Higher-Order Quantification and the Elimination of Abstract Objects

Cian Dorr

Draft: 19th February 2024
The final version will appear in special issue of *Disputatio*

Contents

1 Introduction **2**

2 From first-order to higher-order logic, in five steps **4**

3 Higher-Orderese and 'property' talk **10**
3.1 A view from beyond 10
3.2 The scope of the challenge 14

4 Doing natural-language semantics in Higher-Orderese **16**
4.1 Meaning as category-relative 16
4.2 Composition laws 17
4.3 Four kinds of ambiguity 20

5 Developing a deflationary account of property-talk **22**
5.1 Type-ambiguity and property-talk 26
5.2 Why we need category-relativization 29
5.3 Pervasive type-ambiguity and trivial meanings 30
5.4 Type-ambiguity and Russell's paradox 33

6 The problems of mixed co-ordination and quantification **36**
6.1 Preliminaries to a solution 1: co-ordination and lifting 38
6.2 Preliminaries to a solution 2: sum types 40
6.3 Explaining mixed co-ordination with sum-types 42
6.4 Explaining cross-type quantification with sum-types 45
6.5 Dispensing with sum-types 47
6.6 Could we appeal to sum-types to avoid type-ambiguity? 49

7 Worries about schemas and universal generalizations **51**

8 Type-neutral generality in English **55**

9 Conclusion: in defence of type-ambiguity **58**

Appendix A Systems of classical higher-order logic **61**

Appendix B Adding sum-types **66**

Appendix C Dispensing with sum-types **71**

1 Introduction

Formal languages are interesting objects to study mathematically. To do so, one need not actually *understand* the language one is studying in the same sense in which one understands natural languages. One need not, for example, be in a position to use any of its expressions to say things about the world. Nevertheless, sometimes we *do* understand formal languages in this way—the most important roles of formal languages in philosophy in fact depend on such understanding. Moreover, we can sometimes manage to understand a formal language without the aid of some stipulated translation scheme that maps expressions of the formal language to expressions in some natural language we already understand. Indeed, this is the normal case. For example, even though we draw heavily on parallels with English words like 'and', 'or', and 'if' in introducing the standard connectives of propositional logic, we eventually end up with a grasp of these connectives that is sufficiently independent of any particular English words to make room for serious debates about the extent to which the meanings of those English words match those of their formal analogues.

These general remarks apply in particular to the formal languages of *higher-order logic*. Higher-order languages have played a central role in logic at least since Frege (1879), though their popularity dwindled in the second half of the twentieth century, thanks especially to Quine's influential crusade to establish the primacy of first order logic (Quine 1953, 1970). Their distinctive feature is the presence of variables, and associated quantifiers, in every syntactic category.[1] For example, starting with a sentence $\neg(\mathbf{Fs})$ ('Socrates is not foolish') we can not only generate a first-order generalization by replacing the singular term '**s**' with an existentially quantified variable

$$\exists x.\neg(\mathbf{F}x)$$

[1]This isn't perfectly accurate since there are also variable-free versions of higher-order logic where the job normally done by variable-binding is done by primitive combinators: see Bacon 2023 (§3.5).

but also do the same thing to all of the other constituents of the original sentence:

$$\exists X.\neg(X\mathbf{s})$$
$$\exists O.O(\mathbf{Fs})$$
$$\exists p.\neg p$$
$$\exists p.p$$

In teaching higher-order logic, and in trying to find useful sounds to make while pointing to higher-order expressions on handouts and slides, it is helpful to exploit parallels with certain constructions found in natural languages. For example, one might find it helpful to note the analogy between the four formulae exhibited above and the following English sentences:

> Some property is such that Socrates does not have it.
>
> Some property of propositions is such that the proposition that Socrates is foolish has it.
>
> Some proposition is not true.
>
> Some proposition is true.

But such glosses do not provide the only, or even the main, way into an understanding of higher-order logic. Indeed, much of the promise of higher-order logic as a tool for bringing clarity to the statement of philosophical questions and arguments depends on our *not* understanding its sentences as mere stipulative equivalents of English sentences along these lines. The most important pathway to understanding comes, instead, from learning and becoming comfortable with the logical rules for manipulating its expressions—logical rules which generalize the usual rules of first order logic in a straightforward way, and which can play a similarly central role in the constellation of facts about use in virtue of which the symbols mean what they do.[2]

I will not be arguing in this paper for the intelligibility of the language of higher-order logic. Indeed this isn't really something I would know how to argue for: by far the best way to convince people that something is intelligible is to get them to understand it, and understanding isn't a state apt to be inculcated by arguments. I will also not be arguing for the *interest* or *usefulness* of higher-order logic. I do in fact find many questions that can be formulated in higher-order logic intrinsically interesting, and I have found it to be a useful aid to clear thinking and rigorous reasoning in very much the same ways as first-order logic (which it extends). But again, this perspective is best gained through experience.[3]

[2] Williamson (2003) gives a clear statement of this point.

[3] Hofweber (2022) interprets friends of "Higher Order Metaphysics" as holding that we should *stop* asking certain questions posed in English and start investigating questions posed in Higher-Orderese instead. He complains that no adequate reasons have been given for making such a switch.

The present paper does not depend on our being interested in any questions formulated in Higher-Orderese *for their own sakes*. I will (after §2) be assuming that we understand the language and accept a standard system of classical logic for it. But my interest here is in the light which such understanding may shed on certain classic philosophical questions formulated in English: especially, the question whether there are *abstract objects*, such as properties, relations, and propositions. This question has been central to the debate over "nominalism" in metaphysics, especially in the wake of Quine (1953). It is not obvious how anything we can say using the resources of higher-order logic might bear on it. But I will be arguing in this paper that the perspective offered by higher-order logic points the way towards an illuminating, irenic resolution of the debate formulated in English—a resolution on which the key English sentences turn out to be ambiguous, in such a way that the central claims of both nominalists and their opponents admit of true readings.

The argument will proceed as follows. §2 will be a quick introduction to higher-order languages, emphasizing the idea that they can be understood without relying on natural-language glosses using words like 'property', 'relation', and 'proposition'. §3 will introduce and motivate the thesis of this paper, which is that English sentences using such words do, nevertheless, have the higher-order meanings needed for those glosses to be literally correct. The following two sections will address the main challenge for this thesis by showing how it can be derived from a systematic compositional theory of meaning for English, stated in Higher-Orderese: §4 will introduce a general format for such theories, and §5 will implement a higher-order semantics for 'property'-talk within this theoretical architecture. The remainder of the paper (§6–§9) discuss some further challenges and show how the theory may be extended to address them.

2 From first-order to higher-order logic, in five steps

Happily, the situation for philosophers wishing to learn about higher-order logic with a view to understanding it (as opposed to taking it as an object of mathematical investigation) is better now than it was a few years ago: good resources include Bacon 2023, the introduction to Fritz and Jones 2024, and the papers by Bacon (2024) and Goodman (2024) in that volume. Here, nevertheless, is my version of such a primer, aimed at philosophers already familiar with first-order logic. I will proceed in five steps: if at the end you don't feel that you have ended up in the state of understanding I'm hoping to instil, it may at least be interesting to reflect on where you got off the boat.

Step One: types. When we are studying formal languages that extend first-order

I see a false dilemma here: there is no reason why starting to ask new questions would require ceasing to ask any old ones! Of course one might worry that time devoted to the new questions will subtract from the time available for thinking about the old ones, thus slowing progress on them. But this thought neglects the potential for cross-fertilization between the investigations.

logic, it becomes useful to have a general system of nomenclature for the syntactic categories of the expressions we may find ourselves introducing. The standard approach uses special strings called *types* to label the syntactic categories. For our purposes, the so-called "simple types" will suffice: the smallest set of strings that contains the letters e and t and is such that whenever it contains σ and τ, it also contains $(\sigma \to \tau)$ (i.e., the string derived from them by interpolating an arrow and surrounding the result with parentheses).

In referring to simple types, I will omit parentheses, restoring them from right to left: for example I will write '$t \to e \to t$' instead of '$(t \to (e \to t))$'. To save space, I'll also use a more compact notation where, for any types $\sigma_1, ..., \sigma_n$, $\overline{\sigma_1 ... \sigma_n}$ abbreviates $\sigma_1 \to \cdots \to \sigma_n \to t$ (that is, $(\sigma_1 \to (\cdots(\sigma_n \to t)\cdots)))$. Thus, $\overline{e}$ is $e \to t$; $\overline{ee}$ is $e \to e \to t$; $\overline{\overline{e}}$ is $(e \to t) \to t$; $\overline{\overline{e}e}$ is $(e \to t) \to (e \to t)$, etc.

Simple types are useful for labelling syntactic categories, including some syntactic categories already encountered in first-order languages. An expression of type t is a *formula* (a *sentence* if it lacks free variables); one of type e is a *singular term*. An expression of type $\sigma \to \tau$ is something that combines with one of type σ to make one of type τ. So, for example, in first order languages, the negation symbol $\neg$ is of type $t \to t$ (or $\overline{t}$): it can be written in front of a sentence P to make a new sentence $\neg P$. A monadic predicate F is of type $e \to t$ (or $\overline{e}$): it combines with a singular term a to make a sentence Fa.

What about the conjunction symbol $\wedge$, or a dyadic predicate R? Using a convenient trick (pioneered by Frege (1893) and further developed by Curry and Feys (1958)), we can assign simple types to these by thinking of them as taking their arguments one at a time rather than all together. For example, in a conjunctive sentence $P \wedge Q$, we think of the combination of the first sentence P with the conjunction symbol $\wedge$ not as a meaningless string of symbols, but rather as a well-formed expression of type $t \to t$ (just like $\neg$), which is combined with Q to make the sentence. $\wedge$ itself is thus of type $t \to t \to t$ (or $\overline{tt}$). (To keep the official syntax simple, we treat the infix notation $P \wedge Q$ as an unofficial way of writing the string $((\wedge P)Q)$, which explicitly displays the order of pairwise syntactic combination.) Similarly, we think of a binary (first-order) predicate R as having type $e \to e \to t$ (or $\overline{ee}$), with Rab short for $((Ra)b)$. More generally, an n-ary first-order predicate will have type $\overline{e \cdots e}$, with n e's.

This is reminiscent of what we see in natural languages. In the English sentence 'Phobos orbits Mars', 'orbits Mars' is a meaningful unit in its own right—a verb phrase (VP), syntactically just like 'rotates' in 'Phobos rotates'. Admittedly, the standard pedagogy of first-order logic does not encourage us to see a string like Ra (where R is a binary predicate) as a meaningful expression. If you wanted to insist that it wasn't, you could use a richer type system with types appropriate for "primitively polyadic" expressions; the ascent from first-order to higher-order logic can be carried out just as well in this setting. Here, I have found it more convenient to stick to formal languages where every complex expression has exactly two immediate constituents.

Step Two: first-order lambda-abstraction. Natural languages provide a rich array of generating complex predicates. Among VPs, we not only have 'orbits Mars', but also 'doesn't orbit Mars', 'orbits some planet', 'orbits itself', 'is orbited by Phobos', and 'is orbited by something'. While our binary-branching regimentation of first-order logic does contain some complex predicates—namely, those formed by combining an n-ary predicate with fewer than n arguments—it does not yet provide analogues of the other complex VPs above. The next step along the road to higher-order logic is to add a more flexible device for forming complex predicates that yields analogues for all the above English VPs, and more. The standard such approach is to add *lambda predicates*, constructed according to the following rule:

> If P is a formula (expression of type t) and x is an individual variable, $(\lambda x.P)$ is a monadic predicate (expression of type $\overline{e}$).

The following glosses will help illuminate the intended interpretations of these predicates, given a binary predicate **orbits** and singular terms **mars** and **phobos**. Of course in the light of the remarks in the Introduction, you shouldn't take these as stipulative perfect equivalences:

$(\lambda x.\,\mathbf{orbits}\,x\,\mathbf{mars})$	orbits Mars
$(\lambda x.\neg(\mathbf{orbits}\,x\,\mathbf{mars}))$	doesn't orbit Mars
$(\lambda x.\exists y.\,\mathbf{planet}\,y \wedge \mathbf{orbits}\,x\,y)$	orbits some planet
$(\lambda x.\,\mathbf{orbits}\,x\,x)$	orbits itself
$(\lambda x.\,\mathbf{orbits}\,\mathbf{phobos}\,x)$	is orbited by Phobos
$(\lambda x.\exists y.\,\mathbf{orbits}\,y\,x)$	is orbited by something
$(\lambda x.\,\mathbf{orbits}\,\mathbf{phobos}\,\mathbf{mars})$	is such that Phobos orbits Mars

We would also like to be able to construct *polyadic* predicates, which we can do using the following generalization of the above rule:

> If R is an n-ary predicate (expression of type $\tau = \overline{e \cdots e}$, with n e's), $(\lambda x.R)$ is an $n+1$-ary predicate (expression of type $e \to \tau$).

As an abbreviation, we will omit unnecessary parentheses and collapse strings of lambdas into one: $\lambda xyz.A$ is short for $(\lambda x.(\lambda y.(\lambda z.A)))$.

Having introduced these new predicates, we will need some way of reasoning with them. For example, we want to be able to derive the sentence **orbits mars phobos** from $(\lambda x.\,\mathbf{orbits}\,x\,\mathbf{mars})\,\mathbf{phobos}$ and vice versa. The usual such approach is to lay down a rule or axiom of *beta-conversion* that lets us freely interchange any expression of the form $(\lambda v.A)b$ ("beta-redex") with its "reduct", $A[b/v]$—the result of substituting b for every free occurrence of the variable v in the expression A—provided b is "safe" for v in A, in the sense that this substitution will not result in any of b's free variables becoming bound. It is worth noting that with this rule, every sentence involving predicates formed by lambda-abstraction can be proved to be equivalent to some sentence involving no such predicates, suggesting that this step is not a promising place for someone to dig in their heels and plead unintelligibility.[4]

Once we have λ as a variable-binder, it is natural and conceptually helpful to rethink the syntax of the quantifiers in such a way that it is the *only* variable-binder, so that the quantifiers can be treated not as mere punctuation marks but as meaningful expressions on a par with the propositional connectives (see Stalnaker 1977). We can do this by treating $\forall v.P$ and $\exists v.P$ as shorthand for $\forall(\lambda v.P)$ and $\exists(\lambda v.P)$, where $\forall$ and $\exists$ are of type $(e \to t) \to t$ (or $\overline{\overline{e}}$). The important insight here—that first-order quantifiers are higher-order predicates—-is due to Frege (1879), and already informs his understanding of the more familiar variable-binding notation.

Step Three: complex higher-order predicates. A *predicate type* is any type of the form $\sigma_1 \to \cdots \to \sigma_n \to t$ (or $\overline{\sigma_1 ... \sigma_n}$), with $n \geq 0$; a *higher-order predicate type* is any type of this form where at least one σ_i is not e. By a *higher-order predicate* I just mean an expression whose type is a higher-order predicate type. By this definition, the logical constants $\neg, \wedge, \vee, \to$ are higher-order predicates, as are the quantifiers $\forall$ and $\exists$ (when regimented as type-$\overline{\overline{e}}$ constants). Our decision to use binary-branching syntax means we also have some *complex* higher-order predicates—for example, for any formula P, $\wedge P$ is a complex higher-order predicate of type $t \to t$ (or $\overline{t}$). But our ability to generate complex higher-order predicates is rather artificially limited, just as our ability to generate complex first-order predicates was before the introduction of lambda terms.

This limitation is easily remedied. Instead of having every variable be of type e, we can help ourselves to a separate infinite population of variables for each type σ (with the type of a variable indicated by a superscript when necessary), and adopt the following further generalization of the rule for forming complex predicates:

> When A is an expression of a predicate type τ and v is a variable of type σ, $(\lambda v.A)$ is an expression of type $\sigma \to \tau$.

Of course, once we extend the formation rule in this way, we will want to extend the β-conversion rule to cover higher-order as well as first-order predicates. This requires nothing new, except perhaps a reminder that for a redex $(\lambda v.A)B$ even to be well-formed, the term B must have the same type as the variable v. With this rule, it remains true that every sentence we can form with lambda-abstraction is provably equivalent to some sentence without it.

---

[4] In the setting where we have added nothing to first-order logic besides complex predicates, the effect of the β rule can also be got from an axiom-scheme whose instances are biconditionals:

Extensional β $$(\lambda v_1 \dots v_n.P)a_1 \dots a_n \leftrightarrow P[a_1/v_1, \dots, a_n/v_n]$$

where $v_1 ... v_n$ are all distinct, and any free variables in a_i are safe for v_i in P. However, once we add further resources to the languages, such as modal operators, β becomes strictly stronger. In this paper I will be assuming a higher-order logic that uses the full β rule, mostly for reasons of convenience. See Dorr, Hawthorne, and Yli-Vakkuri 2021 (ch. 1) for a weaker system that only has the extensional version of β-conversion.

To keep things looking neat, we will try to avoid using variables that differ only in their type-superscript. This lets us omit the superscripts on bound variable-occurrences, since they can be recovered by finding the λ that binds them.[5]

Allowing complex higher-order predicates built up by lambda-abstraction lets us introduce new binary connectives like the Sheffer stroke ('neither/nor'):

$$\lambda p^t q^t . \neg p \wedge \neg q.$$

new restricted quantifiers

$$\lambda X^{\bar{e}} . \forall z^e . \, \mathbf{red}\, z \to Xz$$

and new higher-order predicates, of the same type as the quantifiers

$$\lambda X^{\bar{e}} . X \, \mathbf{mars}$$

In fact, by successive application of λ, we can form closed terms in every predicate type.

Terms of all these will of course differ as regards which other terms they can grammatically take as arguments. It is worth emphasizing that, contrary to the impression given by many authors (including to some extent Russell (1908)), these restrictions are not something invented to block some paradox that would arise in a system without them. They are motivated in exactly the same way as the grammatical restrictions we are already familiar with in first-order logic, such as the fact that the negation symbol needs to combine with a formula, whereas an ordinary one-place predicate needs to combine with a singular term.[6]

Step Four: non-logical higher-order predicates. In first-order logic, we are encouraged to introduce novel predicates and singular terms as non-logical constants (e.g. by exploiting our pre-existing understanding of corresponding expressions in natural languages), but we are stuck with a fixed stock of constants of other types, namely the standard propositional connectives and quantifiers.[7] But many other terms of other types seem intelligible. And insofar as they are intelligible, it is hard to see any in-principle problem with adding them to the language of first-order logic (either in its original form, or extended with complex predicates as at the previous steps). Most famously, *modal* logic considers languages that add one or more new constants of type $\bar{t}$, which we are can understand by leveraging our understand-

[5] This convention leads to a visual convergence with the approach more common in theoretical computer science, where variables lack built-in types, but expressions have types only relevant to "environments" that assign types to their free variables, and λ-terms (normally written as $(\lambda v : \sigma . A)$) explicitly specify a type for the variable they bind. This approach brings elegance and suggests important avenues of generalization, at the cost of making many syntactic notions annoyingly environment-relative.

[6] Goodman (2024) is particularly clear on this point.

[7] Standard versions of first-order logic also allow function symbols with types like $e \to e$, but these raise special interpretative issues I don't want to get into.

ing of English modal words like ‘necessarily’. Generalizations such as conditional logics add constants taking further type-t arguments (e.g. of type $\overline{tt}$). But there are many other such additions, inspired by other families of expressions in natural language, that *prima facie* suggest regimentation as predicates taking arguments of types other than e and t. One particularly relevant example for this paper involves talk about *meaning* or *reference*—important topics where the kind of rigor made possible by formalization will be particularly welcome. A natural formalization will provide, for each type σ, an “expressing” predicate $\vdots_\sigma$ (naturally taken to be of type $\overline{e\sigma}$), intended to be the analogue for type σ of the first-order reference predicate in ‘“Mars” refers to Mars’. With the aid of these predicates (along with a quotational device for talking about linguistic expressions), we will be able to formalize not only claims about meaning in formal languages, but also claims about meaning in natural languages like is·self-identical $\vdots_{\overline{e}}\ \lambda x^e.x = x$, everything $\vdots_{\underline{\overline{e}}}\ \lambda X^{\overline{e}}.\forall y^e.Xy$, and everything·is·self-identical $\vdots_t\ \forall x^e.x = x$.

Obviously any particular proposed addition must be individually scrutinized for intelligibility. Natural languages are a rich resource, but things can go badly wrong if we try to naïvely import words from natural language into formal languages in types that do not properly reflect their original linguistic role (see Dorr, Hawthorne, and Yli-Vakkuri 2021: §1.3). But I see no good grounds for a sweeping Quinean scepticism that questions the intelligibility of *any* new non-logical expressions other than first-order predicates.

Step Five: higher-order quantifiers and identity. The last and probably most controversial step is to add new *logical* constants, intended to stand to higher types just as the usual identity symbol and quantifiers of first-order logic stand to type e. In particular: for each type σ, we have new logical constants $=_\sigma$ of type $\overline{\sigma\sigma}$, and $\forall_\sigma$ and $\exists_\sigma$ of type $\overline{\sigma}$. (As in the first-order case, $\forall v^\sigma.P$ abbreviates $\forall_\sigma(\lambda v^\sigma.P)$; $\forall u^\sigma v^\tau.P$ abbreviates $\forall_\sigma(\lambda \overline{u}^\sigma.\forall_\tau(\lambda v^\tau.P))$, etc.)

For certain types $\sigma \neq e$, we can find expressions in natural languages which look *prima facie* like close analogues to the type–σ quantifiers and identity predicate. For higher-order quantification, Prior (1971) considers sentences like ‘He is something I am not—kind’ and ‘However he says things are, thus they are’.[8] For higher-order identity, Dorr (2016) considers sentences like ‘To be a water molecule is to a molecule consisting of two hydrogen atoms and one oxygen atoms’ and ‘For there to be vixens is for there to be female foxes’. Such natural-language counterparts provide one potentially helpful entry-point for a more general understanding of higher-order quantification and identity. However, as usual, it would be a mistake to hold the intelligibility of a formal language hostage to our ability to translate each of its sentences into a natural language (Williamson 2003). There are other routes to understanding a new constant, such as grasping its *inferential role* in the linguistic practice of those who use it: what they are happy to infer from a sentence involving the constant, and

[8] See also Rayo and Yablo 2001. For a competing treatment that assimilates the quantifiers in such sentences to first-order quantifiers, see Moltmann 2013 (ch. 3).

what they are happy to infer it from. In the case of the higher-order quantifiers and identity, these inferential roles are easy to master, since they are exactly the same as those of the first-order quantifiers and first-order identity, modulo the change in type. These roles are very rich: as with other logical constants, there is a sense in which they *completely pin down* the interpretation of the symbols that they characterize (see Harris 1982, Dorr 2014*a*).

3 Higher-Orderese and 'property' talk

As we have noted, English glosses using words like 'property', 'relation', and 'proposition' can be be helpful pedagogical tools for helping people understand expressions in Higher-Orderese. But if we understand Higher-Orderese in the way that makes it a valuable tool, it is not obvious that its sentences are equivalent (in meaning, or even in truth value) to their glosses. Nor is it obvious that they are *not* equivalent: it is a tricky question, with reasonable arguments on both sides. It need not be settled by our linguistic competence in both languages, just as it may be an open question for a bilingual native speaker of English and French whether 'There are elms' means the same as 'Il y a des ormes'.[9]

Moreover, it is a philosophically profitable question to investigate. For words like 'property' are widely used in formulating philosophically interesting theses and questions in English, and play an especially central role in certain canonical debates in metaphysics such as the debate over nominalism (often understood as committed to the claim that there are no properties, relations, or propositions). Once we have attained a sufficiently independent understanding of a higher-order language, reflecting on its semantic relationship to these puzzling parts of natural language may help us resolve, or at least better navigate, these classic disputes.

3.1 A view from beyond

A thought experiment will make the investigation a bit more vivid. Imagine we are alien field linguists (from the exoplanet Lambda Serpentis b, say), whose native language is Higher-Orderese. Having travelled to Earth, we are now trying to figure out what English-speakers are trying to tell us using various English sentences. For a wide swath of ordinary language, this project has been going well, thanks to the convenient presence in our native dialect of Higher-Orderese of constants that look like close analogues of various English nouns, verbs, adjectives, and adverbs. For example, we have type-$\bar{e}$ constants $\mathbf{planet}^{\bar{e}}$ and $\mathbf{twinkles}^{\bar{e}}$ with which we can form the sentence $\neg\exists x^{e}.\, \mathbf{planet}^{\bar{e}}\, x \wedge \mathbf{twinkles}^{\bar{e}}\, x$. And we have found strong support for

[9]Someone might argue that if one understands two sentences A and B that both mean p, one will know that A means p and know that B means p, and thus be in a position to come to know that A and B mean the same thing by a deductive inference. But the phenomenon whereby one can think about the same thing under multiple "guises" means that extending one's knowledge in this way need not be so easy: see Quine 1956, Kripke 1979, Salmon 1986.

the hypothesis that when the English-speakers say ‘No planet twinkles’, they mean that this is the case. We express this hypothesis in our version of Higher-Orderese as follows:

$$\text{no·planet·twinkles} \mathrel{⁝_t} \neg\exists x^e.\, \mathbf{planet}^{\bar{e}}\, x \wedge \mathbf{twinkles}^{\bar{e}}\, x \tag{1}$$

Here, no·planet·twinkles is our analogue of the English quote-name ‘“No planet twinkles”’, and $A \mathrel{⁝_t} P$ is our analogue of the English ‘A means in English that P’.[10] Moreover, we have have found plausible ways of explaining these discoveries about the semantic roles of English sentences using hypotheses about the semantic roles of the words from which they are constructed. Whereas English features a rather bewildering variety of terminology for talking about matters of linguistic meaning—‘means’, ‘refers’, ‘denotes’, ‘expresses’...—on Lambda Serpentis we are a bit more systematic: for each type σ, we have a binary predicate $⁝_\sigma$ taking an argument of type σ. Using these predicates, we have formulated some hypotheses about the constituent words of the above sentence, which one can think of as Higher-Orderese analogues of ‘“Planet” means planet’, ‘“Twinkles” means twinkles’, and ‘“No” means no’:

$$\text{planet} \mathrel{⁝_{\bar{e}}} \mathbf{planet}^{\bar{e}} \tag{2}$$

$$\text{twinkles} \mathrel{⁝_{\bar{e}}} \mathbf{twinkles}^{\bar{e}} \tag{3}$$

$$\text{no} \mathrel{⁝_{\overline{\bar{e}\bar{e}}}} \lambda X^{\bar{e}} Y^{\bar{e}}.\neg\exists z^e.Xz \wedge Yz \tag{4}$$

We have also formulated some general compositional laws, which can be combined with these hypotheses to derive claims about the meanings of complex expressions, such as (1).[11]

But we are perplexed when we try to extend our theory to English expressions like ‘property’, ‘relation’, ‘proposition’, ‘fact’, and so on. We have not found any analogues of these words in our native language. Indeed it is not even obvious which of our variously-typed ‘expressing’ predicates $⁝_\sigma$ we should be using in trying to explain their communicative roles in English.

To flesh out the thought experiment a bit more, let’s imagine that our field linguists are, like many Lambda Serpentian philosophers, adherents of the following thesis:

e-Materialism $\qquad \forall x^e.\, \mathbf{material}^{\bar{e}}\, x$

Here, ‘$\mathbf{material}^{\bar{e}}$’ is a type-$\bar{e}$ constant whose usage is broadly analogous to that of ‘material object’ in English.[12] Of course, e-Materialism does not command universal

[10] Presumably, the binary predicate $⁝_t$ is not actually a primitive constant in our language, but is derived from a ternary predicate with an extra argument for a linguistic community by plugging in our name for the community of English speakers.

[11] See §4.2 below for one possible version of such laws.

[12] We can imagine that philosophers propose slightly different explanations of what it is to be

assent. But its main detractors are proponents of various traditional religious views that no longer have many adherents in educated circles on Lambda Serpentis b.

The sociological situation there thus contrasts interestingly with the situation facing philosophers on Earth who propound theses like

Materialism Everything is a material object.

Materialism is *not* so popular, since many philosophers take it to be refuted by simple arguments like the following:

> Redness isn't a material object; so not everything is a material object.
>
> The fact that there are dogs isn't a material object; so not everything is a material object.
>
> The number two isn't a material object; so not everything is a material object.
>
> The word 'cat' isn't a material object; so not everything is a material object.

Proponents of *e*-Materialism on Lambda Serpentis b never have to address arguments analogous to these, since their dialect of Higher-Orderese has no *type-e* terms that are at all analogous to 'redness', 'the fact that there are dogs', 'the number two', or 'the word "cat"'. The closest analogues to these English expressions in their language are terms of higher types. For example, they just get by with a type $\bar{e}$ predicate $\mathbf{red}^{\bar{e}}$, and have no need for anything like the English nominalizing device '-ness'. In situations where we might utter an English sentence with 'redness' as its grammatical subject, they will reach for a Higher-Orderese sentence in which $\mathbf{red}^{\bar{e}}$ occurs as the argument of a type-$\bar{\bar{e}}$ predicate. Similarly, they may have a system of Arabic numerals that look like ours, but they use them as type-$\bar{\bar{e}}$ terms—for example, treating '2' as equivalent to $\lambda X^{\bar{e}}.\exists y^e z^e.Xy \wedge Xz \wedge y \neq_e z$.[13] And the closest analogue of quotation in their language is a system of special constants which can be used to form terms of type $\bar{e}$. For example, for them, cat (written in the special font) is a term of type $\bar{e}$ which they can truly apply to blobs of ink shaped like this: cat.[14]

Given this background and their Materialist$_e$ sympathies, it's clear why our field linguists will be puzzled by English expressions like 'redness', 'the property of being red', 'the fact that there are dogs', 'the number two', and 'the word "cat"'. They seem to make a meaningful contribution to many straightforwardly meaningful sentences. And in view of their *syntactic* resemblance to expressions like 'Mars', 'this dog', and 'he', it is natural to assume that their semantic role would consist in "referring to objects", i.e. that if they are meaningful at all, then $\exists x^e$.redness $\vdots_e$ x, $\exists x^e$.the·number·two $\vdots_e$

$\mathbf{material}^{\bar{e}}$, with some favouring geometrical vocabulary, some mereological vocabulary, some drawing on theoretical terms drawn from their best theories of physics, etc.

[13] They may have a separate numeral system for each type, distinguished with subscripts, e.g. treating '2_σ' as equivalent to $\lambda X^{\bar{\sigma}}.\exists y^\sigma z^\sigma.Xy \wedge Xz \wedge y \neq_\sigma z$.

[14] The symbol $\cdot$ is a constant of type $\bar{e} \to \bar{e} \to \bar{e}$ such that $X^{\bar{e}} \cdot Y^{\bar{e}}$ holds of blobs of ink consisting of an $X^{\bar{e}}$ blob followed by a $Y^{\bar{e}}$ blob. For example,cat·cat applies to blobs of ink shaped like this: cat cat.

x, etc. Likewise, in view of the syntactic resemblance between words like 'property', 'fact', and 'number' and ordinary nouns like 'planet', and the syntactic resemblance between words like 'instantiates' and ordinary transitive verbs like 'orbits', it is natural to assume that the former have meanings of type $\bar{e}$ and the latter meanings of type $\overline{ee}$, if they are meaningful at all. The problem is that given *e*-Materialism, there don't seem to be any good candidate meanings of these types. No material object looks like a plausible candidate to be referred to by 'redness' or in the extension of 'property', and no pair of material objects seems a plausible candidate to be in the extension of 'instantiates'.

In view of these facts, some of the field linguists —call them "inflationists"—might be led to the view that most English-speakers are in the grip of what, from their point of view, will look like an astonishingly bold and specific metaphysical hypothesis, radically unlike their own Materialist$_e$ outlook. According to this hypothesis, not all individuals are material: instead, there is a vast array of *non*-material individuals—*abstract objects*. Among these are *universals*, such as redness. Material objects (like planets) can *have* or *instantiate* abstract objects (like universals). And this happens according to some quite distinctive laws: for example, that there is a property (redness) that is, necessarily, instantiated by all and only red objects:

$$\exists x^e.\Box\forall y^e.\, \mathbf{red}^{\bar{e}}\, y \leftrightarrow y\, \mathbf{inst}^{\overline{ee}}\, x$$

The closest parallel to this on Lambda Serpentis involves certain religious systems whose believers were, in a somewhat similar way, committed not only to the bare existence of non-material objects, but to some suspiciously specific claims about the natures of these objects and about how they stand towards material objects.

Others—call them "deflationists"—will have misgivings about the attribution of this view to English-speakers. They will find it hard to believe that such a demanding and *prima facie* implausible belief-system could persist on Earth with so little scepticism. Of course, if they visit some philosophy departments, they will eventually find some who seem to be the analogues of atheists, namely *nominalists*, who from time to time can be heard to produce sentences like 'There are no properties', 'There are no abstract objects', 'Everything is material', etc. But even these oddballs seem to have trouble sticking to their convictions: when not safely ensconced in a "philosophy room" (Lewis 1983*b*: x), their use of the problematic vocabulary is more or less indistinguishable from everyone else's.

The deflationists will also be impressed by the extensive parallels between the ways English-speakers use many sentences involving words like 'property' and the use on Lambda Serpentis of certain Higher-Orderese sentences, namely those of which they could serve as "glosses" according to the usual pronunciation schemes. For example, English-speakers' use of 'Some property of Mars applies to Venus'—a sentence which, according to the inflationists, expresses a tendentious claim relating Mars and Venus to the supposed realm of abstract objects—has a lot in common with their own use of the sentence $\exists X^{\bar{e}}.X\, \mathbf{mars}^e \wedge X\, \mathbf{venus}^e$. Just as they would unhesitatingly infer the latter sentence from anything of the form $F\,\mathbf{mars}^e \wedge F\,\mathbf{venus}^e$ (where F is some type $\bar{e}$ term), English speakers seem—with the exception of a few philosophers in those special rooms—to unhesitatingly infer the former sentence from any sentence of the form 'Mars VP and Venus VP' (where VP is any verb phrase). The deflationists see this as weighty evidence that that's what the sentence means in English: i.e.,

$$\text{some·property·of·Mars·holds·of·Venus} \mathrel{\vdots_t} \exists X^{\bar{e}}.X\,\mathbf{mars}^e \wedge X\,\mathbf{venus}^e$$

More generally, they propose that a wide variety of English sentences involving words like 'property' express exactly what the Higher-Orderese sentences of which they could serve as "glosses" express, so that their truth is entirely consistent with *e*-Materialism.

While the deflationary hypothesis has obvious attractions, its proponents face a challenge. Can they derive their proposals about the meanings of the puzzling sentences from proposals about their constituent words, together with simple, principled generalizations about the semantic role of syntactic composition? The challenge is pressing one: if their proposal required some radical departure from the standard vision of semantic compositionality, that would be a serious count against it. In §5, I will attempt to show how this challenge can be met, by sketching a theory, constructed along standard compositional lines, that has the desired claims about the meanings of the puzzling sentences as theorems. The basic idea—not in the least novel!—is that the syntactic similarities we have noted are *not* indicative of any kind of semantic similarity. For example, despite its resemblance to 'Mars', 'the property of being red' doesn't refer to any object, but rather expresses exactly what the verb phrase 'is red' expresses:

$$\text{the·property·of·being·red} \mathrel{\vdots_{\bar{e}}} \mathbf{red}^{\bar{e}}$$

Similarly, despite its resemblance to 'picture of Mars', 'property of Mars' has a higher-type meaning:

$$\text{property·of·Mars} \mathrel{\vdots_{\underline{\bar{e}}}} \lambda X^{\bar{e}}.X\,\mathbf{mars}^e$$

Many other expressions also end up with meanings in semantic types other than what purely syntactic considerations might lead one to expect. And the compositional rules are set up in a way that is general enough to allow these meanings to combine to spit out the desired "deflationary" meanings for sentences.

3.2 The scope of the challenge

Although "property talk" is widespread in philosophy, it is not particularly common in ordinary life; Moltmann (2013) classifies it as part of the "periphery" rather than the "core" of natural language. Given this, the task of providing a semantic account of this particular family of expressions may not seem a very high priority for our field linguists, or for their real-world counterparts in linguistics departments. But there are many phenomena closer to the "core" that generate similar puzzles: namely, families of expressions that seem capable of playing similar *syntactic* roles to ordinary proper names like 'Mars', or to ordinary predicates like 'planet', but where if one took these syntactic similarities seriously as a guide to the type of the expressions' meanings, one would be forced to interpret some wide range of ordinary sentences as expressing claims inconsistent with *e*-Materialism. I have already mentioned expressions like 'number', 'word', 'sentence', 'the number two', and '"cat"'. Here are some more kinds of expressions that can sometimes play a name-like syntactic role (e.g. as the subject of a verb), and which some semanticists have treated as denoting special individuals, of a sort *prima facie* ruled out by *e*-Materialism:

- Bare gerunds ('being polite') and infinitives ('to be polite') (Chierchia 1984), and more complex nominals built around them ('John being polite', 'John's being polite', 'Brutus's stabbing Caesar', 'Brutus's stabbing of Caesar', 'for John to be polite') (Vendler 1967).
- Bare 'that' clauses ('that it is raining') (Cresswell 1973: 165-169) and more complex nominals built around them ('the belief that it is raining', 'John's belief that it is raining', 'a rumor that I have resigned') (Moltmann 2013).[15]
- Plural definites ('John and Mary', 'the cards') (Link 1983).
- "Kind-denoting" uses of definite singular noun phrases ('the Siberian tiger').
- Bare plurals and mass nouns ('dinosaurs', 'dinosaurs with feathers' 'clean water') (Carlson 1977, Liebesman 2011).
- Bare abstract nouns ('redness', 'hostility'), and more complex nominals built around them ('the wisdom of Socrates') (Moltmann 2013).

We could also add a host of ordinary expressions to which many semanticists would assign denotations in types taking individual arguments (such as $\bar{e}$), which would *prima facie* be uninstantiated given *e*-Materialism, leading to trivialization in the truth conditions:

- Count-noun-like uses of gerunds ('Every killing by John was justified').
- Other kinds of nominalizations of verbs ('Every dance was beautiful', 'The examination was sat in the the Exam Schools', 'The treatment was successful').
- Other nouns that don't seem to apply straightforwardly to material objects ('problem', 'error', 'battle', 'meeting', ...).

[15]Rosefeldt (2008) takes the view that 'that' clauses have type-*e* denotations to be a commitment of the orthodox "relational theory" of propositional attitudes. He argues that they have a different type—the analogue in his system of our *t*—while assuming that expressions of the form 'the proposition that...' have only type-*e* denotations.

- Verbs and adverbs, according to the influential tradition of Davidson (1980) and Parsons (1990).

Property-talk can be seen as a minimally distracting test case, where the pressure towards semantic homogeneity is especially strong, and the general shape of a "deflationary" alternative especially clear. For all of the other phenomena just listed, there will be similar possibilities for deflationary treatments compatible with *e*-Materialism, though in each case filling in the details will involve significant further choice points.

4 Doing natural-language semantics in Higher-Orderese

Before turning to the special challenges that words like 'property' raise for the project of giving a rigorous semantic theory for English stated in a higher-order language, we need to get on the table some of the basic architectural principles of such a theory, specifically principles connecting the semantic roles of complex expressions with the semantic roles of their constituents.

4.1 Meaning as category-relative

So far, we have been theorizing using a family of binary "expressing" predicates $\vdots_\sigma$ of type $\overline{\epsilon\sigma}$, where ϵ is the type of quote-names like no·planet·twinkles. (We posited that on Lambda Serpentis this is $\bar{e}$, but from now on I'll just write ϵ for the sake of greater neutrality.) But for the purposes of systematic theorizing about human languages, there is reason to want a more discriminating ideology. Consider for example the word 'fish'. It can function both as a noun and as an intransitive verb, and makes different semantic contributions when it occurs in these different capacities. Any systematic theory of the meanings of complex expressions will need to somehow be sensitive to this, e.g., in order to explain why the sentences 'You fish' and 'You are a fish' mean different things.

To provide a canonical format for characterizing this sort of relativity, I will use a more complicated family of primitive *three*-place semantic predicates $⁞_\sigma$, that add an extra argument place for a *syntactic category*, in order to track the differences that matter for words like 'fish'. It doesn't matter much what these syntactic categories *are*, since their role is just to serve as labels to aid in disambiguation. But for convenience, we will identify syntactic categories with strings—e.g. n for noun and s for sentence—making $⁞_\sigma$ a predicate of type $\overline{\epsilon\epsilon\sigma}$.[16] So we can write down sentences like these (writing the symbol $⁞_\sigma$ between its first and second argument):

$$\mathsf{planet}\ ⁞_{\bar{e}}\ \mathsf{n}\ \mathbf{planet}^{\bar{e}}$$
$$\mathsf{no{\cdot}planet{\cdot}twinkles}\ ⁞_t\ \mathsf{s}\ \neg\exists x^e.\ \mathbf{planet}^{\bar{e}}\, x \wedge \mathbf{twinkles}^{\bar{e}}\, x$$

[16]We can understand the ternary predicates $⁞_\sigma$ as derived from more basic quaternary predicates by a name for the community of English speakers into one argument, analogous to what we suggested above for the binary predicates $\vdots_\sigma$ (see note 10 above).

I will often omit the subscript σ when it can be reconstructed by checking the type of the third argument.[17]

The basic theoretical role of these new predicates is given by the following schema, according to which what an expression means is what it means relative to some category or other:

$$a \mathrel{\vdots}_\sigma x \leftrightarrow \exists c^\epsilon . a \mathrel{\vdots}_\sigma c\, x$$

Of course, this falls well short of pinning down the meaning of $\vdots_\sigma$. But as with other theoretical terms, the hope is the meaning of the new vocabulary will be sufficiently constrained by its role in a theory.

4.2 Composition laws

A nice thing about keeping explicit track of syntactic categories in our semantic ideology is that it lets us formulate some quite simple and general compositional laws. To do so, we can follow Bar-Hillel (1953) and Lambek (1958) in allowing for complex syntactic categories, built up from simpler ones using parentheses and two different directions of slashes.

$$/ : \varepsilon \to \varepsilon \to \varepsilon$$
$$\backslash : \varepsilon \to \varepsilon \to \varepsilon$$

When c and d are strings, c/d is just the string formed by concatenating a left parenthesis, c, the forward slash character, d, and a right parenthesis (in that order); $c\backslash d$ is the same but using the backward slash. The idea is that expressions of categories c/d and $d\backslash c$ are ones that can combine with an expression of category d to yield one of category of c, with the former taking its input on the right and the latter taking its input on the left. In writing long expressions built up using the slashes, we economize on parentheses by treating \ as binding more tightly than /; / as left-associative; and \ as right-associative.[18]

Appealing to complex syntactic categories, we can formulate two simple schemas that can be used to derive claims about the meanings of complex expressions from

[17] The idea of meaning as relativized to something like a syntactic category is pervasive in the literature on categorial grammar. For example, in the system of Morrill (1994), the basic items that occur in derivations are tripartite "assignments" comprising a "prosodic object", a "semantic object", and a "category form" (other authors (Moortgat 1996, Kubota and Levine 2020) call these tripartite things "signs"). In the philosophical literature, Rieppel (2016) argues for a relativization of the generic notion of denotation to something he calls an "expression type", which looks quite like what I am calling a "syntactic category". Note that Rieppel uses 'syntactic category' differently, such that certain expressions that have denotations relative to several different "expression types" nevertheless are assigned to a single "syntactic category". I doubt anything substantive turns on this.

[18] Thus, e.g., $a\backslash b/c$ is $(a\backslash b)/c$, $a/b\backslash c$ is $a/(b\backslash c)$, and $a/b\backslash c\backslash d/e/f$ is $((a/(b\backslash(c\backslash d)))/e)/f$.

claims about the meanings of their simpler parts:

Functional Applicaton (>) $a \vdots_{\sigma\to\tau} c/d\, X \wedge b \vdots_{\sigma} d\, y \to a{\cdot}b \vdots_{\tau} c\, Xy$

(<) $a \vdots_{\sigma\to t} d\backslash c\, X \wedge b \vdots_{\sigma} d\, y \to b{\cdot}a \vdots_{\tau} c\, Xy$

Gloss: when a has a type $\sigma \to \tau$ meaning X and b has a type-σ meaning y, and moreover the categories relative which they have these meanings are such as to license the combination, the result of concatenating a and b in the appropriate order has the type-τ meaning that results from applying X to y.[19]

Here is a illustrative example that shows how we might use Function Application to explain a semantic fact about a sentence by appeal to semantic postulates about its constituent words:

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\text{no} \vdots (\text{s/np}\backslash\text{s})/\text{n}\\ \lambda X^{\bar e} Y^{\bar e}.\neg\exists z^e.Xz \wedge Yz\end{array} \quad \begin{array}{c}\text{planet} \vdots \text{n}\\ \mathbf{planet}^{\bar e}\end{array}}{\begin{array}{c}\text{no}{\cdot}\text{planet} \vdots \text{s/np}\backslash\text{s}\\ (\lambda X^{\bar e} Y^{\bar e}.\neg\exists z^e.Xz \wedge Yz)\,\mathbf{planet}^{\bar e}\end{array}}\,>}{\begin{array}{c}\text{no}{\cdot}\text{planet} \vdots \text{s/np}\backslash\text{s}\\ \lambda Y^{\bar e}.\neg\exists z^e.\,\mathbf{planet}^{\bar e}\, z \wedge Yz\end{array}}\,\text{conv} \quad \begin{array}{c}\text{twinkles} \vdots \text{np}\backslash\text{s}\\ \mathbf{twinkles}^{\bar e}\end{array}}{\dfrac{\begin{array}{c}\text{no}{\cdot}\text{planet}{\cdot}\text{twinkles} \vdots \text{s}\\ (\lambda Y^{\bar e}.\neg\exists z^e.\,\mathbf{planet}^{\bar e}\, z \wedge Yz)\,\mathbf{twinkles}^{\bar e}\end{array}}{\begin{array}{c}\text{no}{\cdot}\text{planet}{\cdot}\text{twinkles} \vdots \text{s}\\ \neg\exists z^e.\,\mathbf{planet}^{\bar e}\, z \wedge \mathbf{twinkles}^{\bar e}\, z\end{array}}\,\text{conv}}\,>$$

Here the premises correspond to our earlier hypotheses about 'planet', 'twinkles', and 'no', but add explicit proposals about the categories relative to which they express what they express. The steps labelled with < or > become valid in H when the appropriate instance of Function Application is added as a supplementary premise. Those labelled with 'conv' are just applications of the β-conversion rule of H. (From now on, I will keep these derivations compact by skipping steps of β-conversion.)

The general structure of derivations like the one given above will look familiar to anyone who knows a bit of natural-language semantics, especially in any of

[19] Some recent work in the tradition of categorial grammar (e.g. de Groote 2001, Worth 2014) uses a different system on which the bearers of meaning—in our system, the first argument of the expressing constant $\vdots$—is not always a string, but can be of a more complicated type such as that of functions from strings to strings—e.g. we might have ($\lambda a^e.a \cdot$ twinkles) $\vdots$ np\s $\mathbf{twinkles}^{\bar e}$. In such systems, we can replace the directional slashes with an all-purpose symbol $\multimap$ and replace the two Function Application schemas with

$$U^{\sigma'\to\tau'} \vdots_{\sigma\to\tau} d\multimap c\, X \wedge v^{\sigma'} \vdots_{\sigma} d\, y \to Uv \vdots_{\tau} c\, Xy$$

Since the phenomena we are concerned with do not require the flexibility such systems confer, I am sticking with the more familiar picture where the bearers of meaning are always strings. But the theory below can straightforwardly be modified to fit the picture where meaning-bearers themselves come in a variety of types.

the traditions influenced by Montague (1974). In these traditions, the central activity is the construction of formal systems that allow derivations roughly like ours. However, the prefatory remarks about the intended theoretical significance of these derivations often diverge quite drastically from the perspective adopted here. Following Montague, the systems in question are widely understood as specifications of a "translation function" from the relevant natural language to a formal language such as that of higher order logic. On this understanding, both the natural-language expression that in my notation appears as the first argument of $\vdots_\sigma$ and the formal-language expression that appears as its last argument are taken to be *mentioned*: the derivation establishes that the translation function maps the former to the latter. For Montague, "semantics" is set-theoretic model theory, and the translation from a natural to a formal language is just a component in an indirect way of specifying the required set-theoretic definitions:

> We could...introduce the semantics of our fragment directly; but it is probably more perspicuous to proceed indirectly, by (1) setting up a certain simple artificial language, that of tensed intensional logic, (2) giving the semantics of that language, and (3) interpreting English indirectly by showing in a rigorous way how to translate it into the artificial language. (Montague 1974: 256)

By contrast, for us, the formal-language expression that appears as the last argument of $\vdots_\sigma$ is *used*, not mentioned: the conclusion of the derivation is itself a sentence in the same formal language. Set-theoretic model-theory plays no role; like Burgess (2008), I find Tarski's decision to commandeer the word 'semantics' for that enterprise extremely unfortunate. The questions our field linguists are interested in are certainly not questions about models, and nor do models play any role in the theories they come up with to answer those questions.

As in the derivation above, we will often help ourselves to constants of Higher-Orderese homophonically corresponding to individual English "non-functional" words (above, 'planet' and 'twinkles'). While we can understand these as genuine constants, reflecting a convenient overlap between the language we are theorizing in and the language we are theorizing about, we could alternatively take them as mere placeholders. On this approach, the conclusion of the above derivation is really short for a universal generalization:

$$\forall X^{\bar{e}} Y^{\bar{e}}.\mathsf{planet} \vdots \mathsf{n}\, X \wedge \mathsf{twinkles} \vdots \mathsf{np\backslash s}\, Y \rightarrow \mathsf{no{\cdot}planet{\cdot}twinkles} \vdots \mathsf{s}\, \neg\exists z^e.Xz \wedge Yz$$

or perhaps the conjunction of this with the claim that the relevant words do have meanings of those types relative to those categories, i.e.:

$$\exists X^{\bar{e}} Y^{\bar{e}}.\mathsf{planet} \vdots \mathsf{n}\, X \wedge \mathsf{twinkles} \vdots \mathsf{np\backslash s}\, Y$$

The earlier lines in the derivation can be understood in parallel fashion. Interpreted like this, the claims we are deriving can be expressed in a language with a fairly restricted range of constants: the logical constants; the expressing constants $\vdots_\sigma$; the string-building operations $\cdot$, /, and \, and names for words and basic categories.

4.3 Four kinds of ambiguity

Our field-linguists won't get far before finding reason to posit cases of *ambiguity* or semantic multiplicity.[20] In principle there are four forms this could take for a given expression a:

(i) a expresses distinct things of a certain type σ relative to distinct categories:

$$\exists x^\sigma y^\sigma c^\epsilon d^\epsilon.\, a \vdots_\sigma c\, x \wedge a \vdots_\sigma d\, y \wedge x \neq_\sigma y \wedge c \neq_\epsilon d$$

(ii) a expresses distinct things of a certain type σ relative to a single category:

$$\exists x^\sigma y^\sigma c^\epsilon.\, a \vdots_\sigma c\, x \wedge a \vdots_\sigma c\, y \wedge x \neq_\sigma y$$

(iii) a expresses things of distinct types σ and τ relative to distinct categories:

$$\exists x^\sigma y^\tau c^\epsilon d^\epsilon.\, a \vdots_\sigma c\, x \wedge a \vdots_\tau d\, y$$

(iv) a expresses things of distinct types σ and τ relative to a single category:

$$\exists x^\sigma y^\tau c^\epsilon.\, a \vdots_\sigma c\, x \wedge a \vdots_\tau c\, y$$

An example of (i) would be the distinct type-$\bar{e}$ meanings of 'fish' as noun (category n) and as verb phrase (np\s). An example of (ii) would be the distinct meanings of 'bat' (as noun). A possible example of (iii) would be the type-$\bar{e}$ meaning of 'study' as a verb phrase (np\s) and its type-$\overline{ee}$ meaning as transitive verb (np\s/np).

There are no similarly uncontroversial examples of type-(iv) ambiguity. Indeed, following Montague (1974), semantic theories using categories like ours have generally disallowed even *distinct* expressions from having meanings of different types relative to the same category. Such systems feature a *category-to-type-correspondence*: each category-label C is mapped to a unique type σ_C, such that no expression has a meaning relative to C in any type other than σ_C.[21] In our formal language, this is captured by the following schema:

Correspondence $\quad \neg(a \vdots_\sigma c\, x \wedge b \vdots_\tau c\, y) \quad$ where $\sigma \neq \tau$

[20] Some authors reserve 'ambiguity' only for cases like 'bat' and 'bank' where the use of a single sound and spelling to play multiple roles is just some arbitrary idiosyncrasy of the sort that would be unlikely to arise independently in causally isolated linguistic communities: they will want to use some more neutral expression like 'ambiguity or polysemy or context-sensitivity' where I have 'ambiguity'.

[21] For a categorial formalism built around such a correspondence, see de Groote 2001. By contrast, Linear Categorial Grammar (Worth 2014) allows in principle for a single category ("tectotype") to be associated with multiple semantic types.

While this may a helpful architectural principle to adopt, it is worth noting that nothing will *break* if we reject Correspondence. For an example of a case where one might be tempted to do so, consider so-called relational nouns, such as ‘author’, ‘friend’, and ‘satellite’. We will need to give these type-$\bar{e}$ meanings if we want treat sentences like ‘Some author dances’ according to the same model as ‘Some planet twinkles’ above. But if these are the only basic postulates we have to work with, some semantic facts about relational nouns will be hard to explain. For example, a type-$\bar{e}$ meaning for ‘satellite’ would make it very hard to understand how ‘Phobos is a satellite of Mars’ gets to mean the same as ‘Phobos orbits Mars’.[22] The standard alternative is to instead make ‘satellite’ type-ambiguous, positing an additional type-$\overline{ee}$ meaning $\lambda x^e y^e.\, \mathbf{orbits}^{\overline{ee}}\, yx$—the same as that of the transitive verb ‘orbits’. If we haven’t signed up to Correspondence, we will then face the question whether to assign the type-$\overline{ee}$ meaning to the same category n as the type-$\bar{e}$ meaning. If we did so, we could derive the appropriate meaning for ‘satellite of Mars’ as follows:

$$\dfrac{\begin{array}{c}\text{satellite} \vdots \text{n}\\ \lambda x^e y^e.\, \mathbf{orbits}^{\overline{ee}}\, yx\end{array} \qquad \dfrac{\begin{array}{c}\text{of} \vdots \text{n\textbackslash n/np}\\ \lambda x^e Y^{\overline{ee}}.\, Yx\end{array} \quad \begin{array}{c}\text{Mars} \vdots \text{np}\\ \mathbf{mars}^e\end{array}}{\begin{array}{c}\text{of·Mars} \vdots \text{n\textbackslash n}\\ \lambda Y^{\overline{ee}}.\, Y\,\mathbf{mars}^e\end{array}}>}{\begin{array}{c}\text{satellite·of·Mars} \vdots \text{n}\\ \lambda z^e.\, \mathbf{orbits}^{\overline{ee}}\, z\,\mathbf{mars}^e\end{array}}<$$

(Of course, we will presumably need other meanings for *of*, such as of⁝n\n/np $\lambda x^e Y^{\bar{e}} z^e.\, Yz \wedge \mathbf{owns}^{\overline{ee}}\, xz$, to handle expressions like ‘sword of King Arthur’.) Lumping the differently-typed meanings into the same category seems natural insofar as we are trying to be guided by syntax as much as we can—it is not obvious that there is any difference in syntactic distribution between nouns like ‘satellite’ that admit relational uses and nouns like ‘sword’ that do not. However, it is probably too much to expect all decisions about category-labels to be justified on entirely non-semantic grounds, and it may ultimately prove better to use some other label for the relational meanings.[23] Moreover,

[22]We would presumably have to posit that of ⁝ n\n/np O, where O is some operation such that $\lambda x^e y^e.O\, x\, \mathbf{satellite}^{\bar{e}}\, y =_{\overline{ee}} \lambda x^e y^e.\, \mathbf{orbits}^{\overline{ee}}\, yx$. The difficulty is in coming up with an operation like this that handles the whole range of cases to be explained.

[23]One problem with the lumping-together proposal is that some relational nouns can arguably have their extra argument supplied by prepositions other than ‘of’. For example, perhaps ‘letter’ in ‘letter to Jane’ also contributes a relation of which ‘Jane’ gets plugged into one argument. Difference of category-label provide a place we can keep track of such lexical differences, e.g. as follows:

$$\dfrac{\text{satellite} \vdots \text{n}/pp_{\text{of}}\; \lambda X^{\bar{e}} y^e.\, X(\lambda z^e.\, \mathbf{orbits}^{\overline{ee}}\, yz) \qquad \dfrac{\text{of} \vdots pp_{\text{of}}/\text{np}\; \lambda x^e Y^{\bar{e}}.\, Yx \quad \text{Mars} \vdots \text{np}\; \mathbf{mars}^e}{\text{of·Mars} \vdots pp_{\text{of}}\; \lambda Y^{\bar{e}}.\, Y\,\mathbf{mars}^e}>}{\text{satellite·of·Mars} \vdots \text{n}\; \lambda y^e.\, \mathbf{orbits}^{\overline{ee}}\, y\,\mathbf{mars}^e}>$$

Similarly we could assign the relational meaning of ‘letter’ the category n/pp_{to}.

since the category-labels are part of the internal bookkeeping of the theory, we always have the option of preserving Correspondence by fine-graining the labels (e.g., replacing the single category n with two categories $\mathsf{n}_{\bar{e}}$ and $\mathsf{n}_{\overline{ee}}$).

One further point about the treatment of ambiguity in this system is worth noting: each time we encounter an occurrence of an ambiguous word, we get to make a fresh choice from among all its meanings that are of an appropriate type and are had relative to an appropriate category. For example, assuming 'bank' is ambiguous (as a noun) between money bank ($\mathbf{mbank}^{\bar{e}}$) and river bank ($\mathbf{rbank}^{\bar{e}}$), we will be able to derive a true meaning for 'Some bank lends and no bank lends':

$$\frac{\text{some} \vdots (\mathsf{s/np\backslash s})/\mathsf{n} \quad \lambda X^{\bar{e}} Y^{\bar{e}}.\exists z^e.Xz \wedge Yz \qquad \text{bank} \vdots \mathsf{n} \quad \mathbf{mbank}^{\bar{e}}}{\text{some}\cdot\text{bank} \vdots \mathsf{s/np\backslash s} \quad \lambda Y^{\bar{e}}.\exists z^e.\,\mathbf{mbank}^{\bar{e}}\,z \wedge Yz}>$$

$$\frac{\text{some}\cdot\text{bank} \vdots \mathsf{s/np\backslash s} \quad \lambda Y^{\bar{e}}.\exists z^e.\,\mathbf{mbank}^{\bar{e}}\,z \wedge Yz \qquad \text{lends} \vdots \mathsf{np\backslash s} \quad \mathbf{lender}^{\bar{e}}}{\text{some}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s} \quad \exists z^e.\,\mathbf{mbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z}>$$

$$\frac{\text{no} \vdots (\mathsf{s/np\backslash s})/\mathsf{n} \quad \lambda X^{\bar{e}} Y^{\bar{e}}.\neg\exists z^e.Xz \wedge Yz \qquad \text{bank} \vdots \mathsf{n} \quad \mathbf{rbank}^{\bar{e}}}{\text{no}\cdot\text{bank} \vdots \mathsf{s/np\backslash s} \quad \lambda Y^{\bar{e}}.\neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge Yz}>$$

$$\frac{\text{no}\cdot\text{bank} \vdots \mathsf{s/np\backslash s} \quad \lambda Y^{\bar{e}}.\neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge Yz \qquad \text{lends} \vdots \mathsf{np\backslash s} \quad \mathbf{lender}^{\bar{e}}}{\text{no}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s} \quad \neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z}>$$

$$\frac{\text{and} \vdots \mathsf{s\backslash s/s} \quad \lambda p^t q^t.q \wedge p \qquad \text{no}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s} \quad \neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z}{\text{and}\cdot\text{no}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s\backslash s} \quad \lambda p^t.p \wedge \neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z}>$$

$$\frac{\text{some}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s} \quad \exists z^e.\,\mathbf{mbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z \qquad \text{and}\cdot\text{no}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s\backslash s} \quad \lambda p^t.p \wedge \neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z}{\text{some}\cdot\text{bank}\cdot\text{lends}\cdot\text{and}\cdot\text{no}\cdot\text{bank}\cdot\text{lends} \vdots \mathsf{s} \quad (\exists z^e.\,\mathbf{mbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z) \wedge \neg\exists z^e.\,\mathbf{rbank}^{\bar{e}}\,z \wedge \mathbf{lender}^{\bar{e}}\,z}>$$

Often (though not always) speakers and hearers prefer *uniform* interpretations, in which repeated occurrences of a word are interpreted in the same way (types and categories permitting), and in which related words (e.g. 'necessary' and 'possible') are interpreted in co-ordinated ways. This notion of uniformity is particularly important when we are interested in logical notions like consequence and contradictoriness (see Dorr 2014*b*). However, providing a rigorous theory of *uniform expressing* will require some significant further adjustments to our semantic ideology (see Dorr unpublished), which are not relevant for our present purposes.

5 Developing a deflationary account of property-talk

Now that we have a clearer picture of the general shape of a semantic theory for English given in Higher-Orderese, we can both sharpen the puzzles that words like 'property', 'relation', 'feature', 'state', 'condition', 'quality', and 'concept' will raise for our field linguists, and begin to see our way to a compositional implementation of the "deflationary" approach to these words. Let's focus on the word 'instantiates' and the expression 'the property of being red'. Both have all the hallmarks of meaningfulness. Moreover, 'instantiates' looks to have the same syntactic distribution as run-of-the-mill transitive verbs like 'orbits', while 'the property of being red' looks very similar to proper nouns like 'Mars'.[24] This suggests that we should assign 'the property of being red' the same category-label np we used for 'Mars', and 'instantiates' the same category-label np\s/np we used for 'orbits'. Assuming $\exists x^e.\mathsf{mars} \mathrel{\vdots} \mathsf{np}\, x$ and $\exists Y^{\overline{ee}}.\mathsf{orbits} \mathrel{\vdots} \mathsf{np\backslash s/np}\, Y$, this means that if we endorse Correspondence, we will have to say that

$$\exists x^e.\mathsf{the{\cdot}property{\cdot}of{\cdot}being{\cdot}red} \mathrel{\vdots} \mathsf{np}\, x$$
$$\exists Y^{\overline{ee}}.\mathsf{instantiates} \mathrel{\vdots} \mathsf{np\backslash s/np}\, Y$$

But what could these x^e and $Y^{\overline{ee}}$ *be*? Given our assumed Materialist$_e$ commitments, there are no good candidates.[25]

The obvious solution is simply *not* to treat the syntactic parallelisms between 'instantiates' and 'orbits' and between 'the property of being red' and 'Mars' as requiring semantic parallelism. This is most easily done by rejecting Correspondence, thus enabling us to endorse derivations like the following:

$$\dfrac{\mathsf{Mars} \mathrel{\vdots} \mathsf{np}\ \mathbf{mars}^e \qquad \dfrac{\dfrac{\mathsf{instantiates} \mathrel{\vdots} \mathsf{np\backslash s/np}\ \lambda X^{\bar{e}}.X \qquad \mathsf{the{\cdot}property{\cdot}of{\cdot}being{\cdot}red} \mathrel{\vdots} \mathsf{np}\ \mathbf{red}^{\bar{e}}}{\mathsf{instantiates{\cdot}the{\cdot}property{\cdot}of{\cdot}being{\cdot}red} \mathrel{\vdots} \mathsf{np\backslash s}\ ((\lambda X^{\bar{e}}.X)\ \mathbf{red}^{\bar{e}})}\ >}{\mathsf{instantiates{\cdot}the{\cdot}property{\cdot}of{\cdot}being{\cdot}red} \mathrel{\vdots} \mathsf{np\backslash s}\ \mathbf{red}^{\bar{e}}}\ \beta}{\mathsf{Mars{\cdot}instantiates{\cdot}the{\cdot}property{\cdot}of{\cdot}being{\cdot}red} \mathrel{\vdots} \mathsf{s}\ (\mathbf{red}^{\bar{e}}\ \mathbf{mars}^e)}\ <$$

Of course we will also have

$$\dfrac{\mathsf{Mars} \mathrel{\vdots} \mathsf{np}\ \mathbf{mars}^e \qquad \mathsf{is{\cdot}red} \mathrel{\vdots} \mathsf{np\backslash s}\ \mathbf{red}^{\bar{e}}}{\mathsf{Mars{\cdot}is{\cdot}red} \mathrel{\vdots} \mathsf{s}\ (\mathbf{red}^{\bar{e}}\ \mathbf{mars}^e)}\ <$$

We will thus conclude that 'Mars instantiates the property of being red' and 'Mars is red' express the same thing.

This looks like a welcome result, given that the former would (outside philosophical contexts) generally be treated as nothing more than an oddly pompous variant of the latter. True, the two sentences are not treated interchangeably by all speakers in all contexts—there are those oddball philosophers. But the fact that some competent speakers fail to treat two sentences as interchangeable is not a good objection to the claim that they express the same thing, in the sense we are concerned with:

[24] This syntactic resemblance is not perfect. Proper names admit count uses ('Every Kennedy was elected') whereas expressions headed by 'the' do not ('*Every the property of resembling Kennedy is instantiated'). Also, 'The planet Mars' is well-formed, unlike '*The planet the property of being red' (and '*The property the property of being red').

[25] If we were just looking at these two words we might consider interpreting 'instantiates' to mean something like *resembles in color* and taking 'the property of being red' to denote some specific red material object. But this sort of proposal will obviously run aground once we look at a wider pattern of usage (e.g. including 'the property of being coloured').

consider ‘Furze has yellow flowers’ and ‘Gorse has yellow flowers’, or ‘Vixens live in dens’ and ‘Female foxes live in dens’.[26]

The rejection of Correspondence isn’t crucial: as before, if one were determined to preserve Correspondence as a matter of principle, one could do so by imposing a more fine-grained system of categories, e.g. np_e for ‘Mars’ and $\mathsf{np}_{\bar{e}}$ for ‘the property of being red’. But this does not look like a particularly illuminating exercise. Given how much theoretical freedom we have in the choice of a system of categories, it seems more disciplined to try to restrict them to the kind of information that might be relevant for a theory of syntactic well-formedness.[27]

‘The property of being red’ is, I suppose, what authors in the tradition of Frege would call a “singular term” (*Eigenname*). The proposal we are considering thus directly contradicts Frege’s notorious doctrine that no singular term can denote the same thing as any predicate.[28] Indeed, it is tempting to describe the proposal as one on which on which ‘is red’ and ‘the property of being red’ are *semantically* identical and merely *syntactically* different. However, one might dispute this characterization on the grounds that not only properties of the form *expressing such-and-such relative to some category or other*, but also *expressing such-and-such relative to category c*, should count as semantic properties of expressions. The question which category a given meaning of an expression is relative to will, after all, be crucial for determining how the expression contributes to the meanings of larger expressions having it as a constituent.[29] On the other hand, the broader interpretation of ‘semantic property’ would arguably make it impossible for *any* two expressions to be “semantically identical but syntactically different”, so the narrower interpretation provides a more interesting understanding of that formula. I don’t see a genuine issue here: both ways of talking are fine so long as we don’t confuse them.

Obviously, if it is true that ‘is red’ and ‘the property of being red’ stand in this intimate semantic relationship, this fact is an instance of a much more general pattern. To capture that pattern, we would need to characterize the general syntactic transformation that turns ‘is red’ into ‘the property of being red’. Roughly, the oper-

[26] I am open to there being other uses of the word ‘meaning’ on which we would want to talk about distinctions of meaning in some of these cases. But I believe the more “worldly” notion corresponding to our higher-order ‘expressing’ constants will be a crucial component of any good theory of meaning in this finer-grained sense.

[27] For the claim that ‘Mars’ and ‘the property of being red’ denote relative to the same syntactic category (“expression type”), while ‘the property of being red’ and ‘is red’ denote the same thing relative to different syntactic categories, see Rieppel 2016.

[28] This was the doctrine that notoriously led Frege to say that ‘the concept horse is not a concept’ (Frege 1892). Many contemporary authors seem to accept the doctrine: for example, Trueman (2021: 49) argues that ‘it *does not make even make sense* to suppose that a term and a predicate might co-refer’. However, I think the right translation of my thesis into Trueman’s terminology is not the thing he is claiming not to make sense, but rather that some term and some predicate ‘predicate-refer’ to the same thing—a claim that he regards as meaningful but false.

[29] Rieppel (2016) understands category-relativity as a generalization of the idea (Wright 1998, Liebesman 2015) that there are importantly different semantic relations of “reference” and “ascription”, such that what ‘the property of being red’ refers to is what ‘is red’ ascribes.

ation starts with a normal present-tense VP, replaces the main verb with its infinitive form (e.g., 'is' with 'be'), adds 'ing' to the stem, and adds 'the property of' in front.[30] Calling this operation $\mathbf{PROP}^{\varepsilon\to\varepsilon}$, we can propose the following schematic law:

Nominalization $$\forall\, a^{\varepsilon} X^{\bar{\sigma}}.a \mathrel{\vdots_{\bar{\sigma}}} \mathsf{np\backslash s}\, X \leftrightarrow \mathbf{PROP}^{\varepsilon\to\varepsilon}(a) \mathrel{\vdots_{\bar{\sigma}}} \mathsf{np}\, X$$

Of course, we would ultimately like to explain Nominalization compositionally from semantic accounts of 'the', 'property', 'of', '-ing', and the present tense. While this is certainly not a trivial task, it does not look any *harder* than the corresponding compositional challenge that would arise if one wanted to posit a type-*e* meaning for 'the property of being red'.[31]

Since expressions like 'the property of being red' take up a lot of space, I will henceforth assume that NP-uses of abstract nouns should be treated analogously: e.g., redness $\vdots$ np $\mathbf{red}^{\bar{e}}$ and wisdom $\vdots$ np $\mathbf{wise}^{\bar{e}}$. A full semantic account of these words will also need to account for uses that aren't syntactically NP-like—e.g. in 'There is more wisdom in your body than in your deepest philosophy' and 'The wisdom of Socrates was celebrated'.[32] But there is no *a priori* reason to expect such a theory to be

[30]Calling this an "operation" is perhaps a bit unnatural—it is more natural to think of 'is red' and 'the property of being red' as both derived via different syntactic operations from the base form 'be red'. Note that there doesn't seem to be anything that stands to 'was red' as 'the property of being red' stands to 'is red': the only candidate I can think of is 'the property of having been red', but this corresponds to 'has been red', not 'was red'. Similarly for VPs with modal auxiliaries like 'will be red', 'might be red', 'can be red', etc.

Moltmann (forthcoming) claims that there are further restrictions on the verbs that can go after 'the property of...', namely that "eventive" verbs (as in 'the property of walking home' and 'the property of remaining sick') and "concrete state" verbs (as in 'the property of sleeping' and 'the property of standing in the corner') are disallowed (she marks them '??'). This seems wrong to me. In English, 'Mary walks home' and 'Mary stands in the corner' only have habitual meanings: roughly, those of 'Mary typically {walks home/stands in the corner}'. And 'Mary has the property of {walking home/standing in the corner}' seem to allow exactly these habitually meanings. Plugging in an eventive verb sounds odd when the habitual meaning is unlikely, as in 'Mary has the property of eating that piece of meat'. But this isn't something special about 'the property of...', since 'Mary eats that piece of meat' is equally odd. (I may not actually be disagreeing with Moltmann here, since she qualifies her claim by admitting that 'the property of...' is acceptable for 'verbs on a dispositional reading'.)

[31]The attempt to integrate Nominalization with a general account of 'the' might make trouble for the principle as stated. If we demand a strictly parallel treatment of the determiners 'the' and 'every', 'the property of being red' will need a denotation of the same type as 'every property Mars instantiates': perhaps $\lambda X^{\underline{\bar{e}}}.X\,\mathbf{red}^{\bar{e}}$ (or something logically equivalent). Handling this will require some adaptations to our semantic architecture to handle quantificational expressions in object position (e.g. along the lines of Moortgat 1996 or Barker and Shan 2014), but the result would not be importantly different from Nominalization for the purposes of this paper.

[32]As Moltmann (2013) points out, there is a strong syntactic parallel between words like 'wisdom' and ordinary mass nouns like 'water', which also admit both bare uses ('Water is vital to all life') and modified ones ('There is more water in this glass than in that thimble', 'The water of this fountain is said to have magical properties'). Moltmann suggests that what stands to 'wisdom' as particular samples of water stand to 'water' are *tropes*—the referents of expressions like 'the wisdom of Socrates'. While she thinks of these as individuals, I would put them in some higher type, though the question

incompatible with the claim that *qua NPs* these words have type-$\bar{e}$ meanings. Those with doubts on this front should substitute 'the property of being *F*' wherever I have '*F*-ness'.[33]

5.1 Type-ambiguity and property-talk

By analogy with our earlier treatment of 'No planet twinkles', we can explain why 'Some planet twinkles' means what it does by treating 'some' as a binary first-order existential quantifier:

(5) $$\text{some} \vdots \text{s/np\s/n} \ \lambda X^{\bar{e}} Y^{\bar{e}}.\exists z^{e}.Xz \wedge Yz$$

But given the deflationary approach, this claim does not help to explain why 'Some property of Venus holds of Mars' means what it does (or means anything at all). 'Property of Venus' and 'holds of Mars' will both have meanings of type $\bar{e}$, which we could perhaps derive as follows:

$$\dfrac{\begin{array}{c}\text{property} \vdots \text{n} \\ \lambda x^{e} Z^{\bar{e}}.Zx\end{array} \quad \dfrac{\begin{array}{c}\text{of} \vdots \text{n\n/np} \\ \lambda x^{e} Y^{\overline{ee}}.Yx\end{array} \quad \begin{array}{c}\text{Venus} \vdots \text{np} \\ \mathbf{venus}^{e}\end{array}}{\begin{array}{c}\text{of·Venus} \vdots \text{n\n} \\ \lambda Y^{\overline{ee}}.Y\,\mathbf{venus}^{e}\end{array}}>}{\begin{array}{c}\text{property·of·Venus} \vdots \text{n} \\ \lambda Z^{\bar{e}}.Z\,\mathbf{venus}^{e}\end{array}}<
\qquad
\dfrac{\begin{array}{c}\text{holds} \vdots \text{np\s} \\ \lambda x^{e} Z^{\bar{e}}.Zx\end{array} \quad \dfrac{\begin{array}{c}\text{of} \vdots \text{(np\s)\np\s/np} \\ \lambda x^{e} Y^{\overline{ee}}.Yx\end{array} \quad \begin{array}{c}\text{Mars} \vdots \text{np} \\ \mathbf{mars}^{e}\end{array}}{\begin{array}{c}\text{of·Mars} \vdots \text{(np\s)\np\s} \\ \lambda Y^{\overline{ee}}.Y\,\mathbf{mars}^{e}\end{array}}>}{\begin{array}{c}\text{holds·of·Mars} \vdots \text{np\s} \\ \lambda Z^{\bar{e}}.Z\,\mathbf{mars}^{e}\end{array}}<$$

---

which type that should be is one I don't want to get into here.

[33]Moltmann (2013) points out some intriguing contrasts between pairs like 'wisdom' and 'the property of being wise' which are *prima facie* problematic for my proposal to treat them as having the same denotation (as NPs). For example, whereas 'John has encountered hostility' is true (on its most natural interpretation) so long as some people have been hostile to John, according to Moltmann 'John has encountered the property of being hostile' lacks this reading—its only possible interpretation is something that could be true only in a 'metaphysical fantasy'. While there is certainly a striking contrast here, I am not convinced that the more humdrum meanings are entirely unavailable with 'the property of…'. Consider: 'The shapes produced by this machine often have interesting mathematical properties. For example, one I have often encountered is the property of having a prime number of sides'. Given these, I suspect that the humdrum meaning is in principle available for 'John has encountered the property of being hostile', but hard to notice because of competition with more natural modes of expression (such as using 'hostility'). Similar points apply to Moltmann's other contrasts. For example, prima facie 'Generosity exists' seems to have a reading tantamount to 'Some people are generous', while 'The property of being generous exists' lacks this reading. But consider 'Many properties of humans also exist in the great apes. For example, the property of caring about your place in society exists in bonobos'.

I do however see some differences specific to gradable adjectives: for example, 'Even people who aren't wise have wisdom (it's just that they have very little of it)' sounds reasonable, unlike 'Even people who aren't wise have the property of being wise'. While these contrasts might motivate assigning 'redness' a different semantic type from 'the property of being red', I don't see how they could motivate giving the former but not the latter a type-*e* denotation.

But there is no way to use these ingredients to derive any prediction about 'Some property of Venus holds of Mars': the meaning of 'some' given by (5) needs to combine with inputs of type $\bar{e}$, not $\underline{\bar{e}}$. The obvious solution is to say that while (5) is true, (6) is also true:

(6) $$\text{some} \vdots \text{s/np\textbackslash s/n}\ \lambda X^{\underline{\bar{e}}} Y^{\underline{\bar{e}}}.\exists z^{\bar{e}}.Xz \wedge Yz$$

In other words, 'some' is *type-ambiguous*: it is both a first order and a second-order quantifier. Which of its semantic roles is relevant to a given occurrence will depend on the meanings of the expressions it combines with. Its role as a second-order quantifier is what lets it combine meaningfully with 'property of Venus' and 'holds of Mars' to derive:

$$\dfrac{\dfrac{\begin{matrix}\text{some} \vdots \text{s/np\textbackslash s/n} \\ \lambda X^{\underline{\bar{e}}} Y^{\underline{\bar{e}}}.\exists Z^{\bar{e}}.XZ \wedge YZ\end{matrix} \qquad \begin{matrix}\text{property·of·Venus} \vdots \text{n} \\ \lambda Z^{\bar{e}}.Z\,\mathbf{venus}^{e}\end{matrix}}{\begin{matrix}\text{some·property·of·Venus} \vdots \text{s/np\textbackslash s} \\ \lambda Y^{\underline{\bar{e}}}.\exists Z^{\bar{e}}.Z\,\mathbf{venus}^{e} \wedge YZ\end{matrix}}> \qquad \begin{matrix}\text{holds·of·Mars} \vdots \text{np\textbackslash s} \\ \lambda Z^{\bar{e}}.Z\,\mathbf{mars}^{e}\end{matrix}}{\begin{matrix}\text{some·property·of·Venus·holds·of·Mars} \vdots \text{s} \\ \exists Z^{\bar{e}}.Z\,\mathbf{venus}^{e} \wedge Z\,\mathbf{mars}^{e}\end{matrix}}>$$

The need to posit type-ambiguity is not confined to determiners like 'some': parallel considerations motivate type-ambiguity for a very wide range of expressions. For example, consider the VP 'is interesting'. It will have to express something of type $\bar{e}$, to explain the meaningfulness of 'Mars is interesting'. But this will not account for the meaningfulness of 'Redness is interesting'. To do that, we can posit that 'is interesting' also expresses something of type $\underline{\bar{e}}$ which can combine with the type-$\bar{e}$ meaning of 'redness'. And of course, given that we are having 'Mars instantiates redness' and 'Redness holds of Mars' turn out semantically equivalent to 'Mars is red', we will presumably also want 'Redness instantiates the property of being interesting' and 'The property of being interesting holds of redness' to be semantically equivalent to 'Redness is interesting'. 'The property of being interesting' will thus need to be type-ambiguous in the same way as 'is interesting'; and we will need additional higher-type meanings for 'instantiates' and 'holds [of]':

$$\text{instantiates} \vdots \text{np\textbackslash s/np}\ \lambda X^{\underline{\bar{e}}}.X \qquad \text{holds} \vdots \text{np\textbackslash s}\ \lambda y^{\bar{e}} X^{\underline{\bar{e}}}.Xy$$

Similarly, 'mentions' will need to both express something of type $\overline{ee}$, to account for 'Aristotle mentions Mars', and something of type $\overline{ee}$, to account for 'Aristotle mentions redness'.

We are not just dealing with two-way type-ambiguities: similar considerations motivate the view that many of the relevant expressions are in fact ambiguous across *infinitely many* types. For example, neither of the two meanings for 'some' discussed above can explain why 'Some property of redness holds of blueness' means what it does. 'Property of redness' and 'holds of blueness' will both have type-$\underline{\underline{\bar{e}}}$ meanings, namely, namely $\lambda X^{\underline{\bar{e}}}.X\,\mathbf{red}^{\bar{e}}$ and $\lambda X^{\underline{\bar{e}}}.X\,\mathbf{blue}^{\bar{e}}$. For ‘some’ to be able to combine *these* meanings, we will need to assign it a third meaning:

(7) $$\text{some} \vdots \text{s/np}\backslash\text{s/n}\ \lambda X^{\underline{\underline{\bar{e}}}} Y^{\underline{\underline{\bar{e}}}}.\exists z^{\underline{\bar{e}}}.Xz \wedge Yz$$

Similarly, for ‘Some property of the property of holding of Mars holds of the property of holding of Venus’, ‘some’ will need a meaning taking arguments of the next type up, and so on. All these meanings can be subsumed under a general schema with an instance for each type σ:

$$\text{some} \vdots \text{s/np}\backslash\text{s/n}\ \lambda X^{\bar{\sigma}} Y^{\bar{\sigma}}.\exists z^{\sigma}.Xz \wedge Yz$$

On similar grounds, we will need to posit infinite type-ambiguity in words like ‘instantiates’, ‘holds [of]’, and ‘property [of]’: consider for example ‘The property of holding of Mars instantiates the property of holding of redness’, ‘The property of holding of redness instantiates the property of holding of the property of holding of Mars’, and so on. Here again, we can subsume all the required meaning-attributions under some simple general schemas:

$$\text{instantiates} \vdots \text{np}\backslash\text{s/np}\ \lambda X^{\bar{\sigma}}.X$$
$$\text{holds} \vdots \text{np}\backslash\text{s}\ \lambda x^{\sigma} Y^{\bar{\sigma}}.Yx$$
$$\text{property} \vdots \text{n}\ \lambda x^{\sigma} Y^{\bar{\sigma}}.Yx$$

The pressure towards infinite type-ambiguity also applies to expressions like ‘is interesting’ and ‘mentions’ for which there is no prospect of writing down comparable general schemas using only the logical constants. For example, ‘is interesting’ will need not only the type-$\bar{e}$ and type-$\underline{\bar{e}}$ meanings discussed above, but a type $\underline{\underline{\bar{e}}}$ meaning (to account for ‘The property of holding of Mars is interesting’), a type $\underline{\underline{\bar{\bar{e}}}}$ meaning (to account for ‘The property of holding of redness is interesting’), and so on. If our field linguists are lucky enough to be theorizing in a higher-order language with a corresponding infinite family of nonlogical constants, they will be able to use these to formulate general schemas analogous to those discussed for ‘some’, ‘instantiates’, etc.:

$$\text{is·interesting} \vdots \text{np}\backslash\text{s}\ \mathbf{int}^{\bar{\sigma}}$$
$$\text{mentions} \vdots \text{np}\backslash\text{s/np}\ \mathbf{mentions}^{\overline{\sigma e}}$$

If they are not in this fortunate situation, it will take more work if they want to identify or constrain the meanings of ‘interesting’ and ‘mentions’ in terms of vocabulary they do have. But of course they can still endorse existentially general schemas like

$$\exists X^{\overline{\sigma}}.\text{is·interesting} \vdots \text{np\\s}\ X$$
$$\exists X^{\overline{\sigma e}}.\text{mentions} \vdots \text{np\\s/np}\ X$$

even if they are not yet in a position to propose any plausible witnesses.[34]

5.2 Why we need category-relativization

Now that we have seen how the deflationary proposal naturally leads to type-ambiguity, we can see why the original architectural decision to work with category-relative expressing predicates was crucial. Without such relativization, it would have been natural to replace our function application schemas with ones that allow meanings to compose in whatever direction makes type-theoretic sense:

Type-Driven FA (>) $a \vdots_{\sigma\to\tau} x \wedge b \vdots_{\sigma} y \wedge \mathbf{wellformed}^{\overline{\varepsilon}}(a{\cdot}b) \to a{\cdot}b \vdots_{\tau} xy$

(<) $a \vdots_{\sigma\to\tau} x \wedge b \vdots_{\sigma} y \wedge \mathbf{wellformed}^{\overline{\varepsilon}}(b{\cdot}a) \to b{\cdot}a \vdots_{\tau} xy$

This idea of *type-driven composition* (introduced by Klein and Sag (1985), and used influentially by Heim and Kratzer (1998)) has several attractions, including the way it lets us outsource the theory of well-formedness to a separate theory of syntax whose inner workings need play no further role in the theory of meaning. But it wouldn't work for us, as we can see by considering the following sentences:

(8) Something is interesting.

(9) The property of being instantiated is interesting.

For each type σ, we will have the following claims about the constituents of these sentences:

(10) $$\text{something} \vdots_{\overline{\overline{\sigma}}} \lambda X^{\overline{\sigma}}.\exists y^{\sigma}.Xy$$

(11) $$\text{the·property·of·being·instantiated} \vdots_{\overline{\overline{\sigma}}} \lambda X^{\overline{\sigma}}.\exists y^{\sigma}.Xy$$

(12) $$\text{is·interesting} \vdots_{\overline{\overline{\sigma}}} \mathbf{int}^{\overline{\overline{\sigma}}}$$

Since the available meanings for 'something' and 'being instantiated' are exactly the same, Type-Driven FA predicts that substituting one for the other will not affect the

[34] I see no special reason why the Lambda Serpentians would not have infinite families of nonlogical constants such as $\mathbf{int}^{\overline{\sigma}}$ and $\mathbf{mentions}^{\overline{\sigma e}}$. Of course, since they are finite beings, they will not learn to use these constants one by one. Rather, they will come to understand them just as they come to understand the quantifiers in each type, by acquiring *general* dispositions of use, which are sensitive to the internal structure of the type that occurs as a superscript to the symbol. Although the dispositions relevant for constants like $\mathbf{int}^{\overline{\sigma}}$ and $\mathbf{mentions}^{\overline{\sigma e}}$ are much messier than those relevant for the quantifiers (which can arguably be codified as natural deduction rules), I see no deep difference of principle.

range of interpretations available for a complex expression, unless it blocks well-formedness. Since (8) and (9) are obviously both well-formed, we get the obviously wrong prediction that they have the same range of interpretations. In both cases, we can apply Type-Driven FA from left to right to get

(13) $$\text{something·is·interesting} \vdots_t \exists y^{\sigma}.\, \mathbf{int}^{\bar{\sigma}}\, y$$

(14) $$\text{the·property·of·being·instantiated·is·interesting} \vdots_t \exists y^{\sigma}.\, \mathbf{int}^{\bar{\sigma}}\, y$$

or from right to left to get

(15) $$\text{something·is·interesting} \vdots_t \mathbf{int}^{\underline{\bar{\sigma}}}(\lambda X^{\bar{\sigma}}.\exists y^{\sigma}.Xy)$$

(16) $$\text{the·property·of·being·instantiated·is·interesting} \vdots_t \mathbf{int}^{\underline{\bar{\sigma}}}(\lambda X^{\bar{\sigma}}.\exists y^{\sigma}.Xy)$$

While (13) and (16) are just what we want, (14) and (15) are plainly false for many types σ. Category-relativization blocks this disaster by forcing the meanings to be combined in such a way that the constituent with the more complex category label—'something' in (8), and 'is interesting' in (9)—takes the other one as argument.[35]

5.3 Pervasive type-ambiguity and trivial meanings

The reasons for positing type-ambiguity are clear in the case of expressions like 'some', 'instantiates', and 'is interesting', for which we can identify useful, discriminating meanings in many types. But once we admit type-ambiguity at all, there is reason to posit it too in many other cases where there only seems to be one type with a natural non-trivial candidate meaning. Consider 'is red'. We certainly want to say that is·red $\vdots$ np\s $\mathbf{red}^{\bar{e}}$, to explain why Mars·is·red $\vdots$ s $\mathbf{red}^{\bar{e}}\, \mathbf{mars}^{e}$. But this will not help with 'The property of being red is red' or 'Redness is red', which are also grammatical sentences of English.

One might of course deny that these sentences are meaningful. Indeed it has often been taken for granted, by proponents and opponents alike, that higher-order accounts of property-talk would require denying meaningfulness in various cases like this.[36] But this is hard to square with the fact that some larger sentences having 'Redness is red' as a constituent seem to be not only meaningful but to express truths: for example, 'Either redness is red or it is not the case that redness is red'

[35]Note that nothing here turns on the choice to use directional slashes: we could instead follow Lewis (1983*a*) and van Benthem (1988) in using directionless slashes and leaving the determination of allowable word orders to a separate component of the theory.

[36]For example, in the context of expounding the view that 'is a fact' expresses $\lambda p^{t}.p$, Prior (1971: 25) assumes that it implies various claims of meaninglessness: 'For "Percy is a fact" (which would mean "It is the case that Percy", if it meant anything), "Percy is a falsehood" (="It is not the case that Percy"), 'Percy is neither a fact nor a falsehood (="It neither is nor is not the case that Percy") are all of them senseless, ungrammatical.' Likewise Bealer (1993), in the passage quoted in §6 below, assumes that the meaningfulness of various sentences suffices to show that the relevant discourse is 'type-free'.

or 'It would be strange to say that redness is red'. To my mind, the need to accommodate such sentences provides a very strong reason to think that 'Redness is red' says something—i.e., that $\exists p.\mathsf{redness{\cdot}is{\cdot}red} \vdots \mathsf{s}\, p$. Given that 'redness' lacks a type-e denotation, this means that 'is red' will need to express something of type $\underline{\bar{e}}$ as well as something of type $\bar{e}$, just as 'is interesting' does.[37]

The central difference between the two cases is that ordinary, non-philosophical usage doesn't provide us with any useful guidance as regards what 'is red' might express in types other than $\bar{e}$. Unlike the question 'Is redness interesting?', the question 'Is redness red?' is one that only a philosopher would ask. Moreover, as is common for such questions, it tends to put hearers into a mode of philosophical speculation (amateur or professional), in which they come out with such a variety of strange speeches that there is no question of being charitable to all of them. Considering this, our field linguists might reasonably maintain that the type $\underline{\bar{e}}$ aspect of the semantics of 'is red' is drastically vague, for want of such clear constraints.[38] But this is not inevitable; they might, instead, decide to take seriously some of the more level-headed of these philosophical speeches, such as the following:

> No, redness is not red! The only red things are material objects, like books and walls. But redness is a property, and no property is a material object.

This suggests that the best candidates of type $\underline{\bar{e}}$ to be expressed by 'is red' are *empty* (i.e., $\forall X^{\underline{\bar{e}}}.\mathsf{is{\cdot}red} \vdots \mathsf{np\backslash s}\, X \;\rightarrow\; \neg\exists y^{\bar{e}}.Xy$). For example, we might propose that $\mathsf{is{\cdot}red} \vdots \mathsf{np\backslash s}\, \lambda X^{\bar{e}}.\bot$. In what follows, I will assume this account, partly because the above speech seems pretty sensible to me, and partly because it might not otherwise be clear that the deflationary approach is even compatible with endorsing it.[39]

One might object to the hypothesis that 'is red' expresses $\lambda X^{\bar{e}}.\bot$ on the grounds that it runs together differences between English sentences which have different "cognitive value", for example 'Redness is red' and 'Yellowness is red'. Even if it is necessarily false both that redness is red and that yellowness is red, couldn't someone *believe*, or *say*, that redness is red but not that yellowness is red? Replying to this objection would take us too far afield into a dialectic about Frege's puzzle and the problem of logical omniscience. Suffice it to say that in my view, examples like 'vixen'/'female fox' and 'attorney'/'lawyer' show that it must be a mistake for semanticists to infer that there is a difference in what things express from the fact that substituting one for the other in the context of an operator like 'believes' or 'says' can take us from a true-seeming sentence to a false-seeming one.

[37] This argument is defended in detail by Magidor (2013: §3.4), in the context of defending the general claim that "category mistakes" are meaningful. Magidor also has a plausible proposal for explaining the oddity of the relevant sentences without denying their meaningfulness.

[38] On the theory of vagueness in Dorr unpublished, that would be to say that it expresses enormously many type-$\underline{\bar{e}}$ things; maybe even *every* type-$\underline{\bar{e}}$ thing.

[39] This substitution of automatic falsehood for meaninglessness corresponds to the second version of Parsons' (1979: 142) system 'PQTB'. The first version, by contrast, treated what Parsons calls 'quasi-grammatical' sentences, such as 'A property runs', as uninterpretable.

The argument against meaninglessness applies in the same way for other types. For example, 'the property of holding of Mars is red' will require a type $\bar{\bar{e}}$ meaning, and so on. And clearly the same argument can be made for any other verb phrase. We are thus led to the conclusion that every meaningful verb phrase has a meaning of type $\overline{\sigma}$, for any type σ.[40] In the case of 'is red', we could plausibly add that in all types $\overline{\sigma} \neq \overline{e}$, its meaning is empty, perhaps $\lambda x^{\sigma}.\bot$.

A parallel argument can be made for various other categories, such as that of transitive verbs. For example, the previous subsection proposed the following schema, providing meanings for 'instantiates' in all types of the form $\overline{\overline{\sigma}\sigma}$:

$$\text{instantiates} \vdots \text{np\textbackslash s/np}\ \lambda X^{\overline{\sigma}}.X$$

But this can't be the full story, since it doesn't generate any meaning for well-formed (though odd) sentences like 'Mars instantiates Venus', 'Redness instantiates Venus', 'The property of holding of Mars instantiates redness', etc. Again, it seems a bad idea to treat these as meaningless—consider the appeal of speeches like 'Mars doesn't instantiate Venus, since Venus isn't a property, and only properties can be instantiated'. So, we will need meanings for 'instantiates' in a lot more types not of the form $\overline{\overline{\sigma}\sigma}$. In fact, it looks like every meaningful transitive verb (type np\s/np) will need to express something of type $\overline{\sigma_1\sigma_2}$, for any two types σ_1 and σ_2. For 'instantiates', we can default to some boring meaning such as $\lambda x^{\sigma_1} y^{\sigma_2}.\bot$ in the case where σ_1 is some type other than $\overline{\sigma_2}$. At a minimum, I will assume that when $\sigma_1 \neq \overline{\sigma_2}$, every type-$\overline{\sigma_1\sigma_2}$ meaning of 'instantiates' is actually empty.[41]

The availability of these all these supplementary meanings for property-theoretic words will rarely be relevant to ordinary, unreflective uses of property talk. But it will matter when we come to certain speeches made by philosophers. For example, consider a nominalist who argues, 'Everything is material. But no property is material. So there are no properties. So no property of Venus holds of Mars'. The first sentence plausibly expresses $\forall x^e.\, \mathbf{material}^{\overline{e}}\, x$. In a typical use, we would expect 'No property of Venus holds of Mars' to mean that $\neg\exists X^{\overline{e}}.X\,\mathbf{venus}^e \wedge X\,\mathbf{mars}^e$. But the context of making a deductive argument generates special pressure towards uniform interpretation across the argument: ceteris paribus, we don't want to convict people of committing fallacies of equivocation unless we must. So when a nominal-

[40] Formally, this can be captured by the schema: $(\exists x^{\rho}.a \vdots_{\rho} \text{np\textbackslash s}\, x) \to (\exists y^{\overline{\sigma}}.a \vdots_{\overline{\sigma}} \text{np\textbackslash s}\, y)$.

[41] For type $\overline{ee}$, there is some temptation to think that 'instantiates' expresses a relation that is in fact empty (assuming e-Materialism is true), but *could* have been non-empty if there had been a sufficient supply of non-material objects, playing roles sufficiently close to the roles of universals (properties-as-individuals) in some reasonably fleshed-out theory about them.

In types of the form $\overline{\sigma_1\sigma_2}$, where $\sigma_1 \neq \sigma_2$, one might consider trying to craft a non-trivial meaning for 'instantiates' based on some relation $\mathbf{C}$ of "counterparthood" between type $\overline{\sigma_1}$ and type $\overline{\sigma_2}$—the idea would be that 'instantiates' expresses $\lambda Y^{\overline{\sigma_2}} x^{\sigma_1}.\exists Z^{\overline{\sigma_1}}.Zx \wedge \mathbf{C}\, ZY$. By saying that, e.g., $\lambda x^{\sigma_1}.x =_{\sigma_1} x$ is a counterpart of $\lambda y^{\sigma_2}.y =_{\sigma_2} y$, this will allow us to give a true reading for 'the property of being self-identical instantiates itself'. But it is not clear to me that the reasons to posit such readings are strong enough to motivate the complexity of such a view.

ist sets forth the above argument, it may be best to think that the quantifiers retain their first-order meanings throughout. That would mean interpreting the conclusion $\neg\exists x^e . R\, \mathbf{venus}^e\, x \wedge R\, \mathbf{mars}^e\, x$, where R is whatever we take 'property' and 'applies' to express in type $\overline{ee}$ (presumably, the converse of whatever 'instantiates' expresses in that type). Given our assumptions, R is empty. So on the operative interpretation, the conclusion is in fact true, despite the unequivocal truth of 'Mars is red and Venus is red'.

5.4 Type-ambiguity and Russell's paradox

To sharpen our understanding of the ways in which the systematic type-ambiguity we have posited can generate multiple readings for a single sentence, it will be useful to see how it plays out in the kinds of sentences that feature in (the property-theoretic version of) Russell's paradox. First, we can consider the following English 'Naïve Property Comprehension' schema:

NPC [NP] [VP] if and only if [NP] instantiates [**PROP**(VP)].

In our system, every well-formed instance of this schema has a true reading, since whatever the VP expresses (of any given type $\overline{\sigma}$) will also be expressed by **PROP**(VP) and by 'instantiates **PROP**(VP)'.

It would be too much to expect that instances of NPC *exclusively* have true readings. For recall that when a sentence contains multiple ambiguous expressions we get to "mix and match" their interpretations, insofar as this is allowed by their categories and their meanings' types. For example, we can generate a false reading for the NPC-instance

(17) Barclays is a bank if and only if Barclays instantiates the property of being a bank

by interpreting the first 'bank' as money bank and the second as river bank. The claims of type-ambiguity characteristic of the deflationary approach will also lead to false interpretations of instances of NPC for similar reasons. For example, suppose that redness is in fact interesting—i.e., $\mathbf{int}^{\underline{\bar{e}}}\, \mathbf{red}^{\bar{e}}$—and consider the following NPC-instance:

(18) Redness is interesting if and only if redness instantiates the property of being interesting.

We can get the left-hand side to express a truth by taking 'redness' to express $\mathbf{red}^{\bar{e}}$ and 'is interesting' to express $\mathbf{int}^{\underline{\bar{e}}}$. We can nevertheless get the right-hand side to be false by taking the second 'redness' still to express $\mathbf{red}^{\bar{e}}$, 'the property of being interesting' to express $\mathbf{int}^{\bar{e}}$, and 'instantiates' to express something of type $\overline{\bar{e}\bar{e}}$. (Recall that in the last section I posited that 'instantiates' expresses empty relations in all types $\overline{\sigma_1\sigma_2}$ where $\sigma_1 \neq \overline{\sigma_2}$.)

Of course, one would expect such false readings of (18) to be much less salient than the true ones. The choice to interpret the two occurrences of 'interesting' in different ways is unforced, and flies in the face of the pressure towards uniformity generated by the parallel structure of the utterance. But one thing that distinguishes type-ambiguity from ambiguity within a type is that the demands of compositionality sometimes make it *impossible* to interpret all the occurrences of a type-ambiguous expression in a sentence uniformly. To make the false readings of NPC-instances more salient, we can consider instances all of whose readings involve some non-uniformity:

(19) The property of being interesting is interesting if and only if the property of being interesting instantiates the property of being interesting.

In the necessarily-true readings, the first and third 'interesting' express $\mathbf{int}^{\overline{\sigma}}$ for some σ, while the second and fourth express $\mathbf{int}^{\underline{\overline{\sigma}}}$, and 'instantiates' expresses $\lambda X^{\underline{\overline{\sigma}}}.X$, so that both sides of the biconditional express $\mathbf{int}^{\underline{\overline{\sigma}}}\,\mathbf{int}\,\overline{\sigma}$. But given the structure of the sentence, the option of assigning the same semantic type $\overline{\sigma}$ to all three occurrences of 'the property of being interesting' (and hence to all but the second occurrence of 'interesting') is also quite salient. On *that* kind of interpretation, the right-hand side will be false, since 'instantiates' will express an empty relation of type $\overline{\overline{\sigma}\overline{\sigma}}$. By contrast, our system provides no way of getting a corresponding false reading for the left-hand-side of (19): the only readings we generate for 'The property of being interesting is interesting' are propositions of the form $\mathbf{int}^{\underline{\overline{\sigma}}}\,\mathbf{int}^{\overline{\sigma}}$, which we may assume to be all true.

Armed with an appreciation of these potential ambiguities, we can consider the NPC-instance that figures in Russell's paradox:

(20) The property of not instantiating oneself does not instantiate itself if and only if the property of not instantiating oneself instantiates the property of not instantiating oneself.[42]

The one extra ingredient we'll need to treat (20) in our semantics is an account of the reflexive pronoun 'itself'. Following Szabolsci (1987), we can treat this as a reflexivizer (W combinator), i.e. accept the following schema:

$$\text{itself} \vdots (\text{np}\backslash\text{s}/\text{np})\backslash(\text{np}\backslash\text{s})\ \lambda X^{\overline{\sigma\sigma}} y^{\sigma}.Xyy$$

This explains the basic facts about sentences where 'itself' occurs as the object of a transitive verb:

$$\frac{\text{Mars} \vdots \text{np}\ \mathbf{mars}^{e} \qquad \dfrac{\text{orbits} \vdots \text{np}\backslash\text{s}/\text{np}\ \mathbf{orbits}^{\overline{ee}} \qquad \text{itself} \vdots (\text{np}\backslash\text{s}/\text{np})\backslash\text{np}\backslash\text{s}\ \lambda X^{\overline{ee}} y^{e}.Xyy}{\text{orbits}\cdot\text{itself} \vdots \text{np}\backslash\text{s}\ \lambda y.\,\mathbf{orbits}^{\overline{ee}}\,yy} <}{\text{Mars}\cdot\text{orbits}\cdot\text{itself} \vdots \text{s}\ (\mathbf{orbits}^{\overline{ee}}\,\mathbf{mars}^{e}\,\mathbf{mars}^{e})} <$$

[42] Let's understand the black-boxed syntactic operation **PROP** in such a way that it does in fact turn 'does not instantiate itself' into 'the property of not instantiating oneself', so that (20) counts as an instance of NPC.

Note that the only meanings of a transitive verb V that are relevant to the interpretation of 'V itself' are those of types of the form $\overline{\sigma\sigma}$. Whatever meanings V may have in types $\overline{\sigma_1\sigma_2}$ where $\sigma_1 \neq \sigma_2$ are distinct will be irrelevant to the semantics of 'V itself'. Given our postulate that the only non-empty meanings of 'instantiate' are those of types $\overline{\sigma\sigma}$, the only relations of type $\overline{\sigma\sigma}$ that are expressed by 'instantiates' are empty. This means that the only meanings of 'does not instantiate itself' in any type $\overline{\sigma}$ are *universal* properties (i.e. are $X^{\overline{\sigma}}$ such that $\forall y^{\sigma}.Xy$). This implies that the only possible interpretations of the left-hand side of (20) are true. By contrast, we can derive a wider range of propositions for the right-hand side of (20):

(21) The property of not instantiating oneself instantiates the property of not instantiating oneself.

Given that 'the property of not instantiating oneself' is the result of applying **PROP** to 'does not instantiates itself', it is also such as to express only universal properties. Thus on any derivable interpretation of (21), the first and second occurrences of 'the property of not instantiating oneself' will have to be interpreted as denoting some universal $X^{\overline{\sigma_1}}$ and $Y^{\overline{\sigma_2}}$ (for some pair of types σ_1 and σ_2). Since 'instantiates' is a transitive verb, it has a meaning for every pair of types; thus, each choice of σ_1 and σ_2 determines a meaning for (21) in which the type-($\overline{\overline{\sigma_1}\,\overline{\sigma_2}}$) meaning of 'instantiates' combines with the two universal properties expressed by the two occurrences of 'not instantiating oneself'. If we choose σ_2 to be $\overline{\sigma_1}$, the relevant meaning of 'instantiates' will be $\lambda x^{\overline{\sigma_1}} Y^{\overline{\overline{\sigma_1}}}.Yx$. This will yield a truth when x and Y are instantiated by universal properties (indeed it is sufficient that Y be universal). So for these choices of σ_1 and σ_2, (21) comes out true, and hence (20) comes out true as well (since as we have already seen, the left hand side only allows true interpretations). On the other hand, if σ_2 is something other than $\overline{\sigma_1}$, the relevant meaning of 'instantiates' will be some empty relation of type $\overline{\overline{\sigma_1}\,\overline{\sigma_2}}$. On this reading, (21) comes out false, and hence so does (20). The general pressure to treat 'instantiates **PROP**(VP)' as equivalent to VP militates in favour of the necessarily-true readings of (20). By contrast, the pull to interpret all three occurrences of 'the property of not instantiating oneself' uniformly—which is perhaps particularly strong in this case—militates in favour of the false ones. Little wonder if we end up confused![43]

I don't want to put forth this way of defusing Russell's paradox as a weighty argument for the existence of the relevant type-ambiguities. In general, the fact that positing some kind of semantic multiplicity would allow us to escape some paradox

[43]The foregoing analysis depends on the assumption that 'instantiates' only expresses empty relations in types $\overline{\sigma_1\sigma_2}$ where $\sigma_1 \neq \overline{\sigma_2}$. The more complex "counterparthood" proposal considered in note 41 allows for it to express non-empty relations in some such types, and thus allows 'instantiates itself' to have non-empty meanings in some types. But the general outline of the account of (20) remains the same: it's true if we interpret the first and third occurrence of 'instantiate' uniformly and interpret the second and fifth occurrence uniformly, false if we interpret the three occurrences inside 'the property of not instantiating oneself' uniformly, and uninterpretable if we insist on interpreting all five occurrences uniformly.

by pleading equivocation does not yet amount to a strong case for the existence of such multiplicity. Still, such diagnoses become more satisfying when the multiplicity is backed up by independent arguments of some sort. And I do have an independent argument: namely, that we should interpret arguments like 'Mars is red; Venus is red; so some property of Mars holds of Venus' as good arguments, in a way that they would *not* be good if the conclusion implied the existence of a non-material individual of a sort that monolingual speakers of Higher-Orderese would never even have been tempted to posit.

6 The problems of mixed co-ordination and quantification

The idea that property-talk in natural languages is a disguised form of higher-order quantification is not at all new. The thought has been around for long enough that there are some standard objections to the proposal. All of them are present in the following passage by George Bealer, which stuck in my mind when I first read it and for a long time convinced me that a type-ambiguity approach was a complete non-starter:

> In my view, there are serious difficulties facing the prosentential theory. The most salient in the present context is perhaps that, syntactically, the theory is rigidly typed. The parts of discourse which the theory is designed to capture are manifestly type-free, however, as the following sorts of examples indicate: 'Some things are neither true nor false; for example, commands, questions, rules of inference, intellectual movements, governments, artistic styles, sensations, events, and, of course, persons and physical objects'. 'Murphy's Law is that whatever can go wrong does go wrong. When I first heard of O'Reiley's Law, I mistakenly thought that it was the same thing as Murphy's Law, but it is not. O'Reiley's Law is the blackjack that O'Reiley keeps behind the bar at his saloon'. 'When I was young, the things I cared most about were things that I could see or feel, but now they are things I can know to be true'. Etc. (Bealer 1993: 9, n. 8)

Similar arguments are made by Chierchia (1982), Menzel (1986), Hofweber (2022), Button and Trueman (2024), and many others. There are at least three challenges to a type-ambiguity approach contained in Bealer's passage. The first we have already addressed: even if 'O'Reiley's Law' is interpreted as type e and 'Murphy's Law' as type t, we can get 'O'Reiley's Law is the same thing as Murphy's Law' to be meaningful (and false) by assigning 'is the same thing as' some empty fallback meaning in type $\overline{et}$.[44] The second challenge involves uses of quantifiers ('some things') which

[44] A follow-up challenge is to say how prefixing 'I mistakenly thought that...' to the identity sentence can turn it into a truth. But this is bound up with general puzzles about attitude verbs which are beyond the scope of the present paper.

seem somehow to span multiple types simultaneously. And the third challenge—which could be raised more simply by a sentence like ‘Questions and persons are neither true nor false’—involves the possibility of *co-ordination*—i.e., linking with words like ‘and’ or ‘or’—of expressions whose meanings differ in type. Button and Trueman (2024: 19) helpfully isolate the third challenge from the second, giving a battery of further examples including ‘Plato loves Socrates and wisdom’ and ‘Mary can see roses, but not the colour red’.

I’ll focus initially on the co-ordination challenge, turning to the one about quantification in §6.4. To avoid getting distracted by the semantics of plurals, let’s use an example using ‘or’ rather than ‘and’:

(22) [Either] Mars or redness is interesting.

We would like this to turn out interchangeable with

(23) Mars is interesting or redness is interesting.

In particular, we would like to be able to derive that (22) and (23) both express

$$\mathbf{int}^{\bar{e}}\,\mathbf{mars}^{e} \vee \mathbf{int}^{\underline{\bar{e}}}\,\mathbf{red}^{\bar{e}}$$

This is straightforward for (23): we can simply treat the two disjuncts separately and combine them by positing that ‘or’ expresses disjunction ($\mathsf{or}\ \vdots_{\overline{tt}}\ (\vee)$). But it is hard to see how to arrange this for (22), given that we only have one occurrence of ‘interesting’ to work with. The one way we have of deriving semantic claims about sentences, namely Function Application, only gives us access to one meaning at a time for any given occurrence of a word.[45]

With other forms of ambiguity, we do get “zeugmatic” readings of sentences like ‘She came home in a sedan chair and a flood of tears’, where ‘in’ seems to be simultaneously functioning in two different ways. But it’s far from obvious how such uses are to be explained, and anyway it does not seem plausible that whatever is going with them also accounts for (22).

The underlying difficulty is not just about the role of ‘and’ and ‘or’. Similar issues can be raised using ellipsis (‘Either Mars is interesting or redness is’) and anaphora (‘Mars is interesting and so is redness’). In all these cases, the type-ambiguity theorist seems to need a single occurrence of a word to somehow simultaneously play two different semantic roles. While I won’t discuss ellipsis and anaphora, the technique I’ll use to account for cases of mixed co-ordination should be easily to integrate with standard accounts of these phenomena as well.

[45] Indeed, with our current toolkit, we cannot derive *any* meaning for (22), even a clearly incorrect one.

6.1 Preliminaries to a solution 1: co-ordination and lifting

To have any hope of addressing the challenge raised by tricky sentences like (22), we had better first make sure we have a way of handling not-so-tricky sentences like (24) and (25):

(24) [Either] Mars or Venus is interesting.

(25) [Either] redness or blueness is interesting.

Semanticists have developed many different accounts of 'or' which can account for sentences like (24); for the analogous sentence using 'and', there is an even more daunting menu of theoretical options. As far as I know, the solution I will propose to the puzzle raised by (22) can be adapted to work with any of these approaches. But for the sake of concreteness, I will adopt perhaps the best-known semantic treatment of sentences like (24), due to Partee and Rooth (1983), generalizing the earlier work of Montague (1974). This solution has two elements. The first is a treatment of the word 'or' as type-ambiguous, indeed infinitely type-ambiguous, which we can capture with the following postulates:

(26)
a. $\text{or} \vdots \text{s}\backslash\text{s}/\text{s}\ (\vee)$
b. $\text{or} \vdots \text{n}\backslash\text{n}/\text{n}\ \lambda X^{\overline{\sigma}} Y^{\overline{\sigma}} z^{\sigma}.Xz \vee Yz$
c. $\text{or} \vdots c\backslash c/c\ R \to \text{or} \vdots (c/d)\backslash(c/d)/(c/d)\ \lambda X^{\sigma\to\tau} Y^{\sigma\to\tau} z^{\sigma}.R(Xz)(Yz)$
d. $\text{or} \vdots c\backslash c/c\ R \to \text{or} \vdots (c\backslash d)\backslash(c\backslash d)/(c\backslash d)\ \lambda X^{\sigma\to\tau} Y^{\sigma\to\tau} z^{\sigma}.R(Xz)(Yz)$

This approach accounts not just for disjunctions of sentences, but verb phrases ('is red or is hot'), transitive verbs ('precedes or succeeds'), quantificational noun phrases ('every star or every planet'), quantificational determiners ('most or all'), and many other expressions. However, since none of the interpretations of 'or' provided by (26) takes type-*e* arguments, it isn't immediately clear how to apply it to (24). But this is easy when we combine (26) with the second element of the solution, namely the following extremely general *type-shifting schema*:

Lift

$$a \vdots_{\sigma} c\, x \to a \vdots_{(\sigma\to\tau)\to\tau} (d/c)\backslash d\ \lambda Y^{\sigma\to\tau}.Yx$$
$$a \vdots_{\sigma} c\, x \to a \vdots_{(\sigma\to\tau)\to\tau} d/(c\backslash d)\ \lambda Y^{\sigma\to\tau}.Yx$$

Using the instance of Lift where $\sigma = e$ and $\tau = t$, we can now provide a plausible account of why (24) means what it does:

or ⁝ s\s/s (∨) (26a)

or ⁝ (s/np\s)\(s/np\s)/(s/np\s) (26c)
$\lambda U^{\bar{e}} V^{\bar{e}} X^{\bar{e}}.VX \vee UX$

Venus ⁝ np
venus^{e}
Lift: Venus ⁝ s/np\s
$\lambda X^{\bar{e}}.X\,\textbf{venus}^{e}$

Mars ⁝ np
mars^{e}
Lift: Mars ⁝ s/np\s
$\lambda X^{\bar{e}}.X\,\textbf{mars}^{e}$

>: or·Venus ⁝ (s/np\s)\(s/np\s)
$\lambda V^{\bar{e}} X^{\bar{e}}.VX \vee X\,\textbf{venus}^{e}$

<: Mars·or·Venus ⁝ s/np\s
$\lambda X^{\bar{e}}.X\,\textbf{mars}^{e} \vee X\,\textbf{venus}^{e}$

is·interesting ⁝ np\s
int$^{\bar{e}}$

>: Mars·or·Venus·is·interesting ⁝ s
$\textbf{int}^{\bar{e}}\,\textbf{mars}^{e} \vee \textbf{int}^{\bar{e}}\,\textbf{venus}^{e}$

The treatment of (25) will be parallel, substituting $\bar{e}$ for e throughout.[46]

Note that while (26) generates interpretations for 'or' in all types of the form $\tau \to \tau \to \tau$ (remembering that e is not a terminal type), these interpretations are not available relative to just any category, or even relative to any category of the form $c\backslash c/c$. In particular, (26) does not imply that 'or' expresses anything relative to the category np\np/np. That's good, since otherwise we could derive the implausible conclusion that 'Redness or blueness is interesting' can mean the same as 'The property of being red or blue is interesting':

or ⁝ np\np/np
$\lambda U^{\bar{e}} V^{\bar{e}} x^{e}.Vx \vee Ux$

blueness ⁝ np
blue$^{\bar{e}}$

>: or·blueness ⁝ np\np
$\lambda V^{\bar{e}} x^{e}.Vx \vee \textbf{blue}^{\bar{e}}\,x$

redness ⁝ np
red$^{\bar{e}}$

<: redness·or·blueness ⁝ np
$\lambda x^{e}.\,\textbf{red}\,x \vee \textbf{blue}^{\bar{e}}\,x$

is·interesting ⁝ np\s
int$^{\bar{e}}$

<: redness·or·blueness·is·interesting ⁝ s
$\textbf{int}^{\bar{e}}\,\lambda x^{e}.\,\textbf{red}^{\bar{e}}\,x \vee \textbf{blue}^{\bar{e}}\,x$

Note that the problem isn't with the claim that 'or' expresses the disjunction operation $\lambda U^{\bar{e}} V^{\bar{e}} x^{e}.Vx \vee Ux$ of type $\bar{e} \to \bar{e} \to \bar{e}$. (26) implies that it *does* express this relative

[46] The combination of Lift and (26) is quite powerful, since it allows 'or' to take arbitrarily wide scope over any other scope-taking elements in a sentence. No matter how deeply embedded a disjunction might be, we can get a widest-scope interpretation by Lifting each disjunct to a complex category that takes all the remaining ingredients of the sentence as arguments, and combining the associated high-type meanings using the relevant meaning of 'or'. Partee and Rooth (1983) believe that at least in the case of 'and', many of these readings are in fact impossible—for example, they claim that 'John caught and ate a fish' cannot mean *John caught a fish and John ate a fish*. As a result, they posit a 'processing strategy' on which 'all expressions are interpreted at the lowest type possible, invoking higher-type homonyms only when needed for type coherence'. In Rooth and Partee 1982 they introduce a very different treatment of 'or', based on dynamic semantics, that can generate wide-scope readings for 'or' (but not 'and') even when everything is interpreted in its lowest possible type. Kubota and Levine (2020: ch. 4), by contrast, argue that there are no in-principle limits to the ability of either 'and' or 'or' to take arbitrarily wide scope.

to the category (np\s)\(np\s)/(np\s), and this will be crucial for explaining why, e.g., 'is red or is blue' means what it does. This further illustrates the reason we needed to work in a system where expressing is relativized to syntactic categories, rather than relying on something like Type-Driven Composition.

6.2 Preliminaries to a solution 2: sum types

With a serviceable general account of noun phrase co-ordination under our belt, let's return to the sentence (22) ('Mars or redness is interesting') which seems to doom the deflationary approach. My approach will be roundabout. In this section, I will introduce a more complex kind of higher-order language than we have been working with so far, with a richer system of types and new primitive term-forming operations (analogous to λ-abstraction and application) that allow for the construction of terms with the new types, along with an extension of our basic higher-order logic with new conversion rules (analogous to β and η conversion) for manipulating these new terms. In §6.3, I will suggest a straightforward extension of the semantic theory we have been developing, using the resources of this extended language, which can account for problematic co-ordinations like 'Either Mars or redness is interesting'. We can imagine this theory being formulated by a different group of field linguists who have come to Earth from the exoplanet Iota Draconis c, whose inhabitants natively speak such an extended higher-order language. §6.4 will extend this theory to cover mixed quantification. Finally, §6.5 will show how the theory can provide a template for an equally adequate account of the same phenomena that can be be stated in the simply-typed higher-order language of Lambda Serpentis b.

The extended language is characterized by an extension of our type-system to include *sum types*. Whenever α and β are types, we have a new type $\alpha + \beta$. We add a new primitive way of forming terms of these types: when A is a term of type α, $\iota^1_\beta A$ is a term of type $\alpha + \beta$, and when B is a term of type β, $\iota^1_\alpha B$ is a term of type $\alpha + \beta$. The intuitive gloss is that insofar as you were thinking of α and β as names for sets (something you really should not do!),you should think of $\alpha + \beta$ as a name for the *disjoint union* of these two sets: a set that contains a representative $\iota^1_\beta x$ for every x of type α, a separate representative $\iota^2_\alpha y$ for every y in β, and nothing else. But this heuristic must be taken with a generous sprinkling of salt, since the interest of this whole project depends on *not* thinking of higher order logic as "set theory in sheep's clothing" (Quine 1970).

In the system I will use, sum types are *non-terminal* types, which means that while we have types like $(e+t) \to t$, we do not have types like $t \to (e+t)$. Such types won't be needed for the theory, and allowing them would complicate the project of translating from the extended language to the simply-typed language.[47]

In addition to ι^1 and ι^2, we need one further term-forming operation. When F

[47] The absence of types like $t \to (e + t)$ explains why we needed to take ι^1_β and ι^2_β are mere punctuation, like the letter λ, rather than as terms in their own right.

is a term of type $\alpha \to \gamma$ and G is a term of type $\beta \to \gamma$, $\langle\!\langle F, G\rangle\!\rangle$ is a term of type $(\alpha+\beta) \to \gamma$. Intuitively, $\langle\!\langle F, G\rangle\!\rangle$ is a function that looks at its argument (of type $\alpha+\beta$) to see whether it is the representative of something from type α or of something from type β, and depending on which case obtains, applies either F or G to whatever the argument represents. This intuitive gloss is reflected by the addition to the logic of new "injection-conversion" rules, playing an analogous rule to β-conversion for the new types. The first rule lets us freely replace any constituent of the form $\langle\!\langle F, G\rangle\!\rangle\iota^1_\beta A$ with FA; the second lets us replace any constituent of the form $\langle\!\langle F, G\rangle\!\rangle\iota^2_\alpha B$ with GB.[48]

(In the official system (see Appendix B), $\langle\!\langle F, G\rangle\!\rangle$ is not actually primitive, but defined in terms of an operation of "delta-application". This brings certain advantages of generality and elegance, but since it's a bit harder to parse I'll refer the reader to the appendix for the details.)

Apart from these additions, the higher-order logic we'll assume for the languages using the extended type system is the same as the one we have already using for languages using only simple types. One feature worth noting is that we don't need to take quantifiers for the new types as primitive: there turns out to be a natural way of defining the new quantifiers out of the old ones, using the new term-forming operations. The details of this are given the Appendix, but the most important clause is the following:

$$\forall_{\alpha+\beta} := \lambda X^{\overline{\alpha+\beta}}.(\forall y^{\alpha}.X\iota^1_\beta y) \wedge (\forall z^{\beta}.X\iota^2_\alpha z)$$

That is: a property of type-$\alpha+\beta$ things is defined to be universal iff it holds of the representative of every type-α thing and of every type-β thing.

Now, if I managed to convince you that you can understand our starting version type theory, metaphysical alarm bells should be ringing. Doesn't the very idea of a sum-type like $e+t$ amount to taking a perspective "outside the type hierarchy" in which one thinks in terms of a big domain of "things", of which type-e things and

[48] To illustrate the application of these conversion rules, here is how we can use them to establish that the left injection $\iota^1_\beta x$ of some type-α thing x is not also the left-injection of something other than x, or the right-injection of anything:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\iota^1_\beta x =_{\alpha+\beta} \iota^1_\beta y]^1 \qquad \dfrac{\dfrac{\dfrac{}{x =_\alpha x}\,{=}\text{Intro}}{(\lambda u^\alpha.x =_\alpha u)x}\,\text{conv}}{\langle\!\langle \lambda u^\alpha.x =_\alpha u, \lambda v^\beta.\bot\rangle\!\rangle\iota^1_\beta x}\,\text{conv}}{\langle\!\langle \lambda u^\alpha.x =_\alpha u, \lambda v^\beta.\bot\rangle\!\rangle\iota^1_\beta y}\,{=}\text{Elim}}{(\lambda u^\alpha.x =_\alpha u)y}\,\text{conv}}{x =_\alpha y}\,\text{conv}}{\iota^1_\beta x =_{\alpha+\beta} \iota^1_\beta y \to x =_\alpha y}\,{\to}\text{Intro 1}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{[\iota^1_\beta x =_{\alpha+\beta} \iota^2_\alpha z]^1 \qquad \dfrac{\dfrac{\dfrac{}{x =_\alpha x}\,{=}\text{Intro}}{(\lambda u^\alpha.x =_\alpha u)x}\,\text{conv}}{\langle\!\langle \lambda u^\alpha.x =_\alpha u, \lambda v^\beta.\bot\rangle\!\rangle\iota^1_\beta x}\,\text{conv}}{\langle\!\langle \lambda u^\alpha.x =_\alpha u, \lambda v^\beta.\bot\rangle\!\rangle\iota^2_\alpha z}\,{=}\text{Elim}}{(\lambda v^\beta.\bot)z}\,\text{conv}}{\bot}\,\text{conv}}{\iota^1_\beta x \neq_{\alpha+\beta} \iota^2_\alpha z}\,\neg\text{Intro 1}$$

type-t things are just two of many varieties? And isn't it integral to the metaphysically interesting interpretation of higher-order logic that it rejects the very idea of such a perspective?

On the other hand, the spirit of the deflationary approach to property-talk in natural language is that when we are looking at a linguistic practice that doesn't neatly fit the syntactic limitations characteristic of simply-typed higher-order languages, we should not be quick to assume that the practitioners are in the grip of some metaphysical picture radically at odds with those popular on Lambda Serpentis b. If our field linguists travel to Iota Draconis, the same charitable sensibilities which suggested a higher-order semantics for English property-talk will lead them to look for a similarly "deflationary" accounts of the language spoken there. And in fact, as I will explain below in §6.5, this project will meet with success. There is a systematic way of "translating" from a higher-order language with sum types to a simply-typed higher-order language, such that the sentences valid in the higher-order logic with sum-types are exactly those that are mapped to sentences valid in simply-typed higher-order logic.

Given the availability of this "reduction", sum-types are not in fact crucial to the argument of this paper. Applying the translation procedure to the semantic theory for English given in a higher-order language with sum-types will yield an equally adequate semantic theory for English in a simply-typed higher-order language. There is thus in principle no need to think at all about extra types, term-forming operations, or conversion rules. However, this apparatus is helpful in getting an intuitive understanding of what's going on, since the new semantic primitives and principles needed to account for the puzzles look more familiar when presented in the extended language.

6.3 Explaining mixed co-ordination with sum-types

For field linguists from Iota Draconis c, whose native dialect of Higher-Orderese includes sum-types, the challenge of extending the deflationist semantics to handle 'Mars or redness is interesting' is not a very hard one.[49] They can extend their type-ambiguous treatment of 'is interesting' to allow it to express not only something of type $\bar{e}$ (namely $\mathbf{int}^{\bar{e}}$) and something of type $\underline{\bar{e}}$ (namely $\mathbf{int}^{\underline{\bar{e}}}$), but something of type $\overline{e + e}$ that behaves like $\mathbf{int}_e$ on arguments of the form $\iota^1_e x^e$ and like $\mathbf{int}_{\bar{e}}$ on arguments of the form $\iota^2_e Y^{\bar{e}}$. In other words, we can say that

(27) $$\text{is·interesting} \,\vdots\, \text{np}\backslash\text{s} \,\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle$$

Here, $\vdots$ is as a constant of type $\varepsilon \to \varepsilon \to ((\bar{e} + \underline{\bar{e}}) \to t)$. To use this new entry for 'interesting', we'll need to find a way to get 'Mars or redness' to express something of the higher type $\underline{\overline{e + e}}$ that can combine with it by function application to make a

[49]For the general utility of sum-types as a way of handling co-ordination of semantically dissimilar expressions, see Morrill 1994 (§6.1) and Carpenter 1997 (§6.2.4).

proposition. There are a few ways we could do this. One is to have a general law that allows for arbitrary words to be shifted into sum-types:

Inject $$a \mathrel{\vdots_{\alpha}} c\, x \to a \mathrel{\vdots_{\alpha+\beta}} c\, \iota^1_{\beta} x \wedge a \mathrel{\vdots_{\beta+\alpha}} c\, \iota^2_{\beta} x$$

Note that we use the same syntactic category c for the new sum-type meaning as for the original one.[50] Using Inject together with Lift, we can derive the meaning we want for (22). (We abbreviate the category s/np\s as $\text{np}^{\uparrow}$.)

$$\frac{\text{Mars} \vdots \text{np} \quad \mathbf{mars}^{e}}{\text{Mars} \vdots \text{np} \quad \iota^1_{\bar{e}}\, \mathbf{mars}^{e}}\ \text{Inject} \qquad \frac{\text{Mars} \vdots \text{np} \quad \iota^1_{\bar{e}}\, \mathbf{mars}^{e}}{\text{Mars} \vdots \text{np}^{\uparrow} \quad \lambda X^{\overline{e+\bar{e}}}. X \iota^1_{\bar{e}}\, \mathbf{mars}^{e}}\ \text{Lift}$$

$$\frac{\text{redness} \vdots \text{np} \quad \mathbf{red}^{\bar{e}}}{\text{redness} \vdots \text{np} \quad \iota^2_{e}\, \mathbf{red}^{\bar{e}}}\ \text{Inject} \qquad \frac{\text{redness} \vdots \text{np} \quad \iota^2_{e}\, \mathbf{red}^{\bar{e}}}{\text{redness} \vdots \text{np}^{\uparrow} \quad \lambda X^{\overline{e+\bar{e}}}. X \iota^2_{e}\, \mathbf{red}^{\bar{e}}}\ \text{Lift}$$

$$\frac{\text{or} \vdots \text{np}^{\uparrow}\backslash\text{np}^{\uparrow}/\text{np}^{\uparrow} \quad \lambda U^{\underline{\overline{e+\bar{e}}}} V^{\underline{\overline{e+\bar{e}}}} X^{\overline{e+\bar{e}}}. VX \vee UX \qquad \text{redness} \vdots \text{np}^{\uparrow} \quad \lambda X^{\overline{e+\bar{e}}}. X \iota^2_{e}\, \mathbf{red}^{\bar{e}}}{\text{or}\cdot\text{redness} \vdots \text{np}^{\uparrow}\backslash\text{np}^{\uparrow} \quad \lambda V^{\underline{\overline{e+\bar{e}}}} X^{\overline{e+\bar{e}}}. VX \vee X \iota^2_{e}\, \mathbf{red}^{\bar{e}}}\ >$$

$$\frac{\text{Mars} \vdots \text{np}^{\uparrow} \quad \lambda X^{\overline{e+\bar{e}}}. X \iota^1_{\bar{e}}\, \mathbf{mars}^{e} \qquad \text{or}\cdot\text{redness} \vdots \text{np}^{\uparrow}\backslash\text{np}^{\uparrow} \quad \lambda V^{\underline{\overline{e+\bar{e}}}} X^{\overline{e+\bar{e}}}. VX \vee X \iota^2_{e}\, \mathbf{red}^{\bar{e}}}{\text{Mars}\cdot\text{or}\cdot\text{redness} \vdots \text{np}^{\uparrow} \quad \lambda X^{\overline{e+\bar{e}}}. X \iota^1_{\bar{e}}\, \mathbf{mars}^{e} \vee X \iota^2_{e}\, \mathbf{red}^{\bar{e}}}\ <$$

$$\frac{\text{Mars}\cdot\text{or}\cdot\text{redness} \vdots \text{np}^{\uparrow} \quad \lambda X^{\overline{e+\bar{e}}}. X \iota^1_{\bar{e}}\, \mathbf{mars}^{e} \vee X \iota^2_{e}\, \mathbf{red}^{\bar{e}} \qquad \text{is}\cdot\text{interesting} \vdots \text{np}\backslash\text{s} \quad \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle}{\text{Mars}\cdot\text{or}\cdot\text{redness}\cdot\text{is}\cdot\text{interesting} \vdots \text{s} \quad \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle \iota^1_{\bar{e}}\, \mathbf{mars}^{e} \vee \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle \iota^2_{e}\, \mathbf{red}^{\bar{e}}}\ >$$

$$\frac{\text{Mars}\cdot\text{or}\cdot\text{redness}\cdot\text{is}\cdot\text{interesting} \vdots \text{s} \quad \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle \iota^1_{\bar{e}}\, \mathbf{mars}^{e} \vee \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle \iota^2_{e}\, \mathbf{red}^{\bar{e}}}{\text{Mars}\cdot\text{or}\cdot\text{redness}\cdot\text{is}\cdot\text{interesting} \vdots \text{s} \quad \mathbf{int}^{\bar{e}}\, \mathbf{mars}^{e} \vee \mathbf{int}^{\underline{\bar{e}}}\, \mathbf{red}^{\bar{e}}}\ \text{conv}$$

The crucial thing here is the new entry (27). It is natural to look for some general principle which could let us derive this from our old (simply-typed) entries is·interesting ⁝ np/s $\mathbf{int}^{\bar{e}}$ and is·interesting ⁝ np/s $\mathbf{int}^{\underline{\bar{e}}}$. The following law, which allows us to combine any two meanings of a type-ambiguous predicate into a meaning taking a sum-type argument, would do the trick:

Combination $$a \mathrel{\vdots_{\alpha\to\gamma}} c\, x \wedge a \mathrel{\vdots_{\beta\to\gamma}} c\, y \to a \mathrel{\vdots_{\alpha+\beta\to\gamma}} c\, \langle\!\langle x, y \rangle\!\rangle$$

Unfortunately, Combination will have problematic consequences if we want a system that can not only deal with the specific controversial claims of type-ambiguity required by a deflationary account of 'property' talk, but with other more familiar cases of ambiguity or semantic multiplicity. For by applying Combination to a single occurrence of an ambiguous expression, we will be able to generate weird "zeugmatic"

[50] By contrast, Morrill (1994) has sum-categories as well as sum-types (as one would expect given that his system conforms to Correspondence), meaning that we can derive meanings of the same type and category for any two expressions whatsoever. This gives rise to worries about overgeneration (Milward 1994, Carpenter 1997: §6.2.4). No analogous problems arise for our rule, because of its limitation to a single category.

readings in which that occurrence functions like multiple occurrences with different interpretations. For example, given is·a·bank $\vdots_{\bar{e}}$ $\mathbf{mbank}^{\bar{e}}$ and is·a·bank $\vdots_{\bar{e}}$ $\mathbf{rbank}^{\bar{e}}$, we'll be able to derive a further entry is·a·bank $\vdots_{\overline{e+e}}$ $\langle\!\langle \mathbf{mbank}^{\bar{e}}, \mathbf{rbank}^{\bar{e}} \rangle\!\rangle$. And using this hybrid entry we'll be able, e.g., to derive a false reading for 'Either Barclays or Willowy is a bank', where 'Barclays' names a certain money-bank and 'Willowy' names a certain river-bank.

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{\dfrac{\begin{array}{c}\text{Barclays} \vdots \text{np}\\ \mathbf{b}^e\end{array}}{\begin{array}{c}\text{Barclays} \vdots \text{np}\\ \iota^2_e\,\mathbf{b}^e\end{array}}\,\text{Inject}}{\begin{array}{c}\text{Barclays} \vdots \text{np}^\uparrow\\ \lambda X^{\overline{e+e}}.X\iota^2_e\,\mathbf{b}^e\end{array}}\,\text{Lift}
\quad
\dfrac{\begin{array}{c}\text{or} \vdots \text{np}^\uparrow\backslash\text{np}^\uparrow/\text{np}^\uparrow\\ \lambda U^{\underline{\overline{e+e}}}V^{\underline{\overline{e+e}}}X^{\overline{e+e}}.VX \lor UX\end{array}
\quad
\dfrac{\dfrac{\begin{array}{c}\text{Willowy} \vdots \text{np}\\ \mathbf{w}^e\end{array}}{\begin{array}{c}\text{Willowy} \vdots \text{np}\\ \iota^1_e\,\mathbf{w}^e\end{array}}\,\text{Inject}}{\begin{array}{c}\text{Willowy} \vdots \text{np}^\uparrow\\ \lambda X^{\overline{e+e}}.X\iota^1_e\,\mathbf{w}^e\end{array}}\,\text{Lift}}{\begin{array}{c}\text{or·Willowy} \vdots \text{np}^\uparrow\backslash\text{np}^\uparrow\\ \lambda V^{\underline{\overline{e+e}}}X^{\overline{e+e}}.VX \lor X\iota^2_e\,\mathbf{w}^e\end{array}}\,>
}{\begin{array}{c}\text{Barclays·or·Willowy} \vdots \text{np}^\uparrow\\ \lambda X^{\overline{e+e}}.X\iota^2_e\,\mathbf{b}^e \lor X\iota^1_e\,\mathbf{w}^e\end{array}}\,<
\quad
\begin{array}{c}\text{is·a·bank} \vdots \text{np}\backslash\text{s}\\ \langle\!\langle \mathbf{mbank}^{\bar{e}}, \mathbf{rbank}^{\bar{e}} \rangle\!\rangle\end{array}
}{\begin{array}{c}\text{Barclays·or·Willowy·is·a·bank} \vdots \text{s}\\ \langle\!\langle \mathbf{mbank}^{\bar{e}}, \mathbf{rbank}^{\bar{e}} \rangle\!\rangle\iota^2_e\,\mathbf{b}^e \lor \langle\!\langle \mathbf{mbank}^{\bar{e}}, \mathbf{rbank}^{\bar{e}} \rangle\!\rangle\iota^1_e\,\mathbf{w}^e\end{array}}\,>
}{\begin{array}{c}\text{Barclays·or·Willowy·is·a·bank} \vdots \text{s}\\ \mathbf{rbank}^{\bar{e}}\,\mathbf{b}^e \lor \mathbf{mbank}^{\bar{e}}\,\mathbf{w}^e\end{array}}\,\text{conv}
$$

Since Barclays isn't a river-bank and Willowy isn't a money-bank, both disjuncts are false, whereas prima facie the sentence should only allow the true readings *Either Barclays or Willowy is a money bank* and *Either Barclays or Willowy is a river bank*.[51] The moral is that we should not try to derive (27) from some general principle applying to all kinds of ambiguity. Rather, we should see it as capturing a distinctive kind of harmony that obtains between the type-$\bar{e}$ and type-$\underline{\bar{e}}$ meanings we have posited for 'is interesting'.[52]

[51]One might hope to avoid this unwelcome consequence by restricting Combination to require α and β to be two different types. But this just postpones the problem. Recall that in our system, 'is a bank' will also have a type $\underline{\bar{e}}$ meaning (presumably something empty like $\lambda x^{\bar{e}}.\bot$), to account for the meaningfulness of 'Redness is a bank'. Using this, we could apply the restricted version of Combination twice, to get first

$$\text{is·a·bank} \vdots \text{np}\backslash\text{s}\ \langle\!\langle \mathbf{mbank}^{\bar{e}}, \lambda x^e.\bot \rangle\!\rangle$$

and then

$$\text{is·a·bank} \vdots \text{np}\backslash\text{s}\ \langle\!\langle \mathbf{mbank}^{\bar{e}}, \langle\!\langle \mathbf{rbank}^{\bar{e}}, \lambda x^e.\bot \rangle\!\rangle \rangle\!\rangle$$

This will generate an unwelcome false reading for 'Barclays, Willowy, or redness is a bank'.

[52]There *are* some special circumstances in which this kind of two-faced interpretation is arguably possible given an appropriate context: see the counterexamples to 'Repetition' in Dorr 2014*b* (§ 8)

6.4 Explaining cross-type quantification with sum-types

Sum-types also help with a related problem. Plausibly,

(28) Everything Mary mentioned is interesting

has a reading where it both entails 'If Mary mentioned Mars, Mars is interesting' and 'If Mary mentioned redness, redness is interesting'. We can secure such readings easily by extending the type-ambiguity we have already postulated in the quantifiers to cover sum-types. For example, we will allow entries like:

$$\text{every} \vdots (\text{s/np}\backslash\text{s})/\text{n} \ \lambda X^{\overline{e+\bar{e}}} Y^{\overline{e+\bar{e}}}.\forall z^{e+\bar{e}}.Xz \rightarrow Yz$$

Similarly, mutatis mutandis, for unary quantifiers:

$$\text{everything} \vdots \text{s/np}\backslash\text{s} \ \forall_{e+\bar{e}}$$

Recall that quantifiers for sum-types are defined, in the background logic, using quantifiers for simple types. In particular,

$$\forall_{e+\bar{e}} := \lambda X^{e+\bar{e}}.(\forall y^e.X\iota^1_{\bar{e}}y) \wedge (\forall z^{\bar{e}}.X\iota^2_e z).$$

Allowing quantifier-words to express sum-type quantification thus gives the effect of quantifying in several types at once. For example:

$$\dfrac{\dfrac{\begin{array}{c}\text{everything} \vdots \text{s/np}\backslash\text{s} \\ (\lambda X^{e+\bar{e}}.(\forall y^e.X\iota^1_{\bar{e}}y) \wedge (\forall z^{\bar{e}}.X\iota^2_e z))\end{array} \qquad \begin{array}{c}\text{is·interesting} \vdots \text{np}\backslash\text{s} \\ \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\bar{e}}_{-} \rangle\!\rangle\end{array}}{\begin{array}{c}\text{everything·is·interesting} \vdots \text{s} \\ (\forall y^e.\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\bar{e}}_{-} \rangle\!\rangle \iota^1_{\bar{e}}y) \wedge (\forall z^{\bar{e}}.\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\bar{e}}_{-} \rangle\!\rangle \iota^2_e z)\end{array}} >}{\begin{array}{c}\text{everything·is·interesting} \vdots \text{s} \\ (\forall y^e.\, \mathbf{int}^{\bar{e}}\, y) \wedge (\forall z^{\bar{e}}.\, \mathbf{int}^{\bar{e}}_{-}\, z)\end{array}} \text{conv}$$

The same treatment will carry over to sentences like (28).[53]

(which also work when we use co-ordination rather than ellipsis). Combination might explain how those cases work; however, we would not want to lose sight of the contrast between these special cases and completely unproblematic examples (22).

[53]The above derivation turned on our decision to treat quantifiers for sum-types as abbreviations. But even if one wanted to treat all quantifiers as primitive, any remotely adequate background metaphysics should at least imply the necessitated biconditional

$$\Box(\forall_{\alpha+\beta}F \leftrightarrow ((\forall y^\alpha.F\iota^1_\beta y) \wedge (\forall z^\beta.F\iota^2_\alpha z)))$$

and hence also

$$\Box(\forall_{e+\bar{e}}\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\bar{e}}_{-} \rangle\!\rangle \leftrightarrow ((\forall y^e.\, \mathbf{int}^{\bar{e}}\, y) \wedge (\forall z^{\bar{e}}.\, \mathbf{int}^{\bar{e}}_{-}\, z)))$$

So the conclusion of the above derivation will at least be true up to necessary equivalence.

This solution to the problem of cross-type quantification is not limited to quantifiers like 'every' and 'some' that commute with conjunction or disjunction. For example, 'Most *F*s are *G*' can plausibly be analyzed as 'any relation that every *F*-and-*G* thing bears to some *F*-and-not-*G* thing fails to be one-to-one':

$$\mathbf{most}^{\cancel{\overline{\alpha}\overline{\alpha}}} := \lambda F^{\overline{\alpha}} G^{\overline{\alpha}}.\forall R^{\overline{\alpha\alpha}}.$$
$$(\forall x^{\alpha}.(Fx \wedge Gx) \to (\exists y^{\alpha}.Fy \wedge \neg Gy \wedge Rxy)) \to (\exists x^{\alpha} y^{\alpha} z^{\alpha}.Rxz \wedge Ryz \wedge y \neq_{\alpha} z)$$

We can use this to derive a cross-type meaning for 'Most essay topics are interesting':

$$\begin{array}{cc} \text{most} \vdots \text{s/np\textbackslash s/n} & \text{essay·topics} \vdots \text{n} \\ \mathbf{most}^{\cancel{\overline{e+\overline{e}}\,\overline{e+e}}} & \langle\!\langle \mathbf{topic}^{\overline{e}}, \mathbf{topic}^{\overline{e}}_{-} \rangle\!\rangle \\ \hline \end{array} >$$

$$\begin{array}{cc} \text{most·essay·topics} \vdots \text{s/np\textbackslash s} & \text{are·interesting} \vdots \text{np\textbackslash s} \\ \mathbf{most}^{\cancel{\overline{e+\overline{e}}\,\overline{e+e}}} \langle\!\langle \mathbf{topic}^{\overline{e}}, \mathbf{topic}^{\overline{e}}_{-} \rangle\!\rangle & \langle\!\langle \mathbf{int}^{\overline{e}}, \mathbf{int}^{\overline{e}}_{-} \rangle\!\rangle \\ \hline \end{array} >$$

$$\begin{array}{c} \text{most·essay·topics·are·interesting} \vdots \text{s} \\ \mathbf{most}^{\cancel{\overline{e+\overline{e}}\,\overline{e+e}}} \langle\!\langle \mathbf{topic}^{\overline{e}}, \mathbf{topic}^{\overline{e}}_{-} \rangle\!\rangle \langle\!\langle \mathbf{int}^{\overline{e}}, \mathbf{int}^{\overline{e}}_{-} \rangle\!\rangle \end{array}$$

We can convert the final denotation to an $\mathscr{L}$-sentence, analogous to what we did earlier with in the case of 'everything'. The result is a paragraph-length universal quantification over quadruples of relations of types $\overline{ee}$, $\overline{e}\overline{e}$, $\overline{e\,e}$, and $\cancel{\overline{e}\overline{e}}$, which does not wear its meaning on its sleeve. But given its equivalence to the more comprehensible $\mathscr{L}^{+}$-sentence, there is no doubt that it will behave as we would wish. For example, if we define exact numerical quantifiers $\exists^{n}_{\sigma}$ in the standard way, the sentence follows from

$$(\exists^{j} x^{e}.\, \mathbf{topic}^{\overline{e}}\, x) \wedge (\exists^{n} x^{e}.\, \mathbf{topic}^{\overline{e}}\, x \wedge \mathbf{int}^{\overline{e}}\, x) \wedge (\exists^{k} y^{\overline{e}}.\, \mathbf{topic}^{\overline{e}}_{-}\, y) \wedge (\exists^{m} y^{\overline{e}}.\, \mathbf{topic}^{\overline{e}}_{-}\, y \wedge \mathbf{int}^{\overline{e}}_{-}\, y)$$

if $n + m > (j + k)/2$, and the two are inconsistent otherwise.

Introducing these new options for the interpretation of English quantifier-words makes a big difference to our linguists' assessment of English-speaking philosophers' debates about doctrines like "materialism" and "nominalism". Imagine an anti-nominalist who puts forth the following argument:

> Some property of Mars applies to Venus—for example, *being a planet*. So, there are properties. But no property is material. So, not everything is material. But of course some things are material—for example, Mars. So, some but not all things are material.

Before, it might have seemed that we would have to say that this anti-nominalist was just getting confused, committing a fallacy of equivocation by failing to track the different semantic roles of the quantifiers in the first four sentences (which involve higher-order quantification) and the penultimate sentence (which involves first-order quantification). But now, we have the wherewithal to interpret the entire argument uniformly, using type $e + \overline{e}$ quantifiers. So interpreted, the argument is sound. Even if the type-$\bar{e}$ denotation of 'material' is universal (as e-Materialism holds), its type $\bar{\bar{e}}$ denotation is empty (maybe $\lambda X.\bot$), so the type-$\overline{e+\bar{e}}$ denotation is neither universal nor empty.

Of course we can still interpret 'Everything is material' as true (assuming e-Materialism) by taking the quantifier to be first order. When *nominalists* come out with this sentence, considerations of charity will tend to favour doing so. One might worry that such charity would be misplaced, analogous to a theist perversely interpreting an atheistic materialist's utterance of 'Absolutely everything is material' to mean 'Absolutely everything that is not God or an angel is material'. But that analogy is not apt, since the different candidate interpretations we are considering involve quantifiers of different types, not differently restricted quantifiers of a single type. And while there is a certain loose sense in which a sum-type quantifier $\forall_{\alpha+\beta}$ might be said to be "more inclusive" than the quantifiers $\forall_\alpha$ and $\forall_\beta$ in terms of which it is defined (at least when $\alpha \neq \beta$), there is no corresponding sense in which we can think of all three as restrictions of a "maximally inclusive" quantifier, since there is no type that could be assigned to such a quantifier.[54]

Notoriously, the presence of ambiguity in language—especially subtle ambiguity—is apt to lead to "merely verbal disputes", in which two parties mistakenly take one another to believe incompatible things, when in fact they are merely focusing on different readings of some ambiguous sentence. Type-ambiguity is no exception.

6.5 Dispensing with sum-types

The work we have done with sum-types in explaining facts about co-ordination and quantification in English might suggest that we are now in for a difficult debate about their intelligibility or metaphysical good standing. Fortunately, we can bypass the need for such a debate, since there is a way of systematically "translating" sentences and theories stated in the extended language $\mathscr{L}^+$ into the simply-typed language $\mathscr{L}$. The details are explained in Appendix C; here, I will just give a brief sketch that illustrates how the translation applies to the small fragment of the semantic theory for the extended language which we used in the previous section to derive an interpretation for a sentence involving cross-type quantification.

(29) $$\mathsf{everything{\cdot}is{\cdot}interesting} \mathrel{\vdots} \mathsf{s}\, (\forall y^{e}.\, \mathbf{int}^{\bar{e}}\, y) \wedge (\forall z^{\bar{e}}.\, \mathbf{int}^{\bar{\bar{e}}}\, z)$$

While this conclusion is a sentence in the original (simply-typed) higher-order language, our derivation of it turned on three premises—two new semantic postulates,

[54] As the author of a paper called 'There Are No Abstract Objects' (Dorr 2007), I can attest that from my present high-order-logic-loving standpoint, the interpretation of my former self as intending only type-e quantification feels correct. The balance of interpretative considerations may however be rather different for nominalists of a more aggressive temperament, who take their view to imply that the beliefs of ordinary folk are rife with error.

plus an instance of Function Application—given in the extended language.

(30) $\quad$ is·interesting $\vdots$ np\s $\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}} \rangle\!\rangle$

(31) $\quad$ everything $\vdots$ s/np\s $\forall_{e+\bar{e}}$

(32) $\quad \forall\, a^{\varepsilon} b^{\varepsilon} c^{\varepsilon} d^{\varepsilon} X^{\underline{\overline{e+\bar{e}}}} y^{\overline{e+\bar{e}}}.\, a \vdots_{\underline{\overline{e+\bar{e}}}}\, c/d\, X \wedge b \vdots_{\overline{e+\bar{e}}}\, d\, y \rightarrow a{\cdot}b \vdots_t c\, Xy$

These sentences contain two new nonlogical constants

$$\vdots_{\overline{e+\bar{e}}} : \epsilon \rightarrow \epsilon \rightarrow ((e+\bar{e}) \rightarrow t) \rightarrow t$$
$$\vdots_{\underline{\overline{e+\bar{e}}}} : \epsilon \rightarrow \epsilon \rightarrow (((e+\bar{e}) \rightarrow t) \rightarrow t) \rightarrow t$$

The translation depends on the observation that these two extended types are "isomorphic" to two simple types:

$$\epsilon \rightarrow \epsilon \rightarrow (e \rightarrow t) \rightarrow (\bar{e} \rightarrow t) \rightarrow t$$
$$\epsilon \rightarrow \epsilon \rightarrow ((e \rightarrow t) \rightarrow (\bar{e} \rightarrow t) \rightarrow t) \rightarrow t$$

By saying that types α and β are isomorphic, I mean that there is a function f from the set of type-α terms to the set of type-β terms and a function g from the set of type-β terms to the set of type-α terms such that $f(g(A))$ is convertible with A for any term A of type α, $g(f(B))$ is convertible with B for any B of type β, and both functions "commute with substitution up to convertibility".[55] When types are isomorphic in this sense, any theoretical work we do with a nonlogical constant of one type could be done equally well using a constant of the other type and converting as necessary. In our translated theory, the two new nonlogical constants $\vdots_{\overline{e+\bar{e}}}$ and $\vdots_{\underline{\overline{e+\bar{e}}}}$ are replaced by simply-typed surrogates $\vdots'_{\overline{e+\bar{e}}}$ and $\vdots'_{\underline{\overline{e+\bar{e}}}}$. We can turn any sentence using the former constants into a sentence using the surrogates just by applying the relevant isomorphism to each surrogate. And once we have turned a sentence into one all of whose *constants* have simple types, it turns out that—thanks to a theorem of Dag Prawitz—that we can always convert the resulting sentence into one all of whose *constituents* have simple types—i.e. a sentence of a simply-typed language. Applying this procedure to the three key premises given above yields the following:

(33) $\quad$ is·interesting $\vdots'_{\overline{e+\bar{e}}}$ np\s $\mathbf{int}^{\bar{e}}\ \mathbf{int}^{\underline{\bar{e}}}$

(34) $\quad$ everything $\vdots'_{\underline{\overline{e+\bar{e}}}}$ s/np\s $(\lambda X^{\bar{e}} Y^{\underline{\bar{e}}}.\forall_e X \wedge \forall_{\bar{e}} Y)$

(35) $\quad \forall\, a^{\epsilon} b^{\epsilon} c^{\epsilon} d^{\epsilon} X^{\underline{\overline{\bar{e}\bar{e}}}} y^{\bar{e}} z^{\underline{\bar{e}}}.\, a \vdots'_{\underline{\overline{e+\bar{e}}}}\, c/d\, X \wedge b \vdots'_{\overline{e+\bar{e}}}\, d\, y\, z \rightarrow a{\cdot}b \vdots_t c\, Xyz$

These three $\mathscr{L}$-sentences can easily be seen to imply (29) in H.[56]

[55] To say that f commutes with substitution up to convertibility is to say that whenever every free variable in B is safe for v in A, $f(A[B/v])$ is convertible with $f(A)[f(B)/v]$.

[56] If we instantiating the universally quantified variables in (35) respectively with everything, is·interesting, s, np\s, $(\lambda X^{\bar{e}} Y^{\underline{\bar{e}}}.\forall_e X \wedge \forall_{\bar{e}} Y)$, $\mathbf{int}^{\bar{e}}$, and $\mathbf{int}^{\underline{\bar{e}}}$, we get something that beta-converts to the conditional (34) $\wedge$ (33) $\rightarrow$ (29).

We can apply a parallel procedure to all of the new semantic constants and to the theory stated in terms of them. That theory is thus re-expressed in simply-typed terms as a theory about novel, primitively polyadic semantic relations. Some expressions primitively stand (relative to a category) in a kind of 'joint expressing' relation to multiple entities of different types, as 'is interesting' does according to (33). Other expressions, have meanings that require multiple arguments, as 'everything' does according to (34). The new rules for syntactic combination, like (35), generalize Function Application in such a way that an expression of the former sort can provide an expression of the second sort with all the arguments it needs. This theoretical architecture is conceptually independent of the idea of sum-types: it could have been invented on Lambda Serpentis. The formulation with sum-types has the advantage that we can stick with the more familiar ideology where our "expressing" predicates all take the same number of arguments, and the relatively familiar Function Application rules. But the existence of the translation means that we are free to adopt this mode of expression without worrying that we will thereby become subject to new intelligibility worries or "ideological costs".

6.6 Could we appeal to sum-types to avoid type-ambiguity?

If one really hated type-ambiguity, one could also use the apparatus of sum-types to account for some central facts about property-talk in English *without* any type-ambiguity, and without having to expand one's view of what individuals there are beyond the confines of e-Materialism. The strategy would be to choose, once and for all, some reasonably inclusive sum type α combining e with various other types such as $\bar{e}$, and say that α is the only type in which expressions of type np express anything, just as for us t is the only type in which expressions of type s express anything. Suppose for simplicity that α is $e + \bar{e}$. Then we will no longer accept that Mars ⁝ np $\mathbf{mars}^e$: instead, we will just say that Mars ⁝ np $\iota^1_e\ \mathbf{mars}^e$, as a basic postulate. Similarly, instead of $\mathbf{red}^{\bar{e}}$, 'is red' will express something of type $\overline{\alpha}$: presumably, $\langle\!\langle \mathbf{red}^{\bar{e}}, F\rangle\!\rangle$ for some empty $F : \bar{e}$. 'The property of being red' and 'redness' will likewise no longer express $\mathbf{red}^{\bar{e}}$ but instead its injection into type α, $\iota^2_e\ \mathbf{red}^{\bar{e}}$. 'Instantiates', finally, will, like any transitive verb, have a meaning of type $\overline{\alpha\alpha}$. By taking this to be $\langle\!\langle\langle\!\langle R, S\rangle\!\rangle, \langle\!\langle \lambda Y^{\bar{e}}.Y, T\rangle\!\rangle\rangle\!\rangle$ for some $R : \overline{ee}$, $S : \overline{ee}$, and $T : \overline{\bar{e}e}$—plausibly empty—we can predict that 'Mars instantiates redness' and 'Mars is red' express the same thing:

$$\dfrac{\text{Mars ⁝ np} \quad \text{is·red ⁝ np\textbackslash s}}{\text{Mars·is·red ⁝ s}} < \qquad \dfrac{\iota^1_e\ \mathbf{mars}^e \quad \langle\!\langle \mathbf{red}^{\bar{e}}, F\rangle\!\rangle}{\mathbf{red}^{\bar{e}}\ \mathbf{mars}^e}$$

$$\dfrac{\text{instantiates ⁝ np\textbackslash s/np} \quad \text{redness ⁝ np}}{\text{instantiates·redness ⁝ s}} > \qquad \dfrac{\langle\!\langle\langle\!\langle R, S\rangle\!\rangle, \langle\!\langle \lambda Y^{\bar{e}}.Y, T\rangle\!\rangle\rangle\!\rangle \quad \iota^2_e\ \mathbf{red}^{\bar{e}}}{\langle\!\langle \mathbf{red}^{\bar{e}}, T\,\mathbf{red}^{\bar{e}}\rangle\!\rangle}$$

$$\dfrac{\text{Mars ⁝ np} \quad \text{instantiates·redness ⁝ s}}{\text{Mars·instantiates·redness ⁝ s}} < \qquad \dfrac{\iota^1_e\ \mathbf{mars}^e \quad \langle\!\langle \mathbf{red}^{\bar{e}}, T\,\mathbf{red}^{\bar{e}}\rangle\!\rangle}{\mathbf{red}^{\bar{e}}\ \mathbf{mars}^e}$$

We can generalize this by claiming that the nominalization of any verb phrase expresses the injection into α of the $\bar{e}$-component of its meaning, ignoring the $\underline{\bar{e}}$-component:

Nominalization* $\quad \forall a^{\varepsilon} X^{\bar{e}}.(\exists Y^{\underline{\bar{e}}}.a \vdots_{\bar{\alpha}} \mathsf{np}\backslash\mathsf{s} \langle\!\langle X, Y \rangle\!\rangle) \leftrightarrow \mathbf{PROP}(a) \vdots_{\sigma} \mathsf{np}\, \iota^2_e X$

This treatment ensures that '[NP] [VP]' and '[NP] instantiates **PROP**([VP])' express the same thing when the NP expresses something of the form $\iota^1_e x^e$. But all bets are off when the NP expresses something of the form $\iota^2_e Y^{\bar{e}}$: '[NP] instantiates **PROP**([VP])' will then express $TY^{\bar{e}}Z^{\bar{e}}$ (where T is the type-$\overline{\bar{e}e}$ component of the meaning of 'instantiate' and Z is the type-$\bar{e}$ component of the meaning of the VP). If T is empty, all such sentences will be false. We could try to do better by finding some appropriate non-empty T such that, e.g., $T(\lambda x^e.x = x)(\lambda x^e.x = x)$, so that 'the property of being self-identical instantiates the property of being self-identical' can be true.[57] But without type-ambiguity (and given our classical background logic) there is of course no hope of getting a true reading for every instance of Naïve Property Comprehension. Nevertheless, one might argue that this approach does *sufficient* justice to our disposition to treat any (present-tense) VP as interchangeable with 'instantiates [VP]'. And by choosing a more complex sum-type to inject everything into—say, $e + t + \overline{e + t} + \overline{e + t + e + t}$—one might hope to keep failures of Naïve Property Comprehension out of view in all but highly theoretical contexts.[58]

The biggest problem with this approach is that it seems excessively arbitrary to pick a single sum-type as the once-and-for-all semantic correlate of the category NP. Whatever sum-type one picks, there will in principle be ways in which our disposition to treat Naïve Property Comprehension as unproblematic could play out that one will be unable to accommodate, although one could have done so if one had picked a more complex sum-type. True, computational limitations mean that in practice we will anyway be terrible at keeping track of discourses about, say, properties of properties of properties of properties of individuals. But it looks wrongheaded to inscribe these limitations into the architecture of our theory of meaning. The type-ambiguity approach, by contrast, lets us treat ascent to an appropriately capacious sum-type as an option that is always in-principle available when needed, with no need to pretend that the option is always exercised. To my eye, this promises more insight into how the linguistic practice actually works, even in simple cases that come nowhere near the computational limitations.

[57] See note 41 for an idea on how this might go.

[58] The view that the Lambda Serpentians should reject type-ambiguity and instead pick some complex sum-type as the one and only semantic correlate of the category NP raises some interesting further questions about the interpretation of the expression 'type e' as used by actual metaphysicians and semanticists. Some explanations of this type link it tightly to English NPs. At least when the Lambda Serpentians are interpreting theoreticians who use 'type e' in this way, they will have good reason to interpret their 'type e' terms using the same complex type they use for NPs. It is a good question whether there is some more independent kind of explanation that we could use to latch on to the Lambda Serpentians' type e—something for which e-Materialism is plausible but nontrivial.

7 Worries about schemas and universal generalizations

Schemas with infinitely many instances play a crucial role in the higher-order semantic theory for English I have sketched. We have both general schemas like Function Application as well as specific schemas characterizing the semantic profiles of words like 'some', 'instantiates', and 'interesting'. At a very abstract level, the use of schemas or some equivalent device is an inevitable part of the development of any sort of rigorous theory. Even if one's theory of some subject matter is finitely axiomatized, its predictive power will depend on some rules by which those axioms can generate infinitely many theorems to which one is also taken to be committed. For example, the capacity of a universally quantified axiom to do useful theoretical work will depend on having in the background something schema-like such as the Universal Instantiation axiom or the ∀Elim rule.[59]

However, even if one is unperturbed by the general practice of theorizing using schemas, one might be uncomfortable to have to rely so heavily on schemas in formalizing one's theory about a subject matter as localized as English semantics. If we imagine the perspective of a theorist investigating such a subject matter, refining their views about it in response to new evidence, and attempting to communicate their new ideas to others engaged in the same project, we can see that they would find it very helpful to have some *sentence* they could assert that would be treated as implying all the infinitely many instances of the schema. An autobiographical sentence like 'I accept all the instances of this schema' doesn't really play the desired role: what's wanted is something that one theorist might assert, while another might only embed under some operator like 'It is not that case that...' or 'I am moderately sure that...' or 'I am tempted to conjecture that...'. In my informal English presentation of the deflationary semantic theory, I have often found it helpful to play fast and loose by writing down sentences that appeared to quantify into the position of a type superscript or subscript.[60] If the Lambda Serpentians have to content themselves with schemas in place of such sentences, their debates about English semantics will have a strangely indirect or performative character that makes them look oddly different from debates in biology or history.

Moreover, extensive reliance on schemas tends to make for theories with a disappointingly limited capacity to explain universally generalized facts about the target subject matter. For example, consider the English sentence 'Everything is self-identical'. One striking fact about this sentence is that everything it expresses is true. That is:

(36) $$\forall p^t.\, \text{everything}\cdot\text{is}\cdot\text{self-identical} \mathrel{\vdots} \text{s}\, p \rightarrow p$$

[59]The difference between schemas and inference rules doesn't matter for present purposes; the job of any of our schemas could be done by a customized inference rule allowing schema-instances for higher types to be inferred from schema-instances for lower types. The observation about the inescapability of such devices is due to Lewis Carroll (1895).

[60]See, e.g., the gloss on the Function Application schema in §4.2.

This is a striking fact that cries out for an explanation. An adequate explanation must surely appeal to the fact that the sentence is the result of combining 'everything' and 'is self-identical', together with some facts about those constituents. But given that both of them are type-ambiguous, it is hard to see how we could achieve this. We could certainly endorse schemas like the following:[61]

$$\text{(37)} \qquad \forall X^{\underline{\overline{\sigma}}}.\text{everything} \vdots \text{s/np\textbackslash s}\, X \to X =_{\underline{\overline{\sigma}}} \forall_{\sigma}$$

$$\text{(38)} \qquad \forall Y^{\overline{\sigma}}.\text{is·self-identical} \vdots \text{np\textbackslash s}\, Y \to Y =_{\overline{\sigma}} \lambda z^{\sigma}.z =_{\sigma} z$$

But these do not imply (36). And the deficit can't be made up by adding further auxiliary premises: since deducibility is compact, any set of sentences that did imply (36) and included all instances of (37) and (38) would be such that some finite subset of it already implied (36).

It might seem that this problem isn't specific to our deflationary approach, but arises even for competing theories that flatly reject type-ambiguity. For how is that very rejection of type-ambiguity to be expressed, if not by schemas such as (39) (corresponding to the idea that $\overline{e}$ is the unique semantic type for verb phrases)?

$$\text{(39)} \qquad \neg\exists a^{\epsilon} x^{\sigma}.a \vdots_{\sigma} \text{np\textbackslash s}\, x \quad [\text{where } \sigma \neq \overline{e}]$$

But adding this schema to our theory will be of no use for explaining (36), given that no finite set of its instances is. However, so long as we reject type-ambiguity, there will be a potential avenue for explaining facts like (36) that does not turn on schemas like (39). Instead, we could invoke "unique decomposition" principles like the following, which limit the interpretations of sentences of the form exemplified by (36) to those that can be derived by Function Application from meanings of their constituents, in the types dictated by the constituents' syntactic categories:

$$\text{(40)} \quad \forall a^{\epsilon} b^{\epsilon} p^{t}.\, a{\cdot}b \vdots_{t} \text{s}\, p \wedge \mathbf{Syn}^{\overline{\epsilon\epsilon}}\, a\, \text{s/np\textbackslash s} \wedge \mathbf{Syn}^{\overline{\epsilon\epsilon}}\, b\, \text{np\textbackslash s}$$
$$\to \exists X^{\underline{\overline{e}}} y^{\overline{e}}.a \vdots_{\underline{\overline{e}}} \text{np\textbackslash s/s}\, X \wedge b \vdots_{\overline{e}} \text{np\textbackslash s}\, y \wedge p = Xy$$

Here $\mathbf{Syn}^{\overline{\epsilon\epsilon}}$ is a constant meaning 'is the (unique) syntactic category of'. In conjunction with the $\sigma = e$ instances of (37) and (38) and the unproblematic syntactic claims $\mathbf{Syn}^{\overline{\epsilon\epsilon}}$ everything s/np\s and $\mathbf{Syn}^{\overline{\epsilon\epsilon}}$ is·self-identical np\s, (40) implies (36). To generalize this approach, one would want to allow somehow for the fact that many sentences are associated with more than one "parse tree"; but this should be feasible, given a background syntactic theory strong enough to spit out a finite list of parse trees for any given string.

Unfortunately, this strategy depends crucially on the rejection of type-ambiguity. So, there is a genuine explanatory and expressive challenge here for the type-ambiguity

[61] (37) is arguably too strong because of the phenomenon of contextual quantifier domain-restriction, but the present challenge would still arise if we weakened it to $\forall X^{\underline{\overline{\sigma}}}.\text{everything} \vdots \text{s/np\textbackslash s}\, X \to \exists Z^{\overline{\sigma}}.X =_{\underline{\overline{\sigma}}} \lambda Y^{\overline{\sigma}}.\forall u^{\sigma}.Zu \to Yu$.

approach, that we should want to meet somehow. One might suppose that meeting it would require some enrichment of the syntax of our higher-order language—for example, one that includes variable-like expressions that can take play the role of a type subscript or superscript, and quantifier-like expressions that stand to them as ordinary quantifiers stand to ordinary variables. But the need can also be met within the confines of the standard higher-order syntax by adding some new constants, governed by an appropriate logic. The constants needed are *disquotational meaning predicates* M_σ for the original higher-order language (i.e. the one without the new constants). And the background logic is some analogue for these predicates of the kind of theory investigated in the literature on so-called "typed" theories of truth in classical logic.[62] Of course, the language of the Lambda Serpentians already has a system of predicates $\vdots_\sigma$ that can be used to talk about the meanings of arbitrary expressions as used by arbitrary communities, including their own. But the new disquotational-expressing predicates M_σ work very differently: for example, truths about what an expression disquotationally-expresses are non-contingent. By calling them "disquotational" I mean that the background logic that gives them their useful role includes as theorems all instances of the following schema, where A is a closed type-σ expression of the original language without the M_σ predicates, and $\ulcorner A \urcorner$ is a character-by-character specification of the string A:

DM $$\forall x^\sigma . \ulcorner A \urcorner \; \mathsf{M}_\sigma \; x \leftrightarrow x =_\sigma A$$

I don't mean that the background logic includes nothing more than the instances of DM. For example, any minimally adequate syntactic theory will be able to formalize the notion of a string being a theorem of H (in the original signature), and I assume the background logic will also include the claim that all theorems of H are disquotationally-true (disquotationally-express only truths):

Reflection $$\forall a^\epsilon . \, \mathbf{TheoremOfH}^{\bar{\varepsilon}} \; a \rightarrow \mathbf{T}^{\bar{\varepsilon}} \; a$$

where the "disquotational truth" predicate $\mathbf{T}^{\bar{\varepsilon}}$ is defined as $\lambda a^\epsilon . (\exists p^t . a \; \mathsf{M}_t \; p) \wedge (\forall p^t . (a \; \mathsf{M}_t \; p) \rightarrow p)$. By Tarski's theorem on the undefinability of truth, there is no way these new predicates M_σ can be *defined* in the old language—for any candidate definition we will be able to find an instance of DM whose negation follows from our old theory. Nevertheless, this seems to me to be a good exhibit for the idea that we can sometimes come to understand new predicates by accepting a theory expressed in terms of them that brings us some important explanatory benefits.

Using these new expressive resources, we can strengthen our semantic theory for English in a way that allows it to be finitely axiomatized (thus eliminating the need to rely on schemas), and moreover allows for plausible lines of explanation for facts like (36). For example, instead of the schemas (37) and (38), we can write down single universally quantified sentences that (given the background theory) imply

[62] For example, the theory CT from Halbach 2011 (§8.6).

all their instances:

(41) $\forall a^{\epsilon}.\, \mathbf{Typ}^{\bar{\varepsilon}}\, a \to \mathbf{T}^{\bar{\varepsilon}}\, \ulcorner \forall X^{\overline{\$a}}.\, \mathsf{everything} \mathrel{⁝} \mathsf{np\backslash s/s}\, X \to X =_{\underline{\overline{\$a}}} \forall_{\$a} \urcorner$

(42) $\forall a^{\epsilon}.\, \mathbf{Typ}^{\bar{\varepsilon}}\, a \to \mathbf{T}^{\bar{\varepsilon}}\, \ulcorner \forall Y^{\overline{\$a}}.\, \mathsf{is{\cdot}self\text{-}identical} \mathrel{⁝} \mathsf{np\backslash s}\, Y \to Y =_{\overline{\$a}} (\lambda z^{\$a}.z =_{\$a} z) \urcorner$

(Here, the convention for corner-quotes is extended to allow for "anti-quotation": a type-ϵ variable preceded by \$ still functions as a variable, analogous to the convention for Greek-lettered variables in Quine's corner-quotes. $\mathbf{Typ}^{\bar{\varepsilon}}$ is a predicate characterizing those strings that are well-formed types in our type system.) The disquotational-meaning predicates also give us an analogue of *existential* as well as universal quantification into type position. This means we can now formulate partial converses of Function Application that limit the meanings of complex English expressions to things obtained by function-application from meanings of its constituents. For example:

(43) $\forall a^{\epsilon} b^{\epsilon} c^{\epsilon} p^{t}.\, \mathbf{Syn}^{\overline{\epsilon\epsilon}}\, a\, \mathsf{s}/c \wedge \mathbf{Syn}^{\overline{\epsilon\epsilon}}\, b\, c \wedge a{\cdot}b \mathrel{⁝}_{t} \mathsf{s}\, p \to$
$\exists f^{\epsilon} Z^{\bar{t}}.\, \mathbf{Typ}^{\bar{\varepsilon}}\, f \wedge \ulcorner \lambda q^{t}.\exists X^{\overline{\$f}} y^{\$f}.\, \$a \mathrel{⁝}_{\overline{\$f}} \mathsf{s}/\$c\, X \wedge \$b \mathrel{⁝}_{\$f} \$c\, y \wedge q =_{t} Xy \urcorner\, \mathsf{M}_{\bar{t}}\, Z \wedge Zp$

(Gloss: for any sentence formed by concatenating an expression *a* of some category s/*c* with an expression *b* of category *c*, and any proposition *p* it expresses, there is a type *f* such that *p* has the property *Z* disquotationally-expressed by 'is a *q* such that *q* is the result of applying something expressed by *a* in type $\bar{f}$ to something expressed by *b* in type f'.) And thanks to the presence of Reflection in the background theory, (41), (42), and (43) will imply (36), via

(44) $\forall p^{t}.\, \mathsf{everything{\cdot}is{\cdot}self\text{-}identical} \mathrel{⁝} \mathsf{s}\, p \to$
$\exists f^{\epsilon} Z^{\bar{t}}.\, \mathbf{Typ}^{\bar{\varepsilon}}\, f \wedge \ulcorner \lambda q^{t}.q = \forall z^{\$f}.z =_{\$f} z \urcorner\, \mathsf{M}_{\bar{t}}\, Z \wedge Zp$

The disquotational meaning theory thus gives us a simulacrum of universal and existential quantification into type position, which we can use to strengthen our theory to allow it to derive important facts about the limits of what can be expressed by certain English sentences.

Of course, schemas—specifically, the disquotational meaning schema DM, or some other collection of schemas that imply it—still play a crucial role in this stronger body of theory. But the hope is that the theory of disquotational-meaning can occupy the same background role that classical higher-order logic was already playing, as something that can be implicitly assumed by all of the parties involved in the empirical project of constructing explanatory theories of meaning for some natural language.

8 Type-neutral generality in English

The semantic theory for quantification presented in section §6.4 (using the formalism of sum-types) generates an infinite array of readings for quantified sentences like (45) and (46):

(45) Every essay topic is interesting.

(46) Some essay topic is interesting.

For any finitely many simple types $\sigma_1, \ldots, \sigma_n$, we can construct a sum-type to generate meanings for (45) and (46) equivalent respectively to (47) and (48):

(47) $$(\forall x^{\sigma_1}.\, \mathbf{topic}^{\overline{\sigma_1}}\, x \rightarrow \mathbf{int}^{\overline{\sigma_1}}\, x) \wedge \cdots \wedge (\forall x^{\sigma_n}.\, \mathbf{topic}^{\overline{\sigma_n}}\, x \rightarrow \mathbf{int}^{\overline{\sigma_n}}\, x)$$

(48) $$(\exists x^{\sigma_1}.\, \mathbf{topic}^{\overline{\sigma_1}}\, x \;\wedge\; \mathbf{int}^{\overline{\sigma_1}}\, x) \vee \cdots \vee (\exists x^{\sigma_n}.\, \mathbf{topic}^{\overline{\sigma_n}}\, x \;\wedge\; \mathbf{int}^{\overline{\sigma_n}}\, x)$$

Under the reasonable assumption that all the relevant propositions are modally independent, this will yield infinite collections of meanings for (45) and (46). In the former collection, we can find for every meaning a strictly stronger one (based on a larger finite collection of types). In the latter, we can find for every meaning a strictly weaker one.

This pattern is somewhat reminiscent of "quantifier relativism", the view that there is no such thing as "absolutely unrestricted quantification". Proponents of this view hold, roughly speaking, that for any possible interpretation of the quantifiers, there is an even broader interpretation of which the first interpretation is a restriction.[63] Our type-ambiguous treatment of English quantification *could* be viewed as an articulation of this notoriously elusive idea. However the picture is very different from the usual vision of quantifier relativism. For us, absolutely unrestricted quantification for any given type is unproblematic; the different interpretations of quantifiers responsible for the stronger-and-stronger meanings for (45) and the weaker-and-weaker meanings for (46) belong to different types, and none is in any straightforward sense a "restriction" of any other.

Nevertheless, the limitations in the readings we can generate for (45) and (46) are not easy to live with. It *feels* like there should be the possibility of using (45) to express something that entails all the conjunctions (47) (and is thus stronger than any of them), and the possibility of using (46) to express something that is entailed by all the disjunctions (48) (and is thus weaker than any of them).

Should we just dismiss the impulse to want such an overarching interpretation as a mistake, illustrating the pitfalls of a way of talking that fails to perspicuously display type distinctions in the syntax? That line would certainly fit standard assumptions about what friends of higher-order logic should and should not take to be intelligible. But we should not be too quick here. The impulse to think that there is an intelligible *type-neutral* generalizations in the vicinity of (45) and (46) is not just a

[63] See Rayo and Uzquiano 2007 for a sample of the literature on this topic.

product of the idiosyncratic expressive choices made in natural language. As we saw in the previous section, it is an impulse that could be expected to arise organically even in a community whose mother tongue was an entirely syntactically standard dialect of Higher-Orderese as its mother tongue. Such speakers would have genuine communicative and explanatory needs that could motivate them to introduce new resources allowing them to, in effect, quantify simultaneously in every type. If these new resources take the form of the disquotational meaning and truth predicates contemplated in the previous section, they will be able to formulate sentences like the following:

(49) $$\forall f^{\epsilon}.\, \mathbf{Typ}^{\overline{\epsilon}} f \to \mathbf{T}^{\overline{\epsilon}} \ulcorner \forall x^{\$f}.\, \mathbf{topic}^{\$f\to t}\, x \to \mathbf{int}^{\$f\to t}\, x \urcorner$$

(50) $$\exists f^{\epsilon}.\, \mathbf{Typ}^{\overline{\epsilon}} f \wedge \mathbf{T}^{\overline{\epsilon}} \ulcorner \exists x^{\$f}.\, \mathbf{topic}^{\overline{\$f\to t}}\, x \wedge \mathbf{int}^{\$f\to t}\, x \urcorner$$

Against the background of the disquotational meaning theory, these will have the desired logical behavior: (49) will imply every sentence of the form $\forall x^{\sigma}.\, \mathbf{topic}^{\sigma\to t}\, x \to \mathbf{int}^{\sigma\to t}\, x$, and (50) will be implied by every sentence of the form $\exists x^{\sigma}.\, \mathbf{topic}^{\sigma\to t}\, x \wedge \mathbf{int}^{\sigma\to t}\, x$.

There is thus no problem imagining a language with the capacity to express the desired maximally strong and weak meanings. Could English itself be such a language? We could generate the maximal readings for (45) and (46) by appealing to some new basic principles that generate meanings for certain quantified sentences *not* in accordance with function application. There are various possible ways this might work: the one I will consider here posits new rules according to which if a sentence expresses the conjunction/disjunction of any finite subcollection of some collection of propositions, it also expresses the conjunction/disjunction of all of them.

To capture this idea in our higher order language, we can introduce a system of "rigidity" predicates $\mathbf{Rig}^{\overline{\sigma}}$: for X of type $\overline{\sigma}$, $\mathbf{Rig}^{\overline{\sigma}}\, X$ implies that X is modally rigid in the way that *sets* or *pluralities* are standardly considered to be.[64] Then we can understand quantification over 'collections' of propositions as quantification over rigid properties of propositions (things of type $\overline{t}$). We can equate the conjunction [disjunction] of a collection of propositions with the proposition that all [some] of them are true, and similarly for properties and relations:

$$\mathbf{Conj} := \lambda X^{\overline{\overline{\sigma_1 \ldots \sigma_n}}} y_1^{\sigma_1} \ldots y_n^{\sigma_n}.\forall Z^{\overline{\sigma_1 \ldots \sigma_n}}.XZ \to Zy_1 \ldots y_n$$
$$\mathbf{Disj} := \lambda X^{\overline{\overline{\sigma_1 \ldots \sigma_n}}} y_1^{\sigma_1} \ldots y_n^{\sigma_n}.\exists Z^{\overline{\sigma_1 \ldots \sigma_n}}.XZ \wedge Zy_1 \ldots y_n$$

In these terms, the new principle can be expressed by the following schemas, where

[64]For details, see Dorr, Hawthorne, and Yli-Vakkuri 2021 (§1.5).

τ is any predicate type:

Conjunctive Closure

$$\forall a^{\epsilon} c^{\epsilon} X^{\overline{\tau}}.\, \mathbf{Rig} X \wedge (\forall Y^{\overline{\tau}}.\mathbf{Rig} Y \wedge \mathbf{Finite}\, Y \wedge Y \subseteq X \rightarrow a \vdots_{\tau} c\ \mathbf{Conj}\, Y) \rightarrow a \vdots_{\tau} c\ \mathbf{Conj}\, X$$

Disjunctive Closure

$$\forall a^{\epsilon} c^{\epsilon} X^{\overline{\tau}}.\, \mathbf{Rig} X \wedge (\forall Y^{\overline{\tau}}.\mathbf{Rig} Y \wedge \mathbf{Finite}\, Y \wedge Y \subseteq X \rightarrow a \vdots_{\tau} c\ \mathbf{Disj}\, Y) \rightarrow a \vdots_{\tau} c\ \mathbf{Disj}\, X$$

(Here, **Finite** is short for a standard higher-order definition of finitude, and $Y \subseteq X$ just means $\forall z^{\tau}.Yz \rightarrow Xz$.) Using these principles, and assuming some moderately coarse-grained background logic (such as Classicism, see Appendix A), we will be able to derive that (45) and (46) do in fact express appropriately maximal propositions.

Note that generating these meanings crucially relies on having new compositional rules, rather than a special new meaning for 'every' and 'some' that could combine by function application. Since there is no "sum of all types", no such meanings could do the same work as Conjunctive and Disjunctive Closure.

Related to this, adding the closure principles will disrupt certain logical relationships that were valid before. Consider for example the following argument:

P1 Some essay topic instantiates the property of being interesting.

P2 Jim mentioned the property of being interesting.

C Some essay topic instantiates some property Jim mentioned.

Assuming we posit a meaning for 'mentioned' of type $\overline{e\alpha}$ for every extended type α, the sum-type-based treatment of quantifiers will give this argument the following nice status: for every proposition p expressed by P1, there are corresponding propositions q and r expressed by P2 and C respectively, such that $\Box(p \wedge q \rightarrow r)$. But if we add the closure principles, we will no longer predict this status. The problem is that while P1 and C will both get new maximally weak readings that behave like the disjunctions of all their old readings, P2 will not get any new readings, since its old readings were not closed under finite conjunction or disjunction. In particular, it will not get a reading whose conjunction with the new reading of P1 entails the new reading of C.[65]

Ultimately, we know from Russell's paradox that there are limits to the extent to which we can vindicate all the modes of inference that seem prima facie compelling when we are dealing with type-neutral interpretations of quantified sentences. Because of the logical surprises they engender, I am ultimately not sure that

[65]It is not entirely obvious that P2 won't get a new reading. In principle one might think that in types of the form $\overline{e\overline{\alpha + \beta}}$, 'mentioned' expresses the conjunctive relation $\lambda x^{e} Y^{\alpha+\beta}.\, \mathbf{mention}^{e\overline{\alpha}}\, x\, (\lambda z^{\alpha}.Y\iota^{1}_{\beta} z) \wedge \mathbf{mention}^{e\overline{\beta}}\, x\, (\lambda z^{\beta}.Y\iota^{2}_{\alpha} z)$. If so, Conjunctive Closure will after all apply to P2, and generate a strong meaning that entails $\mathbf{mention}^{e\overline{\sigma}}\, \mathbf{jim}^{e}\, \mathbf{int}^{\overline{\sigma}}$ for each σ, and thus bridges the gap from the weak reading of P1 to the weak reading of C. But if you like that particular theory of mentioning, just replace 'mentioned' in P2 and C with 'didn't mention' and the problem will come back.

Conjunctive Closure and Disjunctive Closure are sufficiently well-motivated when understood as claims about the semantics of ordinary language. They may, nevertheless, be a good description of certain extensions of our standard linguistic practices that can naturally arise in theoretical contexts when the need to convey certain propositions not semantically expressible according to the ordinary rules becomes pressing.

The foregoing discussion barely scratches the surface of the question whether and how type-neutral generalizations can be expressed in natural languages raises. The approach using the closure principles is certainly not the only one; there may be other, more flexible and systematic ways to generate the desired readings for English sentences like (45) and (46). I include this preliminary discussion here here mainly as a warning against assuming that a deflationary, higher-order approach to property-talk in English will inevitably require resisting the impulse to posit type-neutral readings in such cases.

9 Conclusion: in defence of type-ambiguity

There is a canonical way to argue for the presence of ambiguity (understood as encompassing all forms of semantic multiplicity) in a certain sentence: one identifies an actual or possible scenario in which it is plausible that one speaker sincerely asserts the sentence and another speaker sincerely denies it, even though neither speaker has any false beliefs. I have not given this kind of argument for the sentential ambiguities posited by my favoured approach—e.g., that 'The property of being interesting is interesting' can mean $\mathbf{int}^{\overline{\sigma}}_{-}\,\mathbf{int}^{\overline{\sigma}}$ for any type σ. And it wouldn't be easy to do so, since our usage of this sentence doesn't seem to display any of the usual telltale hallmarks of ambiguity.[66] Ordinary speakers are not, for example, disposed to feel that anything is going wrong with arguments which, according to the theory, involve fallacies of equivocation, such as 'The property of being self-identical is self-identical, so the property of being self-identical has the property of being self-identical, so some properties instantiate themselves'. Relatedly, insofar as ordinary speakers have direct intuitions on the question, they seem to favour non-ambiguity. Opponents of the type-ambiguity approach have seen these facts as constituting strong evidence for their view: 'the proponent of the typed conception must argue, appearances to the contrary, that not only quantifiers but many predicates as well in ordinary language are systematically ambiguous. Intuitively, that just seems wrong' (Menzel 1986 (5), crediting Chierchia 1984).

In response, I want to insist that the question whether to posit some form of semantic multiplicity in a population's language is a theoretical one, to which many kinds of evidence are relevant, among which the specific patterns of usage supposedly diagnostic of ambiguity and its absence play no privileged role. Indeed,

[66] Or at least not the relevant sort of ambiguity—there will of course be variation stemming from different answers to the question 'interesting to whom', different thresholds, etc.

compared to really central aspects of the linguistic practice, such as speakers' tendency to treat 'instantiates the property of being *F*' as interchangeable with 'is *F*', the usage-dispositions constituting the "appearance of non-ambiguity" do not seem particularly firm or central. In support of this comparison, we can note that our overall suite of dispositions includes a disposition to fall into self-contradiction the first time we encounter Russell's paradox. When people are brought face to face with that paradox, the usual result—visible across the philosophical literature—is an uneasy wariness when it comes to expressions like 'instantiates itself', coupled with a continued willingness to rely on Naïve Comprehension in ordinary contexts that seem safely far from the paradox. Often, when trying to explain the paradox, people pick some other word like 'condition' or 'category' that is essentially a synonym of 'property' and appeal to instances of Naïve Comprehension for it even while rejecting them for 'property': 'For some conditions, such as not instantiating oneself, there is no property instantiated by all and only the things that meet that condition'. It is implausible that this reflects some permanent semantic contrast between, say, 'condition' and 'property'. Rather, this looks like an example of the common pattern where we facilitate communication by associating two different words that are ambiguous in the same way with different resolutions of ambiguity for the purposes of some particular discourse.[67]

The widespread view that *vagueness* is or involves a kind of semantic multiplicity provides extra reason to resist the idea that semantic multiplicity should be expected to show up in distinctive hallmarks of usage or speaker intuitions.[68] For almost all expressions are vague: avoiding vagueness normally takes luck and hard work. And while we do often display implicit awareness of vagueness, we can also easy be misled into thinking a word is not vague when is, e.g. because its borderline cases are rare, hard to notice, or modally remote.[69] Of course, the type ambiguities I am defending differ in important ways from those involved in ordinary vagueness.[70] Nevertheless, the picture of semantic multiplicity on which it is as pervasive as vagueness seems friendly to the type-ambiguity approach. If even unnoticed vagueness involves semantic multiplicity, we should give little weight to objections to claims of

[67]For example, both 'book' and 'volume' are plausibly ambiguous (polysemous) between 'informational' and 'physical' readings (Liebesman and Magidor MS), but 'Although there are many volumes on that shelf, there is only one book' and 'Although there are many books on that shelf, there is only one volume' are much better than 'Although there are many books on that shelf, there is only one book'.

[68]See Dorr unpublished for a defence of this view of vagueness, and a more general picture on which semantic multiplicity is a ubiquitous feature of human language.

[69]Relativity and quantum physics present many examples where words that might have been precise if Newtonian physics had been correct turned out to be vague—e.g. position-predicates like 'are closer together than', as applied to particles.

[70]With an ordinary vague word, it is always possible (and typically natural) to interpret all occurrences of the word within a given sentence or discourse uniformly; by contrast, our theory implies that fully uniform resolution of type-ambiguity is sometimes not possible (because of the need to assign types that permit Function Application).

multiplicity that turn on appeals to speakers' intuitions, or on their failure to exercise the kinds of care that would be warranted if one were sensitive to the dangers of equivocation. Rather, we should see multiplicity as the default outcome, and as especially liable to show up in interesting ways when there are internal tensions within our overall package of usage-dispositions, whether or not we have noticed the tensions or devised explicit techniques for managing them. Given that such tensions are clearly present in ordinary property-talk, it is *prima facie* plausible that they would manifest in *some* sort of multiplicity. And while it is not initially obvious what form this might take, the pervasive, but simple and systematic, multiplicity characteristic of the type-ambiguity approach seems to me to be very much the sort of thing one might have expected.

Appendix A Systems of classical higher-order logic

In this appendix I will precisely characterize the syntax of the higher-order language whose intelligibility I defended in §2, along with the system of classical higher-order logic presupposed in the rest of the paper.

The first order of business is to specify the set of *types* to be used. This is slightly more complex than the basic definition of simple types from §2 because I want to exclude types like $e \to e$.

Definition 1. $\mathbb{T}$ (the set of types) and $\mathbb{T}_0$ (the set of "terminal" types) are the smallest pair of sets meeting the following conditions:

$$\frac{}{e \in \mathbb{T}} \qquad\qquad \frac{}{t \in \mathbb{T}_0}$$

$$\frac{\tau \in \mathbb{T}_0}{\tau \in \mathbb{T}} \qquad\qquad \frac{\sigma \in \mathbb{T} \quad \tau \in \mathbb{T}_0}{(\sigma \to \tau) \in \mathbb{T}_0}$$

A $\mathbb{T}$-*typed signature* Σ is a function that maps some set of *constants* (strings excluding certain special characters) to types in $\mathbb{T}$.

Definition 2. For a given $\mathbb{T}$-typed signature Σ, the *higher-order language* $\mathscr{L}(\Sigma)$ is the smallest relation : between strings (terms) and $\mathbb{T}$-types subject to the following conditions:

$$\frac{\sigma \in \mathbb{T}}{x_i^\sigma : \sigma}\ \text{Var} \qquad\qquad \frac{\Sigma(C) = \sigma}{C : \sigma}\ \text{Const}$$

$$\frac{A : \sigma \to \tau \quad B : \sigma}{(AB) : \tau}\ \text{App} \qquad\qquad \frac{A : \tau \quad \tau \in \mathbb{T}_0 \quad \sigma \in \mathbb{T}}{(\lambda x_i^\sigma . A) : \sigma \to \tau}\ \text{Abs}$$

$$\frac{}{\to : t \to t \to t}\ (\to) \qquad\qquad \frac{\sigma \in \mathbb{T}}{\forall_\sigma : (\sigma \to t) \to t}\ (\forall)$$

We write $\mathscr{L}(\Sigma)^\sigma$ for the set of all terms A such that $A :_{\mathscr{L}(\Sigma)} \sigma$.

We could, if we wished, add more logical constants, but for brevity we will make do with $\to$ and $\forall_\sigma$ and treat the rest as defined. The following definitions have the advantage of making it espeically easy to prove that the defined expressions obey the expected logical rules:

$$
\begin{aligned}
\bot &:= \forall p^t.p \\
\top &:= \bot \to \bot \\
\neg &:= \lambda p^t.p \to \bot \\
\wedge &:= \lambda p^t q^t.\forall r^t.(p \to q \to r) \to r \\
\vee &:= \lambda p^t q^t.\forall r^t.(p \to r) \to (q \to r) \to r \\
\leftrightarrow &:= \lambda p^t q^t.(p \to q) \wedge (q \to p) \\
\exists_\sigma &:= \lambda X^{\overline{\sigma}}.\forall p^t.(\forall y^\sigma.Xy \to p) \to p \\
=_\sigma &:= \lambda x^\sigma y^\sigma.\forall Z^{\sigma \to t}.Zx \to Zy \\
\neq_\sigma &:= \lambda x^\sigma y^\sigma.\neg(x =_\sigma y)
\end{aligned}
$$

So much for the syntax of $\mathscr{L}(\Sigma)$. Before characterizing the system of logic we will be using, we need a few auxiliary syntactic notions.

Definition 3. For any term $A : \sigma$ and type-σ variable v, the *substitution function* $[A/v]$ is the function (written in postfix position) such that $v[A/v] = v$, $C[A/v] = C$ if C is a constant or a variable other than v, $(BC)[A/v] = B[A/v]C[A/v]$, $(\lambda u.B)[A/v] = \lambda u.(B[A/v])$ when $u \neq v$, and $(\lambda v.B)[A/v] = (\lambda v.B)$.

Definition 4. FV is the function from $\mathscr{L}(\Sigma)$-terms to sets of variables such that $FV(v) = \{v\}$ for every variable v, $FV(C) = \varnothing$ for every constant C, $FV(AB) = FV(A) \cup FV(B)$, and $FV(\lambda v.A) = FV(A) - \{v\}$. For a set of terms Ξ, $FV(\Xi) = \bigcup\{FV(A) : A \in \Xi\}$.

We read '$v \in FV(A)$' as 'v is free in A'.

Definition 5. Safe is the function from pairs comprising a variable and an $\mathscr{L}(\Sigma)$-term to sets of variables such that $\text{Safe}(v, C)$ for any variable or constant C is the set of all variables, $\text{Safe}(v, AB) = \text{Safe}(v, A) \cup \text{Safe}(v, B)$, and $\text{Safe}(v, \lambda u.A) = \text{Safe}(v, A)$ if $v \notin FV(A)$ and $\text{Safe}(v, A) - \{u\}$ otherwise.[71]

We read '$u \in \text{Safe}(v, A)$' as 'u is safe for v in A' (or 'u is in no danger of being captured if we substitute something for v in A'); we say that a *term B* is safely substitutable for v in A iff $FV(B) \subseteq \text{Safe}(v, A)$ and B is of the same type as v.

Definition 6. $\sim_{\beta\eta}$ ("beta-eta convertibility") is the smallest binary relation $\sim$ on

[71] For a set of terms Ξ, $\text{Safe}(v, \Xi) = \bigcap\{\text{Safe}(v, A) : A \in \Xi\}$.

$\mathscr{L}(\Sigma)$-terms meeting the following conditions:

$$\frac{A : \sigma}{A \sim A}\ \text{Ref} \qquad \frac{A \sim B}{B \sim A}\ \text{Sym} \qquad \frac{A \sim B \quad B \sim C}{A \sim C}\ \text{Trans}$$

$$\frac{A \sim B \quad C \sim D \quad (AC) : \tau}{(AC) \sim (BD)}\ \text{App}_{\sim} \qquad \frac{A \sim B}{(\lambda v.A) \sim (\lambda v.B)}\ \text{Abs}_{\sim}$$

$$\frac{FV(B) \subseteq \text{Safe}(v, A)}{(\lambda v.A)B \sim A[B/v]}\ \beta \qquad \frac{v \notin FV(A)}{(\lambda v.Av) \sim A}\ \eta$$

Here, the first five conditions amount to the requirement that $\sim_{\beta\eta}$ is a *congruence* relation (with respect to the syntactic operations of application and abstraction); the central work is done by β and η. β allows for beta-conversion, as explained in §2. η takes care of an edge case, guaranteeing that every predicate—even constants and variables—can be converted into a lambda-term.[72]

Instead of the axiomatic proof systems used in much recent philosophical work (e.g. Dorr, Hawthorne, and Yli-Vakkuri 2021, Bacon and Dorr 2024), I will present the logic in the form of a natural-deduction system, which may be more familiar. So, I will identify the logic with its derivability relation—a set of ordered pairs comprising a finite set of formulae and a formula.[73]

Definition 7. For a $\mathbb{T}$-typed signature Σ, the *classical higher order logic* $\mathsf{H}(\Sigma)$ is the smallest relation $\vdash$ between finite subsets of $\mathscr{L}(\Sigma)^t$ and members of $\mathscr{L}(\Sigma)^t$ meeting

[72] In combination with β, η also guarantees the convertibility of terms that differ only by a permutation of bound variables ("α-equivalent" terms)—for so long as v isn't free in A and u is safe for v in A, we will have $\lambda u.A \sim_{\beta\eta} \lambda v.(\lambda u.A)v \sim_{\beta\eta} \lambda v.A[v/u]$. For more on the motivation for including η in the definition of convertibility, see Bacon 2023 (§3.3).

[73] Such ordered pairs are known as *sequents*. I will write $\Gamma \rhd P$ as a notational variant of $\langle \Gamma, P \rangle$, and $\Gamma \vdash P$ for $(\Gamma \rhd P) \in \vdash$.

the following conditions:

$$\frac{P : t}{P \vdash P}\ \text{Assumption} \qquad \frac{\Gamma \vdash P \quad Q : t}{\Gamma, Q \vdash P}\ \text{Weakening}$$

$$\frac{\Gamma, P \vdash Q}{\Gamma \vdash P \to Q}\ {\to}\text{Intro} \qquad \frac{\Gamma \vdash P \to Q \quad \Gamma \vdash P}{\Gamma \vdash Q}\ {\to}\text{Elim}$$

$$\frac{\Gamma \vdash P \quad v \notin FV(\Gamma)}{\Gamma \vdash \forall v.P}\ \forall\text{-intro} \qquad \frac{\Gamma \vdash \forall_\sigma F \quad A : \sigma}{\Gamma \vdash FA}\ \forall\text{-Elim}$$

$$\frac{\Gamma \vdash \neg\neg P}{\Gamma \vdash P}\ \text{DNE} \qquad \frac{\Gamma \vdash P \quad P \sim_{\beta\eta} Q}{\Gamma \vdash Q}\ \text{conv}$$

Using this definition, we can easily show that $\mathsf{H}(\Sigma)$ has various other handy properties:

- The standard introduction and elimination rules for all the other logical constants are admissible in $\mathsf{H}(\Sigma)$ when they are defined as above.
- $\mathsf{H}(\Sigma)$ is transitive, in the sense that if $\Gamma \vdash_{\mathsf{H}(\Sigma)} P$ for all $P \in \Delta$ and $\Gamma, \Delta \vdash_{\mathsf{H}(\Sigma)} Q$, then $\Gamma \vdash_{\mathsf{H}(\Sigma)} Q$.
- $\mathsf{H}(\Sigma)$ is substitution-invariant, in the sense that if $\Gamma \vdash_{\mathsf{H}(\Sigma)} P$ and $FV(A) \subseteq \text{Safe}(v, \Gamma \cup \{P\})$, then $\Gamma[A/v] \vdash_{\mathsf{H}(\Sigma)} P[A/v]$.

The logic H is quite weak. For example, while it has $P \leftrightarrow \neg\neg P$ as a theorem for every formula P, it does not have any theorems of the form $P =_t \neg\neg P$. Such questions about "fineness of grain" cry out for some systematic treatment, so it is interesting to consider strengthenings of H which answer them. The most commonly-encountered such strengthening is a system I'll call E (for 'Extensionalism'), which may be characterized as the result of adding the following rule of "extensional substitution" to those of H:

$$\frac{\Gamma, P \vdash Q \quad \Gamma, Q \vdash P \quad \Gamma, \Delta \vdash R[P/x^t] \quad FV(\Gamma) \subseteq \text{Safe}(x^t, R)}{\Gamma, \Delta \vdash R[Q/x^t]}\ \text{Ext}$$

But while E is convenient for some purposes (such as that of formalizing mathematics), it has the problem that on the interpretation we care about, some of its theorems are manifestly false. Here is a representative example, with $\Box$ a nonlogical constant:

$$\dfrac{\dfrac{\vdots}{P \leftrightarrow Q, P \vdash Q}\,\mathsf{H} \quad \dfrac{\vdots}{P \leftrightarrow Q, Q \vdash P}\,\mathsf{H} \quad \dfrac{\vdots}{P \leftrightarrow Q, \Box P \vdash \Box P}\,\mathsf{H}}{\dfrac{P \leftrightarrow Q, \Box P \vdash \Box Q}{\dfrac{P \leftrightarrow Q \vdash \Box P \to \Box Q}{\vdash (P \leftrightarrow Q) \to \Box P \to \Box Q}\,{\to}\text{Intro}}\,{\to}\text{Intro}}\,\text{Ext}$$

But this is no good: for example, if $\Box$ expresses some nontrivial form of necessity, there will clearly be P and Q for which $P \leftrightarrow Q$ and $\Box P$ are true but $\Box Q$ is false.

To stake out a middle ground between H and E, we can weaken the extensional substitution rule by requiring each of P and Q to be derivable from the other *on its own*, without reliance on the side-premises Γ. In other words, we close H under the rule of *substitution of logical equivalents*:

$$\frac{P \vdash Q \quad Q \vdash P \quad \Delta \vdash R[P/x^t]}{\Delta \vdash R[Q/x^t]}\ \text{Subst}$$

The resulting system C (for 'Classicism') is explored in detail in Bacon and Dorr 2024. In it, we can—happily—no longer prove the likes of $(P \leftrightarrow Q) \to \Box P \to \Box Q$. But we can, for example, prove $P =_t \neg\neg P$: just set Q in Subst to be $\neg\neg P$ and R to be $P =_t x$ and note that $P \vdash_{\mathsf{H}} \neg\neg P$ and $\neg\neg P \vdash_{\mathsf{H}} P$.

Even H is fraught with controversy—the rule of beta-conversion, in particular, has been rejected for various metaphysically interesting reasons (Dorr 2016: §5).[74] Nevertheless, as the ubiquitous invocations of β-conversion in the semantic derivations in this paper show, H is a rather convenient logic to have in the background when we are doing natural language semantics. Without β-conversion to simplify things, the only theorems we could derive about the denotations of even moderately complex natural-language sentences involve hard-to-parse towers of lambda-abstractions. Of course, even if we reject full β-conversion, we might hold on to have the Extensional Beta schema discussed in note 4, meaning that the denotations we come up with will typically be *provably coextensive* to the ones we would have derived using beta-conversion; one might argue that this is all we should want given the fine-grained metaphysical views that motivate the rejection of beta-conversion. Still, H is quite helpful if only as a simplifying hypothesis. In fact the same goes for C, which is much more metaphysically controversial than H. Typically when we are trying to account for the meaning of some natural language sentence, we will be happy enough if we can derive an interpretation that is correct "up to logical equivalence". For example, if one theory says that 'no human is immortal' expresses $\neg\exists x^e.\, \mathbf{human}^{\bar{e}}\, x \wedge \mathbf{immortal}^{\bar{e}}\, x$ and another says that it expresses $\forall x^e.\, \mathbf{human}^{\bar{e}}\, x \to \neg\, \mathbf{immortal}^{\bar{e}}\, x$, then even if we think that these propositions are distinct, so long as we accept their logical equuvalence we will be happy to take a "spoils to the victor"

[74]See Dorr, Hawthorne, and Yli-Vakkuri 2021 (ch. 1) for a weaker system H_0 that replaces beta-conversion with Extensional Beta (note 4).

attitude to the question which one ‘no human is immortal’ expresses. Even though this paper hasn’t dealt with the kinds of phenomena where appealing to C would lead to a substantial simplification, I anticipate that many working semanticists will find it easier to adapt their usual practice to the setting of higher-order logic if they allow themselves a relatively coarse-grained logic like C.

Appendix B Adding sum-types

Definition 8. The sets of *extended types* $\mathbb{T}^+$ and *extended terminal types* $\mathbb{T}_0^+$ are the smallest pair of sets satisfying all the conditions in Definition 1, together with

$$\frac{\alpha \in \mathbb{T} \quad \beta \in \mathbb{T}}{\alpha + \beta \in \mathbb{T}}\ (+)$$

Definition 9. For a given $\mathbb{T}^+$-typed signature Σ, the *extended higher-order language* $\mathscr{L}^+(\Sigma)$ over Σ is the smallest relation : between strings and $\mathbb{T}^+$-types subject to all the conditions in Definition 2 together with the following three:

$$\frac{A : \alpha \quad \beta \in \mathbb{T}^+}{\iota^1_\beta A : \alpha + \beta}\ \iota^1 \qquad \frac{A : \beta \quad \alpha \in \mathbb{T}^+}{\iota^2_\alpha A : \alpha + \beta}\ \iota^2 \qquad \frac{A : \alpha + \beta \quad B : \gamma \quad C : \gamma}{\delta A(x_i^\alpha .\, B)(y_j^\beta .\, C) : \gamma}\ \delta$$

The intuition for $\delta A(u^\alpha .\, B)(v^\beta .\, C)$ is that if A is the representative of something of type α it will denote the same thing that B denotes when u^α is assigned to that thing, and if A is the representative of something of type β, it denotes the same thing that C denotes when v^β it is assigned to that thing. The term-building operation $\langle\!\langle\ \rangle\!\rangle$ used in the main text is defined by a combination of lambda-abstraction and delta-application: when $F : \alpha \to \tau$ and $G : \beta \to \tau$, $\langle\!\langle F, G \rangle\!\rangle$ is short for $\lambda x^{\alpha+\beta}.\delta x(y^\alpha .\, Fy)(z^\beta .\, Gz)$, where x, y, z are arbitrarily chosen variables not free in F or G.[75]

The auxiliary syntactic notions we defined for $\mathscr{L}$ all extend naturally to $\mathscr{L}^+$:

Definition 10. The substitution operation $[A/v]$ on $\mathscr{L}^+(\Sigma)$ is defined by the same clauses as in Definition 3, plus:

- $(\iota^1_\alpha B)[A/v] = \iota^1_\alpha(B[A/v])$
- $(\iota^2_\beta B)[A/v] = \iota^2_\beta(B[A/v])$
- $\delta B(x.\, C)(y.\, D)[A/v] = \delta B[A/v](x.\, C[A/v])(y.\, D[A/v])$ when $v \neq x$ and $v \neq y$
- $\delta B(x.\, C)(y.\, D)[A/x] = \delta B[A/x](x.\, C)(y.\, D[A/x])$ when $x \neq y$

[75] If we wanted to $\langle\!\langle\ \rangle\!\rangle$ to be primitive, we could define $\delta A(u^\alpha .\, B)(v^\beta .\, C)$ as $\langle\!\langle \lambda u^\alpha .B, \lambda v^\beta .C \rangle\!\rangle A$. But this works only when the type γ of B and C is a terminal type. Taking $\langle\!\langle\ \rangle\!\rangle$ as primitive would give us no analogue of, for example, the delta-application $\delta A(x^\alpha .\, \iota^2_\beta x)(y^\beta .\, \iota^1_\alpha y)$, which turns $A : \alpha + \beta$ into a term of type $\beta + \alpha$.

- $\delta B(x.\,C)(y.\,D)[A/y] = \delta B[A/y](x.\,C[A/y])(y.\,D)$ when $x \neq y$
- $\delta B(x.\,C)(x.\,D)[A/x] = \delta B[A/x](x.\,C)(x.\,D)$.

Definition 11. FV is defined on $\mathscr{L}^+$ by the same clauses as in Definition 4, plus $FV(\iota^1_\beta A) = FV(A)$, $FV(\iota^2_\alpha A) = FV(A)$, and $FV(\delta A(u.\,B)(v.\,C)) = FV(A) \cup (FV(B) - u) \cup (FV(C) - v)$.

Definition 12. Safe is defined on $\mathscr{L}^+$ by the same clauses as in Definition 4, plus $\text{Safe}(v, \iota^1_\beta A) = \text{Safe}(v, A)$, $\text{Safe}(v, \iota^2_\alpha A) = \text{Safe}(v, A)$, and $\text{Safe}(v, \delta A(x.\,B)(y.\,C)) = \text{Safe}(v, A) \cap \text{Safe}(v, B) \cap \text{Safe}(v, C)$ if $v \neq x$ and $v \neq y$, $\text{Safe}(x, \delta A(x.\,B)(y.\,C)) = \text{Safe}(x, A) \cap (\text{Safe}(x, B) - \{x\}) \cap \text{Safe}(v, C)$ and $\text{Safe}(y, \delta A(x.\,B)(y.\,C)) = \text{Safe}(y, A) \cap \text{Safe}(y, B) \cap (\text{Safe}(y, C) - \{y\})$ if $x \neq y$, and $\text{Safe}(x, \delta A(x.\,B)(x.\,C)) = \text{Safe}(x, A) \cap ((\text{Safe}(x, B) \cap \text{Safe}(x, C)) - \{x\})$.

Definition 13. The conversion relation $\sim_{\beta\eta+}$ is the smallest smallest binary relation $\sim$ on terms of $\mathscr{L}^+(\Sigma)$ obeying all the conditions in the definition of $\sim_{\beta\eta}$ (Definition 6), together with six further conditions:

$$\frac{A \sim B}{\iota^1_\beta A \sim \iota^1_\beta B} \qquad \frac{A \sim B}{\iota^2_\alpha A \sim \iota^2_\alpha B} \qquad \frac{A \sim D \quad B \sim E \quad C \sim F}{\delta A(u^\alpha.\,B)(v^\beta.\,C) \sim \delta D(u^\alpha.\,E)(v^\beta.\,F)}$$

$$\frac{FV(A) \subseteq \text{Safe}(u^\alpha, B)}{\delta(\iota^1_\beta A)(u^\alpha.\,B)(v^\beta.\,C) \sim B[A/u^\alpha]}\ \beta^1_+ \qquad \frac{FV(A) \subseteq \text{Safe}(v^\beta, B)}{\delta(\iota^2_\alpha A)(u^\alpha.\,B)(v^\beta.\,C) \sim C[A/v^\beta]}\ \beta^2_+$$

$$\frac{u^\alpha, v^\beta \notin FV(B) \quad FV(A) \cup \{u^\alpha, v^\beta\} \subseteq \text{Safe}(x^{\alpha+\beta}, B)}{\delta A(u^\alpha.\,B[\iota^1_\beta u^\alpha/x^{\alpha+\beta}])(v^\beta.\,B[\iota^2_\alpha v^\beta/x^{\alpha+\beta}]) \sim B[A/x^{\alpha+\beta}]}\ \eta_+$$

Here, the first three conditions merely guarantee that $\sim_{\beta\eta+}$ will still be a congruence with respect to the new term-forming operations. The next two conditions are the analogues, in the system with delta-application as primitive, of the "injection-conversion" rules $\langle\!\langle F, G\rangle\!\rangle \iota^1_\beta A \sim FA$ and $\langle\!\langle F, G\rangle\!\rangle \iota^2_\alpha B \sim GB$, and yield them as theorems: $\langle\!\langle F, G\rangle\!\rangle \iota^1_\beta A$ is $(\lambda x^{\alpha+\beta}.\delta x(y^\alpha.\,Fy)(z^\beta.\,Gz))\iota^1_\beta A$ (where x, y, z are not free in F, G, or A), which β-converts to $\delta(\iota^1_\beta A)(y^\alpha.\,Fy)(z^\beta.\,Gz)$, which β^1_+-converts to $(Fy)[A/y]$, which is just FA since y isn't free in F; similarly for the other rule. η_+, finally, plays a role analogous to that of η. If A was already of the form $\iota^1_\beta C$, the instance of η_+ would also be an instance of β^1_+ (since then $B[\iota^1_\beta u/x][C/u] = B[\iota^1_\beta C/x] = B[A/x]$) and similarly if A was already of the form $\iota^2_\alpha D$; so η_+ merely ensures that all terms of type $\alpha + \beta$ behave like these paradigm terms, just as η guarantees that all terms of type $\alpha \to \gamma$ behave like the paradigm such terms, namely those of the form $\lambda v^\alpha.A$ where $A : \gamma$.

Note that $\mathscr{L}^+$ does not have any new logical constants beyond those of $\mathscr{L}$: instead, we treat each quantifier $\forall_\alpha$, where α is an $\mathbb{T}^+$-type that is not an $\mathbb{T}$-type, as defined in terms of the basic logical constants already present in $\mathscr{L}$. The definitions are a little involved; I give them below.

Definition 14. The logic $\mathsf{H}^+(\Sigma)$ is defined by the same clauses as Definition 7, with $\sim_{\beta\eta+}$ replacing $\sim_{\beta\eta}$ in the conv rule, and one new rule:

$$\dfrac{\Gamma \vdash P[\iota^1_\beta u^\alpha/z^{\alpha+\beta}] \quad \Gamma \vdash P[\iota^2_\alpha v^\beta/z^{\alpha+\beta}] \quad FV(A) \subseteq \mathsf{Safe}(z^{\alpha+\beta}, P) \quad u^\alpha, v^\beta \notin FV(\Gamma \cup \{P\})}{\Gamma \vdash P[A/z^{\alpha+\beta}]}\ \text{Sum-Subst}$$

Note that since we don't have logical constants $\forall_\alpha$ unless α is a type in $\mathbb{T}$, the quantifier rules in this definition are to be understood as restricted to this case. However, as we will show below when we have defined $\forall_\alpha$ for extended types, H^+ is in fact also closed under $\forall$Intro and $\forall$Elim for other types.

Intuitively, Sum-Subst captures the idea that everything of type $\alpha+\beta$ is either $\iota^1_\beta u^\alpha$ for some u^α or $\iota^2_\alpha v^\beta$ for some v^β. Given the extended quantifier rules, we can prove this straightforwardly from Sum-Subst:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\vdash \iota^1_\beta u^\alpha =_{\alpha+\beta} \iota^1_\beta u^\alpha}\ {=}\text{Intro}}{\vdash \exists x^\alpha.\iota^1_\beta u^\alpha =_{\alpha+\beta} \iota^1_\beta x}\ \exists\text{Intro}}{\vdash (\exists x^\alpha.\iota^1_\beta u^\alpha =_{\alpha+\beta} \iota^1_\beta x) \vee (\exists y^\beta.\iota^1_\beta u^\alpha =_{\alpha+\beta} \iota^2_\alpha y)}\ \vee\text{Intro} \qquad \dfrac{\dfrac{\dfrac{}{\vdash \iota^2_\alpha v^\beta =_{\alpha+\beta} \iota^2_\alpha v^\beta}\ {=}\text{Intro}}{\vdash \exists y^\beta.\iota^2_\alpha v^\beta =_{\alpha+\beta} \iota^2_\alpha y}\ \exists\text{Intro}}{\vdash (\exists x^\alpha.\iota^2_\alpha v^\beta =_{\alpha+\beta} \iota^1_\beta x) \vee (\exists y^\beta.\iota^2_\alpha =_{\alpha+\beta} \iota^2_\alpha y)}\ \vee\text{Intro}}{\dfrac{\vdash (\exists x^\alpha.z^{\alpha+\beta} =_{\alpha+\beta} \iota^1_\beta x) \vee (\exists y^\beta.z^{\alpha+\beta} =_{\alpha+\beta} \iota^1_\alpha y)}{\vdash \forall z^{\alpha+\beta}.(\exists x^\alpha.z =_{\alpha+\beta} \iota^1_\beta x) \vee (\exists y^\beta.z =_{\alpha+\beta} \iota^1_\alpha y)}\ \forall\text{Intro}}\ \text{Sum-Subst}$$

In the presence of the quantifier rules for extended types, Sum-Subst is actually equivalent to this axiom. But Sum-Subst is used in deriving these rules.

Finally, we can define analogues of Classicism and Extensionalism for the extended language in the obvious way: C^+ is the result of closing H^+ under Subst, and E^+ is the result of closing H^+ under Ext. In C^+—and, a fortiori, in H^+—Sum-Subst need not be taken as primitive, since it can be derived from Subst. Let $\bigwedge\Gamma$ be the conjunction of the finite set Γ, x^α and y^β variables not free in Γ, and $z^{\alpha+\beta}$ a variable for which A is safe in P:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\Gamma \vdash \bigwedge\Gamma}\ \mathsf{H}}{\Gamma \vdash \delta A(x^\alpha.\bigwedge\Gamma)(y^\beta.\bigwedge\Gamma)}\ \eta_+ \quad \dfrac{\Gamma \vdash P[\iota^1_\beta x^\alpha/z]}{\bigwedge\Gamma \dashv\vdash P[\iota^1_\beta x^\alpha/z] \wedge \bigwedge\Gamma}\ \mathsf{H}}{\Gamma \vdash \delta A(x^\alpha.P[\iota^1_\beta x^\alpha/z] \wedge \bigwedge\Gamma)(y^\beta.\bigwedge\Gamma)}\ \text{Subst} \quad \dfrac{\Gamma \vdash P[\iota^2_\alpha y^\beta/z]}{\bigwedge\Gamma \dashv\vdash P[\iota^2_\alpha y^\beta/z] \wedge \bigwedge\Gamma}\ \mathsf{H}}{\Gamma \vdash \delta A(x^\alpha.P[\iota^1_\beta x^\alpha/z] \wedge \bigwedge\Gamma)(y^\beta.P[\iota^2_\alpha y^\beta/z] \wedge \bigwedge\Gamma)}\ \text{Subst}}{\dfrac{\Gamma \vdash P[A/z] \wedge \bigwedge\Gamma}{\Gamma \vdash P[A/z]}\ \wedge\text{-Elim}}\ \eta_+$$

In order to define $\forall_\alpha$ where α is an extended type, we will need a cluster of new definitions. These will also be useful in the next section, when we are showing how to eliminate the extended types.

Definition 15. We recursively define two subsets $\mathbb{T}^1$ and $\mathbb{T}^2$ of $(\mathbb{T}_0^+ - \mathbb{T})$; a mapping $\cdot' : \mathbb{T}^1 \to \mathbb{T}_0^+$; mappings $\cdot^*$ and $\cdot_* : \mathbb{T}^2 \to \mathbb{T}_0^+$; for each $\alpha \in \mathbb{T}^1$, functions $\nabla_\alpha : \mathscr{L}^+(\Sigma)^\alpha \to \mathscr{L}^+(\Sigma)^{\alpha'}$ and $\Delta_\alpha : \mathscr{L}^+(\Sigma)^{\alpha'} \to \mathscr{L}^+(\Sigma)^\alpha$; and for each $\alpha \in \mathbb{T}^2$, functions $\nabla^1_\alpha : \mathscr{L}^+(\Sigma)^\alpha \to \mathscr{L}^+(\Sigma)^{\alpha^*}$, $\nabla^2_\alpha : \mathscr{L}^+(\Sigma)^\alpha \to \mathscr{L}^+(\Sigma)^{\alpha_*}$, and $\Delta_\alpha : \mathscr{L}^+(\Sigma)^{\alpha^*} \times \mathscr{L}^+(\Sigma)^{\alpha_*} \to \mathscr{L}^+(\Sigma)^\alpha$, as follows (where in each case the variables x, y are chosen to not be free in the terms A, B):

- If $\tau \in \mathbb{T}$, then $\alpha + \beta \to \tau \in \mathbb{T}^2$, and

$$(\alpha + \beta \to \tau)^* = \alpha \to \tau \qquad (\alpha + \beta \to \tau)_* = \beta \to \tau$$
$$\nabla^1_{\alpha+\beta\to\tau}(A) := \lambda x^\alpha . A\iota^1_\beta x \qquad \nabla^2_{\alpha+\beta\to\tau}(A) := \lambda y^\beta . A\iota^2_\alpha y \qquad \Delta_{\alpha+\beta\to\tau}(B, C) := \langle\!\langle B, C\rangle\!\rangle$$

- If $\tau \in \mathbb{T}$ and $\alpha \in \mathbb{T}^2$, then $\alpha \to \tau \in \mathbb{T}^1$, and

$$(\alpha \to \tau)' = \alpha^* \to \alpha_* \to \tau$$
$$\nabla_{\alpha\to\tau}(A) := \lambda x^{\alpha^*} y^{\alpha_*} . A\Delta_\alpha(x, y) \qquad \Delta_{\alpha\to\tau}(B) := \lambda x^\alpha . B\nabla^1_\alpha(x)\nabla^2_\alpha(x)$$

- If $\tau \in \mathbb{T}$ and $\alpha \in \mathbb{T}^1$, then $\alpha \to \tau \in \mathbb{T}^1$, and

$$(\alpha \to \tau)' := \alpha' \to \tau \qquad \nabla_{\alpha\to\tau}(A) := \lambda x^{\alpha'} . A\Delta_\alpha(x) \qquad \Delta_{\alpha\to\tau}(B) := \lambda x^\alpha . A\nabla_\alpha(x)$$

- If $\beta \in \mathbb{T}^2$, then $\alpha \to \beta \in \mathbb{T}^2$, and

$$(\alpha \to \beta)^* = \alpha \to \beta^* \qquad (\alpha \to \beta)_* = \alpha \to \beta *$$
$$\Delta_{\alpha\to\beta}(A, B) = \lambda x^\alpha . \Delta_\beta(Ax, Bx) \qquad \nabla^1_{\alpha\to\beta}(A) = \lambda x^\alpha . \nabla^1_\beta(Ax) \qquad \nabla^2_{\alpha\to\beta}(A) = \lambda x^\alpha . \nabla^2_\beta(Ax)$$

- If $\beta \in \mathbb{T}^1$, then $\alpha \to \beta \in \mathbb{T}^1$, and

$$(\alpha \to \beta)' = \alpha \to \beta' \qquad \Delta_{\alpha\to\beta}(A) = \lambda x^\alpha . \Delta_\beta(Ax) \qquad \nabla_{\alpha\to\beta}(A) = \lambda x^\alpha . \nabla_\beta(Bx)$$

This is a legitimate recursive definition, since when α is in $\mathbb{T}_1^+$, α' contains fewer occurrences of + than α, and when α is in $\mathbb{T}_2^+$, both α^* and α_* contain fewer occurrences of + than α. Moreover, we can straightforwardly check that for every type in $\mathbb{T}^+$ is either in $\mathbb{T}$, in $\mathbb{T}^1$, in $\mathbb{T}^2$, or of the form $\alpha + \beta$. Thus, repeated application of $\cdot'$, $\cdot^*$, and $\cdot_*$ will reduce any type to either a sum-type or a $\mathbb{T}$-type.

The key property of these definitions is that the Δ and ∇ operations are "inverses", in a sense made precise by the following lemma:

Lemma 16.

$$
\begin{aligned}
\Delta_\alpha(\nabla_\alpha(A)) &\sim_{\beta\eta+} A && \text{for } \alpha \in \mathbb{T}^+_1\\
\nabla_\alpha(\Delta_\alpha(A)) &\sim_{\beta\eta+} A && \text{for } \alpha \in \mathbb{T}^+_1\\
\Delta_\alpha(\nabla^1_\alpha(A), \nabla^2_\alpha(A)) &\sim_{\beta\eta+} A && \text{for } \alpha \in \mathbb{T}^+_2\\
\nabla^1_\alpha(\Delta_\alpha(A, B)) &\sim_{\beta\eta+} A && \text{for } \alpha \in \mathbb{T}^+_2\\
\nabla^2_\alpha(\Delta_\alpha(A, B)) &\sim_{\beta\eta+} B && \text{for } \alpha \in \mathbb{T}^+_2
\end{aligned}
$$

This is proved by an induction on types, appealing to β_+ and η_+ for the base cases of types of the form $\alpha + \beta \to \tau$.

Using this system of isomorphisms, we can introduce the quantifiers for extended types:

Definition 17. The quantifiers $\forall_\alpha$ for $\mathbb{T}^+$-types α that are not in $\mathbb{T}$ are defined recursively as follows:

$$
\begin{aligned}
\forall_{\alpha+\beta} &:= \lambda X^{\overline{\alpha+\beta}}.(\forall y^\alpha.X\iota^1_\beta y) \wedge (\forall z^\beta.X\iota^2_\alpha z) &&\\
\forall_\alpha &:= \lambda X^{\overline{\alpha}}.\forall y^{\alpha^*} z^{\alpha_*}.X\Delta_\alpha(y, z) && \text{when } \alpha \in \mathbb{T}^+_2\\
\forall_\alpha &:= \lambda X^{\overline{\alpha}}.\forall y^{\alpha'}.X\Delta_\alpha(y) && \text{when } \alpha \in \mathbb{T}^+_1
\end{aligned}
$$

Since the types of quantifiers on the right always contain fewer occurrences of + than the ones on the left, this gives a well-formed definition for each $\forall_\alpha$.

Given these definitions, we can establish the following important fact:

Theorem 18. H^+ is closed under $\forall$Intro and $\forall$Elim for all the quantifiers $\forall_\alpha$.

Proof. By induction on the construction of the type-subscript of $\forall_\alpha$. The base case is that of a simple type σ. Then both claims hold since in that case $\forall_\sigma v\, P \sim_\beta \forall v : \sigma.\, P'$. For the induction step, there are three cases, depending on whether the type is a sum-type, a member of $\mathbb{T}^2$, or a member of $\mathbb{T}^1$.

(i) For a sum-type $\alpha + \beta$, we use the following derivations. On both sides, y^α and z^β are variables that don't occur in Γ or P; on the right, $v^{\alpha+\beta}$ is also required not to be free in Γ. Steps labeled 'SI' are justified by the substitution-invariance of H^+.

$$
\dfrac{\dfrac{\dfrac{\Gamma \vdash \forall v^{\alpha+\beta}.P}{\Gamma \vdash \forall y^\alpha.P[\iota^1_\beta y/v] \wedge \forall z^\beta.P[\iota^2_\alpha z/v]}\,\text{Df.}\forall_\alpha\text{, conv}}{\dfrac{\Gamma \vdash \forall y^\alpha.P[\iota^1_\beta y/v]}{\Gamma \vdash P[\iota^1_\beta y/v]}\,\text{IH} \qquad \dfrac{\Gamma \vdash \forall z^\beta.P[\iota^2_\alpha z/v]}{\Gamma \vdash P[\iota^2_\alpha z/v]}\,\text{IH}}\,\wedge\text{Elim}}{\Gamma \vdash P[A/v]}\,\text{Sum-Subst}
\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash P}{\Gamma \vdash P[\iota^1_\beta y/v]}\,\text{SI}}{\Gamma \vdash \forall y^\alpha.P[\iota^1_\beta y/v]}\,\text{IH} \qquad \dfrac{\dfrac{\Gamma \vdash P}{\Gamma \vdash P[\iota^2_\alpha z/v]}\,\text{SI}}{\Gamma \vdash \forall z^\beta.P[\iota^2_\alpha z/v]}\,\text{IH}}{\Gamma \vdash \forall y^\alpha.P[\iota^1_\beta y/v] \wedge \forall z^\beta.P[\iota^2_\alpha z/v]}\,\wedge\text{Intro}}{\Gamma \vdash \forall v^{\alpha+\beta}.P}\,\text{conv, Df.}\forall_{\alpha+\beta}
$$

(ii) For a type $\sigma \in \mathbb{T}^1$, we reason as follows:

$$\dfrac{\dfrac{\dfrac{\Gamma \vdash \forall_\alpha v.P}{\Gamma \vdash \forall y^{\alpha'}.P[\Delta_\alpha(y)/v]}\,\text{Df. }\forall_\alpha\text{, conv}}{\Gamma \vdash P[\Delta_\alpha(\nabla_\alpha(A))/v]}\,\text{IH}}{\Gamma \vdash P[A/v]}\,\text{Lemma 16} \qquad \dfrac{\dfrac{\dfrac{\Gamma \vdash P}{\Gamma \vdash P[\Delta_\alpha(y)/v]}\,\text{SI}}{\Gamma \vdash \forall y^{\alpha'}.P[\Delta_\alpha(y)/v]}\,\text{IH}}{\Gamma \vdash \forall v^\alpha.P}\,\text{conv, Df.}\forall_\alpha$$

(iii) For a type $\alpha \in \mathbb{T}^2$, we reason as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash \forall v^\alpha.P}{\Gamma \vdash \forall y^{\alpha^*} z^{\alpha_*}.P[\Delta_\alpha(y,z)/v]}\,\text{Df. }\forall_\alpha\text{, conv}}{\Gamma \vdash \forall z^{\alpha_*}.P[\Delta_\alpha(\nabla^1_\alpha(A),z)/v]}\,\text{IH}}{\Gamma \vdash P[\Delta_\alpha(\nabla^1_\alpha(A),\nabla^2_\alpha(A))/v]}\,\text{IH}}{\Gamma \vdash P[A/v]}\,\text{Lemma 16} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash P}{\Gamma \vdash P[\Delta_\alpha(y,z)/v]}\,\text{SI}}{\Gamma \vdash \forall z^{\alpha_*}.P[\Delta_\alpha(y,z)/v]}\,\text{IH}}{\Gamma \vdash \forall y^{\alpha^*} z^{\alpha_*}.P[\Delta_\alpha(y,z)/v]}\,\text{IH}}{\Gamma \vdash \forall v^\alpha.P}\,\text{conv, Df.}\forall_\alpha$$

We could equally well have defined H^+ using the full versions of ∀Intro and ∀Elim and leaving out Sum-Subst, since that rule is derivable from ∀Elim given the definition of $\forall_{\alpha+\beta}$.[76] Note however that this depends crucially on the definition of $\forall_{\alpha+\beta}$. If we had chosen instead to take $\forall_\alpha$ as primitive for every $\mathbb{T}^+$-type, Sum-Subst would not have been derivable even assuming ∀Intro and ∀Elim in every type.[77]

Appendix C Dispensing with sum-types

This appendix will describe mappings that turn any $\mathbb{T}^+$-signature Σ into a $\mathbb{T}$-signature $\Sigma^\dagger$ (that includes all the $\mathbb{T}$-typed constants of Σ), and turn any sentence P of $\mathscr{L}^+(\Sigma)$ into a sentence $P^\dagger$ of $\mathscr{L}(\Sigma^\dagger)$, in a way that faithfully preserves logical relationships: $\Gamma \vdash_{\mathsf{H}^+(\Sigma)} P$ iff $\Gamma^\dagger \vdash_{\mathsf{H}(\Sigma^\dagger)} P^\dagger$.[78] The point is to take a theory T in some extended language $\mathscr{L}^+(\Sigma)$ and replace it with a new theory $T^\dagger$ (in $\mathscr{L}(\Sigma^\dagger)$), that has exactly

[76] Where u^α and v^β aren't free in Γ or P and A is safe for $z^{\alpha+\beta}$ in P:

$$\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash P[\iota^1_\beta u^\alpha/z^{\alpha+\beta}]}{\Gamma \vdash \forall u^\alpha.P[\iota^1_\beta u^\alpha/z^{\alpha+\beta}]}\,\forall\text{Intro} \quad \dfrac{\Gamma \vdash P[\iota^2_\alpha v^\beta/z^{\alpha+\beta}]}{\Gamma \vdash \forall v^\beta.P[\iota^1_\beta v^\beta/z^{\alpha+\beta}]}\,\forall\text{Intro}}{\dfrac{\Gamma \vdash (\forall u^\alpha.P[\iota^1_\beta u^\alpha/z^{\alpha+\beta}]) \wedge (\forall v^\beta.P[\iota^1_\beta v^\beta/z^{\alpha+\beta}])}{\Gamma \vdash \forall z^{\alpha+\beta}.P}\,\text{Df.}\forall_{\alpha+\beta}\text{, conv}}\,\wedge\text{Intro}}{\Gamma \vdash (\lambda z^{\alpha+\beta}.P)A}\,\forall\text{Elim}}{\Gamma \vdash P[A/z^{\alpha+\beta}]}\,\text{conv}$$

[77] This can be shown by constructing a simple model, but I will not describe it here.

[78] $\Gamma^\dagger$ here is $\{Q^\dagger : Q \in \Gamma\}$; we lift other functions from formulae to sets of formulae analogously.

the same consequences as T in the simply-typed fragment $\mathscr{L}(\Sigma_0)$ of the original language (where Σ_0 is the restriction of Σ to $\mathbb{T}$-typed constants). Assuming the facts we are trying to *explain* (e.g., meaning-facts of the form $A \vdots_t \mathrm{s}\, P$) can be stated in $\mathscr{L}(\Sigma_0)$, this means that we could explain them just as well by appealing to $T^{\dagger}$, without thereby incurring any potentially problematic commitments to new sentences of $\mathscr{L}(\Sigma_0)$.

The key to this mapping is a result from Prawitz (1965: ch. IV). To state it, we define the obvious syntactic "containment" relations on extended types and terms:

- $\leq_{\mathbb{T}^+}$ is the smallest relation $\leq$ on $\mathbb{T}^+$ such that for all $\alpha, \beta \in \mathbb{T}^+$ and $\gamma \in \mathbb{T}_0^+$: $\alpha \leq \alpha$, $\alpha \leq \alpha \to \gamma$, $\gamma \to \alpha \to \gamma$, $\alpha \leq \alpha + \beta$, and $\beta \leq \alpha + \beta$.

- $\leq_{\mathscr{L}^+(\Sigma)}$ is the smallest relation $\leq$ on $\mathscr{L}^+(\Sigma)$ such that for all terms $A, B \in \mathscr{L}^+(\Sigma)$: $A \leq A$, $A \leq AB$, $B \leq AB$, $A \leq \lambda v.A$, $A \leq \iota^1_\alpha A$, $A \leq \iota^2_\alpha A$, $A \leq \delta A(u.\,B)(v.\,C)$, $B \leq \delta A(u.\,B)(v.\,C)$, and $C \leq \delta A(u.\,B)(v.\,C)$, whenever these are well-formed.

Say that a term A has the *subterm property* iff for any terms $B : \beta$ and $C : \gamma$, if B is contained in A and C is contained in B, then γ is either contained in β, contained in the type of some free variable in B, or contained in the type of some constant in B. Then the result we'll need is the following:

Theorem 19 (Prawitz)**.** There is a "normalization" function $\Downarrow$ on $\mathscr{L}^+(\Sigma)$-terms such that for every $\mathscr{L}^+$-term A: (i) $A \sim_{\beta\eta+} \Downarrow A$; (ii) if A is an $\mathscr{L}(\Sigma_0)$-term, $A \sim_{\beta\eta} \Downarrow A$; (iii) $\Downarrow(A[B/v]) = \Downarrow(A[\Downarrow B/v])$; (iv) $\Downarrow A$ contains no terms matching the left-hand-side of the β, β^1_+, or β^2_+ rules; (v) $FV(\Downarrow A) \subseteq FV(A)$; and (vi) $\Downarrow A$ has the subterm property.[79]

Note that any $\mathbb{T}^+$-type that is contained in an $\mathbb{T}$-type is itself an $\mathbb{T}$-type. So when Σ_0 is a $\mathbb{T}$-signature and A is a closed $\mathscr{L}^+(\Sigma_0)$-term such that its type and the types of all its free variables are in $\mathbb{T}$, the conditions on $\Downarrow$ imply that $\Downarrow A$ belongs to $\mathscr{L}(\Sigma_0)$, since it cannot contain any terms constructed using ι^1,ι^2, or δ. In particular, this is true when A is a *sentence* (type-t term with no free variables) of $\mathscr{L}^+(\Sigma_0)$.

We will use $\Downarrow$ as the last step in our translation procedure. Since our *logical* constants already all have types in $\mathbb{T}$, our remaining task in specifying the translation thus amounts to providing, for each nonlogical constant in the input sentences, a replacement expression built up entirely from $\mathbb{T}$-typed constants.

I will assume here that our initial $\mathbb{T}^+$-signature Σ doesn't have any nonlogical constants in sum-types $\alpha + \beta$ (though it may still have constants in other extended types such as $\overline{\alpha + \beta}$ or $\overline{\underline{\alpha + \beta}}$). This is fine for the application we have in mind, since the semantic theory stated in terms of sum-types didn't need any sum-typed constants.[80]

[79]Prawitz's theorem concerns derivations in intuitionistic propositional logic rather than terms, but the result is essentially the same thanks to the Curry-Howard isomorphism. We could if we wished include further normalization steps which would, e.g., guarantee that $\Downarrow A$ also contains no terms matching the left-hand-side of η or η_+.

[80]If we wanted to allow constants with sum-types, we would need an extra step where we add a sentence letter P_c to $\Sigma^{\dagger}$ for each such constant in Σ, such that the translation of $\exists x^{\alpha}.c =_{\alpha+\beta} \iota^1_\beta x$ will be logically equivalent to P_c.

Given this restriction, we can specify a procedure for eliminating $\mathbb{T}^+$-typed constants by appealing to the "isomorphisms" defined in the previous section. Starting with an $\mathbb{T}^+$-signature Σ, we define two new signatures:

- Σ^* is the smallest extension of Σ such that whenever $c :_{\Sigma^*} \alpha$ and $\alpha \in \mathbb{T}_1^+$, $c' :_{\Sigma^*} \alpha'$, and whenever $c :_{\Sigma^*} \alpha$ and $\alpha \in \mathbb{T}_2^+$, $c^* :_{\Sigma^*} \alpha^*$ and $c_* :_{\Sigma^*} \alpha_*$.
- $\Sigma^\dagger$ is the restriction of Σ^* to $\mathbb{T}$.

We then define a mapping $|\cdot|$ from $\mathscr{L}^+(\Sigma^*)$ to $\mathscr{L}^+(\Sigma^\dagger)$. The interesting clauses are the ones for nonlogical constants:

$$
\begin{aligned}
|c| &\coloneqq c && \text{when } c :_{\Sigma^*} \alpha \text{ and } \alpha \in \mathbb{T}\\
|c| &\coloneqq \Delta_\alpha(|c'|) && \text{when } c :_{\Sigma^*} \alpha \text{ and } \alpha \in \mathbb{T}_1^+\\
|c| &\coloneqq \Delta_\alpha(|c^*|, |c_*|) && \text{when } c :_{\Sigma^*} \alpha \text{ and } \alpha \in \mathbb{T}_2^+
\end{aligned}
$$

All the other clauses are trivial. Thus, for every term A of $\mathscr{L}^+(\Sigma^*)$, $|A|$ is a term of $\mathscr{L}^+(\Sigma^\dagger)$ with the same free variables. When A and all of its free variables have types in $\mathbb{T}$—for example, when A is a sentence—we define its translation $A^\dagger$ to be $\Downarrow|A|$: the $\mathscr{L}(\Sigma^\dagger)$-term that results from applying Prawitz's normalization function to $|A|$.

The adequacy of our translation scheme in this sense follows from the following fact:

Theorem. If all free variables in Γ and P are $\mathbb{T}$-typed, then $\Gamma \vdash_{\mathsf{H}^+(\Sigma)} P$ iff $\Gamma^\dagger \vdash_{\mathsf{H}(\Sigma^\dagger)} P^\dagger$.

Proof. Let T be the set of all sentences of the form $c =_\alpha \Delta_\alpha(c')$ (where $c :_{\Sigma^*} \alpha$ and $\alpha \in \mathbb{T}_1^+$) or $c =_\alpha \Delta_\alpha(c^*, c_*)$ (where $c :_{\Sigma^*} \alpha$ and $\alpha \in \mathbb{T}_2^+$). To prove the theorem, we establish the following five biconditionals:

$$
\begin{array}{ll}
\text{(i)} & \Gamma \vdash_{\mathsf{H}^+(\Sigma)} P \text{ iff } \Gamma \cup T \vdash_{\mathsf{H}^+(\Sigma^*)} P\\
\text{(ii)} & \Gamma \cup T \vdash_{\mathsf{H}^+(\Sigma^*)} P \text{ iff } |\Gamma| \cup T \vdash_{\mathsf{H}^+(\Sigma^*)} |P|\\
\text{(iii)} & |\Gamma| \cup T \vdash_{\mathsf{H}^+(\Sigma^*)} |P| \text{ iff } |\Gamma| \vdash_{\mathsf{H}^+(\Sigma^\dagger)} |P|\\
\text{(iv)} & |\Gamma| \vdash_{\mathsf{H}^+(\Sigma^\dagger)} |P| \text{ iff } \Gamma^\dagger \vdash_{\mathsf{H}^+(\Sigma^\dagger)} P^\dagger\\
\text{(v)} & \Gamma^\dagger \vdash_{\mathsf{H}^+(\Sigma^\dagger)} P^\dagger \text{ iff } \Gamma^\dagger \vdash_{\mathsf{H}(\Sigma^\dagger)} P^\dagger
\end{array}
$$

The left-to-right direction of (i) is trivial. For the other direction, we make the substitutions $[\nabla_\alpha(c)/c']$, $[\nabla_\alpha^1(c)/c^*]$, and $[\nabla_\alpha^2(c)/c_*]$. This turns every member of T into a theorem of $\mathsf{H}^+(\Sigma)$ (which is thus redundant) and leaves Γ and P unaffected.

For (ii), we show (by an easy induction) that $T \vdash_{\mathsf{H}^+(\Sigma^*)} A =_\alpha |A|$ for every $\mathscr{L}^+(\Sigma)$-term $A : \alpha$.

The right-to-left direction of (iii) is trivial. For the other direction, we make the substitutions $\Delta(\alpha)c'/c$ and $\Delta(\alpha)c^*c_*/c$. This turns every axiom of T into a theorem of $\mathsf{H}^+(\Sigma^\dagger)$ and leaves $|\Gamma|$ and $|P|$ unaffected.

For (iv), we use the conversion rule of H^+ to show that conversion preserves derivability.

The right-to-left direciton of (v) is trivial. For the other direction, we show (by induction on the definition of provability in $\mathsf{H}^+(\Sigma^\dagger)$) that every sequent (i.e. ordered pair)

$\Delta \rhd Q$ in $\mathsf{H}^+(\Sigma^\dagger)$ has the following property: for any non-$\mathbb{T}$-typed variables $x_1, ..., x_n$ that include all such variables that are free in Δ or Q, and any $\mathscr{L}^+(\Sigma^\dagger)$-terms $B_1, ..., B_n$ of the same types as $x_1, ..., x_n$ in which all free variables are $\mathbb{T}$-typed and safe for $x_1, ..., x_n$ in Δ and Q, $\Downarrow(\Delta[B_1/x_1] \cdots [B_n/x_n]) \vdash_{\mathsf{H}(\Sigma^\dagger)} \Downarrow(Q[B_1/x_1] \cdots [B_n/x_n])$. The result follows from the $n = 0$ case of this, setting $\Delta = |\Gamma|$ and $Q = |P|$.

The only nontrivial step is the one for Sum-Subst, since all the other rules of H^+ except for the conversion rule are also rules of H, and convertibility is preserved by $\Downarrow$ and the safe substitutions $[B_i/x_i]$. So, suppose $\Delta \rhd Q$ follows by Sum-Subst from sequents $\Delta' \rhd Q'$ and $\Delta'' \rhd Q''$ that have the given property. Then Q is $P[A/z]$ for some variable z of type $\alpha + \beta$ and some A safe for z in P; $\Delta' = \Delta'' = \Delta$; $Q' = P[\iota^1_\beta u^\alpha/z]$, and $Q'' = P[\iota^2_\alpha v^\beta/z]$ for some variables u^α, v^β not free in Δ or P. We may assume that z is free in P, since if isn't, $Q = Q'$ and we are done. Let $x_1, ..., x_n$ be any variables, and $B_1, ..., B_n$ be terms of the same types in which all free variables are $\mathbb{T}$-typed and not free in Δ or $P[A/z]$. Let A^* be $A[B_1/x_1] \cdots [B_n/x_n]$. We can assume that z isn't among $x_1, ..., x_n$ or free in Γ or A; if this assumption is false, just pick some z' that does meet these conditions and replace P with $P[z'/z]$ throughout. Also z isn't free in any B_i since it has a non-simple type. So we have

$$\begin{aligned}\Downarrow(P[A/z][B_1/x_1] \cdots [B_n/x_n]) &= \Downarrow(P[B_1/x_1] \cdots [B_n/x_n][A^*/z]) \\ &= \Downarrow(P[B_1/x_1] \cdots [B_n/x_n][\Downarrow A^*/z]) \\ &= \Downarrow(P[\Downarrow A^*/z][B_1/x_1] \cdots [B_n/x_n])\end{aligned}$$

(since none of $x_1, ..., x_n$ is free in $\Downarrow A^*$ and z isn't free in any of $B_1, ..., B_n$). Since $\Downarrow A^*$ has the subterm property and all of its free variables have $\mathbb{T}$-types, all of its constituents that have non-$\mathbb{T}$ types must have types that are constituents of its type, $\alpha + \beta$. It follows from this that $\Downarrow A^*$ is either $\iota^1_\beta C$ for some $C : \alpha$ or $\iota^2_\alpha D$ for some $D : \beta$. For it can't be a constant, since $\Sigma^\dagger$ only has $\mathbb{T}$-typed constants. It can't be a variable, since all its free variables are $\mathbb{T}$-typed. It can't be a lambda abstraction, since it doesn't have a function type. It can't be an application, since $\alpha + \beta$ isn't a terminal type.[81] And finally, the fact that the output of $\Downarrow$ never contains expressions matching the left-hand-side of β^1_+ or β^2_+ ensures that it can't be a delta-application $\delta C(x.\, D)(y.\, E)$.[82] Suppose without loss of generality that $\Downarrow A^*$ is $\iota^1_\beta C$ (the case where it is $\iota^2_\alpha D$ is parallel). Since all free variables in C are simply typed, safe for v in Δ (since v isn't free in Δ), and safe for v in $P[\iota^1_\beta v/z]$ (since safe for z in P), the induction hypothesis tells us that $\Downarrow(\Delta[C/v][B_1/x_1] \cdots [B_n/x_n]) \vdash_{\mathsf{H}(\Sigma^\dagger)} \Downarrow(P[\iota^1_\beta v/z][C/v][B_1/x_1] \cdots [B_n/x_n])$. But

[81] Even if we allowed terminal sum-types, $\Downarrow A^*$ still couldn't be an application CD, since C would have to have some type $\gamma \to (\alpha + \beta)$ not contained in $\alpha + \beta$.

[82] If it were, then C would also have to be a term of some sum-type $\alpha + \beta$. By the same reasoning it could not be a lambda-abstract or an application, so each must either be of the form $\iota^1_\beta C'$, $\iota^2_\alpha C'$, or $\delta C'(x'.\, D')(y'.\, E')$, where C' is of some even simpler sum-type $\alpha' + \beta'$. We can't have an infinitely descending sequence of delta-applications, so we must eventually find some subterm $\delta C'(x'.\, D')(y'.\, E')$ where C' is of the form $\iota^1_{\beta'} C''$ or $\iota^2_{\alpha''} C''$. But we know that the output of $\Downarrow$ never contains any terms of this form (β_+-redexes).

$\Delta[C/v][B_1/x_1]\cdots[B_n/x_n] = \Delta[B_1/x_1]\cdots[B_n/x_n]$, since v isn't free in Δ. And

$$\begin{aligned}\Downarrow(P[\iota^1_\beta v/z][C/v][B_1/x_1]\cdots[B_n/x_n]) &= \Downarrow(P[\iota^1_\beta C/z][B_1/x_1]\cdots[B_n/x_n]) \\ &= \Downarrow(P[\Downarrow A^*/z][B_1/x_1]\cdots[B_n/x_n]) \\ &= \Downarrow(P[A/z][B_1/x_1]\cdots[B_n/x_n]) \\ &= \Downarrow(Q[B_1/x_1]\cdots[B_n/x_n]).\end{aligned}$$

Hence $\Downarrow(\Delta[B_1/x_1]\cdots[B_n/x_n]) \vdash_{\mathrm{H}(\Sigma^\dagger)} \Downarrow(Q[B_1/x_1]\cdots[B_n/x_n])$, as desired.

To conclude, we should check that the described translation procedure does in fact apply in the way I claimed to the three particular axioms of the extended semantic theory I took as examples in §6.5. Here they are again, written in official prefix form, abbreviating $\overline{e+\bar{e}}$ as α:

(51) $\vdots_{\overline{\alpha}}$ everything s/np\s $\forall_{e+\bar{e}}$

(52) $\vdots_{\alpha}$ is·interesting np\s $\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\bar{e}}_{-} \rangle\!\rangle$

(53) $\forall\, a^\epsilon b^\epsilon c^\epsilon d^\epsilon X^{\overline{\alpha}} y^\alpha .\, (\vdots_{\overline{\alpha}}\ a\, c/d\, X) \wedge (\vdots_\alpha\ b\, d\, y) \to (\vdots_t\ a{\cdot}b\, c\, X y)$

These sentences contain (in addition to several $\mathbb{T}$-typed nonlogical constants), two non-$\mathbb{T}$-typed nonlogical constants $\vdots_\alpha$ and $\vdots_{\overline{\alpha}}$, of types $\overline{\epsilon\epsilon\alpha}$ and $\overline{\epsilon\epsilon\overline{\overline{\alpha}}}$ respectively. α is in $\mathbb{T}^+_2$, so $\overline{\alpha}$, $\overline{\epsilon\alpha}$, $\overline{\epsilon\epsilon\alpha}$, $\underline{\overline{\alpha}}$, $\overline{\epsilon\overline{\overline{\alpha}}}$, and $\overline{\epsilon\epsilon\overline{\overline{\alpha}}}$ are all in $\mathbb{T}^1$. All the types just mentioned are just one reduction-step away from $\mathbb{T}$:

$$\begin{aligned}
\alpha^* &= \bar{e} \\
\alpha_* &= \underline{\bar{e}} \\
\overline{\alpha}' &= \alpha^* \to \alpha_* \to t && = \bar{e} \to \underline{\bar{e}} \to t \\
\overline{\epsilon\alpha}' &= \epsilon \to \overline{\alpha}' && = \epsilon \to \bar{e} \to \underline{\bar{e}} \to t \\
\overline{\epsilon\epsilon\alpha}' &= \epsilon \to \overline{\epsilon\alpha}' && = \epsilon \to \epsilon \to \bar{e} \to \underline{\bar{e}} \to t \\
\underline{\overline{\alpha}}' &= \overline{\alpha}' \to t && = (\bar{e} \to \underline{\bar{e}} \to t) \to t \\
\overline{\epsilon\overline{\overline{\alpha}}}' &= \epsilon \to \underline{\overline{\alpha}}' && = \epsilon \to (\bar{e} \to \underline{\bar{e}} \to t) \to t \\
\overline{\epsilon\epsilon\overline{\overline{\alpha}}}' &= \epsilon \to \overline{\epsilon\overline{\overline{\alpha}}}' && = \epsilon \to \epsilon \to (\bar{e} \to \underline{\bar{e}} \to t) \to t
\end{aligned}$$

When we move to the expanded signature Σ^*, we will add new $\mathbb{T}$-typed constants $\vdots'_\alpha$: $\epsilon\epsilon\overline{\alpha}'$ and $\vdots'_{\overline{\alpha}}$: $\overline{\epsilon\epsilon\overline{\overline{\alpha}}}'$. To apply $|\cdot|$, we simply replace $\vdots_\alpha$ with $\Delta_{\overline{\epsilon\epsilon\alpha}}(\vdots'_\alpha)$ and $\vdots_{\overline{\alpha}}$ with

$\Delta_{\overline{\epsilon\epsilon\overline{\alpha}}}(\vdots'_{\overline{\alpha}})$. The resulting sentences can then be reduced to sentences of $\mathscr{L}$, as follows

$$
\begin{aligned}
|(51)| &= \Delta_{\overline{\epsilon\epsilon\overline{\alpha}}}(\vdots'_{\overline{\alpha}})\ \text{everything}\ \text{s/np\textbackslash s}\ \forall_{\alpha}\\
&\sim_{\beta\eta+} \Delta_{\overline{\epsilon\overline{\alpha}}}(\vdots'_{\overline{\alpha}}\ \text{everything})\ \text{s/np\textbackslash s}\ \forall_{\alpha}\\
&\sim_{\beta\eta+} \Delta_{\underline{\overline{\alpha}}}(\vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s})\ \forall_{\alpha}\\
&\sim_{\beta\eta+} \vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s}\ \nabla_{\overline{\alpha}}(\forall_{\alpha})\\
&\sim_{\beta\eta+} \vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s}\ (\lambda Y^{\bar{e}} Z^{\underline{\bar{e}}}.\forall_{\alpha}(\Delta_{\alpha}(Y,Z)))\\
&\sim_{\beta\eta+} \vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s}\ (\lambda Y^{\bar{e}} Z^{\underline{\bar{e}}}.\forall_{\alpha}\langle\!\langle Y,Z\rangle\!\rangle)\\
&= \vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s}\ (\lambda Y^{\bar{e}} Z^{\underline{\bar{e}}}.(\lambda X^{\alpha}.\forall y^{e}.X\iota^{1}_{e}y \wedge \forall z^{\bar{e}}.X\iota^{2}_{e}z)\langle\!\langle Y,Z\rangle\!\rangle)\\
&\sim_{\beta\eta+} \vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s}\ (\lambda Y^{\bar{e}} Z^{\underline{\bar{e}}}.\forall y^{e}.\langle\!\langle Y,Z\rangle\!\rangle\iota^{1}_{e}y \wedge \forall z^{\bar{e}}.\langle\!\langle Y,Z\rangle\!\rangle\iota^{2}_{e}z)\\
&\sim_{\beta\eta+} \vdots'_{\overline{\alpha}}\ \text{everything}\ \text{s/np\textbackslash s}\ (\lambda Y^{\bar{e}} Z^{\underline{\bar{e}}}.\forall y^{e}.Yy \wedge \forall z^{\bar{e}}.Zz)
\end{aligned}
$$

$$
\begin{aligned}
|(52)| &= \Delta_{\overline{\epsilon\epsilon\alpha}}(\vdots'_{\alpha})\ \text{is·interesting}\ \text{np\textbackslash s}\ \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle\\
&\sim_{\beta\eta+} \Delta_{\overline{\epsilon\alpha}}(\vdots'_{\alpha}\ \text{is·interesting})\ \text{np\textbackslash s}\ \langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle\\
&\sim_{\beta\eta+} \Delta_{\overline{\alpha}}(\vdots'_{\alpha}\ \text{is·interesting}\ \text{np\textbackslash s})\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle\\
&\sim_{\beta\eta+} \vdots'_{\alpha}\ \text{is·interesting}\ \text{np\textbackslash s}\ \nabla^{1}_{\alpha}(\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle)\ \nabla^{2}_{\alpha}(\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle)\\
&\sim_{\beta\eta+} \vdots'_{\alpha}\ \text{is·interesting}\ \text{np\textbackslash s}\ (\lambda y^{e}.\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle\iota^{1}_{e}y)\ (\lambda z^{\bar{e}}.\langle\!\langle \mathbf{int}^{\bar{e}}, \mathbf{int}^{\underline{\bar{e}}}\rangle\!\rangle\iota^{2}_{e}z)\\
&\sim_{\beta\eta+} \vdots'_{\alpha}\ \text{is·interesting}\ \text{np\textbackslash s}\ (\lambda y^{e}.\ \mathbf{int}^{\bar{e}}\ y)\ (\lambda z^{\bar{e}}.\ \mathbf{int}^{\underline{\bar{e}}}\ z)\\
&\sim_{\beta\eta+} \vdots'_{\alpha}\ \text{is·interesting}\ \text{np\textbackslash s}\ \mathbf{int}^{\bar{e}}\ \mathbf{int}^{\underline{\bar{e}}}
\end{aligned}
$$

$$
\begin{aligned}
|(53)| &= \forall a^{\epsilon} b^{\epsilon} c^{\epsilon} d^{\epsilon} X^{\overline{\alpha}} y^{\alpha}.\ (\Delta_{\overline{\epsilon\epsilon\overline{\alpha}}}(\vdots'_{\overline{\alpha}})\ a\ c/d\ X) \wedge (\Delta_{\overline{\epsilon\epsilon\alpha}}(\vdots'_{\alpha})\ b\ d\ y) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ Xy)\\
&\sim_{\beta\eta+} \forall a^{\epsilon} b^{\epsilon} c^{\epsilon} d^{\epsilon} X^{\overline{\alpha}} y^{\alpha}.\ (\Delta_{\overline{\epsilon\overline{\alpha}}}(\vdots'_{\overline{\alpha}}\ a)\ c/d\ X) \wedge (\Delta_{\overline{\epsilon\alpha}}(\vdots'_{\alpha}\ b)\ d\ y) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ Xy)\\
&\sim_{\beta\eta+} \forall a^{\epsilon} b^{\epsilon} c^{\epsilon} d^{\epsilon} X^{\overline{\alpha}} y^{\alpha}.\ (\Delta_{\underline{\overline{\alpha}}}(\vdots'_{\overline{\alpha}}\ a\ c/d)\ X) \wedge (\Delta_{\overline{\alpha}}(\vdots'_{\alpha}\ b\ d)\ y) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ Xy)\\
&\sim_{\beta\eta+} \forall a^{\epsilon} b^{\epsilon} c^{\epsilon} d^{\epsilon} X^{\overline{\alpha}} y^{\alpha}.\ (\vdots'_{\overline{\alpha}}\ a\ c/d\ \nabla_{\overline{\alpha}}(X)) \wedge (\vdots'_{\alpha}\ b\ d\ \nabla^{1}_{\alpha}(y)\nabla^{2}_{\alpha}(y)) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ Xy)\\
&\sim_{\beta\eta+} \forall a^{\varepsilon} b^{\varepsilon} c^{\varepsilon} d^{\varepsilon} X^{\overline{\alpha}} y^{\bar{e}} z^{\underline{\bar{e}}}.\ (\vdots'_{\overline{\alpha}}, a\ c/d\ \nabla_{\overline{\alpha}}(X)) \wedge (\vdots'_{\alpha}\ b\ d\ \nabla^{1}_{\alpha}(\Delta_{\alpha}(y,z))\nabla^{2}_{\alpha}(\Delta_{\alpha}(y,z)))\\
&\qquad\qquad \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ (X\Delta_{\alpha}(y,z)))\\
&\sim_{\beta\eta+} \forall a^{\varepsilon} b^{\varepsilon} c^{\varepsilon} d^{\varepsilon} X^{\overline{\bar{e}\underline{\bar{e}}}} y^{\bar{e}} z^{\underline{\bar{e}}}.\ (\vdots'_{\overline{\alpha}}, a\ c/d\ \nabla_{\overline{\alpha}}(\Delta_{\overline{\alpha}}(X))) \wedge (\vdots'_{\alpha}\ b\ d\ y\ z) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ (\Delta_{\overline{\alpha}}(X)\Delta_{\alpha}(y,z)))\\
&\sim_{\beta\eta+} \forall a^{\varepsilon} b^{\varepsilon} c^{\varepsilon} d^{\varepsilon} X^{\overline{\bar{e}\underline{\bar{e}}}} y^{\bar{e}} z^{\underline{\bar{e}}}.\ (\vdots'_{\overline{\alpha}}, a\ c/d\ X) \wedge (\vdots'_{\alpha}\ b\ d\ y\ z) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ (X\nabla^{1}_{\alpha}(\Delta_{\alpha}(y,z))\nabla^{2}_{\alpha}(\Delta_{\alpha}(y,z))))\\
&\sim_{\beta\eta+} \forall a^{\varepsilon} b^{\varepsilon} c^{\varepsilon} d^{\varepsilon} X^{\overline{\bar{e}\underline{\bar{e}}}} y^{\bar{e}} z^{\underline{\bar{e}}}.\ (\vdots'_{\overline{\alpha}}, a\ c/d\ X) \wedge (\vdots'_{\alpha}\ b\ d\ y\ z) \rightarrow (\vdots_{t}\ a{\cdot}b\ c\ (Xyz))
\end{aligned}
$$

References

Bacon, Andrew (2023), *A Philosophical Introduction to Higher-Order Logics* (London: Routledge).

Bacon, Andrew (2024), 'A Case For Higher-Order Metaphysics', in Peter Fritz and Nicholas K. Jones (eds.), *Higher-Order Metaphysics* (Oxford: Oxford University Press).

Bacon, Andrew, and Dorr, Cian (2024), 'Classicism', in Peter Fritz and Nicholas K. Jones (eds.), *Higher-Order Metaphysics* (Oxford: Oxford University Press): 109–90.

Bar-Hillel, Yehoshua (1953), 'A Quasi-Arithmetical Notation for Syntactic Description', *Language*, 29/1: 47–58.

Barker, Chris, and Shan, Chung-chieh (2014), *Continuations and Natural Language* (Oxford: Oxford University Press).

Bealer, George (1993), 'Universals', *Journal of Philosophy*, 110: 5–32.

Burgess, John (2008), 'Tarski's Tort', in *Mathematics, Models, and Modality: Selected Philosophical Essays* (Cambridge: Cambridge University Press): 149–68.

Button, Tim, and Trueman, Robert (2024), 'A Fictionalist Theory of Universals', in Peter Fritz and Nicholas K. Jones (eds.), *Higher-Order Metaphysics* (Oxford: Oxford University Press).

Carlson, Greg N. (1977), 'A Unified Analysis of the English Bare Plural', *Linguistics and Philosophy*, 1: 413–57.

Carpenter, Bob (1997), *Type-Logical Semantics* (Cambridge, MA: MIT Press).

Carroll, Lewis (1895), 'What the Tortoise Said to Achilles', *Mind*, 14: 278–80.

Chierchia, Gennaro (1982), 'Nominalization and Montague grammar: A Semantics Without Types for Natural Languages', *Linguistics and Philosophy*, 5/3: 303–54.

Chierchia, Gennaro (1984), 'Topics in the Syntax and Semantics of Infinitives and Gerunds', PhD thesis (University of Massachusetts).

Cresswell, M. J. (1973), *Logics and Languages* (London: Methuen).

Curry, Haskell Brooks, and Feys, Robert (1958), *Combinatory Logic* (Amsterdam: North-Holland).

Davidson, Donald (1980), 'The Logical Form of Action Sentences', in *Essays on Actions and Events* (2nd edn., Oxford: Oxford University Press, 2001): 105–48.

de Groote, Philippe (2001), 'Towards Abstract Categorial Grammars', in *Association for Computational Linguistics*, 39th Annual Meeting and 10th Conference of the European Chapter (Toulouse, France): 148–55. doi: `10.3115/1073012.1073045.inria-00100529`.

Dorr, Cian (2007), 'There Are No Abstract Objects', in Theodore Sider, John Hawthorne, and Dean Zimmerman (eds.), *Contemporary Debates in Metaphysics* (Malden, Mass.: Wiley-Blackwell): 32–64.

Dorr, Cian (2014*a*), 'Quantifier Variance and the Collapse Theorems', *The Monist*, 97: 503–70.

Dorr, Cian (2014*b*), 'Transparency and the Context-sensitivity of Attitude Reports', in Manuel García-Carpintero and Genoveva Martí (eds.), *Empty Representations: Reference and Non-existence* (Oxford: Oxford University Press): 25–66.

Dorr, Cian (2016), 'To Be F Is To Be G', in John Hawthorne and Jason Turner (eds.), *Philosophical Perspectives 30: Metaphysics* (Oxford: Blackwell): 1–97.

Dorr, Cian (unpublished), 'The Multiplicity of Meaning', The Hempel Lectures, delivered 2020.

Dorr, Cian, Hawthorne, John, and Yli-Vakkuri, Juhani (2021), *The Bounds of Possibility: Puzzles of Modal Variation* (Oxford: Oxford University Press).

Frege, Gottlob (1879), 'Begriffsschrift: A Formula Language, Modeled Upon That of Arithmetic, for Pure Thought', in *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, ed. and trans. Jean van Heijenoort (Cambridge, MA: Harvard University Press, 1967): 1–82 [Ger. orig. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (Halle: Verlag von Louis Nebert, 1879)].

Frege, Gottlob (1892), 'On Concept and Object', trans. P. T. Geach and Max Black, *Mind*, 60 (1951): 168–80 [Ger. orig. 'Über Begriff und Gegenstand', *Vierteljahrsschrift für wissenschaftliche Philosophie*, 16 (1892): 192–205].

Frege, Gottlob (1893), *Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet* (H. Pohle).

Fritz, Peter, and Jones, Nicholas K. (2024) (eds.), *Higher-Order Metaphysics* (Oxford: Oxford University Press).

Goodman, Jeremy (2024), 'Higher-order Logic as Metaphysics', in Peter Fritz and Nicholas K. Jones (eds.), *Higher-Order Metaphysics* (Oxford: Oxford University Press).

Halbach, Volker (2011), *Axiomatic Theories of Truth* (Oxford: Oxford University Press).

Harris, John H. (1982), 'What's so Logical about the 'Logical' Axioms?', *Studia Logica*, 41: 159–71.

Heim, Irene, and Kratzer, Angelika (1998), *Semantics in Generative Grammar* (Malden, MA: Blackwell).

Hofweber, Thomas (2022), 'The Case Against Higher-Order Metaphysics', *Metaphysics*, 1/5: 29–50.

Klein, Ewan, and Sag, Ivan A. (1985), 'Type-Driven Translation', *Linguistics and Philosophy*, 8/2: 163–201.

Kripke, Saul (1979), 'A Puzzle About Belief', in Avishai Margalit (ed.), *Meaning and Use* (Dordrecht: Reidel): 239–83.

Kubota, Yusuke, and Levine, Robert D. (2020), *Type-Logical Syntax* (Cambridge, MA: MIT Press), `https://direct.mit.edu/books/book/4931/Type-Logical-Syntax`.

Lambek, Joachim (1958), 'The Mathematics of Sentence Structure', *American Mathematical Monthly*, 65/3: 154–70.

Lewis, David (1983*a*), 'General Semantics', in David K. Lewis, *Philosophical Papers*, i (Oxford: Oxford University Press): 189–232. From Lewis, David K., *Synthese*, 22 (1970): 18–67.

Lewis, David K. (1983*b*), *Philosophical Papers* (Oxford: Oxford University Press).

Liebesman, David (2011), 'Simple Generics', *Noûs*, 45/3: 409–42. doi: `10.1111/j.1468-0068.2010.00774.x`.

Liebesman, David (2015), 'Predication as Ascription', *Mind*, 124: 517–69.

Liebesman, David, and Magidor, Ofra (MS), *Copredication and Property Versatility*.

Link, Godehard (1983), 'The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach', in Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow (eds.), *Meaning, Use, and Interpretation of Language* (Berlin: De Gruyter): 303–23.

Magidor, Ofra (2013), *Category Mistakes* (Oxford: Oxford University Press).

Menzel, Christopher (1986), *A Complete, Type-free "Second-order" Logic and Its Philosophical Foundations* (CSLI-86-40, Center for the Study of Language and Information).

Milward, David (1994), 'Non-Constituent Coordination: Theory and Practice', in *COLING 1994 Volume 2: The 15th International Conference on Computational Linguistics*, `https://aclanthology.org/C94-2151`.

Moltmann, Friederike (2013), *Abstract Objects and the Semantics of Natural Language* (Oxford: Oxford University Press).

Moltmann, Friederike (forthcoming), 'Reference to Properties in Natural Language', in Anna Sofia Maurin and Anthony Fisher (eds.), *The Routledge Handbook of Properties* (London: Routledge).

Montague, Richard (1974), 'The Proper Treatment of Quantification in Ordinary English: Selected Papers of Richard Montague', in Richmond H. Thomason (ed.), *Formal Philosophy* (New Haven, CT: Yale University Press): 247–70. From *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, ed. Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes (Dordrecht: D. Reidel, 1973): 221–42.

Moortgat, Michael (1996), 'Generalized Quantifiers and Discontinuous Type constructors', *Natural Language Processing*, 6: 181–208.

Morrill, Glyn (1994), *Type Logical Grammar: Categorial Logic of Signs* (Dordrecht: Kluwer).

Parsons, Terence (1979), 'Type Theory and Ordinary Language', in Steven Davis and Marianne Mithun (eds.), *Linguistics, Philosophy, and Montague Grammar* (Austin: University of Texas Press): 127–51.

Parsons, Terence (1990), *Events in the Semantics of English: A Study in Subatomic Semantics* (Cambridge, MA: MIT Press).

Partee, Barbara, and Rooth, Mats (1983), 'Generalized Conjunction and Type Ambiguity', in Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow (eds.), *Meaning, Use, and Interpretation of Language* (Berlin: De Gruyter).

Prawitz, Dag (1965), *Natural Deduction: A Proof-Theoretical Study* (Stockholm: Almqvist and Wiksell).

Prior, A. N. (1971), *Objects of Thought*, ed. P. T. Geach and A. J. P. Kenny (Oxford: Oxford University Press).

Quine, Willard van Orman (1953), 'On What There Is', in *From a Logical Point of View* (2nd edn., Cambridge, MA: Harvard University Press): 1–19. From *Review of Metaphysics*, 2 (1948): 21–38.

Quine, Willard van Orman (1956), 'Quantifiers and Propositional Attitudes', *Journal of Philosophy*, 53: 177–87.

Quine, Willard van Orman (1970), *Philosophy of Logic* (2nd edn., Cambridge, MA: Harvard University Press, 1986).

Rayo, Agustín, and Uzquiano, Gabriel (2007) (eds.), *Absolute Generality* (Oxford: Oxford University Press).

Rayo, Augustin, and Yablo, Stephen (2001), 'Nominalism through De-Nominalisation', *Noûs*, 35: 74–92.

Rieppel, Michael (2016), 'Being Something: Properties and Predicative Quantification', *Mind*, 125/499: 643–89. doi: `10.1093/mind/fzv173`.

Rooth, Mats, and Partee, Barbara (1982), 'Conjunction, Type Ambiguity and Wide Scope Or', in *Proceedings of the First West Coast Conference on Formal Linguistics*, i, ed. Daniel P. Flickenger, Marlys Macken, and Nancy Wiegand, Stanford Linguistics Association Stanford.

Rosefeldt, Tobias (2008), ''That'-clauses and Non-nominal Quantification', *Philosophical Studies*, 137: 301–33.

Russell, Bertrand (1908), 'Mathematical Logic as Based on the Theory of Types', *American Journal of Mathematics*, 30: 222–62.

Salmon, Nathan U. (1986), *Frege's Puzzle* (Cambridge, MA: MIT Press).

Stalnaker, Robert C. (1977), 'Complex Predicates', *The Monist*, 60: 327–39.

Szabolsci, Anna (1987), 'Bound Variables in Syntax (Are There Any?)', in *Sixth Amsterdam Colloquium Proceedings*, ed. Jeroen Groenendijk, Frank Veltman, and Martin Stokhof, University of Amsterdam.

Trueman, Robert (2021), *Properties and Propositions: The Metaphysics of Higher-Order Logic* (Cambridge: Cambridge University Press).

van Benthem, Johan (1988), 'The Lambek Calculus', in Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler (eds.), *Categorial Grammars and Natural Language Structures* (Studies in Linguistics and Philosophy; Dordrecht: Springer): 35–68.

Vendler, Zeno (1967), *Linguistics in Philosophy* (Ithaca: Cornell University Press).

Williamson, Timothy (2003), 'Everything', in John Hawthorne and Dean Zimmerman (eds.), *Philosophical Perspectives 17: Language and Philosophical Linguistics* (Oxford: Blackwell): 415–65.

Worth, Chris (2014), 'The Phenogrammar of Coordination', in *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics* (*TTNLS*): 28–36, `https://aclanthology.org/W14-14.pdf`.

Wright, Crispin (1998), 'Why Frege Does Not Deserve His Grain of Salt: A Note on the Paradox of 'The Concept Horse' and the Ascription of *Bedeutungen* to Predicates', in Bob Hale and Crispin Wright, *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics* (Oxford: Oxford University Press, 2001): 72–90. From 'Why Frege Does Not Deserve his *Granum Salis*: A Note on the Paradox of 'The Concept Horse' and the Ascription of *Bedeutungen* to Predicates', J. Brandl and P. Sullivan (eds.), *New Essays on the Philosophy of Michael Dummett* (Grazer Philosophische Studien, 55; Amsterdam/Atlanta: Rodopi, 1998): 239–63.