A Philosophical Introduction to Higher-order Logics

Andrew Bacon

February 3, 2023

Contents

Nomenclature

Some common symbols:

H, C, etc.	Names for particular higher-order logics, the minimal system H, Classicism, etc.
$\mathbf{L}, \mathbf{L}', ...$	Name for an arbitrary logic
M, N, P, Q, a, b, c	Terms of arbitrary type, lower case reserved for terms in argument position
A, B, C	Terms of type t
X, Y, Z, x, y, z, w	Variables, lower case reserved for variables appearing exclusively in argument position
$\sigma, \tau, \rho, ...$	Types
Σ	A name for a signature
Λ	The logical signature of higher-order logic
$\mathcal{L}(\Sigma)$	The full λ-language in the signature Σ
$\mathcal{J}(\Sigma)$	An arbitrary λ-language in the signature Σ
$M \sim_{\beta\eta} N$	M and N are $\beta\eta$ equivalent
$CL[\Sigma, \{B, C, K\}]$, etc.	The combinatory language in signature Σ with combinators B, C and K
$\mathbf{ND}[PCW]$, etc.	The natural deduction Curry system containing the rules P, C, and W
$\mathcal{A}, \mathcal{B}, \mathcal{C}...$, $\mathcal{A}^\sigma, \mathcal{B}^\sigma, \mathcal{C}^\sigma...$	Names for applicative structures, names for domains of type σ
$\tilde{f}$	The counterpart of the function f in an applicative structure
$\mathbf{M}, \mathbf{N}, ...$	Models of higher-order languages
$g, h, ...$	Variables assignments
$[\![M]\!]^g$	The denotation of a term M relative to an assignment in a given model
$R, S, T, ...$, $R^\sigma, S^\sigma, T^\sigma...$	Names for logical relations or Kripke logical relations, names for logical or Kripke logical relation at a given type

Some common abbreviations:

$M:\sigma$	'M is a term of type σ', 'M, of type σ,'
$MN_1N_2...N_k$	$(...((MN_1)N_2)...N_k)$
$\lambda x_1x_2...x_n.M$	$\lambda x_1.(\lambda x_2.(...\lambda x_n.M))...)$
$\sigma_1 \to \sigma_2 \to ... \to \sigma_n \to \tau$	$(\sigma_1 \to (\sigma_2 \to (... \to (\sigma_n \to \tau)...)))$
$\forall_\sigma x.A$	$\forall_\sigma(\lambda x.A)$
$A \wedge B$, $a =_\sigma b$, etc	$((\wedge A)B)$, $((=_\sigma a)b)$
$\bar{x}$, $\bar{a}$, $\bar{\sigma}$	$x_1,...,x_n$, $a_1,...,a_n$, $\sigma_1,...,\sigma_n$
$(\bar{\sigma} \to \tau)$	$\sigma_1 \to \sigma_2 \to ... \to \sigma_n \to \tau$
$\lambda\bar{x}.M$, $M\bar{x}$	$\lambda x_1...\lambda x_n.M$, $Mx_1...x_n$
$\bar{x}:\bar{\sigma}$:	$x_1:\sigma$, $x_2:\sigma_2...$ $x_n:\sigma_n$
$\forall\bar{x}.A$	$\forall_{\sigma_1}x_1...\forall_{\sigma_n}x_n.A$ where $\bar{x}:\bar{\sigma}$
$(MN)^n_m$	$\lambda\bar{x}\bar{y}.M\bar{x}(N\bar{y})$ where $\bar{x} = x_1...x_m$, $\bar{y} = y_1...y_n$, $\bar{x}:\bar{\sigma}, \bar{y}:\bar{\rho}, N:\bar{\rho} \to \tau, M:\bar{\sigma} \to \tau \to \theta$
$\wedge_{\bar{\sigma}\to t}, \neg_{\bar{\sigma}\to t}$, etc.	$\lambda XY\bar{x}\bar{y}(X\bar{x} \wedge Y\bar{y})$, $\lambda X\bar{y}.\neg(X\bar{y})$, etc. where $\bar{x}, \bar{y}:\bar{\sigma}$
I^σ	$\lambda x.x$ where $x:\sigma$
$S^{\sigma\tau\rho}$	$\lambda XYz.Xz(Yz)$ where $X:\sigma \to \tau \to \rho$, $Y:\sigma \to \tau$, $z:\sigma$
$K^{\sigma\tau}$	$\lambda xy.x$ where $x:\sigma, y:\tau$
$B^{\sigma\tau\rho}, B'^{\sigma\tau\rho}$	$\lambda XYz.X(Yz)$, $\lambda YXz.X(Yz)$ where $Y:\sigma \to \tau, X:\tau \to \rho, z:\sigma$
$C^{\sigma\tau\rho}$	$\lambda Xyz.Xzy$ where $X:\sigma \to \tau \to \rho, z:\sigma, y:\tau$
$W^{\sigma\tau}$	$\lambda Xy.Xyy$ where $X:\sigma \to \sigma \to \tau, y:\sigma$
$\Box_\top$	$\lambda p.p =_t \top$

Preface

I was motivated to write this book after struggling myself to find the right tools for my purposes in existing philosophical texts. I found, increasingly, that I was reinventing wheels that had previously been invented by computer scientists with completely different applications in mind. These include several logical tools described in this book; for instance, substructural type theory, the concept of a logical relation, and various ideas from category theory. However, because these wheels have been designed for different cars (as it were), and the expository texts written for a different audience, there is no resource which a philosopher can simply consult in order to learn about them in a way that transparently relates them to philosophical concerns. My hope is this book will fill that lacuna.

There are some ways in which this book resembles a textbook, and other ways in which it resembles a monograph. Like a textbook it aims to impart to the reader certain logical tools that I believe to have a great number of applications in philosophy, specifically metaphysics. This represents a majority of the book. On the other hand, I have made no effort to be comprehensive in my coverage of those applications: I have simply taken two topics—modal metaphysics (chapters 7 and 8), and metaphysical structure (chapters 11-13)—that have personally captivated my interest, and to which the tools presented in this book have helped me get things straight in my own head. These cases, I hope, illustrate the power of higher-order logic both as a language for formulating important claims in metaphysics, and as a framework for investigating those questions. But I have followed my own inclinations in deciding what to explore, and this part of the book is in no way representative of the full variety of possible applications, or indeed, existing applications of these tools in the burgeoning literature on higher-order metaphysics. Some of these omissions are mentioned briefly below.[1] My hope is that the reader will be able to take the apparatus in this book and apply them to whatever questions captivate their interest.

A couple of brief remarks on the title of the book are in order. Perhaps a more accurate (but less catchy) name would have been *A Philosophical Introduction to Higher-order Logics and λ-Calculi*. A substantial portion of the book is

[1] The reader interested in the broader field of higher-order metaphysics might wish to consult the recent anthology Fritz and Jones (forthcoming).

devoted to simply typed λ-languages: a very general class of languages that includes many logical languages, such as propositional, first-order and higher-order logic, and even non-logical languages such as programming languages. Pedagogically, the relationship between the λ-calculus and higher-order logic is a bit like the relationship between propositional logic and first-order logic: very few philosophically interesting theories can be formulated in propositional logic alone, but one needs to become reasonably fluent in it before learning first-order logic. I believe the λ-calculus stands in a similar relationship to higher-order logic, and should be studied first.

The other remark regarding the title concerns the use of a plural noun. There is an important difference between propositional and first-order logic on the one hand, and higher-order logics on the other. One cannot consistently extend classical propositional logic with further logical axiom schemas, and the only logical principles one can consistently add to first-order logic with identity make fairly uninteresting statements about how many different things there are. By contrast there are many different ways to consistently extend classical higher-order logic with further purely logical axioms.[2] Although higher-order logics are neutral on questions of mathematical *ontology*—they are formulated without reference to primitive mathematical notions, like number and set membership—some of these logical axioms seem mathematical in nature. There are purely logical statements you can make that settle the continuum hypothesis in the sense that they would imply the continuum hypothesis if there *were* any sets.[3] Similarly, while higher-order logics are neutral about the ontology of propositions, conceived of as certain kinds of abstract individuals, one can formulate purely logical statements that make structural claims about reality: claims that imply that propositions would be structured, if there were any abstract entities that bore an appropriately close relationship to reality (statements like these are explored in chapters 11–13). One can similarly make purely logical statements that correspond in this way to the existence or non-existence of maximally specific propositions, 'possible worlds', or statements that imply that modal reality admits a lot of different possibilities, or not very many at all, or statements that contradict the $\mathsf{S5}$ axiom for a suitably broad kind of necessity (statements like these are explored in chapter 8). The line between logic, mathematics and metaphysics becomes somewhat blurry: logic can constrain mathematical and metaphysical theorizing in new and non-obvious ways. The wider point here is that there are a great many interesting higher-order logics, and so the subject should be studied with this in mind: it is not the study of a single system, like first-order logic, but rather of a class of interestingly different systems. In this

[2]'Purely logical' here means stateable using only logical words (or logical words with the addition of schematic constants) – in this case the truth-functional connectives and the first and higher-order quantifiers.

[3]In higher-order logic one can formulate a general claim that says that any relation satisfying the principles of Zermelo-Fraenkle set theory when taking the place of the membership relation is one in which the continuum hypothesis holds. Given the non-logical assumption that $\in$ — i.e. *set membership* — satisfies the axioms of Zermelo-Fraenkle set theory, we can infer that the continuum hypothesis is true. See Shapiro (1991) p105 for another formulation of a CH-like principle of higher-order logic.

sense the study of higher-order logic is much more like the study of modal logic.

Part I of the book concerns typed languages. Most languages we are familiar with, such as the language of propositional logic, first-order logic, propositional modal logic, and so on, are implicitly typed languages. There are rules about which expressions can be combined with which — one can apply a predicate to an individual constant, but not to a sentence or another predicate, whereas one can apply operators to sentences, but not to predicates, individual constants, other operators, and so on. We also know how to introduce new devices that behave in grammatically novel ways: for instance we could easily add to first-order logic predicate modifiers, that combine with predicates to make other predicates; then we could introduce predicate modifier modifiers that combine with predicate modifiers to make new predicate modifiers, and so on. Type theory systematizes these rules and a typed language is simply a language which fits this schema (chapter 1). There is a particularly important device, λ, that guarantees that there is a predicate corresponding to any open formula parametrized by a variable x, and which makes similar guarantees for expressions of other types. Chapters 2 and 3 introduce the simply typed λ-calculus: a theory concerning typed languages containing the λ device, governed by two central principles β and η. The latter chapter also outlines relations between the λ-calculus and a certain variable-free alternative to it called combinatory logic.

Part II of the book concerns a certain kind of typed language that contains, for each grammatical type, devices that stand to that type as the first-order order quantifiers stand to individual terms. They can bind variables of that type in the same way that the first-order quantifiers bind individual variables, and they are subject to logical laws that are completely parallel to the first-order universal and existential quantifiers. We thus call them 'higher-order quantifiers'. Higher-order quantifiers let us express generality in sentence position, predicate position, operator position, and so on, in the same way that the first-order quantifiers express generality into name position. Chapter 4 outlines these quantificational devices in some detail, and in chapter 5 the key concept of a higher-order logic is introduced. Chapter 6 applies higher-order logic to current theorizing about the granularity of propositions, properties and relations. Chapter 7 and 8 are an extended investigation of the higher-order logic Classicism, introduced in chapter 6, in application to questions in modal metaphysics: In higher-order logic it is possible to define the analogue in reality of an operator expression with a normal modal logic, and many other notions from modal metaphysics—'necessity in the highest degree', entailment, possible world—have a claim to being reducible to pure logic.

Part III of the book we consider weakening the λ formalism. The full λ-calculus is ontologically committal: it contains binary predicate expressions that are converses of other predicates, committing us in the background logic to the existence of converse relations. This is not a first-order ontological commitment but a higher-order one. λ-languages contain many other types of expressions that have ontologically contentious implications. Chapters 9 and 10 explore general λ-calculi that do not have these expressions. Chapters 11–13 consist of extended applications of the machinery to the question of the structure of propositions, properties and relations. Various ideas about the structure of reality discussed in the philosophical literature are formalized in higher-order languages, and some limitative results are presented.

Part IV concerns the model theory of higher-order logics. Chapters 14 and 15 introduce the notion of a model for an arbitrary λ-language and the logical language of higher-order logic respectively. Chapter 16 introduces the key notion of a *logical relation*, an important model theoretic tool which can be used to construct partial quotients of models and partial homomorphisms between models, as well as to establish definability and undefinability results. Chapter 17 introduces concepts from category theory that are useful in the study of type theory, and explores three concrete categories – the category of sets, the category of modalized sets and the category of M-sets – which have particular applications to the study of higher-order logic. Chapter 18 investigates the model theory of Classicism in particular and outlines some model theoretic constructions that can be wielded to settle the consistency of several higher-order logics discussed in chapter 8. The book ends with two appendices on topics relating to part III that are slightly off the main track, but which connect the syntax and semantics of λ-languages with (seemingly disparate) research on non-classical propositional logics.

Many people have made this book possible. My own interest in higher-order logic extends as long as my education in philosophy: it began in my first year as an undergraduate, where I learned about Frege's philosophy of mathematics from Ian Rumfitt, a topic that I continued to pursue under the supervision of Gabriel Uzquiano as a BPhil student. These early teachers instigated my initial fascination with the topic, and helped me appreciate the power of higher-order languages for philosophical theorizing. The catalyst for my present interest in higher-order logic, however, has been from conversations with Mike Caie, Cian Dorr, Peter Fritz, Jeremy Goodman and Harvey Lederman, taking place between 2015 to 2023, concerning applications of higher-order logic to metaphysics. I have learned a great deal from these conversations, and without this stimulus I would not have found my way into the subject to the extent that I presently have.

During the writing of this book, I benefited greatly from a number of people who spotted various mistakes in early drafts of the book, including David Ripley, Chris Scambler, Lingzhi Shi, Shawn Standefer, Juhani Yli-Vakkuri, Arthur Wu and especially Cian Dorr for checking over the final chapter, and Helen Bacon for catching a number of typos. I am also grateful to an anonymous referee for Routledge who, among other things, made some very helpful structural suggestions about the organization of the book. I owe a special debt of gratitude to Jin Zeng, who read an early draft of the book and helped me fix many important errors and inconsistencies.

Introduction

If you are reading this book, you have likely already heard something about its subject matter. Perhaps you have read that higher-order logics are generalizations of first-order logic, or perhaps you have come across some of their applications in philosophy or other disciplines. The aim of this book is to introduce the reader to higher-order languages, their proof theory, model theory, and some of their applications in philosophy. However there is also a distinctively higher-order way of *thinking* about philosophical questions, which is harder to impart, and which goes beyond simple proficiency with these logical tools: it is possible to become well-versed in the higher-order symbolism without having a proper grasp on what the symbolism *means*, or why we are studying it. This informal chapter is intended to introduce readers to typed languages and higher-order generalizations in an casual and imprecise way, give a sense of why they are useful, as well as to introduce some of the main themes of the book along the way.

What is a typed language? §0.1

What are higher-order generalizations? (§0.2 and §0.3)

What is abstraction? (§0.4)

What is the difference between higher-order generalizations and first-order generalizations over properties? (§0.5)

Why are higher-order generalizations useful for doing philosophy? (§0.6)

What is the point of providing a (set-theoretic) model theory for a higher-order language? (§0.7)

Can we translate higher-order sentences into ordinary English? (§0.8)

How should one read this book? (§0.9)

0.1 Typed Languages

When a student is first introduced to logic they will encounter various different *types* of logical expressions. Sentences are fundamental to logic, because they are used to state things and are the premises and conclusions of arguments. But sentences have logical structure—they are composed of subsentential expressions—in virtue of which arguments are valid or invalid. In propositional languages, we encounter devices that can combine with sentences to make other complex sentences, such as the truth-functional connectives, $\wedge, \vee, \neg$, and any other connectives we might add to a propositional language, such as modal or counterfactual connectives. In first-order languages we encounter, in addition to these devices, names—expressions used to *name* things—and predicates that may be combined with names to make sentences. It is possible to think of the universal and existential quantifiers of first-order logic as something that combines with an ordinary predicate to make a sentence: the simplest sorts of quantified statements involve a quantifier combining with a predicate to make a universal or existential statement (e.g. 'somethings smells' = 'something' + 'smells'). On this analysis, quantifiers are *higher-order* predicates.

But why stop here? Just as connectives and first-order predicates can have different arities—$\wedge$ is binary, because it needs two sentences to make a sentence, whereas $\neg$ is unary (it is an 'operator') because it needs only one—we could consider languages that have higher-order predicates that combine with multiple first-order predicates to make sentences. Logicians have studied extensions of first-order logic with binary quantifiers like *most*, and *the*, combining with two predicates, F and G, to make a sentence (e.g. *most Fs are G*). We can also imagine expressions that are mixed between connectives and predicates—*connecticates*—that take a name and a sentence to make a sentence. When studying the logic of knowledge for multiple agents it's natural to theorize with a primitive device meaning 'knows that' taking a name in the first argument, and a sentence in the second.[4] From here the possibilities are boundless. One can similarly imagine adding devices that combine with connectives to make sentences, or with predicates to make predicates, or with any of the devices described above to make any other device described above.

The languages described above are all instances of *typed* languages, which is the topic of part I of this book. Expressions of these languages are divided up into different grammatical categories, *types*, and there are rules concerning how expressions of different types can be combined—you cannot, for instance, combine a name with a binary connective. Almost all of the formal languages you are likely to have encountered already will be typed languages—from the simplest, the language of propositional logic, to any of the richer languages described above. Natural languages are also typed languages—expressions of natural languages belong to different grammatical categories and there are rules about how these expressions can be combined grammatically. But things were not always this way. The two founders of modern logic, Gottlob Frege and Giuseppe Peano, took very different approaches to logical languages: while Frege's system was rigidly typed, Peano had no grammatical distinction like that between predicates, sentences and names—one could put any expression next to any other.[5]

[4] See Prior (1971) p135.

[5] Frege (1879), Peano (1889). We'll return to these systems briefly in section 1.4.

For a while both traditions developed in parallel, with authors like Russell, Hilbert and his colleagues, and later Church following Frege and theorizing in typed languages, and Schönfinkel, Curry, and Fitch following Peano. But ultimately the former approach won out, and research in logic today is primarily conducted in a typed languages.[6]

While the typed framework is, in this general sense, quite orthodox, the most familiar typed languages (first-order languages) do not make use of expressions that go very far up the hierarchy of typed expressions. Yet this extra generality seems to be very useful for philosophical theorizing. We can illustrate this with a couple of examples. In the philosophy of modality one typically distinguishes operator expressions that are used to express modality, like such as 'it's necessary that', 'it *has* to be true that' and so on, from those that do not, such as 'it's not the case that' or 'John said that'. But this is not just a distinction in language — many philosophers believe there is a corresponding distinction in reality. In order to express these metaphysical distinctions in the object language it is therefore necessary to have a device, 'Nec', that combines with an operator to form a sentence. If we had operators $\neg$ and $\Box$ for negation and a particular kind necessity we could the form claims like $\neg(\mathrm{Nec}\,\neg)$ and $\mathrm{Nec}\,\Box$ to say that the latter but not the former is a kind of necessity. Another example: metaphysicians will often talk about what vocabulary is fundamental or 'joint carving' (see also 'perfectly natural' in Lewis (1983) and 'structural' in Sider (2011)). It would be prejudiced to assume prior to inquiry that, say, only first-order predicates could be fundamental or not.[7] These distinctions must be applicable to words of many different grammatical categories: perhaps the predicates of physics (e.g. 'is an electron') are fundamental, or the tense operators are fundamental, or the quantifiers. But, as before, this distinction between expressions in language is to be explained in terms of more basic distinctions in reality, and to express these we need higher-order predicates. Let E be a first-order predicate meaning 'is an electron'. We need a higher-order predicate $\mathrm{Fun}_{\mathrm{pred}}$ that combines with a first-order predicate to make a sentence in order to state the relevant claim about fundamentality, $\mathrm{Fun}_{\mathrm{pred}}\,E$. In order to make an analogous fundamentality claim about necessity, $\mathrm{Fun}_{\mathrm{op}}\,\Box$ we need a higher-order predicate $\mathrm{Fun}_{\mathrm{op}}$ that can combine with an operator to make a sentence. And once we have these higher-order fundamentality predicates we can ask whether they themselves are fundamental, and the analogous questions about reality will require higher-order predicates that combine with predicates of the type of $\mathrm{Fun}_{\mathrm{pred}}$ and $\mathrm{Fun}_{\mathrm{op}}$ to make sentences. Further questions can be asked about the fundamentality of these new predicates, and so higher-order

[6]There are several possible explanations for this. One possible reason is that early untyped logical languages were found to be inconsistent. But I think a more fundamental reason is that they are quite hard to interpret; as we have noted already, natural languages are typed languages and so it is not always possible to translate an expression of an untyped language into an antecedently understood sentence of a natural language. If we explain that the symbol F means *walks* and $\neg$ means *it's not the case that*, then expressions like $F\neg$ don't seem to correspond to anything comprehensible.

[7]Cf Sider (2011) chapter 6, Dorr and Hawthorne (2013).

predicates of ‘arbitrary order’ are needed.[8]

0.2 Generalizations

Apart from the sorts of expressions they admit, the chief difference between propositional languages and first-order languages—the feature that makes first-order languages more *expressive*—is that in the latter one can express generalizations. Specifically, one can express generality in name position. This is a book about higher-order languages, the things we can say in them, and how to reason in and about them. The chief expressive advantage of higher-order languages over first-order languages is simply that they can express more kinds of generalizations: they can express generality in the position of sentences, predicates, connectives, and so on. They can express generality in any grammatical position whatsoever.

In order to understand what this means, we will begin examining ordinary first-order generality in name position. Consider a simple subject-predicate sentence, such as

Socrates is wise.

This would be formalized as Fa. By using the first-order quantifiers we can express generality in the position that ‘Socrates’ occupies by replacing this name with a first-order variable and prefixing with the universal or existential quantifier.

$\forall x$, x is wise.

$\exists x$, x is wise.

For now, we adopt an informal hybrid of English and logical notation. In ordinary English these sentences mean, approximately, ‘everything is wise’ and ‘something is wise’ respectively.

The sense in which these claims express generality in name position is that they bear a special logical relationship to their *instances*. First we identify the position we are generalizing into (the position of the name ‘Socrates’ above) and we form a sentence that has a gap in this position — i.e. that is variable in that position

... is wise

This defines a class of sentences parameterized by the different names we can substitute into the gap. We might call this a ‘sentential function’, for it defines a class of sentences parameterized by the name we plug into the gap. (Early practitioners of higher-order logic, who were not very careful about use and mention, would call this a ‘propositional function’, but they meant the same thing.) These are the instances of the generalization:

[8] See Sider (2011) §7.13.

Socrates is wise, Plato is wise, Aristotle is wise, (...and so on)

The logical relationship is this. From the universal claim any instance can be inferred: from the claim that $\forall x$, x is wise ('everything is wise') one can infer that Socrates is wise, for example. This is often called the 'universal elimination rule'. Conversely, if you can prove an "arbitrary" instance—lets say you can prove that Socrates is wise without making assumptions specific to Socrates—then that reasoning would apply just as well to any other individual, and we may infer the universal claim $\forall x$, x is wise. This is often called the 'universal introduction rule'. A similar pair of elimination and introduction rules characterize the existential quantifier by its relationship to its instances.

The universal and existential sentences do not mention particular individuals by name, but rather generalize from sentences naming particular individuals—'Socrates is wise', 'Aristotle is wise', and so on. It is clear enough, then, what it means to say that the first-order quantifiers express generality with respect to name position, for it is names—not predicates, sentences, or what have you—that vary between the instances.

There are several ways to get a handle on the quantifiers of first-order logic. An approximation, utilized above, is to paraphrase them into English using quantificational idioms like 'everything' and 'something'. But these are imperfect because we often—indeed, some would argue *always*—understand these quantifiers as restricted in some way or other. These restricted uses of the words 'something' and 'everything' are not inferentially connected to their instances in the right way to express the universal and existential generalizations of first-order logic. The professor can say, in an introductory logic course, 'everything will be on the exam' and say something true; it does not therefore follow that Gödel's incompleteness theorems will be on the exam—the professors generalization was restricted to topics covered in this introductory logic class. So the restricted universal does not imply all of its instances—it is not completely *general*.

By contrast, we will understand the universal generalization of first-order logic as being completely general—from a true universal generalization all instances must follow. Another way into understanding generalizations is by their logical relationship to their instances. Someone who already understands the quantifier-free fragment of a first-order language—the sentences that do not contain the symbols '$\forall$' and '$\exists$'—may extend their understanding to the fully quantified language using the rules we described above. These rules in fact *pin down* the quantifiers up to logical equivalence. We'll record this with the following informal theorem (we will have the tools to offer a more precise version of this later in this book).

Theorem 0.2.1 (Harris's theorem for name generalizations)**.** *Suppose $\forall_1$ and $\forall_2$ are two generalizing devices, in the sense that they both satisfy the universal introduction and elimination principles. Consider a sentence of the form ...x..., where the variable x takes the position that a name could ordinarily occupy. Then '$\forall_1 x$...x...' and '$\forall_2 x$...x...' are logically equivalent in the sense that they can be inferred from one another.*

We can illustrate the style of argument with our example. Given $\forall_1 x$, x is wise, we can infer that Socrates is wise, using that fact that we can infer any instance from a general claim (and that $\forall_1$ expresses general claims). Since since we showed that Socrates is wise from assumptions that did not depend on any feature specific to Socrates—it would have worked equally well for an *arbitrary* individual—we can infer $\forall_2 x$, x is wise. The inference from $\forall_2 x$, x is wise to $\forall_1 x$, x is wise can be demonstrated in exactly the same way.[9] A parallel argument is available for the existential quantifiers.

Remark 0.2.1. Some philosophers maintain that there is no use of the word 'everything' and 'something' for which the move from 'everything is F' to 'a is F', and from 'a is F' to 'something is F' are legitimate. Sometimes this stems from the view that all uses of 'everything' and 'something' are restricted—no matter what we do to prime the hearer to a wide reading of these terms, there is always a wider reading available.[10] In this case these inferences fail in the same way that we cannot move from 'everything will be on the exam' to 'Gödel's incompleteness theorems will be on the exam' in our earlier example. Others maintain there is a widest use of 'everything' and 'something', but that we can make true assertions involving fictional names, like 'the ancient Greeks worshiped Zeus' without being able to infer 'the ancient Greeks worshiped something'. These philosophers will instead adopt a weaker-than-classical logic—a 'free logic'—for reasoning about the quantificational words.

These views are usually being motivated by the behaviour of the quantificational idioms of natural language. By contrast we have characterized the symbols $\forall$ and $\exists$ by their inferential role, leaving it open how exactly they relate to the quantificational idioms of natural language. (Of course, practitioners of free logic will often use the same symbols, $\forall$ and $\exists$, in the formal study of these logics, but this is a purely notational point.) It is, of course, a substantive theoretical posit that *any* device satisfies this logical role. Nonetheless, it is a posit that earns its keep: the ability to form unrestricted generalizations is incredibly useful for theoretical disciplines, such as logic and metaphysics, that require a certain level of generality rarely needed in ordinary conversation. Furthermore, nothing the free logicians say contradict this posit: it is one thing to claim, for instance, that the inference from 'the ancient Greeks worshiped Zeus' to 'the ancient Greeks worshiped something' is no good, it is another to claim that there is no generalizing device we could substitute for 'something' which would make the inference OK. On the contrary, it is very hard to say something that contradicts our theoretical posit: the claim 'all uses of 'everything' are restricted' is no good, for to be restricted is usually spelled out in terms of unrestricted quantification. To be restricted is for there to be something (in the unrestricted sense) not in the range of that quantifier. One cannot use a restricted quantifier to spell the notion out either, for every restricted quantifier

[9] The reader should consult Harris (1982) for more discussion, and a more precise formulation of this argument.

[10] Philosophers have been moved to this picture from a variety of motivations, from making sense of ontological disputes, to the paradoxes of set theory. Many of these motivations are discussed in the collection Rayo and Uzquiano (2006).

is unrestricted by its own lights.[11]

0.3 Higher-order generalizations

We have explained how the first-order universal and existential quantifiers, $\forall x$ and $\exists x$, express generality in name position. We might, by analogy, introduce a pair of devices—predicate quantifiers—that let us express generality in predicate position. This time we express generality in the position that 'is wise' occupies in 'Socrates is wise'. So this time we replace this predicate with a 'predicate variable', X, and prefix the result with a universal or existential quantifier to create a generalization in predicate position.

> $\forall X$, Socrates Xs.
>
> $\exists X$, Socrates Xs.

The sense in which these claims express generality in predicate position is completely analogous. They bear a special logical relationship to their *instances*. As before we identify the position we are generalizing into (the position of the predicate 'is wise' above) and we form an sentence that has a gap in this position — i.e. a sentence that is variable with respect to that predicate:

> Socrates ...s

This defines a class of sentences parameterized by the different predicates we can substitute into the gap, and these constitute the instances of the predicate generalization.

> Socrates is wise, Socrates talks, Socrates is old, (...and so on)

The logical relationship is also exactly the analogous. From the universal claim any instance can be inferred: from the claim that $\forall X$, Socrates Xs one can infer that Socrates talks. Conversely, if you can prove an 'arbitrary' instance—lets say you can prove that Socrates talks without making assumptions specific to talking—then that reasoning would apply just as well to any other instance, and we may infer the universal claim $\forall X$, Socrates Xs. The elimination and introduction rules for the predicate existential quantifier are also completely analogous to the existential name quantifier.

Here we use the same symbols, $\forall$ and $\exists$, to express generality in predicate position as we did for generalizing into name position, to emphasize the analogy in their logical behaviour. However, they are not the same devices: in fact, on one way of regimenting them, both sorts of quantifiers belong to grammatical categories, indeed belong to *different* grammatical categories, and so are as different as a predicate is different from a name.

We do not have ready approximations of these generalizations in English—not even the imperfect context sensitive approximations we had for first-order

[11] These sorts of worries are spelled out further in Williamson (2003).

generalizations.[12] But that is OK. The handle we have on these generalizations does not have to go by way of first translating them into an already understood natural language—after all, a child comes at some point to understand first-order quantificational phrases without already having a natural language to translate them into. We can come to understand predicate generalizations in exactly the same way that we came to grips with generality in name position above, namely by their logical relationship to sentences of the language that do not have predicate quantifiers. As with generalizations in name position these predicate generalizations are pinned down, up to logical equivalence, by these relationships.

Theorem 0.3.1 (Harris's theorem for predicate generalizations)**.** *Suppose $\forall_1$ and $\forall_2$ are two devices that express generality in predicate position, in the sense that they both satisfy the universal introduction and elimination principles. Consider a sentence of the form ...X..., where the variable X takes the position that a predicate would ordinarily occupy. Then '$\forall_1 X ... X ...$' and '$\forall_2 X ... X ...$' are logically equivalent in the sense that they can be inferred from one another.*

There is nothing special about the grammatical category of predicates and names. We can, using this same sort of logical analogizing, introduce devices that generalize into the position of sentences, binary connectives, adverbs, or of quantificational phrases themselves. In general, we will call generalizations in positions different from name position *higher-order generalizations*. Given a sentence containing expressions of several grammatical categories, such as

It is possible that Socrates talks

we can form a sentential function by inserting a gap into any one of these expressions—a sentence that is variable in that position:

1. It's possible that ... is wise
2. It's possible that Socrates ...
3. It's possible that ...
4. ... Socrates is wise

As we have seen, a name quantifier lets us make generalizations in the position of the ellipses in 1, logically connecting it to all the possible ways of replacing the ellipses with names, and a predicate generalization is analogously connected to the ways of filling out the ellipses with arbitrary predicates. So it is not hard to spell out the role of a generalization into sentence position, or operator position. A sentential generalization, which we might write '$\forall p$ it's possible that p', is logically connected to the results of replacing the ellipses in 3 with

[12] We do seem to be able to express generality in the position of an adjective—like 'wise' in 'Socrates is wise; so Socrates is something'—but not into the position of the whole predicate 'is wise'.

different sentences, and an operator generalization with the results of replacing the ellipses in 4 with different operator expressions.

Before we move on, let us look at another logical expression that looks like it might be susceptible to this style of logical analogizing: the identity relation. Identity is also pinned down, up to logical equivalence, by a pair of principles. The introduction principle lets us always infer 'a is identical to a' whatever name we substitute for a. The elimination principle (Leibniz's law) lets us infer, from the assumption 'a is identical to b' that 'if ...a... then ...b...', where '...a...' represents a sentence in which a appears in the position of a name.

Theorem 0.3.2 (Harris's theorem for first-order identity). *Suppose* $=_1$ *and* $=_2$ *are two binary predicates that satisfy the identity introduction and elimination principles. Then* $a =_1 b$ *and* $a =_2 b$ *are logically equivalent, whenever* a *and* b *are names, in the sense that they can be inferred from one another.*

Proof. Suppose $a =_1 b$. Since $a =_2 a$ we can infer by Leibniz's law for $=_1$ that $a =_2 b$ replacing the second a with b. The converse direction is proved in the same way. $\square$

First-order identity takes two names as arguments and produces a sentence. However we could imagine that there was also an identity *connective* which, like conjunction, combines with two sentences to form a sentence, that is subject to the same pair of introduction and elimination rules. For any grammatical type whatsoever, we could introduce a binary operation that combines with a pair of expressions of that type to form a sentence, and satisfies the introduction and elimination rules. Some of these may be approximated in English: we might write

> To be wise is to be aware of how little one knows

to gloss the result of applying the predicate version of this operation to 'is wise' and 'is aware of how little one knows', but we do not need to rely too much on our pretheoretic grip given our stipulations pin down the operation up to logical equivalence.

0.4 Abstraction

Apart from the ability to generalize, another key expressive device in the higher-order languages explored in this book is the ability to form new meaningful expressions by *abstraction*. Abstraction is a process by which we can form complex predicates from a uniformly parameterized class of sentences, and by which we can do similar things with expressions of other grammatical types. The process is best illustrated in two steps. Take, for instance, a complex sentence like

> Socrates is wise and Socrates is old.

As we did previously, we can take a particular word, such as 'Socrates', and replace it with a gap forming a sentential function

> ... is wise and ... is old.

This sentential function associates 'Aristotle' with the sentence 'Aristotle is wise and Aristotle is old', it associates 'Plato' with 'Plato is wise and Plato is old', and so on. According to the abstraction hypothesis, one can now introduce a meaningful predicate which, when applied to any name a, yields a sentence synonymous with 'a is wise and a is old'. In English a good candidate would be the complex predicate 'is wise and old', for plausibly 'a is wise and old' is synonymous with 'a is wise and a is old'.

More generally: to every sentential function, associating meaningful names to meaningful sentences, there exists a meaningful predicate which, when applied to a name, is synonymous with the sentence associated to the name by sentential function. This process of abstraction, if it is legitimate, can be applied in other grammatical types. For instance we can form a sentential function by putting a gap where a sentence is:

> ... and Socrates is old

This sentential function associates to each meaningful sentence another sentence — for instance, it associates 'snow is white' with the sentence 'snow is white and Socrates is old'. By abstraction there corresponds to this sentential function a meaningful operator expression that, when applied to a sentence, A, yields a sentence synonymous with 'A and Socrates is old'. Or we could abstract in the position of a predicate

> Socrates ...

giving us a class of sentences—'Socrates walks', 'Socrates talks', etc—paramaterized by predicate we insert into the gap. The theory of abstraction tells us that there is a quantifier phrase that combines with an predicate, F, to make a sentence synonymous with 'Socrates Fs'.

The process of abstraction also allows us to create new binary relations, and expressions of higher arity. One way to do this is to abstract from expressions that aren't sentences to begin with. For instance, by removing 'Socrates' from the complex unary predicate 'diligently studied under Socrates' we obtain a predicate function

> diligently studied under ...

which yields a unary predicate whenever the ellipses are replaced by a name. So by abstraction we obtain a *binary* predicate which, when supplied with a name a yields a unary predicate synonymous with the unary predicate 'diligently studied under a'. We can also abstract on sentential functions of two arguments, like

> ... is wise and ___ is old

which associates each pair of names with a sentence — e.g. ‘Socrates’ and ‘Aristotle’ with ‘Socrates is wise and Aristotle is old’. The result of abstracting in both positions yields a binary relation which when applied to a pair of names a and b is synonymous with the sentence ‘a is wise and b is old’.

The reader may wonder if this process of abstraction is well-defined. One way it could fail to be well-defined would be if there could be two non-synonymous predicates that, when each applied to any name a yielded a sentence synonymous with ‘a is old and a is wise’. If this sort of thing can happen, we need a further stipulation about the behaviour of abstracted expressions. If we abstract from the sentential function

... talks

there may be several non-synonymous predicates which, like the predicate ‘talks’, yields a sentence synonymous with ‘a talks’ when applied to a. If there is a choice here, we might as well take a synonym of ‘talks’ to be the abstracted predicate. We might make the further stipulation that abstraction in degenerate cases like these are always synonymous with the predicate you started with. It turns out, given this further stipulation (and a suitably strong principle governing substitution of synonyms) that we can prove an analogue of Harris's theorem for abstraction (proposition 3.3): any two methods of abstraction that satisfy our stipulations will always yield synonymous abstracted expressions.

Another way abstraction could fail to be well-defined is if no predicate can be introduced into the language that satisfies the stipulations. How could this process of abstraction fail to produce such a predicate? Let's look at an example. The logical atomists liked to treat logical words like ‘not’, ‘and’ and ‘all’, not as stand-alone expressions but as “syncategorematic”. Here's an analogy: when we combine a name with a predicate by applying the former to the latter we do so by prefixing the predicate to the name Fa. A notational variant that makes it clear that application is the mode of combination would be to write something like $\text{app}(F, a)$. In adopting this notation we aren't treating ‘app’ as a meaningful word on its own—it does not stand to anything in reality as the words F and a do—it is simply indicating how F and a should be combined. In either notation, application is syncategorematic. This is how the logical atomists thought about the logical words. Take our example ‘Socrates is wise and Socrates is old’. According to the atomists, the word ‘and’ does not stand for anything in reality in the same way that the word ‘Socrates’ does, rather it indicates how the two sentences ‘Socrates is wise’ and ‘Socrates is old’ should be combined — by conjoining them as opposed, say, to applying one to the other. Words like ‘and’ and ‘not’, therefore, do not correspond to meaningful operators and connectives in their own right, any more than application does. If definitions by abstraction were legitimate, by contrast, we could introduce a self-standing negation operator by abstraction on the sentential function ‘not ...’, and by similar methods we could introduce a conjunction connective. The logical atomists rejected these stand-alone connectives, essentially, because it committed them to a richer *higher-order* ontology: they reject certain higher-order generalizations into operator position and connective position that follow, by existential generalization, from true sentences involving these expressions.[13] For instance

$\forall p$, for it to not be the case that p is for it to not be the case that p

entails the existential

$\exists X$, $\forall p$, for Xp is for it to not be the case that p

Note that this existential doesn't imply the existence of any individuals — it is not committal about 'ontology' as it is usually studied, relating to question formulated using the first-order existential quantifier. However, once one has accepted the higher-order framework there is, in addition to the study of ontology, a parallel domain of inquiry this time concerned with existential quantification into the positions of predicates, sentences and other expressions, which we might call 'higher-order ontology' by analogy.

Remark 0.4.1. Indeed, the reader may have noticed that if there are no restrictions on abstraction we could abstract from the two place sentential function obtained by replacing 'Socrates' and 'talks' with gaps in 'Socrates talks'. The result is a sentential function taking a name and a predicate as an argument

... ___s

The resulting binary operation when applied to a name and a predicate, a and F yields something synonymous with 'a Fs' — i.e. a stand alone operation of application.

The idea of unrestricted abstraction is closely related to other puzzles in metaphysics. The *standard* formalism for notating abstracted predicates, which will be presented in chapter 3, actually provides us with *two* ways to abstract a binary predicate from the sentential function:

... loves ___

These two abstracted binary predicates are not merely notational variants of one another. According to the naïve theory of abstraction they are non-synonymous mutual 'converses'. The idea of a converse seems intimately bound up with the notion of order in which the relation receives its arguments. But what does this really mean? It is especially obscure if reality itself does not have a left-to-right ordering mirroring the ordering of a predicate in a standard linear notation. The higher-order commitment to converses is a hotly contested issue in metaphysics. But by admitting unrestricted kinds of abstraction we can introduce binary predicates and prove higher-order existential statements like the following, using higher-order existential generalization:

$\exists R$ for John to R Mary is for Mary to love John

[13] Defenses of this position on the connectives can be found in Russell (1918), Wittgenstein (1922) and Ramsey (1927). Russell and Ramsey were often explicit in expressing these ideas in higher-order terms.

A higher-order 'ontological commitment' we obtain from having meaningful terms in the language corresponding to converses, and our prior commitment to higher-order existential generalization.

The first half of this book takes it for granted that this unrestricted process of abstraction is in good standing, and takes the reader through the many sorts of things we can do with it. In part III of the book, however, we consider languages in which the method of abstraction is restricted in certain ways that allow us to be more neutral about matters of higher-order ontology without relinquishing the introduction and elimination rules for the quantifiers.

0.5 Some things that higher-order generalizations are *not*

Higher-order logic is *useful.* We'll provide some examples shortly, but before we do that let's briefly look at some other sorts of sentences that are often employed by philosophers to do similar work. These alternatives are sometimes confused with higher-order generalizations, so we will use this as an opportunity to point out some important distinctions and common pitfalls.

The most common mistake is to identify a higher-order generalization with a first-order generalization restricted to an appropriate sort of abstract object. So let's begin by looking at the similarities and differences between a predicate generalization and a first-order generalization restricted to properties. Properties are a special sort of individual. There is some philosophical dispute both about whether they exist at all (nominalists say they don't, platonists say they do), and about their nature if they do. These include questions like: 'are they abstract objects?', 'do they have locations?', 'are there logically complex properties?', or 'are properties reducible to sets?'. But the crucial point is that we use first-order generalizations to theorize about them. The view that all properties are located—i.e. $\forall x$ if x is a property, then x is located—is generalizing into the position of a name, not a predicate. From this generalization we can infer that if wisdom is a property, then wisdom is located, that if Socrates is a property, then Socrates is located, and so on, by substituting the particular names 'wisdom', 'Socrates', and so on, for the variable x. Its instances differ over the names that replace the variable x.

In order to theorize about properties we need some primitives for talking about them. Presumably we need a first-order predicate 'is a property' to draw the distinction between things that are properties and those that are not. We also need to help ourselves to a binary first-order predicate 'instantiates' relating an individual to property. Now there are supposed to be connections between sentences involving names for properties, like 'wisdom', and sentences involving predicates, like 'is wise'. For instance:

> Socrates instantiates wisdom if and only if Socrates is wise.

These biconditionals connect questions about which properties Socrates instantiates to the truth of the instances of a predicate generalization. This suggests that we might be able to come up with a first-order property-theoretic sentence that has the same logical power as a predicate generalization. Namely, we could use

> $\forall x$, if x is property then Socrates instantiates x
>
> $\exists x$, x is property and Socrates instantiates x

to express the same sort of generality as we did with '$\forall X$, Socrates Xs' and '$\exists X$, Socrates Xs'. The general strategy for paraphrasing higher-order generalizations using first-order generalizations is this: replace every predicate variable X with the open predicate 'instantiates x' where x is a first-order variable, and replace all occurrences of '$\forall X$...' and '$\exists X$...' with restricted first-order quantifiers '$\forall x$ if x is a property...' and '$\exists x$, x is a property and...'.

If these sentences have the logical role of a higher-order generalization then they should satisfy the quantifier introduction and elimination inferences. For instance, for the universal elimination rule we need the following inference to be a good one:

> $\forall x$, if x is property then Socrates instantiates x
>
> Therefore, Socrates is wise

The dual inference for the existential (existential introduction) is:

> Socrates is wise.
>
> Therefore, $\exists x$, x is a property and Socrates instantiates x.

These inferences are guaranteed, provided that for each instance—Socrates is wise, Socrates is happy, etc—there is a corresponding property—wisdom, happiness, etc.—such that Socrates instantiating it is equivalent to the corresponding instance.

But these inferences have a poor reputation. Whereas 'Socrates is wise' all by itself suffices for the truth of the existential predicate generalization '$\exists X$, Socrates Xs', it does not suffice for the truth of '$\exists x$, x is a property and Socrates instantiates x'. In order for the latter to be true, there have to be these further special abstract objects—properties. For the nominalist, then, these first-order generalizations do not have the logical role of a predicate generalization: the former would be vacuously true even when Socrates isn't wise, and the latter vacuously false even when Socrates is wise.

Bertrand Russell famously revealed a more fundamental problem with these inferences. An instance of existential generalization is:

> $\forall y$, y doesn't instantiate itself if and only if y doesn't instantiate itself.
>
> Therefore, $\exists X$, $\forall y$, y Xs if and only if y doesn't instantiate itself.

Of course, the premise is unassailable, so by the existential introduction rule for predicate generalizations, the conclusion should be true. The proposed first-order replacement of this conclusion is, by contrast, inconsistent:

> $\exists x$, x is a property and $\forall y$, y instantiates x if and only if y doesn't instantiate itself.

The reason is this. Suppose x is instantiated by all and only those things that do not instantiate themselves. Then x must instantiate itself if and only if it does not instantiate itself.

It follows, then, that first-order generalizations over properties cannot be used to approximate generalizations into predicate position: a first-order generalization will always overlook some instances of the second-order generalization.

It is worth noting that some authors use the word 'type theory' and 'higher-order logic' somewhat differently to the way it is used in this book. According to this alternative use of these terms, type theory and higher-order logic refer to a particular framework for theorizing about first-order properties. Consider, for example, the following excerpt from the Stanford Encyclopedia of Philosophy entry on properties (Orilia and Swoyer (2020) §7.3.)

> Type theory has never gained unanimous consensus and its many problematic aspects are well-known [...]. Just to mention a few, the type-theoretical hierarchy imposed on properties appears to be highly artificial and multiplies properties ad infinitum (e.g., since presumably properties are abstract, for any property P of type n, there is an abstractness of type n+1 that P exemplifies). Moreover, many cases of self-exemplification are innocuous and common [...]. For example, the property of being a property is itself a property, so it exemplifies itself. There also seem to be transcendental relations. A transcendental relation like *thinks about* is one that can relate quite different types of things: Hans can think about Vienna and he can think about triangularity. But typed theories cannot accommodate transcendental properties without several epicycles.

The term 'type theory', as it is used in this book, is merely a theory systematizing the rules that describe how names, predicates, operators, sentences, and so forth combine. The sorts of limitations it imposes — for instance, that one cannot apply a name to an operator — are completely common sense and implicit in the logical languages the reader is likely already familiar with. By contrast, Orilia and Swoyer are using the phrase 'type theory' to refer to a very specific first-order theory of properties: one in which properties are stratified into different levels—levels resembling but distinct from the grammatical types of expressions—and in which the instantiation relation cannot hold between properties of different levels.

A related, and common refrain is that higher-order logic is really a property theory or set theory in 'sheeps clothing', to use Willard van Orman Quine's metaphor.[14] Quine essentially understands a sentence of the form '$\exists X$, Socrates Xs' as asserting that Socrates instantiates a property (or belongs to some set), and so understood these sentences of course do engender a commitment to

[14] Quine (1970); see the section heading in chapter 5.

properties. In this book, however, we interpret sentences like '$\exists X$, Socrates Xs' as a higher-order generalization. Do higher-order generalizations commit you to properties in any interesting sense? The term 'ontological commitment' has a fair amount of baggage in philosophy, which I wont attempt to untangle here.[15] But whatever we mean by it, entailments seem to transmit ontological commitments:

> If A entails B and B commits you to Fs (sets, unicorns, etc.), then A commits you to Fs as well.

This lets us infer, for instance, that

> '$\exists x$, x is wise' commits us to nothing that 'Socrates is wise' does not already commit us to.

The reason is simple: by the rules governing existential generalization, 'Socrates is wise' entails '$\exists x$, x is wise', so everything the latter entails the former entails, by the transitivity of entailment. Thus, by entirely analogous reasoning

> '$\exists X$, Socrates Xs' commits us to nothing that 'Socrates is wise' does not already commit us to.

In particular, this tells us that if '$\exists X$, Socrates Xs' commits us to sets or properties, then so does 'Socrates is wise'. But what does the claim that Socrates is wise commit us to ontologically? Presumably it commits us to wise people, and to Socrates. But it does not commit us to sets, properties or other abstract objects. After all, a nominalist can obviously classify people as wise or not wise without giving up their nominalism. If sentences like 'Socrates is wise' do commit you to sets, then very little is free of this ontological commitment. Quine's quip that second-order logic is set theory in disguise may be true, but it is hardly a rebuke. By these lights biology, astronomy, and so on, are set theory in disguise too, as, indeed, is any theory that makes assertions in subject predicate form.

Let us end this section by discussing another class of statements that are closely related to higher-order generalizations: conjunctions and disjunctions. Like the universal claim '$\forall x$, x is wise' , the infinite conjunction

> Socrates is wise and Aristotle is wise and Plato is wise and ...

bears a tight logical relationship to the list of instances above. Similarly the infinite disjunction 'Socrates is wise or Aristotle is wise or ...' is secured by any disjunct, much as the corresponding existential is secured by any instance. However, there is an important logical disanalogy. The infinite conjunction above is logically weaker than the universal claim $\forall x$, x is wise. This point essentially goes back to Russell (who is in turn drawing on Hume): you will not

[15] A particular treatment of the notion, that seems faithful at least to the way Quine uses the term, states that a claim, P, commits you to Fs just in case P entails the existential sentence $\exists x Fx$. This analysis substantiates the claim that entailments transmit ontological commitments.

be able to infer the universal claim without an extra quantificational premise to the effect that Socrates, Aristotle, Plato and so on, are all the things there are.[16] Indeed, it is arguably possible for the conjunction to be true while the universal false: suppose there had been new individuals that do not in fact exist. Perhaps, at such a possibility, all the actually existing individuals are wise, securing the truth of the conjunction, but the new individuals may fail to be wise, making the corresponding universal generalization false. I have illustrated this point with a first-order generalization, but this point applies just as forcefully to higher-order generalizations into other positions. Having said this, however, it is worth noting that disjunctions and higher-order existentials are in other ways more alike than the latter is to existential quantification over properties. The disjunction 'Socrates is wise or Socrates talks or Socrates walks...', like '$\exists X$ Socrates Xs', clearly does not entail the existence of any properties.

0.6 Higher-order generalizations in philosophy

A central contention of this book is that higher-order generalizations are useful for philosophical theorizing. I think the best way to appreciate this point is to examine a few representative debates in philosophy.[17]

Let's begin with an important theme of 20th century philosophy: the notion of metaphysical necessity as it is found, most notably, in Kripke's work.[18] A principal component of Kripke's conception of metaphysical necessity is that it is supposed to be 'necessity in the highest degree'. What is the proper way to formalize this claim? Instead of answering this question directly, it is best to begin by saying more about the role the claim plays in philosophical theorizing. Among other things, we can appeal to material conditionals like the following:

> If it's metaphysically necessary that this table is mostly made of wood, then it's physically necessary that this table is mostly made of wood.

The slogan that 'metaphysical necessity is necessity in the highest-degree' is not itself an interesting philosophical thesis without concrete implications like the one above. It is significant to the extent that we can draw these consequences from it, for it is these consequences that are used in philosophical applications, not the slogan itself. So a proper formalization of the thesis must entail conditionals like this one.

The above conditional is just one example of many. Any sentence obtained from this sentence by replacing 'it's physically necessary that' with some other operator expressing a kind of necessity should be true as well. Similarly, any sentence obtained by replacing 'this table is mostly made of wood' with another sentence should also be true. We can express this using the higher-order predicate of operators Nec introduced earlier: every sentence obtained by replacing the ellipses below with an *arbitrary* operator expression should be true

[16] Russell (1918) pp69-70.
[17] The examples and discussion in this section draw heavily on Bacon (forthcominga).
[18] Kripke (1980).

If Nec... and it's metaphysically necessary that this table is mostly made of wood, then ... this table is mostly made of wood.

Sentences of this form are all instances of a higher-order generalization which universally generalizes into the position of the ellipses above. Similarly, the result of replacing 'this table is mostly made of wood' with any other sentence in the above will also yield a truth, and sentences of this form are all instances of a further higher-order generalization, this time in sentence position. Thus, the claim that metaphysical necessity is the broadest kind of necessity is partially captured by a doubly quantified higher-order generalization. Writing '$\Box$' for metaphysical necessity:

$$\forall X(\mathrm{Nec}\, X \to \forall p(\Box p \to Xp))$$

where X is an operator variable, and p a sentence variable.[19]

Observe that the question of whether there are notions like, logical necessity, that outstrip metaphysical necessity is quite orthogonal to the debate between nominalism and platonism. In principle the nominalist and the platonist could accept or reject the thesis that metaphysical necessity in the highest degree. So we cannot properly capture this idea with a first-order generalization restricted to a special sort of abstract entity, a *necessity*, for the truth of *that* sort of generalization is not orthogonal to the nominalism-platonism dispute. Moreover, depending on how we respond to Russell's paradox, a generalization over abstract objects may also fail to have the concrete implications that constitute the claim's philosophical role.

Another area where higher-order generalizations are useful is in the formulation of supervenience theses. For concreteness, we can take the thesis that the mental supervenes on the physical. This thesis ought to imply that if Mary is in pain we should be able to fill out the ellipses below with a predicate stated in physical terms that applies to Mary, and in a way that makes the conditional come out true:

It's metaphysically necessary that if Mary ... then Mary is in pain

So it seems like the proper way to spell this out is with a higher-order existential generalization into predicate position. More generally, the supervenience thesis suggests there ought to be such a filling out of the ellipses above whenever we replace 'is in pain' with any other mental predicate, suggesting we are looking for a $\forall\exists$ sentence with alternating higher-order quantifiers. Corresponding to the distinction between predicates that use only physical language and those that don't, we might posit a corresponding distinction in reality: let Phys be a higher-order predicate that combines with a first-order predicate in order to

[19]We will return to the proper formalization of 'necessity in the highest degree' in chapter 7. The above formalization says, putting it roughly, that the extension of metaphysical necessity is included in the extension of any other necessity. But this is too weak: presumably we require this inclusion to hold necessarily. This inclusion must hold of metaphysical necessity, but to be suitably strong the inclusion must be necessary in any other sense of necessary: i.e. the proper formalization of the claim is: $\forall Y(\mathrm{Nec}\, Y \to Y\forall X(\mathrm{Nec}\, X \to \forall p(\Box p \to Xp)))$.

make this distinction, let Ment do the analogous thing for 'mental', and let m be a name for Mary. So the supervenience thesis can be captured by a higher-order generalization:

$$\forall X(\mathrm{Ment}\,X \wedge Xm \rightarrow \exists Y(\mathrm{Phys}\,Y \wedge Ym \wedge \Box(Ym \rightarrow Xm)))$$

Often supervenience theses are formulated in terms of properties. Could Mary and Nora instantiate the same physical properties while instantiating different mental properties? The nominalist thinks not: no two people can instantiate different mental properties because there are no properties to instantiate. Yet the question of supervenience still seems like a substantive one: the nominalist might still wonder whether Mary and Nora could be physically alike—that Mary's c-fibres are firing iff Nora's are, and so on—yet differ over whether they are in pain. This nominalist can express such counterexamples with a higher-order generalization: $\forall X(\mathrm{Phys}\,X \rightarrow (Xm \leftrightarrow Xn))$ conjoined with the claim that Mary is in pain but Nora isn't $Fm \wedge \neg Fn$. Nominalism is an extreme view about the existence of properties, but property theoretic formulations fail to track the issues of substance on many less extreme views. Perhaps the best solution to Russell's paradox and related paradoxes is to reject the existence of certain properties concerning intentional notions, including many mental properties. Property theoretic formulations of supervenience are therefore entangled in all of these extraneous issues, whereas the higher-order formulations are closer to the philosophical action.

Metaphysicians will often talk about 'qualitatively indiscernable' situations or objects. Two objects are indiscernable when they cannot be distinguished without making reference to particular individuals. Qualitative predicates include, for instance, 'is round', 'is heavier than something', 'self-instantiates', whereas non-qualitative predicates include 'is Jupiter shaped', 'is heavier than Jack', 'instantiates wisdom'. Leibniz, for instance, thought that qualitatively indiscernable objects had to be the very same object. By contrast, Max Black thought that two identically proportioned iron balls, in otherwise empty space, would be qualitatively indiscernible yet distinct (Black (1952)). What is the philosophical cash-value of the claim that a and b are qualitatively indiscernable? From this claim, we can infer any of the following biconditionals

a is round if and only if b is round.

a is heavier than something if and only if b is heavier than something.

a doesn't instantiate itself if and only if b doesn't instantiate itself.

By now, I hope, it is clear that the proper formalization of qualitative indiscernability involves a higher-order generalization:

$$\forall X(\mathrm{Qual}\,X \rightarrow (Xa \leftrightarrow Xb))$$

The claim that a and b instantiate the same qualitative properties does not suffice, for it does not, without taking a stand on orthogonal issues, let us infer the above biconditionals.

0.7 Semantics and model theory of higher-order languages

It does not appear to be possible to state the intended meanings of sentences expressing first-order generalizations—that is, to furnish a first-order language with a "semantics"—without using generalizations in the statement of these meanings. It does not seem promising, for example, to describe the semantics for a first-order language in a propositional meta-language. For the same sorts of reasons, we should not expect to be able to describe the semantics of sentences expressing higher-order generalizations in a language that only has first-order generalizations. In this case it is less obviously impossible — much of contemporary mathematics can be stated in the first-order language of set theory, so one has a lot more to work with than in a propositional language. However, for reasons we have already discussed, first-order generalizations over sets seem no more promising than first-order generalization over properties as approximations of higher-order generalizations.[20]

Nonetheless, first-order set-theory is useful for a very different enterprise. Suppose that you want to know whether your favorite higher-order theory contains a contradiction: can you derive a contradiction from your favored principles using the background logical rules? Or perhaps we just want to see whether your favorite principles let you prove some other undesirable consequence that falls short of a contradiction. It is natural, then, to want to have general mathematical tools for probing when one sentence is derivable from some others.

A natural strategy for showing that a sentence cannot be derived from some others using a given set of axioms and rules is to find a property that the assumptions and axioms have, that is preserved by the rules, but which is not had by the sentence in question. For if the axioms and the assumptions have the property, and the property is preserved by the rules then everything derivable from the assumptions will have the property; any sentence that does not have the property will not be derivable from the assumptions. In order for this strategy to be practically useful, we need to isolate a class of properties which are automatically preserved by the logical axioms and rules, and the standard answer to this demand is *model theory*. Model theory typically proceeds by defining a class of mathematical structures, *models* (usually made out of sets, but this is not essential) and defining a special relation between models and sentences, *truth in*, in such a way that for each model M, the logical axioms and rules automatically preserve *being true in* M.

[20] The naïve strategy for paraphrasing higher-order generalizations in predicate position with first-order quantification over sets is to mimic the strategy we adopted with properties, this time replacing each predicate variable, X, with the open predicate 'belongs to x'. This strategy may initially seem *less* promising, because sets, unlike properties, are extensional. More sophisticated paraphrases avoid commitments to extensionalism—e.g. to accommodate intensionality we could quantify instead over functions from worlds to sets—but ultimately they are all subject to the same problems we discussed in relation to the property-theoretic approximations of higher-order generalizations: they take a stand on platonism, whereas higher-order generalizations do not, and do not play the same logical role due to Russell's paradox.

Now model theory should be distinguished sharply from the project, mentioned at the outset, of describing the intended meanings of a language. As a case in point, the model theory for propositional logic consists of certain functions that map sentences to the numbers 1 and 0, and a sentence is true in one of these models when it is mapped to 1. The originators of this method, Emil Post and Ludwig Wittgenstein, did not think that the actual meanings of sentences were numbers (or truth values). Nonetheless, the truth-value method is perfectly suited to establishing what follows from what in propositional logic: it is sound and complete — a sentence is derivable from premises if and only if it is true in every model that makes the premises true.[21] Part IV of this book is devoted to the model theory of higher-order logics; some further remarks on higher-order semantics and its relation to model theory may be found in section 15.5.

0.8 Glossing higher-order generalizations in English

Having distinguished higher-order quantification from first-order quantification over properties and sets, we face a more practical problem of glossing claims like '$\exists X$ Socrates Xs' in ordinary English, which is not a higher-order language. As we have already suggested, the handle we have on these generalizations does not have to go by way of first translating them into English. However, sometimes it is easier to say something suggestive in English that brings a sentence of higher-order logic to mind than to squint at a series of logical symbols.

Some higher-order generalizations arguably do have natural language analogues. For instance, Arthur Prior suggests that in the inference from 'She met him in Paris' to 'She met him somewhere', we are replacing a prepositional phrase with the quantificational phrase 'somewhere', and so 'somewhere' is best understood as expressing a higher-order generalization into the position of a prepositional phrase, and not a generalization into name position.[22] On the other hand, while we do not seem to have something that expresses generality in predicate position in English, once we have the concept of a predicate generalization we can say things like 'Socrates somethings' or 'Socrates whatevers' and make it clear enough what one means. Using the analogy with the prepositional phrases like *somehow*, *however*, *somewhere*, *wherever* and *anywhere* Prior suggests we simply concoct terms out of the analogous sentential 'wh'-word, *whether*, to play the role of sentential quantification: *somewhether*, *anywhether* and *everywhether*. In order to make quantificational statements we need some natural language device that does the work of sentential variables. Since pronouns play the role of first-order variables in English, we need some-

[21]This method of showing logical independence between sentences using unintended interpretations was used extensively in logic and in geometry by David Hilbert and his followers. Hilbert reportedly once remarked "One must be able to say at all times–instead of points, straight lines, and planes–tables, chairs, and beer mugs".

[22]Prior (1971) chapter 3.

thing that stands to sentences as pronouns stand to nouns: prosentences (Grover (1992)). Again, by analogy with the way in which we have prepositional phrases like *there*, which can be bound by quantificational expressions like *anywhere* — as in *if Alice met Bob anywhere, she met Carol there too* — Prior suggests the word *thether*. Thus we might write 'if anywhether then thether' to capture the higher-order generalization $\forall p(p \to p)$, or 'anywhether Alice thinks that, it's necessary that' for $\forall p(Tp \to \Box p)$.

None of these solutions are particularly elegant, and none of them are fully general. A more common practice for glossing sentences of higher-order logic in the philosophical literature is to use a first-order property generalization, like 'Socrates has some property' to *indicate*, without being *synonymous* with, a higher-order generalization; in this case, '$\exists X$ Socrates Xs'. This has the disadvantage that it can lead some to think that higher-order logic is a framework for theorizing about first-order properties. However, having clarified the true relationship between sentences higher-order generalizations and property-theoretic generalizations no confusion should arise from this practice.

0.9 How to read this book

This book should certainly not be your introduction to logic. Familiarity with the syntax and model theory of propositional and first-order logic will be assumed. Apart from this, however, there are no hard prerequisites. There are parts of the book where the reader would benefit from knowing some elementary modal logic, especially chapters 7 and 8 and to a lesser extent chapters 17 and 18. While the book is self-contained in this respect, and defines basic notions like that of a *normal modal logic* and a *Kripke frame*, the book does not attempt to teach these concepts. Chapters 1-4 and 6 of Cresswell and Hughes (1996) will be more than sufficient; there are also routes through the book that do not require any modal logic.

While this book is intended to take the reader progressively through some logical and meta-theoretical concepts in higher-order logic, chapters do not always depend on all the chapters prior to them. There are consequently a couple of routes through the book that one can focus on depending on your goals. One route focuses on the higher-order logic *Classicism*, its applications to modal metaphysics and its model theory, and the other on general λ-languages and their applications to structured theories of propositions.

- Classicism and modal metaphysics: The reader should read part I of the book for an introduction to the simply typed λ-calculus, and part II for higher-order logic, the system Classicism, and the modal notions and principles that can be expressed in Classicism. Part IV of the book covers model theory for a variety of languages (chapters 14-18), but the reader should pay special attention to chapter 18 on the model theory of Classicism.

- Structure: The reader should read part I of the book for the simply typed

λ-calculus, and chapters 4 and 5 from part II for higher-order logic. They may then skip ahead to part III to learn about general lambda languages and the applications to structured theories of reality. The chapters on model theory (chapters 14-17) and appendix A and B will equip the reader with the necessary tools for modeling these theories.

This book also contains exercises. It is recommended that the reader have a pen and paper at hand while working through this book, so that exercises can be attempted as the reader proceeds. There are two sorts of exercises in this book. Those labeled *comprehension checks* are not supposed to be very challenging, they simply offer the reader an opportunity to apply a concept they have just encountered and can be attempted in the head. *Exercises*, by contrast, are longer and sometimes more challenging and will require pen and paper. Exercises are also an opportunity to explore concepts and principles that we have not had space to cover fully in the book. Many of the exercises contain statements of important facts, and should be considered as much a resource for such facts as the numbered theorems, propositions and lemmas found throughout the book.

0.10 Other resources

Higher-order logic and type theory is a vast subject with practitioners belonging to a variety of disciplines: mathematical logic[23], intuitionistic mathematics[24] and category theory[25], theorem proving[26], the foundations of programming languages[27], linguistics[28], and in philosophy, the philosophy of mathematics[29], semantics and metaphysics[30]. As a result any book on higher-order logic will by necessity contain many omissions. In the case of this book, these omissions are for the most part well-covered elsewhere. The reader may consult the references provided here for further details.

Higher-order logic originated with mathematicians, and for a while higher-order logic, not set theory, was the preferred language for foundational mathematics. Actual higher-order mathematics has a long and prestigious lineage, tracing back to Frege (1879), and continuing with Peano (1889)[31], Whitehead

[23] See the references in the next paragraph.

[24] See especially intuitionistic type theories, originating with Martin-Löf (1972), and continuing recently as homotopy type theory Univalent Foundations Program (2013).

[25] Bell (1982), Lambek and Scott (1988), MacLane and Moerdijk (2012).

[26] See for instance Gordon and Melham (1993), Brown (2007).

[27] Mitchell (1996).

[28] Carpenter (1997), Jacobson (1999), Heim and Kratzer (1998), Barker and Shan (2014).

[29] Linnebo (2013), Fine (2002a), Studd (2013), Scambler (2021), Goodsell (2022) for combine modal and higher-order resources in the philosophy of mathematics. Other papers in the philosophy of logic and mathematics involving higher-order logic include: Burgess (2005), Wright and Hale (2001), Boolos (1984), McGee (1997), Shapiro (1987), Rayo and Williamson (2003), Uzquiano (1999).

[30] See, for instance, Dorr (2016), Bacon (2020), Jones (2018),Trueman (2020), and the citations below.

[31] In Peano's original axiomatization of arithmetic the induction principle was formalized as

and Russell (1910-1913), Zermelo (1908)[32], Hilbert and Ackermann (1928)[33], Bernays and Schönfinkel (1928), Tarski (1931)[34], Carnap (1947), Church (1940), Kreisel (1967), Henkin (1950), Montague (1965), Montague (2014), Friedman (1975c), Simpson (2009).[35] Opposing the higher-order mathematicians were Skolem and Gödel, who propounded the use of first-order logic in mathematics, an attitude that Quine advocated for more generally in philosophy.[36] I will end these prefatory remarks by listing some topics in higher-order philosophy that have not been treated as directly in this book as I would have liked—I include with them some references to further resources in the footnotes: plural logic[37], ramified type theory[38], the metaphysics of grounding[39], higher-order contingentism[40], and applications of higher-order logic to propositional attitudes[41]. While discussions of these topics are limited, my hope is that this book will equip the reader with the necessary tools explore these topics on their own.

---

a single second-order generalization (in modern presentations it is presented in a first-order context by an axiom schema with infinitely many instances). There is more discussion of Frege and Peano in section 1.4.

[32]This paper contained the first axiomatization of modern set theory. Modern mathematicians often use 'Zermelo Fraenkel set theory', or 'ZF', to refer to a first-order theory, however Zermelo's original system was second order.

[33]While Hilbert and Ackermann (1928) is often cited as the first modern treatment of first-order logic, the book covers in its four chapters propositional logic, the monadic predicate logic (a version of propositional logic where the letters are interpreted as monadic predicates), predicate logic, and higher-order logic.

[34]Tarski's original definition of truth, and the original foundational papers in the discipline that would later become model theory were formulated in higher-order languages, as opposed to the first-order language of set theory.

[35]The reader should note that there is a very fuzzy distinction (which I haven't tried to draw) between 'first-order' higher-order mathematics, as it were, and mathematical work on the meta-theory of higher-order logic, exemplified in Church (1940), Henkin (1950). A selection of themes in higher-order mathematics can be found in Bell (2022), Väänänen (2012).

[36]For a good overview of how the default framework in mathematics moved from higher-order logic to first-order logic see Moore (1988).

[37]Plural logic was an important stepping stone to the rehabilitation of higher-order notions in post-Quinean philosophy; see, e.g., Boolos (1984), Schein (1993).

[38]A venerable list of philosophers have expressed sympathy towards the ramified theory of types, including Bertrand Russell, Alonzo Church, David Kaplan, Saul Kripke, and Harold Hodes. There is a brief discussion of ramified type theories in section 11.1. See Church (1976), Hodes (2015), Hatcher (1982) for a modern presentation of the ramified theory. Some critical discussion can be found in Bacon et al. (2016), Uzquiano (forthcoming).

[39]Higher-order logic is especially pertinent here in relation to the Russell-Myhill paradox, and other puzzles of ground. The reader should consult Fine (2010), Krämer (2013), Fritz (forthcominga), Fritz (forthcomingb), Fritz (2020), Litland (2020), Goodman (2022), for a start into this literature.

[40]See for instance Fine (1977), Williamson (2013), Stalnaker (2012), Fritz and Goodman (2016). Further references are given in section 8.4.

[41]These papers include Bacon and Russell (2019), Caie et al. (2020), Yli-Vakkuri and Hawthorne (2021). The literature on Prior's paradox is also relevant here: see Prior (1961), Priest (1991), Rapaport et al. (1988), Tucker and Thomason (2011), Bacon et al. (2016), Bacon (2021), Bacon and Uzquiano (2018), Uzquiano (2021).