distinct from it, we say that it is *properly contained* in the latter.

The distinction between 'bound' and 'free' variable-occurrences can be drawn more or less in the usual fashion, but sometimes it matters not only *whether* a given variable-occurrence is bound, but also *by which* operator-occurrence it is bound, if it is bound at all. In the case of the language described in Section 5, this relation can be defined as follows:

Definition B.4. An occurrence o of a variable v is *bound* by an occurrence o' of a variable-binding operator (i.e. either 'λ' or '$\exists$') iff there exists a term-occurrence o^* such that the following three conditions are satisfied:

(i) o^* starts with o' and contains o.

(ii) o' immediately precedes a list of variable-occurrences that has as one of its elements an occurrence of v.

(iii) For any occurrence o'' of any variable-binding operator: if o'' immediately precedes a list of variable-occurrences that has as one of its elements an occurrence of v, then there exists no term-occurrence that is properly contained in o^*, starts with o'', and contains o.

The complicated third condition is here needed to ensure that, in those cases where a variable-occurrence stands in the scope of two or more operator-occurrences, it counts as being bound by the *innermost* applicable operator-occurrence. For example, the third and fourth occurrences of 'x' in '$\lambda x\,(P(x) \wedge \exists x\, Q(x))$' are bound by the '$\exists$', not by the initial '$\lambda$'.

We say that a variable-occurrence is *bound* iff it is bound by some operator-occurrence, and *free* otherwise. Variable-occurrences that are free in this sense are also said to be free *simpliciter*. By contrast, we say that a variable-occurrence is *free in* a term-occurrence o iff it is contained in o and not bound by any operator-occurrence in o; whereas it is said to be *bound in o* iff it is contained in o and bound by some operator-occurrence in o. For instance, the second occurrence of 'x' in '$\exists x\, P(x)$'

is free in the containing occurrence of '$P(x)$' but not free *simpliciter*. This terminology extends in a natural way to occurrences of terms. Thus, a term-occurrence o is bound or free *simpliciter* according as it does or does not contain a bound variable-occurrence that is free in o. And a term-occurrence o is *free in* a term-occurrence o' iff (i) o' contains o and (ii) o contains no variable-occurrence that is bound in o' but free in o.