# Intensional Type Theory for Higher-Order Contingentism

Peter Fritz
Jesus College, Oxford University

Thesis for a DPhil in Philosophy
submitted in Hilary Term 2015

supervised by
Timothy Williamson and Cian Dorr

# Contents Overview

Intensional Type Theory for Higher-Order Contingentism
Peter Fritz, Jesus College, Oxford University
Thesis for a DPhil in Philosophy, submitted in Hilary Term 2015

**Abstract.** Things could have been different, but could it also have been different what things there are? It is natural to think so, since I could have failed to be born, and it is natural to think that I would then not have been anything. But what about entities like propositions, properties and relations? Had I not been anything, would there have been the property of being me? In this thesis, I formally develop and assess views according to which it is both contingent what individuals there are and contingent what propositions, properties and relations there are. I end up rejecting these views, and conclude that even if it is contingent what individuals there are, it is necessary what propositions, properties and relations there are.

Call the view that it is contingent what individuals there are first-order contingentism, and the view that it is contingent what propositions, properties and relations there are higher-order contingentism. I bring together the three major contributions to the literature on higher-order contingentism, which have been developed largely independently of each other, by Kit Fine, Robert Stalnaker, and Timothy Williamson. I show that a version of Stalnaker's approach to higher-order contingentism was already explored in much more technical detail by Fine, and that it stands up well to the major challenges against higher-order contingentism posed by Williamson.

I further show that once a mistake in Stalnaker's development is corrected, each of his models of contingently existing propositions corresponds to the propositional fragment of one of Fine's more general models of contingently existing propositions, properties and relations, and vice versa. I also show that Stalnaker's theory of contingently existing propositions is in tension with his own theory of counterfactuals, but not with one of the main competing theories, proposed by David Lewis.

Finally, I connect higher-order contingentism to expressive power arguments against first-order contingentism. I argue that there are intelligible distinctions we draw with talk about "possible things", such as the claim that there are uncountably many possible stars. Since first-order contingentists hold that there are no possible stars apart from the actual stars, they face the challenge of paraphrasing such talk. I show that even in an infinitary higher-order modal logic, the claim that there are uncountably many possible stars can only be paraphrased if higher-order contingentism is false. I therefore conclude that even if first-order contingentism is true, higher-order contingentism is false.

# Preface

First-order contingentism is the claim that it is contingent what individuals there are; higher-order contingentism is the claim that it is contingent what propositions, properties and relations there are. This thesis explores higher-order contingentism using intensional type theory.

The thesis is composed of seven chapters, grouped into three parts. Apart from two exceptions, the chapters are written as independent papers, and can be read in any order. Each chapter is titled like the paper it comprises, and is, in other chapters, referred to as an independent (unpublished) paper, although the chapter number in this thesis is indicated. E.g., the first chapter is cited in later chapters as "Fritz and Goodman (unpublished c, here ch. 1)". Two chapters are not independent papers: chapters 3 and 7. These were written as the second and third parts of a paper in three parts, titled "Higher-Order Contingentism", the first part of which appears here as chapter 1. Chapters 3 and 7 consequently assume familiarity with chapter 1, but neither assumes familiarity with the other. The order in which the seven chapters are arranged here is a natural one, but readers most interested in later chapters are invited to start there.

Part 1 of this thesis, titled "Higher-Order Contingentism", consists of only one chapter, the first part of the aforementioned three-part paper. It is called "Higher-Order Contingentism, Part 1: Closure and Generation". In this chapter, the main philosophical ideas behind higher-order contingentism are presented, the formal framework of intensional type theory is developed, and a number of closely related theories of higher-order contingentism are formulated. One of its main aims is to bring together previous work on higher-order contingentism by Kit Fine, Robert Stalnaker and Timothy Williamson. The material in chapter 1 was developed together with Jeremy Goodman, who wrote the first two sections; the later sections were written by me. With the exception of formal definitions and proofs, we contributed equally to the ideas throughout this part.

Part 2 focuses on propositions. It is titled "Propositional Contingentism", a label which will be used for the claim that it is contingent what propositions there are. Chapter 2, titled "Propositional Contingentism", is concerned with models for the contingent existence of propositions developed by Robert Stalnaker, correcting a mistake in Stalnaker's work and exploring these models in more detail. Chapter 3 is the second part of the paper comprising chapters 1, 3 and 7, called "Higher-Order Contingentism, Part 2: Patterns of Indistinguishability". With certain restrictions, it shows that Stalnaker's models of contingently existing propositions match the propositional fragment of the models of contingently existing propositions, properties and relations developed in the first chapter. Chapter 4 is titled "Counterfactuals and Propositional Contingentism". It shows that propositional contingentism is in tension with Stalnaker's theory of counterfactuals, but not with David Lewis's theory of counterfactuals. This paper is also co-authored with Jeremy Goodman; all of it was written by myself, but some of it is based on material which was developed in discussion with Jeremy Goodman. Chapter 5 is titled "Logics for Propositional Contingentism". It investigates the propositional modal logic with propositional quantifiers which is characterized by the models of higher-order contingentism explored in chapter 3, proving that this logic is not recursively enumerable, and briefly mentioning a natural syntactic restriction.

Part 3, titled "Contingentism and Paraphrase", is concerned with a challenge faced by first-order contingentists, namely how to make sense of talk of possible objects. Chapter 6, titled "Counting Incompossibles", argues that first-order contingentists must provide a way of paraphrasing such talk, and argues that the only promising way of trying to do so uses higher-order logic. This paper was also co-authored with Jeremy Goodman. With the exception of the appendix, which I developed and wrote, this is a joint work in all of its aspects. Chapter 7 is the third part of the paper comprising chapters 1, 3 and 7, called "Higher-Order Contingentism, Part 3: Expressive Limitations". It shows that unless higher-order contingentism is false, higher-order logic is also unable to provide a way of paraphrasing talk of possible objects.

The two last chapters show that if first-order contingentism is true, then higher-order contingentism is false. If first-order contingentism is false, then presumably higher-order contingentism is false as well. So overall, I conclude that whether first-order contingentism is true or false, higher-order contingentism is false.

I draw this conclusion with some reluctance. What is later called the Fine-Stalnaker view of higher-order contingentnism strikes me as ingenious and – especially when combined with the assumption of intensionalism, the claim that necessarily co-extensive relations are identical – beautiful. However, even if the conclusion of this thesis is correct and higher-order contingentism is false, I hope it will become clear that the kind of detailed investigation of higher-order contingentism undertaken here is necessary for a proper understanding and assessment of the view.

Much of this thesis consists of work in formal logic. Mathematically, all of it is relatively straightforward, but going through it is a laborious task. I therefore expect the readers of this thesis and any articles which I will publish on the basis of it to focus on the philosophical arguments and some of the more basic formal parts. Yet, as the author, it is hard not to feel disappointed about the fact that the time and effort spent on producing the various definitions and proofs in this thesis is more likely inversely than directly correlated with the time and effort spent on digesting them. I therefore ask the reader to indulge me in giving one particular construction – tucked away in chapter 7, but a source of much joy to myself – its moment in the sun.

One of the formalizations of the theories of higher-order contingentism developed below starts with a possible world structure consisting of a set of possible worlds and, for each of these worlds, a domain of individuals and a domain of relations for each type of relations. From these materials, another domain function for relations of the different types is generated, using a formalization of the principle that at every world, a relation exists if and only if it is determined by the individuals and generating relations at this world. In the simplest non-trivial cases, the domain function for relations in the generating structure is empty; the generated structures are then generated solely on the basis of a set of worlds and an individual domain function. Under certain conditions, the construction of *projective generation* allows us to connect two generating structures of this simple kind, and to project the higher-order domain function generated by the first onto the second in such a way that the two structures generated in this way satisfy the same sentences of higher-order modal logic.

Crucial to the definition of projective generation is Definition 7.2.9; crucial to establishing that it serves its intended purpose is Lemma 7.2.10. Pairs of structures constructed using projective generation are used to prove the two limitative results established in chapter 7. In particular, projective generation is used to prove that even in higher-order logic, no paraphrase of talk of possible objects is possible, which leads to the rejection of higher-order contingentism. Thus in a sense, projective generation, and in particular the definition and lemma just highlighted, constitute the nucleus of the argument against higher-order contingentism which this thesis advances.

This thesis bears my name, but that is a poor summary of who is responsible for what it contains. Without interference, I would have written a very different thesis on a different topic. I was introduced to expressive power challenges to views similar to first-order contingentism in 2007 at a lecture by Tobias Rosefeldt discussing Lewis (2004). I learned about Timothy Wiliamson's work on higher-order contingentism and expressive power challenges at a Graduate Workshop on Philosophical Implications of Second-Order Modal Logic at the University of London in March 2010; my presentation at this workshop eventually turned into Fritz (2013). I spent more time thinking about first- and higher-order contingentism from October 2011 until June 2012 as Timothy Williamson's research assistant, assisting him with his work on Williamson (2013). Nicholas Jones told me in 2012 that Robert Stalnaker had written a book on higher-order contingentism – Stalnaker (2012) – and discussing it with Jeremy Goodman led us to develop our own first attempts at formalizing higher-order contingentism. We eventually discovered to our amazement that Kit Fine had already developed such a formalization more than 35 years earlier, in Fine (1977b). Most of all, it is due to Jeremy Goodman's ingenuity and enthusiasm that our joint work on higher-order contingentism, and the work I did by myself which it motivated, ended up occupying all of my time, and so became my thesis project.

Over the last three and a half years I was supervised, in different capacities, by Timothy Williamson, Cian Dorr and Kit Fine. Timothy Williamson was my DPhil supervisor at Oxford throughout this period, Cian Dorr became my second DPhil supervisor after my second year, and Kit Fine supervised me during the nine months I spent as a visitor at NYU in the academic year 2013–2014. I have been extraordinarily fortunate to be taught be them; I thank them, and am deeply grateful, for all the ways in which they have supported me in writing this thesis, and in my research in general.

I have thoroughly enjoyed conceiving the research project that makes up this thesis with Jeremy Goodman, and was very lucky to have, in him, someone with whom I could discuss even the most recherché questions that came up in writing it. I thank him for all I have learned from him in our collaboration, and for his understanding whenever the ratio between our ideas and what we had worked out in detail and written down caused me to despair.

For questions and comments, I thank the audiences at presentations I gave on this material, some of which jointly with Jeremy Goodman, at the Ockham Society in Oxford; the 6th Cambridge Graduate Conference on the Philosophy of Logic and Mathematics; Logica 2013 at the Hejnice Monastery in the Czech Republic; the eidos Fine Conference in Varano Borghi, Italy; the 17th Oxford Philosophy Graduate Conference; the Workshop on Logical and Modal Space in New York; the CSMN Workshop on Modality, Meaning, and Metaphysics in Oslo; the Oxford-Bristol Graduate Workshop in Logic and Philosophy of Mathematics in Oxford; the CRNAP Philosophy of Language Workshop in Oxford; the Semantics and Philosophy in Europe 7 Conference in Berlin; the Advances in Modal Logic 2014 Conference in Groningen; and philosophical colloquia at Humboldt University Berlin and Goethe University Frankfurt.

I am also grateful to the speakers and participants at the Workshop on Logical and Modal Space in New York in April 2014, which Jeremy Goodman and I organized, as well as the participants of the reading group we held in preparation for the workshop, the New York Institute of Philosophy for funding the workshop, and Cian Dorr and Kit Fine for serving as faculty sponsors. In particular, I thank Kit Fine, Robert Stalnaker and Timothy Williamson for their responses to the presentations on topics related to the present thesis which were given by Jeremy Goodman and myself.

In addition to those already mentioned, many more have helped me by discussing the topics of this thesis with me, and by reading and commenting on drafts of its chapters. Thank you: Andrew Bacon, Ralf Bader, Tim Button, Fabrice Correia, Daniel Deasy, Catharine Diehl, Daniel Dohrn, Paul Egré, Nicholas Jones, Stephan Leuenberger, Paolo Maffezioli, Robert Michels, Beau Madison Mount, Gabriela Aslı Rino Nesin, Tobias Rosefeldt, Jeff Russell, Nat Tabris, Lisa Vogt, Natalia Waights Hickman, and, especially, Harvey Lederman. I apologize to all those whose names should be, but are not, on this list.

Finally, I am glad to acknowledge the financial support I have received for my work on this thesis: During my visit at NYU, through a scholarship for post-graduate research abroad of the German Academic Exchange Service, and during my DPhil studies in Oxford, through a Scatcherd European Scholarship, a doctoral studentship of the Arts and Humanities Research Council, and the Hanfling Scholarship.

TeXShop counted 70,622 words in this thesis.

# Detailed Contents

# Part I

# Higher-Order Contingentism

# Chapter 1

# Higher-Order Contingentism, Part 1: Closure and Generation

*with Jeremy Goodman*

**Abstract.** This paper is a study of *higher-order contingentism* – the view, roughly, that it is contingent what properties and propositions there are. We explore the motivations for this view and various ways in which it might be developed, synthesizing and expanding on work by Kit Fine, Robert Stalnaker, and Timothy Williamson. Special attention is paid to the question of whether the view makes sense by its own lights, or whether articulating the view requires drawing distinctions among possibilities that, according the view itself, do not exist to be drawn. The paper begins with a non-technical exposition of the main ideas and technical results, which can be read on its own. This exposition is followed by a formal investigation of higher-order contingentism, in which the tools of variable-domain intensional model theory are used to articulate various versions of the view, understood as theories formulated in a higher-order modal language. Our overall assessment is mixed: higher-order contingentism can be fleshed out into an elegant systematic theory, but only at the cost of abandoning some of its original motivations.

## 1.1 Introduction

*Higher-order contingentism* is the view, roughly, that it is a contingent matter what propositions, properties, and relations there are. This paper aims to clarify, motivate, develop and critically assess higher-order contingentism.[1]

Propositions draw distinctions among possibilities by being true in some of them but not in others. Properties of individuals (hereafter understood as including polyadic relations) draw distinctions both among possibilities and among possible individuals by applying to certain possible individuals, but not others, in certain possibilities, but not in others. Higher-order properties (again, understood as including polyadic relations) likewise draw distinctions both among possibilities and among possible propositions and possible properties of lower types. Higher-order contingentism therefore raises the prospect of it being a contingent matter which such distinctions there are to be drawn.

The view does not *entail* that it is a contingent matter which distinctions among possibilities and possibilia there are to be drawn. This is because the relation between higher-order entities and the modal distinctions they draw may be many-one, as will be the case if necessarily equivalent propositions, and necessarily co-extensive properties, sometimes fail to be identical. In other work we discuss the possibility of accepting a hyperintensional theory of properties and propositions that, although a version of higher-order contingentism, nevertheless entails that it is a non-contingent matter what modal distinctions there are to be drawn.[2] But we will not discuss such eccentric views here. Instead, we take as a working hypothesis the widely held assumption that, if it is contingent what higher-order entities there are, then it is likewise contingent what distinctions among possibilities and possibilia there are to be drawn – this assumption will normally be left implicit. In section 1.3.4 we show how to make this assumption formally precise without begging the question of whether properties and propositions are individuated hyperintensionally.

[1]The present paper is the first of a trio: "Higher-order contingentism, Part 2: Patterns of indistinguishability" and "Higher-order contingentism, Part 3: Expressive limitations" explore further technical issues.

[2]See Fritz and Goodman (unpublished b, here ch. 6, section 6.4).

Our initial characterization of higher-order contingentism was 'rough' in the following respect. Although we have spoken of propositions, and of properties of different 'orders'/'types', such talk should be understood as our a way of communicating in English claims we think are more perspicuously formulated in a higher-order language. For example, talk of 'propositions' is our way of communicating claims we will go on to formalize using quantification into sentence position; talk of 'properties of individuals' is our way of communicating claims we will formalize using quantification into predicate position; and talk of properties 'of higher types' (e.g., 'properties of propositions') is our way of communicating claims we will formalize using quantification into corresponding syntactic positions (e.g., the position of sentential operators). Here is not the place to discuss the status of such higher-order languages as tools for metaphysical theorizing.[3] Although much is contentious, it is widely held that such languages, whatever their appropriate interpretation, are the best available tools for bringing formal discipline to the metaphysical questions in the vicinity of our informal remarks. While we believe such languages can be understood on their own terms, those who question their intelligibility can treat higher-order quantification as tacit first-order quantification over an appropriately abundant yet stratified realm of abstract objects.

With three significant exceptions, there has been hardly any systematic discussion of higher-order contingentism. The exceptions are Fine (1977b), Stalnaker (2012, chapters 1 and 2 and appendix A), and Williamson (2013, sections 6.3–6.4), from which the name of the view is taken. Fine's paper, which is by far the most systematic development of the view, seems to have gone almost completely unnoticed: Williamson cites it only in passing, and Stalnaker does not cite it at all. Its neglect is understandable: the paper is very difficult, containing only a brief philosophical preamble followed by an extensive but compressed technical development, in which the theory of higher-order contingentism is treated as a means to various other philosophical ends, and general principles are laid down without the examples needed to give a feel for their intuitive content. Stalnaker gives a more leisurely

[3]For a defense of such expressive resources, see Prior (1971) and Williamson (2013, section 5.9).

philosophical exposition, but the details of the view are likewise presented in an extremely condensed appendix in which he sketches a formal model of the contingent existence of propositions; conversely to Fine, purported examples of higher-order contingency are adduced without any general principles offered in their support. As for Williamson, the structure of his argument against higher-order contingentism is not immediately transparent and it is at any rate highly condensed at crucial stages. A main aim of the present paper is to clarify, synthesize, and expand on the views and arguments of these three authors.

There are three main reasons to care about higher-order contingentism. The first is intrinsic interest – not only of the thesis itself, but also of the particular higher-order contingentist theories we will be discussing, which offer mathematically elegant and philosophically enticing pictures of modal reality. The second reason is that higher-order modal languages are a rich and fruitful tool for metaphysical theorizing, so having a better understanding of higher-order contingentism is of foundational importance. The third and perhaps most significantly reason is that the question of higher-order contingentism bears strongly on the question of *first-order contingentism* – the view that it is a contingent matter what *individuals* there are. Its negation – *first-order necessitism* – is a radical, some would say incredible, position. Yet many have thought that first-order contingentism leads inexorably to higher-order contingentism, making the tenability of the latter a necessary condition on the tenability of the former. Others have taken the opposite perspective, and argued that first-order contingentists had better be higher-order *necessitists*. They are moved by the worry that, without all of the modal distinctions the latter theory guarantees, first-order contigentism falls victim to a charge of expressive inadequacy. We discuss these issues in other work.[4]

We will begin with a non-technical exposition of our technical results and their philosophical significance. In section 1.2.1 we explain the motivations for higher-order contingentism. In section 1.2.2 we consider Williamson's ar-

[4]Goodman (in preparation a) discusses the bearing of first-order contingentism on higher-order congintentism. Fritz and Goodman (unpublished b, here ch. 6, section 6.2) raises an expressive power challenge to higher-order contingentism, drawing on the technical results of Part 3; see section 1.2.6 below for discussion.

gument against the view. In section 1.2.3 we consider how best to develop the view into a systematic theory, drawing on the ideas of Fine and Stalnaker. In section 1.2.4 we point out a fundamental problem for such theories – namely, that their formulation requires drawing modal distinctions that, according to the theories themselves, do not exist to be drawn – and explain how the Fine-Stalnaker view must be modified in order avoid such incoherence. In section 1.2.5 we show that the modification in question has serious philosophical costs. In section 1.2.6 we take up the issue of expressive power. The remainder of the paper is an extended model-theoretic investigation of higher-order contingentism, in which the previously informally stated arguments and theories are articulated more precisely and in much greater detail.

## 1.2 Informal exposition

### 1.2.1 Motivating higher-order contingentism

The basic idea behind higher-order contingentism is that contingency in what individuals there are leads to contingency in what properties and propositions there are. Consider the proposition that you exist and the property of being identical to you. Now suppose you had never been born. The first-order contingentist thinks that, had you never been born, there would have been no such thing as you. The higher-order contingentist thinks that, moreover, there would have been no such thing as the proposition that you exist or the property of being identical to you. Indeed, most of them think that there would have been no proposition necessarily equivalent to the proposition that you exist, nor any property necessarily equivalent to being identical to you. In this sense the proposition that you exist and the property of being identical to you each draws a modal distinction that, had you never been born, there would not have been to be drawn. (Stalnaker (2012, p. 17) gives a similar example involving the possibility of Kripke having seven sons.) Even Williamson, who rejects higher-order contingentism, agrees that, on the assumption that there could have been no such thing as you (which he denies), had that happened there would have been no proposition necessarily equivalent to your existing.

The motivation for these judgments is not immediately clear. Neither Stalnaker nor Williamson offers any general principle in their support. Crucially, neither author assumes that every individual is such that, necessarily, there is a property necessarily equivalent to being identical to that individual only if there is such a thing as that individual. Consider a spare handle and blade that could easily have been (but never will be) joined to form a knife. It is natural to think that, although nothing there is is possibly such a knife, nevertheless the property of being a thing that actually would have been a knife made from the handle and the blade had the two been joined is a property that, had such a knife been made, would have been necessarily equivalent to being identical to that knife. Since there actually is such a property, it is not the case that any property that possibly 'singles out' an individual must actually single out some individual (where a property $F$ *singles out* $x$ just in case being $F$ is necessarily equivalent to being identical to $x$). In other words, it is not in general true that every haecceitity is the haecceity of something, where $F$ is a *haecceity* just in case it is possible that there be an individual that it singles out. (An analogous point applies concerning propositions that are possibly such that some individual's existence is necessarily equivalent to their truth.) Absent such a principle, it is not immediately clear *why* there shouldn't be a haecceity of you had you never been born. This is not to deny the plausibility of the thought that there would have been no such haecceity had you not been born. But absent a general principle underlying such judgments, one might wonder how seriously we should ultimately take them.

### 1.2.2 Williamson's challenge

Williamson argues for *higher-order necessitism* (the negation of higher-order contingentism) by laying down three independent general principles as conditions of adequacy on any higher-order modal logic and then arguing that the best explanation of these principles is a principle of unrestricted comprehension that entails not only them but also higher-order necessitism. We think this argument has very little dialectical force. Indeed, we have trouble imagining how any argument of this form could be dialectically effective. In order to motivate the general principles of which unrestricted comprehension is supposed to be the best explanation, one must appeal to considerations acceptable by the lights of higher-order contingentists. But why shouldn't those considerations, whatever they are, allow the higher-order contingentist to claim that the principles are already sufficiently accounted for, and so are not in need of any further explanation? Perhaps adducing a great many seemingly independent principles all of which can be derived from unrestricted comprehension would put significant pressure on higher-order contingentists, but, in our judgment, three such principles is not very many.

Furthermore, the third principle Williamson adduces is not accompanied by any serious argument that it ought to be accepted by higher-order contingentists. (The principle is in a certain sense a 'modalization' of the claim that, for any complete order on a domain, any definable condition that has an upper bound in the domain has a least upper bound in the domain.[5]) Although, as we will see later, there is a certain spin one can put on this principle whereby it might form the basis of an interesting argument against higher-order contingentism, all Williamson says in its defense is that it is needed for 'second-order modal mathematics', which to the best of our knowledge is not an extant research program.

As regards the other two principles Williamson adduces, we agree that higher-order contingentists should accept them, since they are exceedingly plausible and are needed to license patterns of modal reasoning in which we regularly engage. The first principle is an 'extensional' comprehension schema Comp$^-$. A comprehension schema is a principle that, for a schemat-

[5]The principle is that

(17) $\Box\forall X(\Diamond\exists y\Box\forall x(Xx \to \Diamond\psi) \to \Diamond\exists y(\Box\forall x(Xx \to \Diamond\psi) \wedge \Box\forall z(\Box\forall x(Xx \to \Diamond\psi[z/y]) \to \Diamond\psi[y, z/x, y])))$

should entail

(18) $\Diamond\exists y\Box\forall x(\varphi \to \Diamond\psi) \to \Diamond\exists y(\Box\forall x(\varphi \to \Diamond\psi) \wedge \Box\forall z(\Box\forall x(\varphi \to \Diamond\psi[z/y]) \to \Diamond\psi[y, z/x, y]))$

in the sense that the schema $\Box\forall x\Box\forall y\Box((17) \to (18))$ should be valid; see Williamson (2013, p. 287) and the end of section 1.5.1 below.

ically specified condition of a certain sort, says that there is a property the having of which corresponds to satisfying the condition. Such principles differ as regards what sorts of conditions are allowed and what sort of correspondence is ensured. Comp$^-$ places no restriction whatsoever on the allowable conditions, but ensures only the weakest sort of correspondence: material equivalence. In the case of conditions with no free variables, the principle entails only that there is some proposition materially equivalent to the condition, which is of course trivial. In the case of conditions with a single free first-order variable, the principle is more substantive: it say, in effect, that for any (perhaps empty or singleton) plurality of individuals there is a property that is satisfied by each of them and by nothing else.[6] A parallel point applies concerning pluralities of ordered pairs of individuals and conditions with two free first-order variables. The principle is 'extensional' because it tells us nothing about the behavior in counterfactual circumstances of the propositions and properties whose existence it guarantees. As such, it doesn't guarantee our ability to draw any interesting modal distinctions.

The more interesting principle for our purposes is Williamson's second condition of adequacy on any reasonable higher-order modal logic, Comp$_C$. Unlike Comp$^-$, Comp$_C$ only ensures the existence of properties and propositions corresponding to conditions that, roughly, are specified in terms of parameters all of which exist, where 'existence'-talk is shorthand for existential quantification of the appropriate type. But unlike Comp$^-$, Comp$_C$ ensures intensional, not merely extensional, correspondence: that is, it ensures the existence of a property necessarily equivalent to satisfying the relevant condition (or, in the case of conditions with no free variables, the existence of a proposition necessarily equivalent to the condition's being the case). In a slogan: Comp$_C$ guarantees the existence of properties and propositions corresponding to conditions that are specified only in terms of existent individuals, properties and propositions. (This gloss slides over the issue of hyperintensionality; a precise statement of the schema is given in section 1.5.1.) For example, it guarantees that every proposition has a negation, in the sense

[6]Note that although this gloss on Comp$^-$ uses plural quantifiers, the formal language defined in section 1.3.1 does not contain such quantifiers.

that, for every proposition, there is another proposition that is necessarily equivalent to the first proposition not being the case. The same goes for any definable operation on properties and propositions.

It might seem that $\text{Comp}_C$ is strictly stronger than $\text{Comp}^-$, in which case there would seem to be no question of the two principles crying out for a unifying explanation. Mustn't any extensional distinction 'among' things there are be 'specifiable' in terms of those very things? The answer is negative if we assume a finitary language, although it becomes somewhat subtle if we move to an infinitary language.[7] At any rate, let us we grant for the sake of argument that $\text{Comp}^-$ and $\text{Comp}_C$ should count as independent principles with respect to whatever criteria are relevant for theory choice. It still does not seem to us that they cry out for a unifying explanation. While the motivations just given for them were admittedly imprecise slogans, they certainly seem to establish, at least to the reasonable satisfaction of higher-order contingentists, that $\text{Comp}^-$ and $\text{Comp}_C$ have no air of mystery or coincidence in need of explanation by Williamson's stronger comprehension principle Comp (which is both unrestricted in the manner of $\text{Comp}^-$ and intensional in the manner of $\text{Comp}_C$, and from which higher-order necessitism immediately follows). So Williamson's appeals to theoretical unity provide little motivation for higher-order necessitism.

But there is a better challenge Williamson might have raised instead. Perhaps $\text{Comp}^-$ and $\text{Comp}_C$ articulate well-motivated *sufficient* conditions for the existence of properties and propositions of certain sorts. But contingency in what properties or propositions there are requires not just the existence of properties and proposition but also their possible *non*-existence. Yet so far we have been given no general *necessary* condition on the existence of

[7]In a finitary language $\text{Comp}_C$ fails to entail $\forall x \Box \exists F \forall y (Fy \leftrightarrow \Diamond Rxy)$, which is an instance of $\text{Comp}^-$. In an infinitary language, $\text{Comp}_C$ does entail $\text{Comp}^-$ relative to the class of models we define in section 1.3 (Proposition 1.5.7). This result relies on the fact that the language in question allows for formulas with as many free variables as there are entities of a given type in the domain of any world in any model. As such, it is arguably an artifact of the fact that the models in question are set-sized, since allowing for formulas with as many variables as there actually are individuals comes dangerously close to violating Cantor's theorem; see Fritz and Goodman (unpublished b, here ch. 6, section 6.1.5).

properties and propositions that delivers the verdict that there is any such contingency, as we emphasized in the previous section. The challenge for the higher-order contingentists, then, is to give some account of why it is that, had you never been born, there would have been no haecceity of you. Having considered and rejected the principle that nothing could have had a haecceity in circumstances where there was not that thing, it is not at all clear what this principle should be.[8]

### 1.2.3 The Fine-Stalnaker view

Fine's paper provides an answer to this challenge: necessarily, there are all *and only* the modal distinctions that can be drawn in terms of existing individuals and certain necessarily existing qualitative properties of individuals. The key question for the view is how to understand a modal distinction's being 'able to be drawn' in certain terms, and hence how the view is going to vindicate our purported pre-theoretical higher-order-contingentist judgments about cases. To answer this question, consider the following example. (The example will be couched using talk of possible worlds – we will return shortly to the question of whether such talk is legitimate in the present context.)

Consider a world $w_1$ in which there exist two hydrogen atoms $a_1$ and $a_2$ in otherwise empty space. Suppose further that none of the particles composing $a_1$ and $a_2$ actually exists, and that the atoms are, respectively, in different energy states $E_1$ and $E_2$. Now consider a second world $w_2$, qualitatively identical to $w_1$, but in which $a_1$ is in energy state $E_2$ and $a_2$ is in energy state $E_1$. The higher-order contigentist will deny that there are *actually* any propositions that distinguish $w_1$ from $w_2$, or properties that are haecceities of $a_1$ or of $a_2$. This result is predicted by Fine's view, since we seem to have no way to distinguish $w_1$ from $w_2$ or $a_1$ from $a_2$ in qualitative terms or in terms of individuals that actually exist.

Fine formalizes this idea using the resources of variable-domain inten-

[8] It is not clear that Williamson is in a dialectical position to make this challenge, since he himself argues for the contingent existence of haecceities on the assumption of first-order contingentism, and so is arguably under the same pressure to give a general theory licensing such judgments as higher-order contingentists are.

sional model theory for a modal higher-order language, described in section 1.3. The central notion, described in section 1.4.1, is that of an *automorphism* of modal space: a function that permutes all 'worlds' and all 'possible individuals' in a model such that, if it maps a world $w$ to $v$ and a possible individual $x$ to $y$, then $x$ exists in $w$ if and only if $y$ exists in $v$.[9] Such an automorphism is *fixed on* a world $w$ just in case it maps $w$ to itself and maps every possible individual that exists at $w$ to itself.[10] For any automorphism, we can ask whether it *preserves* a certain property, where properties are modeled as *intensions* – functions from worlds to the set of tuples of entities that at those worlds satisfy the property. For illustration, we will consider propositions (which we may think of as zero-place properties) and monadic properties of individuals, although the notion is well defined for properties of all types. An automorphism preserves a proposition $p$ just in case, if it maps a world $w$ to a world $v$, then $p$ is true in $w$ if and only if it is true in $v$; it preserves a monadic property of individuals $F$ just in case, if it maps a world $w$ to $v$ and a possible individual $x$ to $y$, then $x$ has $F$ at $w$ if and only if $y$ has $F$ at $v$. Fine's idea is that the properties (modeled as intensions) that exist at a world $w$ are those that are preserved by all automorphisms fixed on $w$ that preserve certain qualitative properties of individuals. Let us assume that there is an automorphism of modal space fixed on the actual world that preserve the relevant qualitative properties of individuals yet permutes $w_1$ and $w_2$ and $a_1$ and $a_2$. It then follows from Fine's proposal, by the definition of preservation, that no actually existing proposition is true at $w$ but not at $v$ (or vice versa), and no actually existing property applies to $a_1$ at $w_1$ but not to $a_2$ at $w_2$ (and hence no actually existing property is a haecceity of $a_1$).

In the interest of conveying Fine's basic idea, the above description is somewhat sloppy in failing to sharply distinguish models from the modal reality they model. The model theory is described precisely in sections 1.3 and 1.4. As for the picture of modal reality the proposal is meant to capture, consider the following analogy. Let the vertices of a cube represent points in

[9]Related, but less developed, ideas can be found in Stalnaker (2012, appendix A) and Williamson (2013, section 6.7).

[10]This notion differs slightly from the one captured by the predicate FIX in section 1.5.3, which also incorporates the 'preservation' condition to be described presently.

modal space. For any vertex $v$, we can distinguish certain of the vertices in terms of their geometric relations to $v$. But we cannot distinguish all such vertices – e.g., we cannot in this way distinguish two vertices adjacent to $v$. The distinctions among vertices we can draw in this way are exactly those that remain invariant when we rotate the cube about it axis through $v$. This fact corresponds to the idea that the automorphisms of modal space fixed on a point in it preserve the exactly the modal distinctions that can be drawn at that point. Finally, which such distinctions we can draw clearly depends on our choice of $v$. This fact corresponds to higher-order contingency.

We can see Fine's view as a reductive account of higher-order being: there are exactly those properties and propositions that respect the identities of all existing individuals and of certain (perhaps all) qualitative properties. Call this the *qualitative generation* view. The view admits of variations. For example, we might single out a slightly differently class of properties for preservation. It turns out, for example, that it makes a difference to the resulting higher-order modal logic if we allow third- and higher-order entities to figure in the 'generating base', or if we allow that the 'generating' properties can themselves have contingent being (as Stalnaker suggests concerning qualitative properties) – see section 1.4.4 for discussion.

More drastic departures from the Finean picture are also possible. Suppose, following Stalnaker, that we reject the qualitative/non-qualitative distinction, or at any rate reject the idea that qualitative properties, or any other independently specifiable class of properties, serves as a 'generating' basis in terms of which we can formulate a reductive theory of higher-order being. We might still appeal to automorphisms of modal space in fleshing out a theory of higher-order contingency. Suppose we simply eliminate the restriction to qualitative properties in the specification of which properties and propositions there are. The resulting view says that there are at a world $w$ exactly those properties and propositions that are preserved by all automorphisms of modal space that are fixed on $w$ and preserve *all* the properties and propositions there are at $w$. This view is non-reductive, since which properties and propositions there are at a world is specified in terms of the properties and propositions there are at that world. But it is far from trivial. For example, the natural model-theoretic implementation of this idea validates $\text{Comp}^-$ and $\text{Comp}_C$, as we show in section 1.5.1. Call this approach the *higher-order closure* view. We will henceforth refer to the general automorphism-based approach as the *Fine-Stalnaker view* of higher-order contingency, keeping in mind that Fine endorses qualitative generation while Stalnaker seems to accept only the idea of higher-order closure.

By failing to be reductive, the higher-order closure view is distinctively less impressive as an answer to the challenge to explain the existence of higher-order contingency. After all, unlike the qualitative generation view, which entails higher-order contingentism on the assumption that modal space has the relevant qualitative symmetries and contingently existing individuals, the higher-order closure view will be accepted by higher-order necessitists too. Perhaps by showing that higher-order contingentism can be developed in a systematic way, the higher-order closure view lends that thesis greater credibility, but it does not, on its face, offer a story of why there would have been no haecceity of you had you never been born. We will return to this point below.

### 1.2.4 Taking higher-order contingency seriously

It is time to be more careful in distinguishing models from the reality they model. We explained the Fine-Stalnaker view by glossing certain set theoretic structures as representing modal space, certain elements of those structures as representing possible individuals, and certain set theoretic constructs – intensions – as representing properties and propositions. But the higher-order contingentist denies that there really *are* all the individuals, properties, and propositions that there could be. So how are they to make sense of the models in terms of which they articulate their view?

There are certain familiar strategies for making respectable quantification over 'merely possible' individuals, properties, and propositions. For example, existential quantification over 'possible individuals' can be eliminated by paraphrasing 'for some possible individual $x \ldots$' as 'possibly, some individual

$x$ is *actually* such that ...'. The same goes for quantification over possible properties and propositions.[11] This much is familiar.

But such manoevers do not allow us to make sense of the claim that there are properties and propositions corresponding to *exactly those intensions* that are preserved by *all automorphisms of modal space* satisfying certain conditions. In specifying the class of models, we quantified over all intensions in, and all automorphism of, those models. But not every intension in a model need be in the higher-order domain of some world of the model. A similar point applies to automorphisms. Permutations of possible individuals correspond to binary relations, which we are modeling as certain sorts of intensions. But not every permutation of the possible individuals in a model need correspond to an intension in the higher-order domain of some world of the model. Similarly, permutations of possible worlds can be thought of as binary relations on 'world propositions', relations which we are also modeling intensionally. But not every permutation of the worlds of a model need correspond to an intension in the higher-order domain of some world of the model. Informally: not all intensions *in* a model, and automorphisms *of* a model, will even *possibly* exist *according to the model*. After all, they correspond to the very sort of modal distinctions that higher-order contingentism claims have contingent being. So even if higher-order contingentists can help themselves to quantification over *possible* entities of all types, it is not clear that they have enough modal distinctions to make sense of their own model theory.

The severity of the problem only becomes apparent when we descend from the realm of model theory to consider how we might capture the Fine-Stalnaker view in our higher-order modal language. We can easily define what it is for a binary relation among individuals to be a permutation of all possible individuals, and what it is for a binary relation on propositions to be a permutation of all possible world-propositions. We can also define, for properties of all types, what it is for them to be preserved by a pair of such permuta-

[11] The strategy goes back to Fine (1977a). Quantification over 'possible worlds' and what is true at them can be replaced with appropriately modalized quantification over 'world propositions' and talk of what propositions they necessitate. See section 1.5.2 for the details.

tions. We can then take the infinite conjunction of such preservation claims for each type, thereby expressing the claim that an automorphism preserves the existing entities of all types. This allows us to capture the higher-order closure version of the Fine-Stalnaker view in the form of an object language comprehension schema that says, as regards any condition $\phi(v_1, \ldots, v_n)$: there is a property intensionally equivalent to satisfying $\phi(v_1, \ldots, v_n)$ just in case, for any possible permutations of possible individuals and of possible worlds that together constitute an automorphism of modal space fixed on the actual world that preserves all existing properties of all types, that automorphism preserves $\phi(v_1, \ldots, v_n)$. Call this principle $\text{Comp}_{FS}$; we characterize it more precisely in section 1.5.3.

As we will prove in section 1.5.4, $\text{Comp}_{FS}$ is not valid on the class of models described in the previous section. This is for the reason described above: in defining that class of models we quantified over automorphisms of the models that did not even possibly exist according to the models. Its invalidity strikes us as unacceptable: it shows that we have not taken higher-order contingency seriously enough in formulating the view, since we have tacitly appealed to distinctions among possibilities that, according to the view itself, do not (and could not) exist to be drawn. The model theory needs revision.

Fortunately, the model theory can be straightforwardly modified so as to validate $\text{Comp}_{FS}$. Say that a model is *internally closed* just in case, for any world $w$, an intension is in the domain of $w$ just in case it is preserved by all automorphisms of the model that are ($i$) fixed on $w$, ($ii$) preserve all intensions in the domain of $w$, and ($iii$) *possibly exist according to the model* – where, as above, an automorphism of a model possibly exists according to it just in case the intensions corresponding to the world-permutation and to the possible-individual-permutation are each in the domain of some world of the model. The resulting model theory validates $\text{Comp}_{FS}$. In fact, it turns out to be a proper restriction of the class of models from the previous section, and hence validates $\text{Comp}^{-}$ and $\text{Comp}_{C}$ (as we show in sections 1.6.1 and 1.6.2).

(A different respect in which the model theory may appear unfaithful to the letter of higher-order contingentism is that, in modeling properties and propositions as intensions, we seem to rule out there being hyperintensional differences among them. In section 1.3.4 we suggest a way around this problem by interpreting our higher-order quantifiers as restricted to entities that fail to draw hyperintensional distinctions.)

### 1.2.5 Implications of internal closure

We have argued that proponents of the Fine-Stalnaker view ought to accept $\text{Comp}_{FS}$ as an object-language expression of their view, and, as a result, should restrict the class of admissible models to those that are internally closed. Let us now consider the philosophical ramifications of this restriction.

Luckily, there turn out to be a range of non-trivial internally closed models, so the view does not immediately collapse. However, as we show in section 1.6.4, the view does not have any realistic models that respect the requirement that properties hold only among existent entities, variously known as ‘serious actualism’, the ‘falsehood principle’, and the ‘being constraint’.[12] This is because there are pairs of incompossible possible individuals that are modally indistinguishable at some worlds.[13] Given the Fine-Stalnaker view, as captured by $\text{Comp}_{FS}$, there must therefore be a possible permutation of all possible individuals that possibly maps one of the individuals to the other. But no possible relation can possibly relate incompossible individuals if we assume that, necessarily, all relations are existence entailing. The effect of the claim that all relations are existence entailing is that there won’t be enough non-trivial possible permutations of possible individuals for $\text{Comp}_{FS}$ to achieve the desired effect: without the relevant permutations, the possible individuals will turn out to be distinguishable after all, contradicting the spirit of the view. We conclude that higher-order contingentists must deny that all relations are existence-entailing.[14]

[12]This principle is accepted by Stalnaker but rejected by Fine; it is trivially true according to Williamson’s necessitist view, but he also assumes it in exploring various forms of contingentism.

[13]An example involving possible identical twins is given in Fritz and Goodman (unpublished b, here ch. 6, section 6.2.3).

[14]One who accepted this principle concerning properties of individuals but not for higher-level properties could avoid this problem by characterizing automorphisms in terms

A second consequence of the move to internally closed structures is that the idea of qualitative generation no longer makes sense as a reductive theory of higher-order being. This is because there are non-isomorphic internally closed models that are *internally generated* by the same pattern of qualitative properties holding among the same individuals at the same worlds – see section 1.6.5. This 'non-functionality' of internal qualitative generation is a consequence of the self-referential character of the definition of internal closure: which automorphisms possibly exist according to a model is a matter of which world-proposition-permuting and possible-individual-permuting intensions are in the domain of some world of the model, which in turn depends on which automorphisms of the model possibly exist according to it. This result seems to stymie Fine's reductive ambitions. Indeed, the resulting class of models fails to validate $\text{Comp}_C$, further confirming that qualitative generation cannot be the whole story of higher-order being. Perhaps the view could at least account for the mere *fact* of higher-order contingency, if not for its exact contours, since there are patterns of qualitative structure that generate only models with variable higher-order domains (and, moreover, generate at least one such model). But we lack any systematic characterization of the class of such models, and the results we do have suggest that it is far from natural.

A third challenge raised by $\text{Comp}_{FS}$ is that it fails to hold in very simple models one might have thought should surely be compatible with Fine's and Stalnaker's general vision about the nature of higher-order contingency. For example, as we show in section 1.6.7, we cannot have a four-world model satisfying $\text{Comp}_{FS}$ in which any three worlds are completely indistinguishable from the perspective of the other. Intuitively, that hypothesis might seem to be perfectly consistent with the general vision. But such thinking fails to take such higher-order contingency seriously. *In such a model*, there would be no

of permutations of possible haecceities rather than of possible individuals; see section 1.6.4. One motivation for such a view would be to reconcile propositional contingentism with the existence of a property of propositions intensionally equivalent to negation: i.e., to be able to accept both $\neg\Box\forall p\Box\exists q\Box(p \leftrightarrow q)$ and $\exists O\Box\forall p\Box(Op \leftrightarrow \neg p)$, which given $\text{Comp}_C$ are inconsistent with the relevant instance of the higher-order being constraint, i.e. $\forall O\Box\forall p\Box(Op \rightarrow \exists q\Box(p \leftrightarrow q))$.

non-trivial permutations of worlds, and so any definable condition would be preserved by all automorphisms of modal space (since every condition is preserved by the identity automorphism), and so there would be no higher-order contingency according to $\text{Comp}_{FS}$, contradicting our assumption. One has the sinking feeling that the appealing vision of higher-order contingency suggested by Fine and Stalnaker's motivating remarks in fact presuppose higher-order necessitism. The worry, in a nutshell, is that articulating that vision requires drawing modal distinctions that, according to the view itself, are not there to be drawn. Although the $\text{Comp}_{FS}$-based articulation of the view does not fall prey to that sort of incoherence, perhaps it is only at the cost of sacrificing the original motivating idea.[15]

(Someone with Stalnakerian inclinations might at this point simply reject $\text{Comp}_{FS}$ and claim that their view, though it cannot be expressed in the object language, can be 'shown' by adopting the class of closed structures as one's model theory in some sort of intrumentalist spirit. We doubt the coherence of such ladder-kicking-away, but here is not the place to press the point.)

### 1.2.6 Williamson's challenge revisited

We do not claim that these challenges constitute decisive reason to reject the Fine-Stalnaker view of higher-order contingency. But they certainly constitute a much stronger reason for concern than Williamson's contention that higher-order contingentism cannot be developed into a unified theory. As far as purely structural virtues of strength and elegance of the sort Williamson emphasizes are concerned, we think the Fine-Stalnaker view does quite well. Indeed, we think it is quite a beautiful theory.

The challenges we have raised for the view only emerge when we consider relations among possible individuals and possible propositions. They do not arise if we consider only monadic second-order modal logic (on which Williamson's criticism is focussed) or a modal theory of propositions (the

[15] Fine (pc) informs us that he in fact conceived of his project as a way for a higher-order *necessitist* to make sense of higher-order contingentism. However, there is no mention of this motivation in his paper.

focus of Stalnaker's discussion). In our minds, this confirms the utility of higher-order modal languages for bringing discipline to modal metaphysics. Without being able to formulate $\text{Comp}_{FS}$ in the object language, it would be much less clear that the apparent mismatch between the model theory and the intuitive picture it was developed to model was a serious problem rather than a representationally insignificant feature of the models.

Let us now return to Williamson's contention that any adequate higher-order modal logic must validate a certain modalization of the claim that, for any complete order on a domain, any definable condition that has an upper bound in the domain has a least upper bound in the domain. This modalized claim is not valid on the class of higher-order closed models, and we conjecture that it is not valid on the class of internally closed models either. The Fine-Stalnaker view therefore supports our original suspicion that higher-order contingentists had no reason to accept the claim in the first place.

Why did Williamson make such an odd demand of higher-order contingentists? We would like to suggest that his demand can be seen as a gesture in the direction of what is in fact a very serious challenge for the view. Being a complete order is an example of a generalized quantifier – one that combines with a single formula and binds two of its free variables. Generalized quantifiers allow us to make claims about the pattern of instantiation of some conditions among the individuals there are. Moreover, we seem to be able to make sense of parallel 'modalized' claims about the pattern of satisfaction of conditions among all the individuals *there could be* – for example, if we say that there are infinitely many possible people. Contingentists cannot take such discourse at face value: they presumably think that there are really only finitely many things that are possible people, because there are only ever going to be finitely many people and there isn't anything that, though never a person, could have been one. So they face a challenge of expressive power: to make sense of such claims in a way consistent with their background metaphysical commitments. We noted in section 1.2.4 that this challenge can be met for first-order-definable generalized quantifiers. But for other generalized quantifiers it is a difficult technical question whether this challenge can be met in a sufficiently systematic way. Crucially, just because a generalized quantifier can be given a certain higher-order definition, it does *not* follow, assuming higher-order contingentism, that its modalized counterpart can be defined in a higher-order modal language.

Part 3 proves that this expressive power challenge cannot be met for the modalized generalized quantifier 'there are uncountably many possible ...', even in a highly infinitary higher-order modal language, relative to the class of internally closed models. Fritz and Goodman (unpublished b, here ch. 6) explore the surrounding philosophical issues, and argue that such expressive limitations give us very strong reason to reject higher-order contingentism. But such highly technical objections cannot even get off the ground until we have a well developed and well motivated model theory for higher-order contingentism. That is the aim of the remainder of this paper.

## 1.3 Variable Domain Type Theory

In this section, we lay out the variable domain intensional type theory in which we formalize the Fine-Stalnaker view. It is based on the constant domain type theory in Gallin (1975) and the use of variable domains in Kripke (1963).

### 1.3.1 Syntax

We use a relational type theory rather than a functional one. In such a setting, every expression of a complex type combines with a finite sequence of expressions of a lower type to form a sentential expression. The only simple type is $e$, expressions of which denote individuals. Complex types are sequences of simpler types; e.g., the type of expressions combining with two terms of type $e$ to form a sentential expression is $\langle e, e \rangle$. Consequently, the type of sentential expressions – expressions combining with no terms to form sentential expressions – is $\langle \rangle$.

We include infinitary devices in our language, allowing us to form the conjunction $\bigwedge \Phi$ of any set of formulas $\Phi$ and to bind any set of variables $V$ using a universal quantifier $\forall V$. (For a standard reference on such infinitary languages, see Dickmann (1985). For set-theoretic details on how to precisely formulate an appropriate theory of syntax, see Karp (1964).) More formally, we define:

**Definition 1.3.1.** Types *are defined inductively by stipulating that* $e$ *is a type and that any finite sequence of types is a type. Let* $T$ *be the set of types.*

For any type $t$ and natural number $n$, we write $t^n$ for the $n$-tuple $\langle t, \ldots, t \rangle$.

We construct our formal language $\mathcal{L}$ relative to a signature. The vocabulary of this language consists of the non-logical constants supplied by a signature, a proper class of variables for each type, and the symbols $=$, $\neg$, $\bigwedge$, $\Box$ and $\forall$. We call the last five elements the *logical vocabulary*, and the rest the *non-logical vocabulary*. Here and in the following, we indicate tuples by bars, e.g., writing $\bar{x}$ for $\langle x_1, \ldots, x_n \rangle$, and relying on context to determine the length $n$.

**Definition 1.3.2.** *A* signature *is a function* $\sigma$ *on the set of types which maps every type to a set, the set of* non-logical constants *of that type.*

*Let* $\sigma$ *be a signature.* Expressions of $\mathcal{L}(\sigma)$ *of the various types are defined inductively using the following rules, calling an expression a* formula *if it is of type* $\langle\rangle$*:*

- *Every variable of some type is an expression of that type.*
- *Every non-logical constant in* $\sigma$ *of some type is an expression of that type.*
- *If* $\varepsilon$ *and* $\eta$ *are expressions of type* $e$*, then* $\varepsilon = \eta$ *is a formula.*
- *If* $\varepsilon$ *is an expression of type* $\langle t_1, \ldots, t_n \rangle$ *and for all* $i \leq n$*,* $\eta_i$ *is an expression of type* $t_i$*, then* $\epsilon\bar{\eta}$ *is a formula.*
- *If* $\varphi$ *is a formula, then* $\neg\varphi$ *and* $\Box\varphi$ *are formulas.*
- *If* $\Phi$ *is a set of formulas, then* $\bigwedge \Phi$ *is a formula.*
- *If* $\varphi$ *is a formula and* $V$ *is a set of variables (of any type), then* $\forall V \varphi$ *is a formula.*

*Closed formulas are called* sentences *(where a formula is closed if it has no free variables, and free variables are defined as usual).*

Note that all complex expressions are formulas. As usual, we treat other standard operators as syntactic abbreviations. In particular, we write $\bigvee \Phi$ for $\neg \bigwedge\{\neg\varphi : \varphi \in \Phi\}$, $\exists V \varphi$ for $\neg\forall V \neg\varphi$ and $\Diamond\varphi$ for $\neg\Box\neg\varphi$. Similarly, we use obvious notational variations such as writing $\bigwedge_{i\leq n} \varphi_i$ for $\bigwedge\{\varphi_1, \ldots, \varphi_n\}$, or $\forall x\varphi$ for $\forall\{x\}\varphi$. A binary conjunction $\bigwedge\{\varphi, \psi\}$ is written more familiarly as $\varphi \wedge \psi$, and other standard Boolean operators of finite arity are defined using conjunction and negation as usual. We sometimes indicate the type of an expression by writing it in the index of its first occurrence in a formula, as in $\forall x^e x = x$.

### 1.3.2 Models

We capture necessity and possibility by quantification over a set of elements which are informally interpreted as representing possible worlds. For simplicity, we assume that the correct logic of necessity is **S5**, and more generally, that we can adequately model modal reality without adding an accessibility relation to our models. Like Kripke, we capture contingency what individuals there are by specifying a domain of individuals for each world. Of course, the abstract structures defined below need not in fact contain possible worlds and fill their domains with merely possible individuals, but for familiarity and simplicity, we will engage in this sloppy way of speaking about the model theory. Giving a complete account of the methodology of possible worlds model theory for contingentists, as discussed by Williamson (2013, chapter 3), is not the topic of this paper.

Since we are formalizing a higher-order contingentist view, our models also have to represent the contingent existence of relations (including 0-ary and 1-ary relations), for which we use domain functions as well. We first define them formally and then discuss their informal interpretation in section 1.3.4. The most perspicuous way of constructing such domains is by first fixing a set of worlds and a set of individuals, constructing a maximally inclusive set of intensions on them, and then specifying constraints on admissible domain functions which map each world to a subset of this set. Interpreting a variable $X$ of a complex type $\bar{t}$ using a higher-order entity $o$, we will need $o$ to tell us the truth-value of combining $X$ with expressions of types $t_1, \dots, t_n$ in any world. We can do so by letting $o$ be a function mapping each world to the set of tuples of entities of types $t_1, \dots, t_n$ to which it is supposed to apply at that world.

We thus bundle the choice of a set of worlds and individuals into a frame $\mathfrak{F}$, and inductively construct the set $\iota(\mathfrak{F})(t)$ of entities of type $t$ that can be constructed out of these materials. To specify this set, we adopt the following notational conventions: For any set $X$, $\mathcal{P}(X)$ is the power set of $X$. For any natural number $n$ and sets $X_1, \dots, X_n$, $\Pi_{i \leq n} X_i$ is the Cartesian product of these sets, i.e., the set of sequences $\langle x_1, \dots, x_n \rangle$ such that $x_i \in X_i$ for all $i \leq n$. For sets $X$ and $Y$, we write $X^Y$ for the set of total functions from $Y$ to $X$.

**Definition 1.3.3.** *A* frame *is a tuple* $\mathfrak{F} = \langle W, I \rangle$ *such that* $W$ *and* $I$ *are sets. Define* $\iota(\mathfrak{F})$ *to be the function on the set of types such that*

$$\iota(\mathfrak{F})(e) = I$$

*and for all types* $t = \bar{t}$

$$\iota(\mathfrak{F})(t) = (\mathcal{P}(\Pi_{i \leq n} \iota(\mathfrak{F})(t_i)))^W$$

*For any set of types* $U$*, we write* $\iota_{\mathfrak{F}}^U$ *for* $\bigcup_{t \in U} \iota(\mathfrak{F})(t)$*; in the case of singletons, we omit set-brackets, writing* $\iota_{\mathfrak{F}}^t$ *for* $\iota_{\mathfrak{F}}^{\{t\}}$*.*

As usual, we sloppily identify a set of one-tuples (a one-place relation in the mathematical sense) with the corresponding set of elements. We call an element of $\iota_{\mathfrak{F}}^t$ for a complex type $t$ an *intension*, and the relation it maps a world to its *extension* at that world. Note that the definition of frames allows for $W$ and $I$ to be empty. Intensions of type $\langle\rangle$ – the type of propositions – are functions from worlds to 0-ary relations. There are two 0-ary relations, $\emptyset$ and $\{\langle\rangle\}$; if we take the first to correspond to falsity and the second to correspond to truth, we can think of such intensions as corresponding to functions from worlds to truth-values, and so as usual as corresponding to sets of worlds.

In the second step of defining the models of our semantics, we define structures by enriching a frame with a domain assignment, which tells us for every type and world which intensions of this type are in the domain of this world:

**Definition 1.3.4.** *A* domain assignment *for a frame* $\mathfrak{F} = \langle W, I \rangle$ *is a function* $D$ *mapping each type* $t$ *to a function mapping each* $w \in W$ *to a subset of* $\iota^t_{\mathfrak{F}}$*; i.e.,* $D(t)(w) \subseteq \iota^t_{\mathfrak{F}}$*. For any set of types* $U$ *and* $V \subseteq W$*, we write* $D^U_V$ *for* $\bigcup_{t \in U} \bigcup_{w \in V} D(t)(w)$*; in the case of singletons, we omit set-brackets, e.g., writing* $D^t_w$ *for* $D^{\{t\}}_{\{w\}}$*.*

*A* structure *is a tuple* $\mathfrak{S} = \langle W, I, D \rangle$ *such that* $\langle W, I \rangle$ *is a frame and* $D$ *is a domain assignment for* $\mathfrak{F}$*.*

We sometimes refer to the union of the domains of all worlds (of some specific types or one specific type) as the *outer domain* (of those types/this type).

As noted in section 1.2.5, Williamson and Stalnaker both endorse the *being constraint*, the claim that necessarily everything could not have a property or stand in a relation without existing. (See Williamson (2013, section 4.1), Stalnaker (2012, p. 139) and Stalnaker (2003 [1994]).) In the present setting, the natural type-theoretic generalization of this constraint corresponds to the following condition on domain functions: an intension of a complex type may only be in the domain of a world if at any world $w$, its extension relates only entities in the domains of $w$. As we will see later, the Fine-Stalnaker view of higher-order contingency is in conflict with this constraint. We therefore also consider a less restrictive class of structures, in which extensions are only required to be relations on the *outer* domains. Borrowing terminology from free logic, we call the first class of structures *negative*, and the second class of structures *positive*. Note that not all structures are positive. In particular, positive semantics rules out intensions whose extensions contain intensions which are not in the domain of any world. It would also be interesting to investigate the version of the model theory which allows all structures as defined above, but we don't do so here. There are also interesting positions intermediate in strength between positive and negative semantics, and we briefly consider one such option in section 1.6.4.

We capture the being constraint and its positive weakening by the notions of negative and positive *support*, which give rise to the notion of negative and positive structures. (Note that this terminology is unrelated to that of the support of a permutation, which will be introduced below.) To be able to talk about the two versions of the semantics in a succinct way, we use the two signs $-$ and $+$ as parameters, standing for negative and positive semantics, over which we quantify using $\times$ as a meta-language variable:

**Definition 1.3.5.** *For any sign $\times$, frame $\mathfrak{F} = \langle W, I\rangle$, domain assignment $D$ for $\mathfrak{F}$, type $t$ and $o \in \iota^t_{\mathfrak{F}}$, we define $D$* $\times$supporting *$o$, written $D \boxtimes o$, as follows:*

If $t = e$, then $D \boxtimes o$ iff $o \in D^e_W$

If $t = \bar{t}$, then

$$D \boxminus o \textit{ iff } o(w) \subseteq \Pi_{i \leq n} D^{t_i}_w \textit{ for all } w \in W$$

$$D \boxplus o \textit{ iff } o(w) \subseteq \Pi_{i \leq n} D^{t_i}_W \textit{ for all } w \in W$$

*Define $\langle W, I, D\rangle$ to be a* $\times$structure *if $D \boxtimes o$ for all $o \in D^T_W$.*

Philosophically, distinct worlds represent different ways for things to be, so in any plausible structure, any distinct worlds differ in some way with respect to intensions in the domain assignment. Inspired by standard terminology in propositional modal logic (see, e.g., Blackburn et al. (2001, p. 308)) and the related but importantly different notion in Fine (1977b, pp. 148), we call such structures differentiated:

**Definition 1.3.6.** *A structure $\langle W, I, D\rangle$ is* differentiated *if for all distinct $w, v \in W$, $D^T_w \neq D^T_v$ or there is a type $t \neq e$ and $o \in D^t_W$ such that $o(w) \neq o(v)$.*

In the following, we will mostly restrict ourselves to differentiated structures. For the constant domain case, Gallin (1975) contains a very thorough discussion of non-differentiated structures, showing that for model-theoretic purposes, non-differentiated structures can safely be ignored. Since it is clear that analogous considerations apply to the present case, we omit such a discussion here. We will briefly return to issues connected to differentiation in section 1.5.2.

We are now finally ready to define the notion of a model in our semantics. One aspect of our definition worth highlighting is that we follow Kripke (1963) in singling out one world as the actual one, and that we additionally take all non-logical constants to be interpreted as expressing entities in the domain of the actual world. As noted in footnote 14, negative semantics and propositional contingentism are inconsistent with there being a property of propositions intensionally equivalent to negation, so the restriction to *non-logical* constants is essential.

**Definition 1.3.7.** *For any signature $\sigma$, a* model for $\sigma$ *is a tuple* $\mathfrak{M} = \langle W, I, D, V, w \rangle$ *such that* $\langle W, I, D \rangle$ *is a differentiated structure,* $w \in W$*, and* $V$ *is a function on* $\bigcup_{t \in T} \sigma(t)$ *mapping every element of* $\sigma(t)$ *to an element of* $D^t_w$*. We say that such a model is is* based on $\langle W, I, D \rangle$*. For any sign* $\times$*, a* $\times$model for $\sigma$ *is a model for* $\sigma$ *based on a* $\times$*structure.*

When it is clear from context, we omit the specification of a signature and simply speak of a model.

### 1.3.3 Truth

To define truth of a formula relative to a model, we have to interpret the variables. We do so using an assignment function, a function which maps variables to entities of the corresponding type. In the present setting, it is best to let this function be partial; on the one hand, this allows us to use a function in the sense of a set of tuples, rather than a proper class, and on the other hand, it allows us to have assignments for structures with empty sets of individuals. We can then give the usual recursive truth-conditions for formulas relative to a model, a world and an assignment. As a relation symbol may take a complex formula as an argument, it is best to recursively define the intension expressed by an expression relative to a model and assignment function, from which we can straightforwardly derive a definition of truth. Since assignments are partial, we have to require the domain of the assignment function to contain the free variables of the expression we are evaluating, in which case we call it admissible for the formula. To give this definition, we adopt the convention of writing, for any – possibly partial – function $f$ from a set $A$ to a set $B$, $\mathrm{dom}(f)$ for the *domain of* $f$, the set $\{x \in A : f(x) = y \text{ for some } y \in B\}$.

**Definition 1.3.8.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure. An* assignment for $\mathfrak{S}$ *is a partial function mapping variables of type* $t$ *to elements of* $D^t_W$. *An assignment* $a$ *is* admissible for an expression $\varepsilon$ *if its domain includes all free variables in* $\varepsilon$*; it is* admissible for a class of expression *if it is admissible for all members of the class.*

We are now ready to define the interpretation relation. This will be done relative to a sign and a signature, but to minimise notational clutter, we fix a sign $\times$ and a signature $\sigma$ for the rest of this section and leave the relativisation implicit. The relativisation to a sign is important as identity is to be 'existence entailing' (in the sense made precise in the clause for identity in the following definition) in the negative but not the positive semantics.

**Definition 1.3.9.** *We define a function* $[\![\cdot]\!]$ *mapping each expression* $\varepsilon$ *of type* $t$ *of* $\mathcal{L}(\sigma)$*, model* $\mathfrak{M} = \langle W, I, D, V, w\rangle$*, and assignment* $a$ *for* $\mathfrak{S} = \langle W, I, D\rangle$ *admissible for* $\varepsilon$ *to an element* $[\![\varepsilon]\!]_{\mathfrak{M},a}$ *of* $\iota^t_{\langle W,I\rangle}$ *such that for all* $u \in W$*:*

$[\![v]\!]_{\mathfrak{M},a} = a(v)$ *(*$v$ *being a variable)*

$[\![c]\!]_{\mathfrak{M},a} = V(c)$ *(*$c$ *being a non-logical constant in* $\sigma$*)*

$$[\![\varepsilon = \eta]\!]_{\mathfrak{M},a}(u) = \begin{cases} \{\langle\rangle : [\![\varepsilon]\!]_{\mathfrak{M},a} = [\![\eta]\!]_{\mathfrak{M},a} \text{ and } [\![\varepsilon]\!]_{\mathfrak{M},a} \in D^e_u\} & \text{if } \times = - \\ \{\langle\rangle : [\![\varepsilon]\!]_{\mathfrak{M},a} = [\![\eta]\!]_{\mathfrak{M},a}\} & \text{if } \times = + \end{cases}$$

$[\![\varepsilon\bar{\eta}]\!]_{\mathfrak{M},a}(u) = \{\langle\rangle : \langle[\![\eta_i]\!]_{\mathfrak{M},a} : i \leq n\rangle \in [\![\varepsilon]\!]_{\mathfrak{M},a}(u)\}$

$[\![\neg\varphi]\!]_{\mathfrak{M},a}(u) = \{\langle\rangle\}\backslash[\![\varphi]\!]_{\mathfrak{M},a}(u)$

$[\![\bigwedge \Phi]\!]_{\mathfrak{M},a}(u) = \bigcap_{\varphi\in\Phi}[\![\varphi]\!]_{\mathfrak{M},a}(u)$

$[\![\Box\varphi]\!]_{\mathfrak{M},a}(u) = \bigcap_{u'\in W}[\![\varphi]\!]_{\mathfrak{M},a}(u')$

$[\![\forall\Lambda\varphi]\!]_{\mathfrak{M},a}(u) = \bigcap_{b \text{ in } X}[\![\varphi]\!]_{\mathfrak{M},b}(u)$, *where $X$ is the class of assignments for $\mathfrak{S}$ such that*

$$b(x) = a(x) \text{ for all } x \in \mathrm{dom}(a)\backslash\Lambda, \text{ and}$$
$$b(x) \in D_u^T \text{ for all } x \in \Lambda$$

*We derive the definition of a formula $\varphi$ or class of formulas $\Phi$ being* true *relative to $\mathfrak{M}$, $u \in W$ and an assignment $a$ for $\langle W, I\rangle$ admissible for $\varphi/\Phi$ as follows:*

$$\mathfrak{M}, u, a \vDash \varphi \text{ iff } \langle\rangle \in [\![\varphi]\!]_{\mathfrak{M},a}(u).$$

$$\mathfrak{M}, u, a \vDash \Phi \text{ iff } \mathfrak{M}, u, a \vDash \varphi \text{ for all } \varphi \text{ in } \Phi.$$

*Finally, we derive truth in a model for a sentence $\varphi$ or class of sentences $\Phi$:*

$$\mathfrak{M} \vDash \varphi \text{ iff } \mathfrak{M}, w, \emptyset \vDash \varphi.$$

$$\mathfrak{M} \vDash \Phi \text{ iff } \mathfrak{M} \vDash \varphi \text{ for all } \varphi \text{ in } \Phi.$$

From this definition of truth, we derive notions of consequence and validity in the standard way. As the interpretation function $[\![\cdot]\!]$ is relative to a sign and a signature, so are truth, consequence and validity, but as above, this is left implicit.

**Definition 1.3.10.** *Let $C$ be a class of models. A sentence or class of sentences $\Gamma$ being a* consequence of a sentence or class of sentences $\Pi$ over $C$, *written $\Pi \vDash_C \Gamma$, is defined as follows:*

$\Pi \vDash_C \Gamma$ *iff $\mathfrak{M} \vDash \Gamma$ for all models $\mathfrak{M}$ in $C$ such that $\mathfrak{M} \vDash \Pi$.*

*We derive $\Gamma$ being* valid on $C$, *written $\vDash_C \Gamma$, as $\emptyset \vDash_C \Gamma$. In the case where $C$ is the class of all models, we simply omit $C$.*

Note that this is a multiple conclusion consequence relation according to which the truth of the premises guarantees the truth of *all* conclusions, rather than *some* conclusion.

The notion of truth relative to a model allows us to define a notion of two models being equivalent in the sense of satisfying the same closed formulas:

**Definition 1.3.11.** *Let $\mathfrak{M}$ and $\mathfrak{M}'$ models. $\mathfrak{M}$ and $\mathfrak{M}'$ are* equivalent, *written $\mathfrak{M} \equiv \mathfrak{M}'$, if for all sentences $\varphi$ of $\mathcal{L}$, $\mathfrak{M} \vDash \varphi$ if and only if $\mathfrak{M}' \vDash \varphi$.*

### 1.3.4 Interpreting the Higher-Order Quantifiers of $\mathcal{L}$

Models assign domains of intensions to worlds; how does this correspond to the contingent existence of relations? Consider intensions of type $\langle\rangle$, i.e., functions from worlds to sets of 0-tuples, which, as usual, we can identify with sets of worlds. One might propose to understand these intensions as corresponding to unique 0-ary relations, i.e., propositions, but it is hard to see how this would go unless necessarily equivalent propositions are identical. (Recall that we use talk of propositions, properties and relations in English to gesture at claims which could be said more precisely in a language with explicit higher-order quantifiers, such as $\mathcal{L}$.) We want to stay neutral on the question whether necessarily equivalent propositions are identical, so we won't take an intension of type $\langle\rangle$ as corresponding to a single proposition. Instead, we understand such an intension as corresponding to all the propositions which are true in exactly those worlds which the intension maps to $\{\langle\rangle\}$. And we understand such an intension being in the domain of a world as representing that there is a proposition at this world to which the intension corresponds.

In addition to allowing that there are distinct but necessarily equivalent propositions, we want to allow at the same time that propositions are identical just in case they have the same properties. So we also don't want to commit to the claim that necessarily equivalent propositions have the same properties. But the following sentence is easily seen to be valid on the class of all models:

$$\forall p^{\langle\rangle} q^{\langle\rangle} (\Box(p \leftrightarrow q) \rightarrow \forall F^{\langle\langle\rangle\rangle}(Fp \leftrightarrow Fq))$$

We must therefore interpret the higher-order quantifiers in $\mathcal{L}$ as somehow restricted. In this section, we will define a condition we call 'hereditary intensionality', and argue that possible world models can be used to interpret $\mathcal{L}$ if its quantifiers are understood as restricted to hereditarily intensional relations.

Given the notion of hereditary intensionality, which we will define shortly, we can explain in general which intensions correspond to which relations: An intension $o$ of type $t$ corresponds to all the hereditarily intensional relations

$R$ of type $t$ such that at each world, $R$ applies to a sequence of hereditarily intensional relations just in case $o$ applies to the corresponding sequence of intensions. To define which relations are hereditarily intensional, we simultaneously define the notion of two relations being *hereditarily intensionally equivalent*: A relation is hereditarily intensional if it does not possibly distinguish between any two sequences of possible hereditarily intensional entities which are pairwise hereditarily intensionally equivalent; two relations are hereditarily intensionally equivalent if they necessarily apply to the same sequences of possible hereditarily intensional entities. To ensure that these definitions are well-founded, we define all possible individuals to be hereditarily intensional, and individuals to be hereditarily intensionally equivalent just in case they are identical.

To make these informal definitions more perspicuous, we introduce a second language $\mathcal{U}$, whose quantifiers are interpreted unrestrictedly. To keep the quantifiers of $\mathcal{L}$ and $\mathcal{U}$ apart, we write the quantifier of $\mathcal{U}$ as $\forall^{\mathcal{U}}$; otherwise, the languages are exactly the same. In $\mathcal{U}$, we can formally define hereditary intensionality and hereditary intensional equivalence as follows:

$$\textsc{hi}(\varepsilon^e) := \top$$

$$\varepsilon^e \overset{\textsc{hi}}{\sim} \eta^e := \varepsilon = \eta$$

$$\textsc{hi}\left(\varepsilon^{\bar{t}}\right) := \Box\forall^{\mathcal{U}}x_1^{t_1}\ldots\Box\forall^{\mathcal{U}}x_n^{t_n}\Box\forall^{\mathcal{U}}y_1^{t_1}\ldots\Box\forall^{\mathcal{U}}y_n^{t_n}$$
$$\Box\left(\bigwedge\nolimits_{i\leq n}(\textsc{hi}(x_i)\wedge\textsc{hi}(y_i)\wedge x_i \overset{\textsc{hi}}{\sim} y_i)\to(\varepsilon\bar{x}\leftrightarrow\varepsilon\bar{y})\right)$$

$$\varepsilon^{\bar{t}} \overset{\textsc{hi}}{\sim} \eta^{\bar{t}} := \Box\forall^{\mathcal{U}}x_1^{t_1}\ldots\Box\forall^{\mathcal{U}}x_n^{t_n}\Box\left(\bigwedge\nolimits_{i\leq n}\textsc{hi}(x_i)\to(\varepsilon\bar{x}\leftrightarrow\eta\bar{x})\right)$$

With this definition, we can specify a recursive translation $\cdot^*$ from $\mathcal{L}$ to $\mathcal{U}$ which tells us exactly how to understand sentences formulated using the restricted quantifier $\forall$ of $\mathcal{L}$ using the unrestricted quantifier $\forall^{\mathcal{U}}$ of $\mathcal{U}$. All recursion clauses of this definition are trivial, apart from the clause for $\forall$, which simply replaces $\forall$ by $\forall^{\mathcal{U}}$ restricted to HI:

$$(\forall V\varphi)^* := \forall^{\mathcal{U}}V\left(\bigwedge\nolimits_{v\in V}\textsc{hi}(v)\to\varphi^*\right)$$

It may be worth illustrating the definition of hereditary intensionality using some specific cases. Consider a property of individuals $F^{\langle e\rangle}$. $F$ is hereditarily intensional just in case for all possible individuals $x^e$ and $y^e$ which are hereditarily intensionally equivalent, necessarily $Fx$ if and only if $Fy$. $x$ and $y$ are hereditarily intensionally equivalent just in case they are identical, so $F$ is trivially hereditarily intensional. This reasoning generalizes to any relation among individuals, including propositions, so necessarily all such relations are trivially hereditarily intensional.

Consider now a property of properties of individuals $\Xi^{\langle\langle e\rangle\rangle}$. $\Xi$ is hereditarily intensional just in case for all possible properties of individuals $F^e$ and $G^e$ which are hereditarily intensionally equivalent, necessarily $\Xi F$ if and only if $\Xi G$. $F$ and $G$ are hereditarily intensionally equivalent just in case necessarily, they apply to the same possible individuals, so it is clearly not trivial for $\Xi$ to be hereditarily intensional. If there are distinct properties $F$ and $G$ which necessarily apply to the same possible individuals, then there is the property of being identical to $F$, and so a property of properties of individuals which is not hereditarily intensional.

In the following, we will normally leave the restriction to hereditarily intensional relations implicit in the informal discussion. As we have just seen, this implicit restriction is only non-vacuous for types beyond the types of relations among individuals. Finally, it will be useful to be able to state in $\mathcal{L}$, rather than $\mathcal{U}$, that two relations are hereditarily intensionally equivalent. Since the definition of $\overset{\text{HI}}{\sim}$ in $\mathcal{U}$ involved only quantifiers restricted to HI, this is straightforward:

$$\varepsilon^e \overset{\text{HI}}{\sim} \eta^e := \ \varepsilon = \eta$$

$$\varepsilon^{\bar{t}} \overset{\text{HI}}{\sim} \eta^{\bar{t}} := \ \Box\forall x_1^{t_1} \ldots \Box\forall x_n^{t_n}\Box(\varepsilon\bar{x} \leftrightarrow \eta\bar{x})$$

## 1.4 Generation and Closure

With a general framework of variable domain type theory in place, we can start to formalize the different versions of the Fine-Stalnaker view introduced in section 1.2.3. We first formalize the intuitive idea of permutations of modal space, and then use this to formalize both the qualitative generation and higher-order closure views.

### 1.4.1 Automorphisms

In the possible-worlds model theory defined in the previous section, a permutation of modal space is naturally understood to determine a permutation of the possible worlds and the outer domains of that structure. (Mathematically, a *permutation of a set* is a bijection from this set to itself.) Clearly, we want the permutation to respect types: an entity of a given type should only be mapped to one of the same type. Permutations should also preserve facts about what applies to what: if an intension applies to a sequence of entities in a world, then its image (under the permutation) should apply to the image of the sequence in the image of the world, and *vice versa.* It turns out that for any permutation of worlds and permutation of the outer domain of individuals, there is only one way of permuting the elements of the outer domains of higher types that satisfies this constraint. It is therefore most convenient to take such pairs to represent permutations of modal space. Extending terminology from algebra, we call them automorphisms.

**Definition 1.4.1.** *An* automorphism of a frame $\mathfrak{F} = \langle W, I\rangle$ *is a tuple* $\langle f, g\rangle$ *such that* $f$ *is a permutation of* $W$ *and* $g$ *is a permutation of* $I$*. Let* $\mathrm{aut}(\mathfrak{F})$ *be the set of automorphisms of* $\mathfrak{F}$*.*

*Writing* $\circ$ *for the composition of functions,* $\cdot^{-1}$ *for taking the inverse of an injection, and* $\mathrm{id}_X$ *for the identity function on a given domain* $X$*, we extend these notions in the straightforward way to automorphisms:*

$$\langle f, g\rangle \circ \langle f', g'\rangle = \langle f \circ f', g \circ g'\rangle$$

$$\langle f, g\rangle^{-1} = \langle f^{-1}, g^{-1}\rangle$$

$$\mathrm{id} = \langle \mathrm{id}_W, \mathrm{id}_I\rangle$$

*In the following, we usually omit* $\circ$*, writing* $ff'$ *for* $f \circ f'$*.*

We will now note that $\mathrm{aut}(\mathfrak{F})$ always forms a group with $\circ$. The definition of a group and an exposition of the basic group-theoretic tools we will use in the following can be found in standard references such as Rotman (1995) and Dixon and Mortimer (1996). We adopt the convention of writing $S_X$ for the *symmetric group on* $X$, i.e., the set of permutations of $X$ with the operation of function composition.

**Proposition 1.4.2.** *For any frame* $\mathfrak{F}$*,* $\mathrm{aut}(\mathfrak{F})$ *forms a group under* $\circ$*.*

*Proof.* Let $\mathfrak{F} = \langle W, I \rangle$. The claim is immediate from the fact that $S_W$ and $S_I$ are groups. □

To extend the automorphisms of a frame to functions on intensions via the constraint motivated above, we make use of actions. An *action* $\alpha$ of a group $G$ on a set $X$ is a homomorphism from $G$ to $S_X$; i.e., a function mapping each element of $G$ to a permutation of $X$ such that for all $g, f \in G$, $\alpha(gf) = \alpha(g)\alpha(f)$. If it is clear from context that we are concerned with a specific action $\alpha$ of $G$ on $X$, we simply say that $G$ *acts on* $X$, and write $g.x$ for $\alpha(g)(x)$. In our present application, we introduce an action for every type that maps every automorphism of a frame to a permutation of the intensions of the relevant type based on the frame. In the service of uniform notation, we do the same for worlds and individuals.

**Definition 1.4.3.** *Let* $\mathfrak{F} = \langle W, I \rangle$ *be a frame. We define a function* $\alpha^W_{\mathfrak{F}} : \mathrm{aut}(\mathfrak{F}) \to W^W$ *and a function* $\alpha^t_{\mathfrak{F}} : \mathrm{aut}(\mathfrak{F}) \to \left(\iota^t_{\mathfrak{F}}\right)^{\iota^t_{\mathfrak{F}}}$ *for every type such that for all* $\xi = \langle f, g \rangle \in \mathrm{aut}(\mathfrak{F})$*:*

$$\alpha^W_{\mathfrak{F}}(\langle f, g \rangle)(w) = f(w)$$

$$\alpha^e_{\mathfrak{F}}(\langle f, g \rangle)(o) = g(o)$$

$$\alpha^{\bar{t}}_{\mathfrak{F}}(\xi)(o)(\alpha^W_{\mathfrak{F}}(\xi)(w)) = \left\{\left\langle \alpha^{t_i}_{\mathfrak{F}}(\xi)(o_i) : i \leq n \right\rangle : \bar{o} \in o(w)\right\}$$

By a straightforward induction on the complexity of types, it can be verified that $\alpha^t_{\mathfrak{F}}$ is well-defined.

**Proposition 1.4.4.** *For any frame* $\mathfrak{F} = \langle W, I \rangle$*,* $\alpha^W_{\mathfrak{F}}$ *and* $\alpha^t_{\mathfrak{F}}$*, for some type* $t$*, are actions of* $\mathrm{aut}(\mathfrak{F})$ *on* $W$ *and* $\iota^t_{\mathfrak{F}}$*, respectively.*

*Proof.* Routine. □

In keeping with our notation for actions, we write $\xi.w$ for $\alpha^W_{\mathfrak{F}}(\xi)(w)$, and similarly for $\alpha^e_{\mathfrak{F}}$ and $\alpha^t_{\mathfrak{F}}$, relying on the context to supply the omitted parameters. In the following, we will often need to extend a function on a set $X$ to apply to subsets of $X$ or sequences of elements of $X$. To keep notation simple, we do so implicitly in the obvious way. I.e., for any $Y \subseteq \mathrm{dom}(f)$, $f(Y) = \{f(x) : x \in Y\}$, and for any $\langle x_1, \ldots, x_n\rangle \in \mathrm{dom}(f)^n$, $f(\langle x_1, \ldots, x_n\rangle) = \langle f(x_1), \ldots, f(x_n)\rangle$. (If $X$ contains sets or tuples of some of its elements, this notation introduces ambiguity; in the following, the intended meaning will always be clear from context.) We adopt exactly the same convention for actions. With this, we can write the last condition of Definition 1.4.3 more concisely as follows: $(\xi.o)(\xi.w) = \xi.(o(w))$.

We can now define an automorphism of a *structure* to be an automorphism of the underlying frame which respects the domain assignment of the structure; i.e., for every type, we require the elements in the domain at any world to be mapped to the elements in the domain at the world to which the original world is mapped. We can prove that these form a subgroup of the automorphisms of the frame.

**Definition 1.4.5.** *An* automorphism $\xi$ of a structure $\mathfrak{S} = \langle W, I, D\rangle$ *is an automorphism of* $\langle W, I\rangle$ *such that for all types* $t$ *and* $w \in W$, $\xi.D^t_w = D^t_{\xi.w}$. *Let* $\mathrm{aut}(\mathfrak{S})$ *be the set of automorphisms of* $\mathfrak{S}$.

**Proposition 1.4.6.** *For any structure* $\mathfrak{S} = \langle W, I, D\rangle$, $\mathrm{aut}(\mathfrak{S})$ *is a subgroup of* $\mathrm{aut}(\langle W, I\rangle)$.

*Proof.* It suffices to prove that $\xi\zeta^{-1} \in \mathrm{aut}(\mathfrak{S})$ for all $\xi, \zeta \in \mathrm{aut}(\mathfrak{S})$. Consider any type $t$ and $w \in W$. Since $\zeta \in \mathrm{aut}(\mathfrak{S})$, $\zeta.D^t_{\zeta^{-1}.w} = D^t_{\zeta\zeta^{-1}.w}$, so $D^t_{\zeta^{-1}.w} = \zeta^{-1}.D^t_w$. Further $\xi.D^t_{\zeta^{-1}.w} = D^t_{\xi\zeta^{-1}.w}$, so $\xi\zeta^{-1}.D^t_w = D^t_{\xi\zeta^{-1}.w}$. Hence $\xi\zeta^{-1} \in \mathrm{aut}(\mathfrak{S})$. □

### 1.4.2 Generation and Closure of Structures

To capture the idea of qualitative generation, we need a way of generating a structure from a choice of individuals and relations. To keep the construction as flexible as possible, we don't limit this choice to relations among individuals, and allow it to vary between worlds. We can thus use a structure to specify the materials from which we generate, with the domain assignment of higher orders containing the intensions corresponding to the relations from which we generate.

So, given a structure, one might suggest generating another structure by letting the domain function of the latter contain an intension at a given world if and only if it is mapped to itself by all automorphisms of the original structure which map the world and all elements of its domains to themselves. However, structures generated in this way need not satisfy the being constraint or its positive weakening. In fact, they will only do so under very special circumstances, as for all types, the trivial intension which applies to all tuples of intensions of the relevant type at all worlds is mapped to itself by all automorphisms, but will in many cases not satisfy the relevant constraint. Therefore, we have to distinguish between negative and positive generation, and in each case require an intension to be negatively or positively supported by the domain function to be included in the domain function of the generated structure.

Note that we do not require the generating structure to be negative or positive. This is not merely for technical generality, but also philosophically motivated. E.g., the generating relations might include a modality of nomological necessity, which in the present type hierarchy would be represented by an intension $\mu$ of type $\langle\langle\rangle\rangle$. Presumably, the extension of $\mu$ at any world would include the propositional intension true in all worlds. Yet the propositional intension true in all worlds might well not be in domain of any world in the relevant generating structure, and so this structure might be neither positive nor negative.

For the formal definition of qualitative generation, note that for a group $G$ acting on a set $X$ and $x \in X$, we write $G_x$ for the set of elements of $G$ which map $x$ to itself; this is called the *stabilizer* of $x$, and can be shown to form a subgroup of $G$.

**Definition 1.4.7.** *Let* $\mathfrak{S} = \langle W, I, B\rangle$ *be a structure. For any* $w \in W$*, define*

$\text{fix}(\mathfrak{S}, w) = \text{aut}(\mathfrak{S})_w \cap \bigcap_{o \in B^T_w} \text{aut}(\mathfrak{S})_o.$

*For any sign* $\times$*, define the* structure $\times$generated by $\mathfrak{S}$*, written* $\otimes\mathfrak{S}$*, to be the structure* $\langle W, I, D\rangle$ *such that for all* $w \in W$*,* $D^e_w = B^e_w$ *and for all types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*:*

$o \in D^t_w$ *iff* $D \boxtimes o$ *and* $\xi.o = o$ *for all* $\xi \in \text{fix}(\mathfrak{S}, w)$

A few observations regarding this definition are straightforward, but helpful to note: First, since subgroups are closed under intersection, $\text{fix}(\mathfrak{S}, w)$ is a subgroup of $\text{aut}(\mathfrak{S})$. Second, the condition defining the domains of complex type can also be stated as follows:

$o \in D^t_w$ iff $D \boxtimes o$ and $\text{fix}(\mathfrak{S}, w) \subseteq \text{aut}(\mathfrak{S})_o$

Finally, we register that the addition of the support requirement has the desired effect and that generated structures are differentiated:

**Proposition 1.4.8.** *For every sign* $\times$ *and structure* $\mathfrak{S}$*,* $\otimes\mathfrak{S}$ *is a differentiated* $\times$*structure.*

*Proof.* That $\otimes\mathfrak{S}$ is a $\times$structure is immediate by construction; that it is differentiated follows from the fact that every world $w$ contains its world-proposition (the propositional intension which is only true at $w$). □

While this formalization of the qualitative generation view allows us to generate one structure from another, the formalization of the higher-order closure view should impose a constraint on structures. With the tools established so far, it is straightforward to state this constraint – we want to require the domain function for any type to map each world to the set of intensions of the relevant type which are supported by the domain function and mapped to themselves by all automorphisms of the structure which map all entities in the domains of the given world to themselves. As with generation, being closed must be relativized to a sign, as the notion of support is relative to a sign:

**Definition 1.4.9.** *For any structure* $\mathfrak{S} = \langle W, I, D\rangle$ *and sign* $\times$*,* $\mathfrak{S}$ *is* $\times$closed *if for all* $w \in W$*, types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*:*

*$o \in D^t_w$ iff $D \boxtimes o$ and $\xi.o = o$ for all $\xi \in \mathrm{fix}(\mathfrak{S}, w)$*

The definitions of generation and closure are evidently closely related:

**Proposition 1.4.10.** *Let $\times$ be a sign. A structure $\mathfrak{S}$ is* $\times$closed *just in case $\mathfrak{S} = \otimes\mathfrak{S}$.*

*Proof.* Immediate by the definitions of closure and generation. □

The definitions of generation and closure constrain only higher-order domains. One might consider a view on which the existence of individuals obeys a similar constraint: if an individual can be singled out from a certain world, it must exist in it. It would be straightforward to adapt the formal definitions of generation and closure to formalize such a strengthened version of the generation and closure thoughts. In the following, we don't consider these variants.

### 1.4.3 A structure is closed if and only if it is generated

Proposition 1.4.10 tells us that a structure being closed is for it to be generated by itself, and therefore that all closed structures are generated by some structure. The converse of this implication is harder to see, but we show in this section that it also holds: any generated structure is closed. To carry out this proof, we first establish a number of lemmas.

**Lemma 1.4.11.** *Let $\mathfrak{S} = \langle W, I, D\rangle$ be a structure and $\xi \in \mathrm{aut}(\mathfrak{S})$.*

*(i) For all $o \in \iota^T_{\langle W,I\rangle}$, $\xi.\mathrm{aut}(\mathfrak{S})_o = \mathrm{aut}(\mathfrak{S})_{\xi.o}$.*

*(ii) For all $w \in W$, $\xi.\mathrm{fix}(\mathfrak{S}, w) = \mathrm{fix}(\mathfrak{S}, \xi.w)$.*

*Proof.* (i) follows from a general principle about stabilizers. So does the fact that $\xi.\mathrm{aut}(\mathfrak{S})_w = \mathrm{aut}(\mathfrak{S})_{\xi.w}$, for any $w \in W$, given which (ii) follows from (i). □

**Lemma 1.4.12.** *Let $\times$ be a sign, $\mathfrak{S} = \langle W, I, B\rangle$ a structure and $\otimes\mathfrak{S} = \langle W, I, D\rangle$. For all types $t$,*

(i) *For all* $\xi \in \mathrm{aut}(\mathfrak{S})$ *and* $o \in \iota^{t}_{\langle W,I\rangle}$, $D \boxtimes o$ *iff* $D \boxtimes \xi.o$.

(ii) *For all* $\xi \in \mathrm{aut}(\mathfrak{S})$ *and* $w \in W$, $\xi.D^{t}_{w} = D^{t}_{\xi.w}$.

*Proof.* By induction on types. The base case is trivial, so consider a complex type $\bar{t}$. (i): For any $\xi \in \mathrm{aut}(\mathfrak{S})$ and $o \in \iota^{\bar{t}}_{\langle W,I\rangle}$, it is routine to show using induction hypothesis (ii) that $D \boxtimes o$ entails $D \boxtimes \xi.o$. The converse direction follows using $\xi^{-1}$. (ii): For any $\xi \in \mathrm{aut}(\mathfrak{S})$ and $w \in W$, $o \in \xi.D^{\bar{t}}_{w}$ iff $\xi^{-1}o \in D^{\bar{t}}_{w}$, which is the case iff $D \boxtimes \xi^{-1}.o$ and $\mathrm{fix}(\mathfrak{S}, w) \subseteq \mathrm{aut}(\mathfrak{S})_{\xi^{-1}.o}$. By induction hypothesis (i), the former is the case iff $D \boxtimes o$. By Lemma 1.4.11, the latter is the case iff $\mathrm{fix}(\mathfrak{S}, \xi.w) \subseteq \mathrm{aut}(\mathfrak{S})_{o}$. Thus $o \in \xi.D^{\bar{t}}_{w}$ iff $o \in D^{\bar{t}}_{\xi.w}$. □

**Lemma 1.4.13.** *Let* $\times$ *be a sign and* $\mathfrak{S} = \langle W, I, B\rangle$ *a structure.*

(i) $\mathrm{aut}(\mathfrak{S}) \subseteq \mathrm{aut}(\otimes\mathfrak{S})$.

(ii) *For all* $w \in W$: $\mathrm{fix}(\mathfrak{S}, w) \subseteq \mathrm{fix}(\otimes\mathfrak{S}, w)$.

*Proof.* (i) follows from Lemma 1.4.12 (ii). For (ii), let $\otimes\mathfrak{S} = \langle W, I, D\rangle$, and consider any $w \in W$ and $\xi \in \mathrm{fix}(\mathfrak{S}, w)$. Then $\xi.w = w$, and by (i), $\xi \in \mathrm{aut}(\otimes\mathfrak{S})$. For any $o \in D^{T}_{w}$, $\mathrm{fix}(\mathfrak{S}, w) \subseteq \mathrm{aut}(\mathfrak{S})_{o}$, so $\xi.o = o$. Hence $\xi \in \mathrm{fix}(\otimes\mathfrak{S}, w)$. □

**Proposition 1.4.14.** *For any sign* $\times$, *any structure is* $\times$*closed if and only if it is* $\times$*generated by some structure.*

*Proof.* By Proposition 1.4.10, any $\times$closed structure is $\times$generated by some structure. So consider any structure $\otimes\mathfrak{S} = \langle W, I, D\rangle$ $\times$generated by a structure $\mathfrak{S} = \langle W, I, B\rangle$. To prove that $\otimes\mathfrak{S}$ is closed, we show for any $w \in W$, type $t \neq e$ and $o \in \iota^{t}_{\langle W,I\rangle}$ that $o \in D^{t}_{w}$ iff $D \boxtimes o$ and $\xi.o = o$ for all $\xi \in \mathrm{fix}(\otimes S, w)$. Since $\otimes\mathfrak{S}$ is a $\times$structure, $o \in D^{t}_{w}$ entails that $D\boxtimes o$, hence the left-to-right direction is immediate. So assume that $D\boxtimes o$ and $\xi.o = o$ for all $\xi \in \mathrm{fix}(\otimes\mathfrak{S}, w)$. We can prove that $o \in D^{t}_{w}$ by showing that $\xi.o = o$ for all $\xi \in \mathrm{fix}(\mathfrak{S}, w)$. This follows from Lemma 1.4.13 (ii). □

In the present formalization, the ideas of qualitative generation and higher-order closure thus coincide in the sense of determining the same class of structures.

A natural relation among structures on a given frame orders them according to containment of higher-order domains; this is easily seen to be a partial order. Proposition 1.4.14 establishes that the generation operator $\otimes$ on structures is idempotent ($\otimes\otimes\mathfrak{S} = \otimes\mathfrak{S}$). One might conjecture that with respect to the order mentioned, it is also extensive (writing $\sqsubseteq$ for the order: if $\mathfrak{S} \sqsubseteq \mathfrak{S}'$ then $\otimes\mathfrak{S} \sqsubseteq \otimes\mathfrak{S}'$) and increasing ($\mathfrak{S} \sqsubseteq \otimes\mathfrak{S}$), and thus a closure operator on the class of structures ordered by it. However, $\otimes$ has neither of these further properties, as can be shown using relatively simple finite structures.

### 1.4.4 Closed and Finely Generated Models

We call a model closed or generated just in case it is based on a closed or generated structure.

**Definition 1.4.15.** *For any sign* $\times$*, a model is* $\times$closed *just in case it is based on a* $\times$*closed structure.* $C\times$ *is the class of* $\times$*closed models.*

Fine (1977b) essentially works with +closed models, apart for two differences. The first difference is that Fine includes extensional types in his type hierarchy; we won't consider such types in the following. We can understand the second difference as consisting of three restrictions on the choice of relations from which higher-order domains are generated. First, he requires this choice to be constant across worlds. Second, he requires it to contain only relations among individuals. Third, he requires for any two distinct worlds with the same individuals that there be a generating relation which allows us to distinguish between these two worlds using only individuals which exist at those worlds. (It is this last existence condition which makes this a stronger restriction than being differentiated.) We can easily capture these restrictions in the present setting:

**Definition 1.4.16.** *Let a structure* $\mathfrak{S} = \langle W, I, D\rangle$ *be* Finely generated *if there is a structure* $\mathfrak{S}' = \langle W, I, B\rangle$ *such that* $\mathfrak{S} = \oplus\mathfrak{S}'$ *and:*

*(i) For all types* $t$ *and such that* $t \notin \{e, e^n : n < \omega\}$, $B^t_W = \emptyset$.

*(ii) For all types $t \neq e$ and $w, v \in W$, $B^t_w = B^t_v$.*

*(iii) For all distinct $w, v \in W$ such that $B^e_w = B^e_v$, there is an $F \in B^{\bar{e}}_W$ and $\bar{o} \in B^e_w$ such that not $\bar{o} \in F(w)$ iff $\bar{o} \in F(v)$.*

*A model is* Finely generated *if it is based on a Finely generated structure.*

Before turning to comprehension principles, we note that not every +closed structure is Finely generated. In fact, each of conditions (i) and (ii) alone restricts the class of structures that can be generated: in both cases, there are closed structures which are not generated by any structure satisfying the constraint. Since there are in particular *finite* such structures, and we can characterize every finite structure relative to the class of closed models up to isomorphism using a single sentence, this means that the classes of closed models satisfying one of these additional constraints validate sentences which are invalid on the class of all closed models. In contrast, it is easy to see that (iii) imposes no restriction on the class of generated structures, since every generating structure can be turned into one that satisfies (iii) and generates the same structure as the initial one, by adding to each world its world-proposition.

We only sketch proofs of the facts that (i) and (ii) restrict the class of generated structures. The proof for (i) is based on the fact that there are patterns of indistinguishability which cannot be reduced to the pairwise indistinguishability of worlds, which is discussed in more detail in Fritz (unpublished d, here ch. 2). We can use this observation for the present application by considering structures in which there necessarily exists one and the same individual. For (ii), consider a structure in which there are two worlds with the same individuals, a third world with no individuals which is able to distinguish them, and a fourth world which is unable to distinguish them. Such a structure can only be generated by a structure which does not have constant higher-order domains.

# 1.5 Comprehension

## 1.5.1 Williamson on Comprehension Principles

Williamson (2013, chapter 6) argues that a good higher-order contingentist theory has to entail adequate comprehension principles. On the one hand, these comprehension principles must be sufficiently strong. E.g., on the abundant conception of higher-order entities we are concerned with here, they should allow us to demonstrate that these entities are closed under complementation and intersection. On the other hand, the comprehension principles it satisfies must not be too strong. E.g., Williamson claims that to be true to the contingentist spirit, the comprehension principles must not entail the existence of the haecceity of every possible individual, where the haecceity of an individual is a property which necessarily applies to that individual and that individual alone.

More systematically, Williamson argues that a good higher-order contingentist theory should entail two restricted comprehension principles he calls $\text{Comp}^-$ and $\text{Comp}_C$, while it should not entail the unrestricted comprehension principle he calls Comp. We show that on the formalizations of the Fine-Stalnaker view presented above, the view stands up well to these constraints – they validate the first two principles and invalidate the third.

Since Williamson thinks that contingentists should accept the being constraint, his formulations of these principles are stated in a way that fits negative semantics. Here, we consider both negative and positive semantics, and consequently distinguish between negative variants $-\text{Comp}_C$ and $-\text{Comp}$, and positive variants $+\text{Comp}_C$ and $+\text{Comp}$. We can use the same version of $\text{Comp}^-$ in both cases. As usual, these comprehension principles are schematic, and can be instantiated using an arbitrary formula; consequently, they are relative to the choice of a signature. To simplify the discussion in this section, we fix an arbitrary signature $\sigma$.

To state the comprehension principles, we first introduce a formula $E^{\bar{v}}_{\varphi}$ for any formula $\varphi$ and sequence of variables $\bar{v}$ which expresses that the values of all constants in $\varphi$ and free variables not among $\bar{v}$ in $\varphi$ exist; to do so, we simply let it be the conjunction of $\exists v(v \overset{\text{HI}}{\sim} \varepsilon)$ for all $\varepsilon$ which are a constant in $\varphi$ or a free variable not among $\bar{v}$ in $\varphi$, where $v$ is a variable of the same type as $\varepsilon$. We can now define the comprehension principles as follows:

Comp$^-$: $\exists V \forall \bar{v}(V\bar{v} \leftrightarrow \varphi)$

$-$Comp$_C$: $E^{\bar{v}}_{\varphi} \rightarrow \exists V \Box \forall \bar{v}(V\bar{v} \leftrightarrow \varphi)$

$+$Comp$_C$: $E^{\bar{v}}_{\varphi} \rightarrow \exists V \Box \forall v_1 \dots \Box \forall v_n \Box (V\bar{v} \leftrightarrow \varphi)$

$-$Comp: $\exists V \Box \forall \bar{v}(V\bar{v} \leftrightarrow \varphi)$

$+$Comp: $\exists V \Box \forall v_1 \dots \Box \forall v_n \Box (V\bar{v} \leftrightarrow \varphi)$

An instance of one of these comprehension schemas is obtained by letting $\varphi$ be any formula not containing free occurrences of $V$, letting $V$ be a variable of some type $\langle t_1, \dots, t_n \rangle$, and $v_1, \dots, v_n$ be variables of types $t_1, \dots, t_n$, and prefixing the formula with any string of universal quantifiers and $\Box$ operators which renders the resulting formula closed. For the purposes of validity and consequence, we identify a schema with the class of its instances.

Before we show that both Comp$^-$ and Comp$_C$ are valid on closed models, it should be noted that Williamson's discussion of these comprehension principles is couched in terms of unrestricted quantifiers. But since he takes the restriction to hereditarily intensional relations to be vacuous (Williamson, 2013, p. 266), this is dialectically unproblematic.

To prove the validity of the two comprehension principles on closed models, we introduce a way of specifying variants of assignments:

**Definition 1.5.1.** *Let* $\mathfrak{S} = \langle W, I, D \rangle$ *be a structure and* $a$ *an* assignment for $\mathfrak{S}$. *For any variable* $v$ *of type* $t$ *and* $o \in D^t_W$, *we let* $a[o/v]$ *be the assignment such that* $\text{dom}(a[o/v]) = \text{dom}(a) \cup \{v\}$, $a[o/v](v) = o$, *and* $a[o/v](x) = a(x)$ *for all* $x \in \text{dom}(a) \setminus \{v\}$. *We extend this notation to finite sequences by defining* $a[\bar{o}/\bar{v}] = a[o_1/v_1] \dots [o_n/v_n]$.

**Proposition 1.5.2.** *For each sign* $\times$, $\vDash_{C\times}$ Comp$^-$.

*Proof.* Consider any $\varphi$, $V$ and $\bar{v}$ with which an instance of Comp$^-$ can be obtained, $\times$closed model $\mathfrak{M} = \langle W, I, D, V, w\rangle$, $v \in W$ and assignment $a$ admissible for $\varphi$. Define $o$ such that $o(u) = \emptyset$ for all $u \neq v$ and $o(v) = \{\bar{r} \in \Pi_{i\leq n} D_v^{t_i} : \mathfrak{M}, v, a[\bar{r}/\bar{v}] \vDash \varphi\}$. By construction, $\mathfrak{M}, v, a[o/V] \vDash \forall\bar{v}(V\bar{v} \leftrightarrow \varphi)$. Note that $D \boxtimes o$. Let $\mathfrak{S} = \langle W, I, D\rangle$, and consider any $\xi \in \text{fix}(\mathfrak{S}, v)$. Since $\xi.o' = o'$ for all $o' \in D_v^T$ and $\xi.v = v$, $\xi \in \text{fix}(\mathfrak{S}, o)$. As $\mathfrak{S}$ is $\times$closed, it follows that $o \in D_v^T$. Hence $\mathfrak{M}, v, a \vDash \exists V\forall\bar{v}(V\bar{v} \leftrightarrow \varphi)$. Since $v$ and $a$ were chosen arbitrarily, any closure of this formula by universal quantifiers and $\Box$ operators is true in $\mathfrak{M}$ as well. $\Box$

To prove the validity of Comp$_C$, we first introduce a notation for the intension expressed by an open formula.

**Definition 1.5.3.** *Let $\times$ be a sign, $\mathfrak{M} = \langle W, I, D, V, w\rangle$ a model and $\bar{t}$ a sequence of types. For any formula $\varphi$, sequence of variables $\bar{x}$ of types $\bar{t}$, and assignment $a$ for $\langle W, I, D\rangle$ admissible for $\varphi$, define $\varphi(\bar{x})^\times_{\mathfrak{M},a}$ to be the element of $\iota^{\bar{t}}_{\langle W,I\rangle}$ such that for all $v \in W$:*

$$\varphi(\bar{x})^-_{\mathfrak{M},a}(v) = \{\bar{o} \in \Pi_{i<n} D_v^{t_i} : \mathfrak{M}, v, a[\bar{o}/\bar{x}] \vDash \varphi\}$$

$$\varphi(\bar{x})^+_{\mathfrak{M},a}(v) = \{\bar{o} \in \Pi_{i<n} D_W^{t_i} : \mathfrak{M}, v, a[\bar{o}/\bar{x}] \vDash \varphi\}$$

Further, we introduce a way of applying an automorphism of a structure to a model based on it and an assignment for it:

**Definition 1.5.4.** *Let $\mathfrak{M} = \langle W, I, D, V, w\rangle$ be a model, $a$ an assignment for $\mathfrak{S} = \langle W, I, D\rangle$ and $\xi \in \text{aut}(\mathfrak{S})$. Define $\xi.\mathfrak{M} = \langle W, I, D, \xi.V, \xi.w\rangle$, where $\xi.V$ is the function on $\text{dom}(V)$ such that $(\xi.V)(c) = \xi.(V(c))$ for all $c \in \text{dom}(V)$, and define $\xi.a$ to be the function on $\text{dom}(a)$ such that $(\xi.a)(v) = \xi.(a(v))$ for all $v \in \text{dom}(a)$.*

With this, we can show that truth of a formula is invariant under applying an automorphisms to the parameters:

**Lemma 1.5.5.** *For any model $\mathfrak{M} = \langle W, I, D, V, w\rangle$, $v \in W$, formula $\varphi$, assignment $a$ admissible for $\varphi$ and $\xi \in \text{aut}(\langle W, I, D\rangle)$,*

$$\mathfrak{M}, v, a \vDash \varphi \text{ iff } \xi.\mathfrak{M}, \xi.v, \xi.a \vDash \varphi.$$

*Proof.* By induction on the complexity of $\varphi$. $\square$

**Proposition 1.5.6.** *For each sign* $\times$*,* $\vDash_{C\times} \times\mathrm{Comp}_C$.

*Proof.* Consider any $\varphi$, $V$ and $\bar{v}$ with which an instance of $\times\mathrm{Comp}_C$ can be obtained, $\times$closed model $\mathfrak{M} = \langle W, I, D, V, w\rangle$, $v \in W$ and assignment $a$ admissible for $\varphi$ such that $\mathfrak{M}, v, a \vDash E^{\bar{v}}_{\varphi}$. Let $o = \varphi^{\times}_{\mathfrak{M},a}(\bar{v})$, so by construction, if $\times = -$, $\mathfrak{M}, v, a[o/V] \vDash \Box\forall\bar{v}(V\bar{v} \leftrightarrow \varphi)$, and if $\times = +$, $\mathfrak{M}, v, a[o/V] \vDash \Box\forall v_1 \ldots \Box\forall v_n\Box(V\bar{v} \leftrightarrow \varphi)$. We show that $o \in D^t_v$. Since $D \boxtimes o$ and $\mathfrak{S} = \langle W, I, D\rangle$ is $\times$closed, it suffices to show that $\mathrm{fix}(\mathfrak{S}, v) \subseteq \mathrm{aut}(\mathfrak{S})_o$. So consider any $\xi \in \mathrm{fix}(\mathfrak{S}, v)$. To show that $\xi \in \mathrm{aut}(\mathfrak{S})_o$, we have to show that $\xi.o = o$, which is equivalent to the claim that $\xi.(o(u)) = o(\xi.u)$ for all $u \in W$. So let $u \in W$. We distinguish two cases.

*Case 1:* $\times = +$. Then we can prove $\xi.(o(u)) = o(\xi.u)$ by showing that for all $\bar{r} \in \Pi_{i\leq n} D^{t_i}_W$, $\mathfrak{M}, u, a[\bar{r}/\bar{v}] \vDash \varphi$ iff $\mathfrak{M}, \xi.u, a[\xi.\bar{r}/\bar{v}] \vDash \varphi$.

*Case 2:* $\times = -$. Then we can prove $\xi.(o(u)) = o(\xi.u)$ by showing that for all $\bar{r} \in \Pi_{i\leq n} D^{t_i}_u$, $\mathfrak{M}, u, a[\bar{r}/\bar{v}] \vDash \varphi$ iff $\mathfrak{M}, \xi.u, a[\xi.\bar{r}/\bar{v}] \vDash \varphi$.

It suffices to show the first claim, as it entails the second. So consider any $\bar{r} \in \Pi_{i\leq n} D^{t_i}_W$. Since the values of all constants and free variables of $\varphi$ apart from those in $\bar{v}$ under $a$ are in $D^T_v$, they are mapped to themselves by $\xi$, so $\mathfrak{M}, \xi.u, a[\xi.\bar{r}/\bar{v}] \vDash \varphi$ is the case iff $\xi.\mathfrak{M}, \xi.u, \xi.(a[\bar{r}/\bar{v}]) \vDash \varphi$. Hence the claim to be proven follows by Lemma 1.5.5.

Therefore, if $\times = -$, $\mathfrak{M}, v, a \vDash E^{\bar{v}}_{\varphi} \to \exists V\Box\forall\bar{v}(V\bar{v} \leftrightarrow \varphi)$, and if $\times = +$, $\mathfrak{M}, v, a \vDash E^{\bar{v}}_{\varphi} \to \exists V\Box\forall v_1 \ldots \Box\forall v_n\Box(V\bar{v} \leftrightarrow \varphi)$. As in the proof of Proposition 1.5.2, the claim to be prove follows. $\square$

$\mathrm{Comp}_C$ is a very strong comprehension principle. In the present infinitary setting, we can even show that it entails $\mathrm{Comp}^-$.

**Proposition 1.5.7.** *For each sign* $\times$*,* $\times\mathrm{Comp}_C \vDash \mathrm{Comp}^-$.

*Proof.* Consider any $\varphi$, $V$ and $\bar{v}$ with which an instance of $\mathrm{Comp}^-$ can be obtained, model $\mathfrak{M} = \langle W, I, D, V, w\rangle$ such that $\mathfrak{M} \vDash \times\mathrm{Comp}_C$, $v \in W$ and assignment $a$ admissible for $\varphi$. We prove that $\mathfrak{M}, v, a \vDash \exists V\forall\bar{v}(V\bar{v} \leftrightarrow \varphi)$. Let $X = \{\bar{o} \in \Pi_{i\leq n} D^{t_i}_v : \mathfrak{M}, v, a[\bar{o}/\bar{v}] \vDash \varphi\}$. Let $\nu$ be an injection from $D^T_v$

to the class of variables, such that for all types $t$ and $o \in D_v^t$, $\nu(o)$ is of type $t$, and define $\psi = \bigvee_{\bar{o}\in X} \bigwedge_{i\leq n} v_i = \nu(o_i)$. Then by $\times\mathrm{Comp}_C$, $\mathfrak{M}, v, \nu^{-1} \vDash \exists V\Box\forall\bar{v}(V\bar{v} \leftrightarrow \psi)$. So there is an $o \in D_v^t$ such that $o(v) \cap \Pi_{i\leq n} D_v^{t_i} = X$ and $\mathfrak{M}, v, a[o/V] \vDash \forall\bar{v}(V\bar{v} \leftrightarrow \varphi)$, from which the claim follows. $\Box$

$\mathrm{Comp}_C$ also entails that the intensions at any world are closed under conjunction in the sense made precise in Williamson (2013, p. 281). More generally, we can adapt the proof in Fine (1977b, pp. 154–155) to show that the higher-order domain of any world in any closed structure forms a complete atomic Boolean algebra.

According to the Fine-Stalnaker view of higher-order contingency, higher-order entities have to satisfy certain non-trivial conditions to exist at a world. It is therefore unsurprising that the unrestricted comprehension principle Comp is not valid on the class of closed models.

Here and in the following, it will be helpful to have a concise way of specifying the intensions of type $\langle\rangle$ which are true in a certain set of worlds or a certain world (a world-proposition). We therefore give a general definition:

**Definition 1.5.8.** *Let $\mathfrak{F} = \langle W, I\rangle$ be a frame. For any $X \subseteq W$ and $w \in W$, define $X_{\mathfrak{F}}^{\langle\rangle}$ and $w_{\mathfrak{F}}^{\langle\rangle}$ to be the elements of $\iota_{\mathfrak{F}}^{\langle\rangle}$ such that for all $v \in W$:*

$$X_{\mathfrak{F}}^{\langle\rangle}(v) = \{\langle\rangle : v \in X\}$$

$$w_{\mathfrak{F}}^{\langle\rangle}(v) = \{\langle\rangle : v = w\}$$

We call such an intension the *representation* of the relevant set of worlds or world in the relevant frame. Specifying the type in the index is necessary in general, as we will later introduce representations of other kinds of entities, but in cases where it is clear from context, we will drop it.

**Proposition 1.5.9.** *For each sign $\times$, $\nvDash_{C\times} \times\mathrm{Comp}$.*

*Proof.* Let $\mathfrak{S} = \langle W, I, B\rangle$ be the structure such that $W = \{1, 2, 3\}$, $I = \emptyset$ and $B$ is the domain assignment such that $B_W^T = \emptyset$. Let $\otimes\mathfrak{S} = \langle W, I, D\rangle$. It is easy to check that $1_{\langle W,I\rangle}$, the intension which is only true in 1, is in $D_1^{\langle\rangle}$ but not in $D_2^{\langle\rangle}$, and therefore for any model $\mathfrak{M}$ based on $\mathfrak{S}$, $\mathfrak{M} \nVdash \forall p\Box\exists q\Box(q \leftrightarrow p)$.

Since this is an instance of $\times$Comp, this comprehension principle is invalid on $C\times$. □

As far as the three comprehension principles discussed so far are concerned, the Fine-Stalnaker view holds up well to the challenges raised in Williamson (2013) – any version of the model theory considered here satisfies both restricted comprehension principles but invalidates the unrestricted comprehension principle. As pointed out in section 1.2.2, there is one remaining constraint proposed by Williamson which the Fine-Stalnaker view does not meet, namely to validate the inference from (17) to (18) (see footnote 5). In Williamson (2013, appendix 6.7), it is shown that this inference is not licensed by Comp$^-$ and $-$Comp$_C$, and it is routine to generalize this proof to show that on $C-$, not every instance of (17) entails the corresponding instance of (18). We conjecture the same holds for the other classes of models we have been and will be considering. As described in section 1.2.5, we don't take this to be a successful argument against the Fine-Stalnaker view, although it does point to what we take to be the most pressing challenge for the view, namely the inexpressibility of certain modalized generalized quantifiers; see Part 3.

### 1.5.2 World-Propositions

Both forms of the Fine-Stalnaker view address Williamson's challenge concerning unification, since they are based on a single unified principle governing the existence of relation which validates both Comp$_C$ and Comp$^-$ without validating Comp. Given Proposition 1.5.7, one might wonder whether a theory which consisted solely of all instances of Comp$_C$ would not also meet these demands. Williamson seems to think not, since he points out that contingentists might have to postulate Comp$^-$ in addition to Comp$_C$, as Comp$^-$ is not *derivable* from Comp$_C$. (See (Williamson, 2013, p. 285); note that Williamson does not make the notion of derivability formally precise.)

However, Williamson's appeal to a notion of derivability also shows that the unrestricted comprehension principle Comp he advocates may be less powerful than he himself makes it out to be, since from the finitary instances of Comp, we cannot derive (in a certain natural proof system) the claim that necessarily, there is a true proposition which strictly entails every proposition or its negation. Formally, this claim can be stated as follows:

Atom: $\Box\exists p^{\langle\rangle}(p \wedge \forall q^{\langle\rangle}(\Box(p \to q) \vee \Box(p \to \neg q)))$

That Atom is not derivable from finitary Comp was proven in Gallin (1975, p. 113, Theorem 15.1), adapting the technique of Boolean-valued models from the literature on forcing in set theory. (In the context of propositional quantifiers, such a principle was missing in an axiomatization suggest by Kripke (1959), as noted in Kaplan (1970).)

Gallin's result is essentially an *underivability* result in a finitary language. If we instead consider the model-theoretic consequence relation among sentences of the infinitary language $\mathcal{L}$, it seems safe to conjecture that Atom does in fact follow from $\text{Comp}_C$ (at any world, take the conjunction of all existing true propositions and negations of existing false propositions) and therefore from Comp.

The upshot is that neither a model-theoretic nor a proof-theoretic notion of entailment does the work required to make Williamson's unification challenge pressing: On the first, $\text{Comp}^-$ and Atom follow from $\text{Comp}_C$ (and all of these follow from Comp), so in this sense neither the higher-order contingentist nor the higher-order necessitist faces any disunity. On the second, $\text{Comp}^-$, $\text{Comp}_C$ and Atom are pairwise independent, and the Fine-Stalnaker view is in this sense disunified, but so is the view of the higher-order contingentist who advocates Comp, since Atom does not follow from it. In the second case, one might appeal to the number of independent principles favoring higher-order necessitism, but this would clearly require a more systematic way of selecting the principles used to compare the two theories.

We now turn to a second issue concerning world-propositions. In the next section, we will consider ways of capturing the Fine-Stalnaker view using object-language comprehension principles. To do so, it will be crucial to be able to rely on every world containing its world-proposition. This is a stronger condition than being differentiated; following Fine (1977b, p. 163), we adopt the following terminology:

**Definition 1.5.10.** *A structure* $\langle W, I, D\rangle$ *is* world-selective *if* $w^{\langle\rangle}_{\langle W,I\rangle} \in D^{\langle\rangle}_w$ *for all* $w \in W$. *A model is* world-selective *if it is based on a world-selective structure.*

Verifying Atom does not suffice for being world-selective. Given the expressive limitations of $\mathcal{L}$ which will be discussed in Part 3, it even seems plausible that there are models which are not world-selective despite verifying $\text{Comp}_C$ (which semantically entails Atom over the class of all models) and the fact that all models are based on differentiated structures. In the presence of $\text{Comp}_C$, we can impose the condition of being world-selective by adding the following strengthened version of Atom:

Atom*: $\Box\exists p^{\langle\rangle}(p\wedge\bigwedge_{\bar{t}\text{ types}}\Box\forall X^{\bar{t}}\Box\forall x_1^{t_1}\ldots\Box\forall x_n^{t_n}(\Box(p\rightarrow X\bar{x})\vee\Box(p\rightarrow\neg X\bar{x})))$

**Proposition 1.5.11.** *Let* $\times$ *be a sign and* $\mathfrak{M}$ *a* $\times$ *model verifying* $\times\text{Comp}_C$. $\mathfrak{M}$ *is world-selective if and only if* $\mathfrak{M}\vDash$ Atom*.

*Proof.* The 'only if' direction is immediate. For the 'if' direction, assume $\mathfrak{M} = \langle W, I, D, V, w\rangle$ satisfies Atom*, and consider any $v \in W$. Let $P \in \iota^{\langle\rangle}_{\mathfrak{F}}$ witness the existential claim of Atom* at $v$. Consider any $u \in W\backslash\{v\}$. By differentiation, either $D^T_v \neq D^T_u$ or there is a type $t \neq e$ and $o \in D^t_W$ such that $o(v) \neq o(u)$. In the first case, let $o$ be a member of exactly one of $D^T_v$ and $D^T_u$. Then by $\times\text{Comp}_C$, $D^{\langle\rangle}_W$ contains a propositional intension which is true in a world if and only if $o$ exists at it, and so by Atom*, $P$ is false in $u$. In the second case, let $o \in D^{\bar{t}}_W$ and $\bar{o} \in \Pi_{i\leq n}D^{t_i}_W$ such that not $\bar{o} \in o(v)$ iff $\bar{o} \in o(u)$. Using Atom*, $o$ and $\bar{o}$ witness that $P$ is false in $u$. Therefore $P = w^{\langle\rangle}_{\langle W,I\rangle}$, as required. □

### 1.5.3 A Comprehension Principle for Closure

Williamson discusses higher-order contingentist theories in the form of comprehension principles; in contrast, we have discussed the Fine-Stalnaker view in terms of a class of models. However, in the very rich type-theoretic setting we are working in, this is not the only option. We can also try to capture the ideas behind qualitative generation and higher-order closure with object-language comprehension principles. Structurally, the resulting principles are analogous to $\mathrm{Comp}_C$; we only replace the existence condition for parameters $E_\varphi$ with a condition stating that $\varphi$ has the property of invariance under the relevant permutations. For simplicity, we will first consider the case of +closure in detail, and return to negative semantics in section 1.6.4 and to generation in section 1.6.5.

To be able to formulate a comprehension principle for +closure, we successively define the relevant concepts. We first have to define how to express quantification over possible entities. Following Fine (1977a), we start with defining what it is for a proposition to be a world-proposition – a proposition which is true in only one world. We can then define what it is for a formula to be true at a world:

$$\textsc{world}(w^{\langle\rangle}) := \Diamond\forall p^{\langle\rangle}(p \leftrightarrow \Box(w \to p))$$

$$\textsc{at}(w^{\langle\rangle}, \varphi) := \textsc{world}(w) \wedge \Box(w \to \varphi)$$

Note that these definitions do not express the intended conditions on all models. E.g., a propositional intension $o$ may satisfy WORLD while being true in more than one world if the propositional domain is sparse enough. This is where the assumption of being world-selective comes into play: from now on, we will only consider world-selective models. As shown above, since all models are by definition based on differentiated structures, this can be enforced in the object language using $\mathrm{Comp}_C$ and $\mathrm{Atom}^*$.

World-propositions allow us to define quantifiers ranging over possible entities, sometimes called *outer quantifiers*, for which we use $\Pi$ and $\Sigma$. The idea behind the definitions of these complex quantifiers is to bind a variable $w$ to the actual world; then we can – in the case of the universal quantifier – express outer quantification by "necessarily for all", using AT and $w$ to evaluate the complement clause of the quantified claim in the world at which we started. For any variable $v$ of type $t$ and formula $\varphi$, we therefore define:

$$\Pi v^t\varphi := \exists w^{\langle\rangle}(\textsc{world}(w) \wedge w \wedge \Box\forall v\,\textsc{at}(w, \varphi))$$

$$\Sigma v^t\varphi := \exists w^{\langle\rangle}(\textsc{world}(w) \wedge w \wedge \Diamond\exists v\,\textsc{at}(w, \varphi))$$

Without being very precise about it, we will assume that the variable $w$ used in the definition of such an outer quantification does not occur in the relevant formula $\varphi$. As they will occur repeatedly, we introduce abbreviations for propositional outer quantifiers restricted to worlds-propositions:

$$\Pi^{\textsc{world}} w^{\langle\rangle} \varphi := \Pi w^{\langle\rangle}(\textsc{world}(w) \rightarrow \varphi)$$

$$\Sigma^{\textsc{world}} w^{\langle\rangle} \varphi := \Sigma w^{\langle\rangle}(\textsc{world}(w) \wedge \varphi)$$

Since this definition straightforwardly generalizes to finite strings of outer quantifiers, we will write any such string $\Pi v_1 \ldots \Pi v_n$ as $\Pi v_1 \ldots v_n$. The fact that such definitions cannot be extended to infinite sets of variables will play a crucial role in Part 3.

To be able to capture closure in the formal language, we have to be able to talk about permutations of worlds and permutations of possible individuals. Such permutations are themselves higher-order entities. In the present setting, we may treat them as functional binary relations. Since there is also no type for worlds, we can understand a permutation of worlds to be a functional relation among world-propositions. Thus to talk about permutations of worlds and permutations of possible individuals, we will use variables of type $\langle\langle\rangle,\langle\rangle\rangle$ and $\langle e, e\rangle$. Given the fact that intensions can have different extensions at different worlds, that there are different ways of understanding permutations as relations. Here, we take the simplest option of requiring representations of permutations to have the same extension at every world. We will reconsider this decision in section 1.6.3. Formally, we define what it is for a relation of the relevant type to be a permutation as follows:

$$\begin{aligned}
&\textsc{wperm}(X^{\langle\langle\rangle,\langle\rangle\rangle}) := \\
&\quad \Pi w^{\langle\rangle} v^{\langle\rangle}(\Diamond Xwv \rightarrow (\textsc{world}(w) \wedge \textsc{world}(v)))\wedge \\
&\quad \Pi^{\textsc{world}} w( \\
&\qquad \Sigma^{\textsc{world}} v \Box(\Pi^{\textsc{world}} u Xwu \leftrightarrow v \overset{\textsc{hi}}{\sim} u)\wedge \\
&\qquad \Sigma^{\textsc{world}} v \Box(\Pi^{\textsc{world}} u Xuw \leftrightarrow v \overset{\textsc{hi}}{\sim} u) \\
&\quad )
\end{aligned}$$

$\textsc{iperm}(X^{\langle e,e\rangle}) := \Pi x^e(\Sigma y^e\Box(\Pi z^e Xxz \leftrightarrow y = z) \wedge \Sigma y^e\Box(\Pi z^e Xzx \leftrightarrow y = z))$

Next, we define for any such permutations $X^{\langle\langle\rangle,\langle\rangle\rangle}$ and $Y^{\langle e,e\rangle}$ what it is for them to map one entity to another. As in the model-theoretic case, we do so by induction on the complexity of types:

$\textsc{map}(X^{\langle\langle\rangle,\langle\rangle\rangle}, Y^{\langle e,e\rangle}, v^e, u^e) := \; Yvu$

$\textsc{map}(X^{\langle\langle\rangle,\langle\rangle\rangle}, Y^{\langle e,e\rangle}, V^{\bar{t}}, U^{\bar{t}}) := \; \Pi z_1^{t_1} \dots z_n^{t_n} z_1'^{t_1} \dots z_n'^{t_n} \Pi^{\textsc{world}} wv$

$\quad \big(\big(Xwv \wedge \bigwedge_{i \leq n} \textsc{map}(X, Y, z_i, z_i')\big) \rightarrow (\textsc{at}(w, V\bar{z}) \leftrightarrow \textsc{at}(v, U\bar{z}'))\big)$

Using these definitions, we can express what it is for $X$ and $Y$ to constitute an automorphism; $X$ must represent a permutation of worlds, $Y$ a permutation of individuals, and for any type $t$ and pair of worlds $w$ and $v$ such that $Xwv$, $X$ and $Y$ must map the domain of type $t$ at $w$ to the domain of type $t$ at $v$:

$\textsc{aut}(X^{\langle\langle\rangle,\langle\rangle\rangle}, Y^{\langle e,e\rangle}) :=$

$\quad \textsc{wperm}(X) \wedge \textsc{iperm}(Y) \wedge$

$\quad \Pi^{\textsc{world}} wv(Xwv \rightarrow \bigwedge_{t \in T}($

$\quad\quad \textsc{at}(w, \forall V^t \Pi U^t(\textsc{map}(X, Y, V, U) \rightarrow \textsc{at}(v, \exists T^t T \overset{\textsc{hi}}{\sim} U))) \wedge$

$\quad\quad \textsc{at}(v, \forall U^t \textsc{at}(w, \exists V^t(\textsc{map}(X, Y, V, U))))$

$\quad ))$

Recall that we are only interested in the possible automorphisms of modal space that respect the identities of the actual world and the entities in it. We formalize this as follows:

$\textsc{fix}(X^{\langle\langle\rangle,\langle\rangle\rangle}, Y^{\langle e,e\rangle}) :=$

$\quad \textsc{aut}(X, Y) \wedge \Pi^{\textsc{world}} w^{\langle\rangle}(w \rightarrow Xww) \wedge \bigwedge_{t \in T} \forall V^t \textsc{map}(X, Y, V, V)$

We now need to formulate in $\mathcal{L}$ the claim that if FIX holds of $X$ and $Y$, then $X$ and $Y$ preserve the intension expressed by an open formula. That is, we have to write down the condition for an intension expressed by a formula $\varphi$ (abstracting over a given sequence of variables $\bar{v}$) to be preserved by $X$ and $Y$, i.e., to be mapped to itself by them. We can do so as follows:

$$\textsc{pres}(X^{\langle\langle\rangle,\langle\rangle\rangle}, Y^{\langle e,e\rangle}, \varphi, \bar{v}) := \Pi z_1^{t_1} \dots z_n^{t_n} z_1'^{t_1} \dots z_n'^{t_n} \Pi^{\textsc{world}} wv$$
$$\big(\big(Xwv \wedge \textstyle\bigwedge_{i\leq n} \textsc{map}(X, Y, z_i, z_i')\big) \rightarrow (\textsc{at}(w, \varphi[\bar{z}/\bar{v}]) \leftrightarrow \textsc{at}(v, \varphi[\bar{z}'/\bar{v}]))\big)$$

As usual, $\chi[\bar{x}/\bar{y}]$ is the result of replacing free occurrences of $y_i$ in $\chi$ by $x_i$, for all $i \leq n$.

We can now put the pieces together to formulate the higher-order closure view as a comprehension principle. As discussed in section 1.2.4, it is natural to formulate it by quantifying over *possible* permutations, and we follows this option here. In section 1.6.6, we will briefly explore a variant view which is formulated using quantification over existing permutations, rather than possible permutations, and conclude that this variant has highly implausible consequences. We therefore state the comprehension principle as follows, where instances of this schema are obtained as from the other comprehension principles (*FS* stands for "Fine-Stalnaker"):

+Comp$_{FS}$:=

$$\Pi XY(\textsc{fix}(X, Y) \rightarrow \textsc{pres}(X, Y, \varphi, \bar{v})) \rightarrow \exists V \Box \forall v_1 \dots \Box \forall v_n \Box (V\bar{v} \leftrightarrow \varphi)$$

With +Comp$_{FS}$, we have a way of stating the Fine-Stalnaker view in its higher-order closure form in the formal object language (at least in a positive setting, where we don't impose the being constraint).

### 1.5.4 Invalidity

How does +Comp$_{FS}$ relate to $C+$, the class of world-selective +closed models? It turns out that it doesn't define it; among world-selective +models, those verifying +Comp$_{FS}$ are not all and only those that are +closed. This would not be too surprising if only the left-to-right direction failed, i.e., if some world-selective +model verifies +Comp$_{FS}$ without being +closed. This kind of mismatch might be explained by the limited expressivity of $\mathcal{L}$ and therefore might not constitute a problem for the view under consideration; we will in fact establish exactly this kind of expressive limitations in a discussion of $\times$Comp$_C$ in Part 3. However, we will now show that right-to-left direction fails: there are world-selective +closed models which do not verify

+Comp$_{FS}$. Since +Comp$_{FS}$ is simply a formalized statement of the higher-order closure version of the Fine-Stalnaker view in a positive semantics, this result therefore also shows that the model theory of (world-selective) +closed models is inadequate for its intended purpose, since it fails to validate the formalized statement of the view.

Abstractly, it is not hard to see how $C+$ might fail to validate +Comp$_{FS}$. The semantic condition of +closure quantifies over all permutations of worlds and individuals of a model; in contrast, when we evaluate +Comp$_{FS}$ on a model, the initial two outer universal quantifiers get interpreted as quantification over permutations whose representations exist *at some world of the model.* In this section, we show that this abstract possibility is realized, by exhibiting a world-selective +closed models which does not verify +Comp$_{FS}$.

We begin by retracing the steps in the sequence of definitions that allowed us to write down +Comp$_{FS}$ and draw out their semantic consequences. The last step of this gives us the condition imposed by +Comp$_{FS}$. Since specifying the semantic conditions expressed by the syntactic constructions defined above is a routine exercise, we only give the details for the last step, the condition expressed by +Comp$_{FS}$. To be able to state it, we first define semantically how permutations of worlds and individuals can be represented as intensions of types $\langle\langle\rangle,\langle\rangle\rangle$ and $\langle e,e\rangle$, extending the notation of Definition 1.5.8. Since permutations are functions and we understand functions as functional relations, we state the definition more generally for relations among worlds and individuals:

**Definition 1.5.12.** *Let $\mathfrak{F} = \langle W, I\rangle$ be a frame. For any $R \subseteq I^2$ and $Z \subseteq W^2$, define $R_{\mathfrak{F}}^{\langle e,e\rangle} \in \iota_{\mathfrak{F}}^{\langle e,e\rangle}$ and $Z_{\mathfrak{F}}^{\langle\langle\rangle,\langle\rangle\rangle} \in \iota_{\mathfrak{F}}^{\langle\langle\rangle,\langle\rangle\rangle}$ such that for all $w \in W$:*

$$R_{\mathfrak{F}}^{\langle e,e\rangle}(w) = R$$

$$Z_{\mathfrak{F}}^{\langle\langle\rangle,\langle\rangle\rangle}(w) = \{\langle v_{\mathfrak{F}}, u_{\mathfrak{F}}\rangle : \langle v, u\rangle \in Z\}$$

As with the representations defined in Definition 1.5.8, we drop the type index when it is clear from the context. It will be useful to note that in all four cases of representations in these two definitions, the function mapping the relevant set of entities to their representations is injective. Furthermore, we state the following very useful lemma:

**Lemma 1.5.13.** *Let* $\mathfrak{F} = \langle W, I\rangle$ *be a frame and* $\xi = \langle f, g\rangle \in \mathrm{aut}(\mathfrak{F})$.

*(i) For any* $X \subseteq W$, $\xi.X_{\mathfrak{F}} = (\xi.X)_{\mathfrak{F}}$.

*(ii) For any* $w \in W$, $\xi.w_{\mathfrak{F}} = (\xi.w)_{\mathfrak{F}}$.

*(iii) For any* $R \subseteq I^2$, $\xi.R_{\mathfrak{F}} = (g.R)_{\mathfrak{F}}$.

*(iv) For any* $Z \subseteq W^2$, $\xi.Z_{\mathfrak{F}} = (f.Z)_{\mathfrak{F}}$.

*Proof.* Routine. □

We now state the model-theoretic condition expressed by $+\mathrm{Comp}_{FS}$:

**Lemma 1.5.14.** *Let* $\mathfrak{M} = \langle W, I, D, V, w\rangle$ *be a world-selective +model,* $\mathfrak{S} = \langle W, I, D\rangle$ *and* $\mathfrak{F} = \langle W, I\rangle$.

$\mathfrak{M} \vDash +\mathrm{Comp}_{FS}$ *iff for all formulas* $\varphi$*, variables* $\bar{v}$ *of types* $\bar{t}$*,* $w \in W$ *and assignments* $a$ *for* $\mathfrak{S}$ *admissible for* $\varphi$*, the following condition holds:*

*If for all* $f \in S_W$ *and* $g \in S_I$ *such that* $f_{\mathfrak{F}} \in D_W^{\langle\langle\rangle,\langle\rangle\rangle}$, $g_{\mathfrak{F}} \in D_W^{\langle e,e\rangle}$ *and* $\langle f, g\rangle \in \mathrm{fix}(\mathfrak{S}, w)$, $\langle f, g\rangle.\varphi(\bar{v})^+_{\mathfrak{M},a} = \varphi(\bar{v})^+_{\mathfrak{M},a}$*, then* $\varphi(\bar{v})^+_{\mathfrak{M},a} \in D_w^{\bar{t}}$.

*Proof.* Routine. □

This lemma shows that the condition expressed by $+\mathrm{Comp}_{FS}$ may differ from +closure, since it only concerns intensions expressible using some formula $\varphi$. Moreover, as advertised above, unlike +closure it concerns only permutations of worlds and individuals that are in the outer domain of $\langle\langle\rangle,\langle\rangle\rangle$ and the outer domain of $\langle e, e\rangle$, respectively, rather than all such permutations.

We now exhibit a world-selective +closed model that fails to satisfy the condition imposed $+\mathrm{Comp}_{FS}$. The idea behind the model is to take a +closed model of the sort described in section 1.2.5, based on four worlds and no individuals with a minimal higher-order domain function, in which, from any world, all other worlds look exactly alike. We then prove that only the identity permutation of worlds exists in the outer domain of the model, and derive from this fact that the structure does not validate $+\mathrm{Comp}_{FS}$. We start with some notation and lemmas.

We use cycle notation to specify permutations; e.g., in the context of $S_\omega$, we write (02)(345) for the permutation which maps 0 to 2 and 2 to 0; 3 to 4, 4 to 5 and 5 to 3; and all other natural numbers to themselves. Let $G$ be a group. For all $g \in G$, the function mapping each $f \in G$ to $gfg^{-1}$ is a permutation of $G$; this is called *conjugation by* $g$. Moreover, the function mapping each $g \in G$ to the conjugation by $g$ is an action, which we call *conjugation*. Unless indicated otherwise, this is the only action of a group on itself considered below, so we will write $g.f$ for $gfg^{-1}$ without further elaborations.

To construct our example, let $\mathfrak{S}^4 = \langle W, I, B \rangle$ be the structure such that $W = \{1, 2, 3, 4\}$, $I = \emptyset$ and $B$ is the domain assignment such that $B^T_W = \emptyset$. Let $\mathfrak{F}^4 = \langle W, I \rangle$ and $\oplus\mathfrak{S}^4 = \langle W, I, D \rangle$. Since this structure doesn't contain any individuals, automorphisms are simply given by a permutation of worlds; thus, for any $f \in S_W$ let $f^\emptyset = \langle f, \mathrm{id}_\emptyset \rangle$. As $B$ is empty, for any world $w$, $\mathrm{fix}(\mathfrak{S}^4, w)$ includes all automorphisms mapping $w$ to itself:

**Lemma 1.5.15.** *For all* $w \in W$, $\mathrm{fix}(\mathfrak{S}^4, w) = \mathrm{aut}(\mathfrak{S}^4)_w = \{f^\emptyset : f \in (S_W)_w\}$.

*Proof.* Immediate. □

However, only the trivial permutation among worlds is represented in the domain of any world in $\oplus\mathfrak{S}^4$:

**Lemma 1.5.16.** *If* $f \in S_W$ *and* $f_{\mathfrak{F}^4} \in D_W^{\langle\langle\rangle,\langle\rangle\rangle}$ *then* $f = \mathrm{id}$.

*Proof.* Consider any $f \in S_W \backslash \{\mathrm{id}\}$ and $w \in W$; we show that $f_{\mathfrak{F}^4} \notin D_w^{\langle\langle\rangle,\langle\rangle\rangle}$. Since $f \neq \mathrm{id}$, there is a $v \in W$ such that $f(v) \neq v$. Let $u \in \{v, f(v)\} \backslash \{w\}$ and $u' \in W \backslash \{w, v, f(v)\}$; such elements clearly exist. It is routine to show that then, $(uu').f \neq f$, so by Lemma 1.5.13 (iv), $(uu')^\emptyset.f_{\mathfrak{F}^4} \neq f_{\mathfrak{F}^4}$. By Lemma 1.5.15, $(uu')^\emptyset \in \mathrm{fix}(\mathfrak{S}^4, w)$, so $f_{\mathfrak{F}^4} \notin D_w^{\langle\langle\rangle,\langle\rangle\rangle}$. □

Using world-propositions, it is easy to show that there is higher-order contingency in $\oplus\mathfrak{S}^4$:

**Lemma 1.5.17.** *For all* $w, v \in W$, *if* $w \neq v$ *then* $v_{\mathfrak{F}^4} \notin D_w^{\langle\rangle}$.

*Proof.* Let $w \neq v$, and consider any $g \in (S_W)_w$ such that $g(v) \neq v$. Thus $g(v)_{\mathfrak{F}^4} \neq v_{\mathfrak{F}^4}$. By Lemma 1.5.13 (ii), $g^{\emptyset}.v_{\mathfrak{F}^4} = g(v)_{\mathfrak{F}^4}$, so $g^{\emptyset}.v_{\mathfrak{F}^4} \neq v_{\mathfrak{F}^4}$. By Lemma 1.5.15, $g^{\emptyset} \in \text{fix}(\mathfrak{S}^4, w)$, so $v_{\mathfrak{F}^4} \notin D_w^{\langle\rangle}$. □

It follows straightforwardly with the last two lemmas and Lemma 1.5.13 that +Comp$_{FS}$ is not true in any model based on $\oplus\mathfrak{S}^4$, and therefore not valid on $C+$. Since all +closed structures are world-selective, this establishes the central claim of this section.

**Theorem 1.5.18.** $\nvDash_{C+}$ +Comp$_{FS}$.

*Proof.* Let $\mathfrak{M}$ be a model based on $\oplus\mathfrak{S}^4$. Consider any $w, v \in W$ such that $w \neq v$. For all $\xi \in \text{fix}(\mathfrak{S}^4, v)$, $\xi.v = v$, so by Lemma 1.5.13 (ii), $\xi.v_{\mathfrak{F}^4} = v_{\mathfrak{F}^4}$. Trivially, $D \boxplus v_{\mathfrak{F}^4}$, so $v_{\mathfrak{F}^4} \in D_v^{\langle\rangle}$. Consider any assignment $a$ mapping $p^{\langle\rangle}$ to $v_{\mathfrak{F}^4}$. By Lemma 1.5.16, the only $f \in S_W$ such that $f_{\mathfrak{F}^4} \in D_W^{\langle\langle\rangle,\langle\rangle\rangle}$ is id, so by Lemma 1.5.14, $\mathfrak{M}$ only verifies +Comp$_{FS}$ if $v_{\mathfrak{F}^4} \in D_w^{\langle\rangle}$. As shown in Lemma 1.5.17, this is not the case, so $\mathfrak{M} \nvDash$ +Comp$_{FS}$. □

## 1.6 Internal Closure and Internal Generation

Given that +Comp$_{FS}$ is not valid on world-selective +closed models, a defender of the higher-order closure variant of the Fine-Stalnaker view faces a choice: They can either modify the model theory so as to validate the comprehension principle +Comp$_{FS}$ or come up with an alternative object-language statement of their view. In this section, we consider the first option, and develop a more restrictive class of models, which we call *internally +closed*, on which +Comp$_{FS}$ is valid. The second option is considered in Part 3.

### 1.6.1 Internal Closure

The definition of internal closure is basically dictated by Lemma 1.5.14: we want to express the condition expressed by +Comp$_{FS}$, and generalize it in the natural way to remove the language-dependence. Modifying the definition of closure, we have to restrict the quantification over permutations to those whose representation is in the domain of some world of the model. We start by introducing some terminology, defining variants of aut and fix, restricted to pairs of permutations existing at some world. These definitions will be phrased slightly more generally than required here, so that we will be able to use them again when we investigate analogous adaptations of the notion of qualitative generation in section 1.6.5.

**Definition 1.6.1.** *Let* $\mathfrak{S} = \langle W, I, B\rangle$ *and* $\mathfrak{S}' = \langle W, I, D\rangle$ *be structures,* $\mathfrak{F} = \langle W, I\rangle$ *and* $\langle f, g\rangle \in \text{aut}(\mathfrak{F})$*.*

> $f$ *is* possible in $\mathfrak{S}'$ *iff* $f_{\mathfrak{F}} \in D_W^{\langle\langle\rangle,\langle\rangle\rangle}$.
>
> $g$ *is* possible in $\mathfrak{S}'$ *iff* $g_{\mathfrak{F}} \in D_W^{\langle e,e\rangle}$.
>
> $\langle f, g\rangle$ *is* possible in $\mathfrak{S}'$ *iff* $f$ *and* $g$ *are* possible in $\mathfrak{S}'$*.*

*For any* $w \in W$*, define* $\text{aut}(\mathfrak{S})|\mathfrak{S}'$ *to be the set of elements of* $\text{aut}(\mathfrak{S})$ *which are possible in* $\mathfrak{S}'$*, and* $\text{fix}(\mathfrak{S}, w)|\mathfrak{S}'$ *to be the set of elements of* $\text{fix}(\mathfrak{S}, w)$ *which are possible in* $\mathfrak{S}'$*.*

We can now state the condition of being internally closed by replacing $\text{fix}(\mathfrak{S}, w)$ in the definition of closure by $\text{fix}(\mathfrak{S}, w)|\mathfrak{S}$:

**Definition 1.6.2.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure.* $\mathfrak{S}$ *is* internally +closed *if for all* $w \in W$*, types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*,*

$$o \in D^t_w \text{ iff } D \boxplus o \text{ and } \xi.o = o \text{ for all } \xi \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}.$$

*A model is* internally +closed *just in case it is based on an internally +closed structure.*

## 1.6.2 Comprehension and Comparison to Closure

It is straightforward to prove that the restriction to internally +closed models solves the problem described in the last section:

**Proposition 1.6.3.** $+\text{Comp}_{FS}$ *is valid on the class of world-selective internally +closed models.*

*Proof.* With Lemma 1.5.14, this is immediate by definition. □

It is also straightforward to show that internal +closure implies +closure, but not *vice versa*.

**Proposition 1.6.4.** *Any internally +closed model is +closed.*

*Proof.* Consider any model based on an internally +closed structure $\mathfrak{S} = \langle W, I, D\rangle$, $w \in W$, type $t \neq e$ and $o \in \iota^t_{\langle W,I\rangle}$. If $o \in D^t_w$, then since $\mathfrak{S}$ is internally +closed, $D \boxplus o$. Also, since $o \in D^t_w$, $\xi.o = o$ for all $\xi \in \text{fix}(\mathfrak{S}, w)$. If $D \boxplus o$ and $\xi.o = o$ for all $\xi \in \text{fix}(\mathfrak{S}, w)$, then since $\text{fix}(\mathfrak{S}, w)|\mathfrak{S} \subseteq \text{fix}(\mathfrak{S}, w)$, $\xi.o = o$ for all $\xi \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}$. Since $\mathfrak{S}$ is internally +closed, $o \in D^t_w$. So $\mathfrak{S}$ is +closed, and so the model with which we started is +closed as well. □

**Proposition 1.6.5.** *There is a +closed model which is not internally +closed.*

*Proof.* We established in Theorem 1.5.18 that there is a +closed model which does not verify $+\text{Comp}_{FS}$; the claim follows with Proposition 1.6.3. □

As a corollary to Proposition 1.6.4, we immediately obtain that $\text{Comp}^-$ and $+\text{Comp}_C$ are valid on internally +closed structures. We can also extend the result of Proposition 1.5.9 to internally +closed structures and prove that they do not validate unrestricted comprehension:

**Proposition 1.6.6.** +Comp *is not valid on the class of internally +closed models.*

*Proof.* We show that for $\mathfrak{S} = \langle W, I, B\rangle$ used in the proof of Proposition 1.5.9, $\oplus\mathfrak{S} = \langle W, I, D\rangle$ is internally +closed. To do so, we show that for all $w \in W$, $\text{fix}(\oplus\mathfrak{S}, w)|\oplus\mathfrak{S} = \text{fix}(\oplus\mathfrak{S}, w)$. Note that by Lemma 1.4.13, $\text{fix}(\oplus\mathfrak{S}, w) = \{\langle f, \emptyset\rangle : f \in (S_W)_w\}$, so it suffices to prove that $f_{\langle W,I\rangle} \in D^{\langle\langle\rangle,\langle\rangle\rangle}_W$ for all $f \in (S_W)_w$. Letting $W = \{w, v, u\}$, note that $(S_W)_w = \{\text{id}, (vu)\}$. The claim is straightforward for id. We show that $(vu)_{\langle W,I\rangle} \in D^{\langle\langle\rangle,\langle\rangle\rangle}_w$. To do so, it suffices to show that $g.(vu) = (vu)$ for all $g \in (S_W)_w$. The claim is trivial for $g = \text{id}$, so it only remains to show that $(vu).(vu) = (vu)$, which is straightforward. □

### 1.6.3 Cumulative Representations of Permutations

When we defined $+\text{Comp}_{FS}$, we chose to regiment talk of permutations of worlds and individuals as talk of binary relations among world-propositions and individuals which have the same extension at every world. This choice wasn't forced on us; we could also have chosen to allow relations which vary in extension across worlds, but in some other way uniquely specify a way of permuting worlds or individuals. One might wonder whether this choice influenced the resulting theory; in particular, one might wonder whether there is some other way of formalizing in $\mathcal{L}$ quantification over permutations of modal space which gives rise to a different notion of internal closure. One might even hope for such a revision to establish the equivalence with closure. Exemplarily, we consider one natural such revision and show that it determines the same class of structures.

From a model-theoretic perspective, what gives rise to the the mismatch between closure and internal closure is that in some structures, some permutations of worlds and individuals are not represented in the domain of any world. The most natural thought is therefore to look for weaker notions of representation, and the most straightforward way of implementing this is to require a representation of a permutation of individuals to *possibly* relate one individual to another if and only if the first is mapped to the second by the permutation, and weaken the requirement for world permutations analogously. On this proposal, we no longer have a unique representative for a given permutation, but this is of course part of the desired generality.

**Definition 1.6.7.** *Let $\mathfrak{F} = \langle W, I\rangle$ be a frame.*

> *For any $R \subseteq I^2$ and $o \in \iota_{\mathfrak{F}}^{\langle e,e\rangle}$, $o$* cumulatively represents *$R$* in *$\mathfrak{F}$ if $\bigcup_{w\in W} o(w) = R$.*
>
> *For any $Z \subseteq W^2$ and $o \in \iota_{\mathfrak{F}}^{\langle\langle\rangle,\langle\rangle\rangle}$, $o$* cumulatively represents *$Z$* in *$\mathfrak{F}$ if $\bigcup_{w\in W} o(w) = \{\langle w_{\mathfrak{F}}, v_{\mathfrak{F}}\rangle : \langle w, v\rangle \in Z\}$.*

We can now define a condition analogous to internal closure using cumulative representations of permutations, and then show that this change makes no difference – the notions coincide. Since this condition will be less important in what follows, the definitions are somewhat condensed.

**Definition 1.6.8.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure and* $w \in W$*. Define* $\mathrm{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$ *to be the set of* $\langle f, g\rangle \in \mathrm{fix}(\mathfrak{S}, w)$ *such that there is an* $f_c \in D_W^{\langle\langle\rangle,\langle\rangle\rangle}$ *which cumulatively represents* $f$ *and a* $g_c \in D_W^{\langle e,e\rangle}$ *which cumulatively represents* $g$*.*

Since every representation of a permutation cumulatively represents it, it is immediate that in general $\mathrm{fix}(\mathfrak{S}, w)|\mathfrak{S} \subseteq \mathrm{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$.

**Definition 1.6.9.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure.* $\mathfrak{S}$ *is* cumulatively internally +closed *if for all* $w \in W$*, types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*,*

$$o \in D^t_w \text{ iff } D \boxplus o \text{ and } \xi.o = o \text{ for all } \xi \in \mathrm{fix}(\mathfrak{S}, w)|_c\mathfrak{S}.$$

**Lemma 1.6.10.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure,* $\mathfrak{F} = \langle W, I\rangle$*,* $f \in S_W$*,* $g \in S_I$*,* $f_c$ *a cumulative representation of* $f$*,* $g_c$ *a cumulative representation of* $f$*, and* $\xi \in \mathrm{aut}(\mathfrak{S})$*. Then*

$$\text{If } \xi.f_c = f_c \text{ then } \xi.f_{\mathfrak{F}} = f_{\mathfrak{F}},$$

$$\text{If } \xi.g_c = g_c \text{ then } \xi.g_{\mathfrak{F}} = g_{\mathfrak{F}}.$$

*Proof.* We present only the case for $f$; the case for $g$ can be proven in exactly the same way. Let $\xi = \langle h, i\rangle$. Assume $\xi.f_{\mathfrak{F}} \neq f_{\mathfrak{F}}$. Then by Lemma 1.5.13 (iv), $(\xi.f)_{\mathfrak{F}} \neq f_{\mathfrak{F}}$, so by the injectivity of the function mapping every permutation to its representation, $h.f \neq f$. So there is a $w \in W$ such that $hfh^{-1}(w) \neq f(w)$, and therefore $v = h^{-1}(w)$ is such that $hf(v) \neq fh(v)$. Since $f_c$ cumulatively represents $f$, there is a $u \in W$ such that $\langle v_{\mathfrak{F}}, f(v)\rangle \in f_c(u)$. So $\langle \xi.v_{\mathfrak{F}}, \xi.(f(v)_{\mathfrak{F}})\rangle \in \xi.f_c(\xi.u)$. By Lemma 1.5.13 (ii), it follows that $\langle h(v)_{\mathfrak{F}}, (hf(v))_{\mathfrak{F}})\rangle \in \xi.f_c(\xi.u)$. Since $fh(v) \neq hf(v)$ and $f_c$ cumulatively represents $f$, $\langle h(v)_{\mathfrak{F}}, (hf(v))_{\mathfrak{F}})\rangle \in f_c(\xi.u)$. So $\xi.f_c \neq f_c$. □

**Proposition 1.6.11.** *A structure is internally +closed if and only if it is cumulatively internally +closed.*

*Proof.* That every internally +closed structure is cumulatively internally +closed can be proven analogously to Proposition 1.6.4. So consider any cumulatively internally +closed structure $\mathfrak{S} = \langle W, I, D\rangle$, $w \in W$, type $t \neq e$ and $o \in \iota^t_{\langle W,I\rangle}$. Let $\mathfrak{F} = \langle W, I\rangle$. Since $\text{fix}(\mathfrak{S}, w)|\mathfrak{S} \subseteq \text{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$, it is immediate that if $o \in D^t_w$ then $D \boxplus o$ and $\xi.o = o$ for all $\xi \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}$. For the converse direction, assume $o \notin D^t_w$ and $D \boxplus o$. Then there is a $\langle f, g\rangle \in \text{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$ such that $\langle f, g\rangle.o \neq o$. So there are $f_c \in D^{\langle\langle\rangle,\langle\rangle\rangle}_W$ and $g_c \in D^{\langle e,e\rangle}_W$ such that $f_c$ cumulatively represents $f$ and $g_c$ cumulatively represents $g$. As $\mathfrak{S}$ is cumulatively internally +closed, (i) $v_{\mathfrak{F}} \in D^{\langle\rangle}_v$ for all $v \in W$, and (ii) $\xi.f_c = f_c$ and $\xi.g_c = g_c$ for all $\xi \in \text{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$. By (i), $D \boxplus f_{\mathfrak{F}}$ and $D \boxplus g_{\mathfrak{F}}$. With Lemma 1.6.10, it follows from (ii) that $\xi.f_{\mathfrak{F}} = f_{\mathfrak{F}}$ and $\xi.g_{\mathfrak{F}} = g_{\mathfrak{F}}$ for all $\xi \in \text{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$. Hence $\langle f, g\rangle \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}$. □

This result shows that in defining internal closure, it makes no difference whether we quantify only over representations of permutations, or include intensions which merely cumulatively represent permutations. Similar results can be obtained for other ways of understanding quantification over permutations in our type-theoretic setting. E.g., we could understand permutations to be individuals, which can be applied to worlds or individuals using an application relation. We can show that every structure internally closed on such an understanding is internally closed as defined above.

### 1.6.4 Internal Closure in Negative Semantics

So far, we have worked in a positive semantics. Were we instead to assume the being constraint, we could formulate a comprehension principle analogous to +Comp$_{FS}$, prove that it is invalid on −closed models, and conclude that −closed models are an inadequate semantics for the higher-order closure view. Since doing so is straightforward, let us move straight to defining the model-theoretic condition of internal −closure.

In the setting of negative semantics, the choice of requiring the representation of a permutation to have a constant intension immediately creates problems: whenever there is any variation in the first-order domain function, no permutation of individuals counts as possible in the relevant structure.

Quantification over automorphisms in the closure constraint is then vacuous, and thus forces the domains of all worlds to contain all intensions which are supported by the domain function, which entails, among other things, propositional necessitism ($\Box\forall p^{\langle\rangle}\Box\exists q^{\langle\rangle}\Box(p \leftrightarrow q)$). It is therefore natural to instead formulate internal $-$closure using cumulatively representing intensions. Of course, there is no reason to think that the analog to Proposition 1.6.11 in a negative setting holds. We therefore define:

**Definition 1.6.12.** *A structure* $\mathfrak{S} = \langle W, I, D\rangle$ *is* internally $-$closed *if for all* $w \in W$*, types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*,*

$$o \in D^t_w \textit{ iff } D \boxminus o \textit{ and } \xi.o = o \textit{ for all } \xi \in \mathrm{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$$

This definition does not make being internally $-$closed incompatible with contingency in what individuals there are. However, it is not hard to see that we get similar problems in cases in which there are incompossible individuals which are indistinguishable from some world. The being constraint ensures that we cannot map any individual to one incompossible with it using an automorphism which has a possible cumulative representation, but the definition of internal $-$closure may require there to be such intensions in order for the individuals to be possibly indistinguishable. The following proposition illustrates this problem using a very simple example.

**Proposition 1.6.13.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure such that* $W = \{1, 2, 3\}$*,* $I = \{a, b\}$ *(for some distinct* $a$ *and* $b$*) and* $D$ *such that* $D^e_1 = \emptyset$*,* $D^e_2 = \{a\}$*,* $D^e_3 = \{b\}$*, and* $o \notin D^{\langle e\rangle}_1$*, where* $o$ *is the element of* $\iota^{\langle e\rangle}_{\langle W,I\rangle}$ *such that* $o(1) = o(3) = \emptyset$ *and* $o(2) = \{a\}$*.* $\mathfrak{S}$ *is not internally* $-$*closed.*

*Proof.* We distinguish two cases. *Case 1:* There is a $g \in S_I$ which has a cumulative representation $g_c$ in $D^{\langle e,e\rangle}_W$ such that $g(a) = b$ or $g(b) = a$. Then there are $w, v \in W$ such that $o \in D^{\langle e,e\rangle}_w$, and $\langle a, b\rangle \in o(v)$ or $\langle b, a\rangle \in o(v)$. Thus not $D \boxminus o$, and so $\mathfrak{S}$ is not internally $-$closed.

*Case 2:* There is no $g \in S_I$ such that $g(a) = b$ or $g(b) = a$ which has a cumulative representation in $D^{\langle e,e\rangle}_W$. Consider any $\langle f, g\rangle \in \mathrm{fix}(\mathfrak{S}, 1)|_c\mathfrak{S}$. Then $g$ has a cumulative representation in $D^{\langle e,e\rangle}_W$, so $g = \mathrm{id}_I$. Since $\langle f, g\rangle \in \mathrm{aut}(\mathfrak{S})$, it follows that $f = \mathrm{id}_W$. So $\langle f, g\rangle.o = o$. As $D \boxminus o$ and $o \notin D_1^{\langle e\rangle}$, $\mathfrak{S}$ is not internally $-$closed. □

The constraints on the structure described in this proposition fit the higher-order closure view perfectly: at world 1, we might have no materials to distinguish individual $a$ at world 2 from individual $b$ at world 3, so the property which only applies to individual $a$ at 2 should not exist at world 1.

There is no obvious way of weakening the notion of cumulative representation even further to allow this structure. Thus if there are incompossible indistinguishable possible individuals, the higher-order closure view is incompatible with the being constraint, destabilizing Stalnaker's overall position. And from a contingentist perspective, it is plausible that there are incompossible indistinguishable possible individuals: consider the knives that could have been made from a merely possible handle and two qualitatively indistinguishable merely possible blades.

There is a way of modifying the being constraint imposed by negative semantics to solve the problem. Both Stalnaker and Williamson only discuss the being constraint as applied to relations among individuals. If we weaken the being constraint to only apply to such relations, then we can solve the problem pointed out in this section by representing permutations of possible individuals not using relations among individuals but using relations among haecceities of individuals (properties necessarily equivalent to being identical to those individuals). We won't explore this option further. Instead, we now turn to adapting the notion of internal closure to the case of generation in a positive setting. The issues raised will not essentially depend on the choice of positive semantics.

### 1.6.5 Internal Generation

Like the definition of closure, we can revise the definition of generation by restricting our quantification over automorphisms to those which are possible in the structure, in the sense defined above. We could also formulate this proposal in the form of a comprehension principle, although this would require extending the syntax to include a connective which expresses the condition of being a 'generating' relation. We omit a formal discussion of these issues and move straight to the model-theoretic condition of internal generation.

In the definition of generation, we are dealing with two structures, the generated one and the generating one. We restrict the quantification over permutations to permutations which are possible in the generat*ed* structure, as there is no reason why we should only be able to appeal to generating permutations in stating the view. Imposing this restriction fundamentally changes the nature of generation. The qualitative generation view is intended as a reductive account of the existence of relations (whereas the higher-order closure view is simply a constraint on what relations there are). Adding a restriction to possibly existing permutations creates trouble for generation understood in this way, because the existence condition for higher-order entities becomes dependent on the generated higher-order domains. Thus we can no longer assume that for a given distribution of individuals and intensions, there is a unique structure generated by it. Consequently, we define internal +generation as a relation rather than a function. The natural way to do so is as follows:

**Definition 1.6.14.** *Let* $\mathfrak{S} = \langle W, I, B\rangle$ *and* $\mathfrak{S}' = \langle W, I, D\rangle$ *be structures.* $\mathfrak{S}'$ *is* internally +generated by $\mathfrak{S}$ *if for all* $w \in W$*,* $D^e_w = B^e_w$ *and for all types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*,*

$$o \in D^t_w \textit{ iff } D \boxplus o \textit{ and } \xi.o = o \textit{ for all } \xi \in \mathrm{fix}(\mathfrak{S}, w)|\mathfrak{S}'.$$

*A model is* internally +generated *just in case it is based on an internally +generated structure.*

We show that this relation of internal +generation is not functional: one structure can internally +generate two distinct structures. To specify these structures, we make use of the cycle notation for permutations introduced above, as well as writing, for any element $g$ of a group $G$, $\langle g\rangle$ for the subgroup of $G$ generated by $g$, which consists of the combinations of $g$ with itself using composition and inverses.

**Proposition 1.6.15.** *There are structures* $\mathfrak{S}$*,* $\mathfrak{S}'$ *and* $\mathfrak{S}''$ *such that* $\mathfrak{S}'$ *and* $\mathfrak{S}''$ *are distinct and both are internally +generated by* $\mathfrak{S}$*.*

*Proof.* Let $\mathfrak{S} = \langle W, I, D\rangle$ be the structure given by $W = \{0, 1, 2, 3, 4\}$ and $I = D^T_W = \emptyset$. Using addition modulo 5, let for each $w \in W$

$$g^w = (w + 1\ w + 2\ w + 3\ w + 4) \text{ and } G^w = \langle g^w\rangle,$$

$$h^w = (w + 1\ w + 2\ w + 4\ w + 3) \text{ and } H^w = \langle h^w\rangle.$$

Let $\mathfrak{F} = \langle W, I\rangle$. As in section 1.5.4, we abbreviate $\langle f, \emptyset\rangle$, for any $f \in S_W$, as $f^\emptyset$. Let $\mathfrak{S}' = \langle W, I, D'\rangle$ and $\mathfrak{S}'' = \langle W, I, D''\rangle$ be the structures such that for all $w \in W$, types $t \neq e$ and $o \in \iota^t_{\mathfrak{F}}$,

$$o \in {D'}^t_w \text{ iff } D' \boxplus o \text{ and } f^\emptyset.o = o \text{ for all } f \in G^w,$$

$$o \in {D''}^t_w \text{ iff } D'' \boxplus o \text{ and } f^\emptyset.o = o \text{ for all } f \in H^w.$$

$\mathfrak{S}'$ and $\mathfrak{S}''$ are distinct since $(1234)_{\mathfrak{F}} \in {D'}^{\langle\langle\rangle,\langle\rangle\rangle}_0$ and $(1234)_{\mathfrak{F}} \notin {D''}^{\langle\langle\rangle,\langle\rangle\rangle}_0$.

To show that both $\mathfrak{S}'$ and $\mathfrak{S}''$ are internally +generated by $\mathfrak{S}$, it suffices to show that for all $w \in W$ and $f \in S_W$, $f^\emptyset \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}'$ iff $f \in G^w$, and $f^\emptyset \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}''$ iff $f \in H^w$. We will now establish this for $G$; the case of $H$ is parallel.

First, note that $f$ is possible in $\mathfrak{S}'$ iff there is a $v \in W$ such that $f_{\mathfrak{F}} \in {D'}^{\langle\langle\rangle,\langle\rangle\rangle}_v$, which in turn is the case iff $D' \boxplus f_{\mathfrak{F}}$ and $g^\emptyset.f_{\mathfrak{F}} = f_{\mathfrak{F}}$ for all $g \in G^v$. The former is the case for any $f \in S_W$, and the latter is the case iff $g.f = f$ for all $g \in G^v$; it is routine to show that this is the case iff $f \in G^v$.

Since all domains of $\mathfrak{S}$ are empty, $f^\emptyset$ is trivially an automorphism of $\mathfrak{S}$ which maps all elements of $D^T_w$ to themselves. So $f^\emptyset \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}'$ iff $f(w) = w$ and both $f$ and $\text{id}_\emptyset$ are possible in $\mathfrak{S}'$. $\text{id}_\emptyset$ is possible in $\mathfrak{S}'$, and as we have shown, $f$ is possible in $\mathfrak{S}'$ iff there is a $v \in W$ such that $f \in G^v$. If $f(w) = w$, this is the case iff $f \in G^w$. So $f^\emptyset \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}'$ iff $f \in G^w$. □

We take it that the main attraction of the qualitative generation view over the higher-order closure view is the reductive account of what relations there are that it supposedly affords. As we just saw, taking higher-order contingency seriously shows that this supposed difference is chimerical. Qualitative generation still offers an explanation of the source of higher-order contingency, since some structures internally +generate only structures with variable higher-order domains (and, moreover, generate at least one such structure). However, there is a further, more serious, problem for the qualitative generation view as formalized by the class of internally +generated models: it fails to validate $+\mathrm{Comp}_C$. The basic observation which drives the proof is that the automorphisms of $\mathfrak{F}$ determined by the members of $G^w$ and $H^w$ are only required to be automorphisms of the generat*ing* structure, but not automorphisms of the respective generat*ed* structures. Consequently, the domains of these generated structures fail to include some intensions which are definable in terms of existing entities.

**Proposition 1.6.16.** $+\mathrm{Comp}_C$ *is not valid on the class of internally +generated models.*

*Proof.* Let $\mathfrak{F} = \langle W, I\rangle$, $\mathfrak{G}' = \langle W, I, D'\rangle$ and the functions $g$ and $G$ on $W$ be defined as in the proof of Proposition 1.6.15, and let $\mathfrak{M}$ be a model based on $\mathfrak{G}'$. Define

$$\varphi := \exists Y^{\langle\langle\rangle,\langle\rangle\rangle}\Sigma^{\text{WORLD}}wv(\text{WPERM}(Y) \wedge \neg(w \overset{\text{HI}}{\sim} v) \wedge Xwv \wedge Xvw \wedge Ywv \wedge Yvw)$$

Let $a$ be an assignment such that $a(X) = (g^0g^0)_{\mathfrak{F}} = (13)(24)_{\mathfrak{F}}$. $\mathfrak{M}$ is world-selective, so WPERM expresses the intended condition on it. So $\varphi^+_{\mathfrak{M},a}$, the propositional intension expressed by $\varphi$ relative to $\mathfrak{M}$ and $a$, is true in the worlds in which there is a representation of a world permutation $f$ such that $f(1) = 3$ and $f(3) = 1$, or $f(2) = 4$ and $f(4) = 2$. This is of course the case for 0, where (13)(24) is represented, and otherwise only for 1, where $g^1g^1 = (03)(24)$ is represented, and 4, where $g^4g^4 = (02)(13)$ is represented. Since $a(X) = (13)(24)_{\mathfrak{F}}$ exists in 0, the truth of $+\mathrm{Comp}_C$ in $\mathfrak{M}$ would entail that $\varphi^+_{\mathfrak{M},a} = \{0, 1, 4\}_{\mathfrak{F}}$ exists in 0, but this is not the case since $(1234)^{\emptyset}.\{0, 1, 4\}_{\mathfrak{F}} = \{0, 1, 2\}_{\mathfrak{F}} \neq \{0, 1, 4\}_{\mathfrak{F}}$. So $+\mathrm{Comp}_C$ is not true in $\mathfrak{M}$. □

In response to these problems, one might propose to adapt as one's model theory the class of internally +generated models which are also internally +closed. Since every internally +closed model is internally +generated (since it internally +generates itself), this is the class of internally +closed models. $+\mathrm{Comp}_C$ is valid on the class of internally +closed models, so this proposal solves the problem posed by Proposition 1.6.16, albeit in a rather *ad* *hoc* fashion. The proposal would gain some appeal if we could prove that when restricting the generated structures to internally +closed ones, internal +generation is a total function. Whether this holds is an open question given what we have established so far: On the one hand, it might well be possible to strengthen Proposition 1.6.15 to show that there are distinct internally +closed structures which are internally +generated by the same structure. On the other hand, it might well be that some structures do not internally +generate any internally +closed structure; in fact, we have not even established that every structure internally +generates *some* structure (whether internally +closed or not). We leave these questions open.

### 1.6.6 Existing Representations of Permutations

We now return to the option of formulating a version of higher-order closure which quantifies over *all* permutations instead of *all possible* permutations. A corresponding variant of $+\mathrm{Comp}_{FS}$ is defined by replacing the outer quantifier binding $X$ and $Y$ by $\forall$. It is easily seen that this variant entails $+\mathrm{Comp}_{FS}$ as defined above: over world-selective models, $\Pi XY\varphi$ entails $\forall XY\varphi$, so $\forall XY\varphi \to \psi$ entails $\Pi XY\varphi \to \psi$. This strengthening of $+\mathrm{Comp}_{FS}$ corresponds to the following strengthening of internal closure:

**Definition 1.6.17.** *Let $\mathfrak{S} = \langle W, I, D\rangle$ be a structure. For any $w \in W$, define* $\mathrm{fix}(\mathfrak{S}, w)|\mathfrak{S}, w$ *to be the set of $\langle f, g\rangle \in \mathrm{fix}(\mathfrak{S}, w)$ such that $f_{\langle W,I\rangle} \in D_w^{\langle\langle\rangle,\langle\rangle\rangle}$ and $g_{\langle W,I\rangle} \in D_w^{\langle e,e\rangle}$.*

$\mathfrak{S}$ *is* strongly internally +closed *if for all $w \in W$, types $t \neq e$ and $o \in \iota^t_{\langle W,I\rangle}$,*

$$o \in D^t_w \textit{ iff } D \boxplus o \textit{ and } \xi.o = o \textit{ for all } \xi \in \mathrm{fix}(\mathfrak{S}, w)|\mathfrak{S}, w.$$

*A model is* strongly internally +closed *just in case it is based on a strongly internally +closed structure.*

It turns out that strong internal +closure is implausibly restrictive: it rules out patterns of indistinguishability which naturally fit the intuitive motivations for the higher-order closure view, and which are plausibly instantiated if it is true. As an example, consider three electrons. Each of them could have existed without the others, in qualitatively identical circumstances. Further, it could have been that none of them exist, and that consequently, no distinctions can be drawn among them. This is ruled out by strong internal +closure, as we now show more precisely.

To do so, we first have to clarify the relevant kind of indistinguishability. It is not enough that no two of the three worlds can be distinguished from the fourth, which we might call *pairwise* indistinguishability. Rather, any way of permuting them must be allowed, which we might call *collective* indistinguishability. To make this idea precise, we let, for any permutation $g$ of a set $X$, the *support of* $g$, written $\mathrm{supp}(g)$, be the set of elements moved by $g$, i.e., $\mathrm{supp}(g) = \{x \in X : g(x) \neq x\}$.

**Definition 1.6.18.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure,* $X \subseteq W$ *and* $w \in W$*. We define* the members of $X$ to be collectively indistinguishable from $w$ *if for all* $f \in S_W$ *such that* $\mathrm{supp}(f) \subseteq X$*, there is a* $g \in S_I$ *such that* $\langle f, g\rangle \in \mathrm{fix}(\mathfrak{S}, w)$*.*

The difference between pairwise and collective indistinguishability is also relevant for the two kinds of models for the contingent existence of propositions proposed in Stalnaker (2012, Appendix A); see Fritz (unpublished d, here ch. 2) and Part 2.

To be able to appeal to it again below, we first prove part of the result on strong internal +closure as a lemma:

**Lemma 1.6.19.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be a structure,* $X \subseteq W$ *and* $w \in W$ *such that* $|X| \geq 3$*. If the members of* $X$ *are collectively indistinguishable from* $w$*, then for every* $v \in X$ *and* $f \in S_W$ *such that* $f_{\langle W,I\rangle} \in D_w^{\langle\langle\rangle,\langle\rangle\rangle}$*,* $f(v) = v$*.*

*Proof.* Assume for contradiction that there are $v \in X$ and $f \in S_W$ such that $f_{\langle W,I\rangle} \in D_w^{\langle\langle\rangle,\langle\rangle\rangle}$ and $f(v) \neq v$. Let $u \in X\backslash\{v, f(v)\}$. Since $\{v, u\} \subseteq X$, it follows from the fact that the members of $X$ are collectively indistinguishable from $w$ that there is an $h \in S_I$ such that $\langle (vu), h\rangle \in \mathrm{fix}(\mathfrak{S}, w)$. So $\langle (vu), h\rangle.f_{\langle W,I\rangle} = f_{\langle W,I\rangle}$, and thus by Lemma 1.5.13 (iv), $(vu).f = f$. But $(vu).f(u) = f(v)$ which is distinct from $f(u)$ as $v \neq u$. ↯. □

**Proposition 1.6.20.** *Let $\mathfrak{S} = \langle W, I, D\rangle$ be a structure, $X \subseteq W$ and $w \in W$ such that $|X| \geq 3$ and the members of $X$ are collectively indistinguishable from $w$. $\mathfrak{S}$ is not strongly internally +closed.*

*Proof.* Consider any $v \in X$. By Lemma 1.6.19, for every $f \in S_W$ such that $f_{\langle W,I\rangle} \in D_w^{\langle\langle\rangle,\langle\rangle\rangle}$, $f(v) = v$. So for every $\xi \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}, w$, $\xi.v_{\langle W,I\rangle} = v_{\langle W,I\rangle}$. If $\mathfrak{S}$ is strongly internally +closed, then $v_{\langle W,I\rangle} \in D_w^{\langle\rangle}$, contradicting the assumption that the members of $X$ are collectively indistinguishable from $w$. So $\mathfrak{S}$ is not strongly internally +closed. □

### 1.6.7 Constraints of Internal Closure

Proposition 1.6.20 shows that strong internal +closure is implausibly restrictive. Can a similar argument be given against internal +closure? It is easy to see a structure $\mathfrak{S}$ satisfying the condition of Proposition 1.6.20 may be internally +closed. To adapt this proposition to internal +closure, the condition on $\mathfrak{S}$ has to be strengthened; the follow is a natural way of doing so:

**Proposition 1.6.21.** *Let $\mathfrak{S} = \langle W, I, D\rangle$ be a structure and $X \subseteq W$ such that $|X| \geq 4$ and for all $w \in W$, the members of $X\backslash\{w\}$ are collectively indistinguishable from $w$ and $X_{\langle W,I\rangle} \in D_w^{\langle\rangle}$. $\mathfrak{S}$ is not internally +closed.*

*Proof.* Assume for contradiction that $\mathfrak{S}$ is internally +closed, and consider any $w \in W$ and $v \in X\backslash\{w\}$. Let $\xi = \langle f, g\rangle \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}$; then there is a $u \in W$ such that $f_{\langle W,I\rangle} \in D_u^{\langle\langle\rangle,\langle\rangle\rangle}$. So by Lemma 1.6.19, $f(x) = x$ for all $x \in X\backslash\{u\}$. Since $\xi \in \text{fix}(\mathfrak{S}, w)$ and $X_{\langle W,I\rangle} \in D_w^{\langle\rangle}$, $\xi.X_{\langle W,I\rangle} = X_{\langle W,I\rangle}$, and so $\xi.x = x$ for all $x \in X$. So in particular $\xi.v_{\langle W,I\rangle} = v_{\langle W,I\rangle}$. Thus with the assumption that $\mathfrak{S}$ is internally +closed, $v_{\langle W,I\rangle} \in D_w^{\langle\rangle}$, which contradicts the assumption that the members of $X\backslash\{w\}$ are collectively indistinguishable from $w$. □

The structure $\mathfrak{S}^4$ used in the proof of Theorem 1.5.18 is a simple instance of this more general result. $\mathfrak{S}^4$ is a paradigmatic instance of the kind of structures with which higher-order contingentists motivate their views. One might therefore take these results to be signs of trouble for the Fine-Stalnaker view.

But it is not obvious how to make this worry more concrete. In particular, it is not obvious how to adapt the concrete example used in section 1.6.6 to the present case: plausibly, any three possible electrons are compossible, so the indistinguishability of possible electrons does not lead to the patterns of indistinguishability ruled out by Proposition 1.6.21. We therefore leave it open whether the restrictions on patterns of indistinguishability imposed by internal +closure are too strong.

# Part II

# Propositional Contingentism

# Chapter 2

# Propositional Contingentism

**Abstract.** According to propositional contingentism, it is contingent what propositions there are. This paper presents two ways of modeling contingency in what propositions there are using two classes of possible worlds models. The two classes of models are shown to be equivalent as models of contingency in what propositions there are, although they differ as to which other aspects of reality they represent. These constructions are based on recent work by Robert Stalnaker; the aim of this paper is to explain, expand, and, in one aspect, correct Stalnaker's discussion.

## 2.1 Introduction

Propositional contingentism is the view that it is contingent what propositions there are. Many of those who have held this view have been motivated by an argument roughly along the following lines:

I could have failed to be.

Had I not been, there would not have been the proposition that I am me.

Therefore, the proposition that I am me could have failed to be.

Early instances of such arguments can be found in Prior's writings, e.g., in Prior (1967, pp. 150–151), where he gives such an argument for contingency in what facts there are. Later examples of at least tentative endorsements of such arguments can be found in Fine (1977b), Adams (1981), Fitch (1996), Bennett (2005), David (2009), Speaks (2012), Stalnaker (2012), and Nelson (2014). Relatedly, Williamson (2002, 2013, chapter 6) endorses the second premise of the argument on the assumption of the truth of the first, which he rejects. An exception in the literature is Lindström (2009), who argues for propositional contingentism on the basis of a puzzle about possible world semantics due to Kaplan (1995). Some, like Williamson, deny the first premise, but few have explicitly denied the second premise; examples are Plantinga (1983) and Bealer (1993, 1998).

This indicates that propositional contingentism is widely regarded as an interesting and plausible view. Yet, while some aspects of propositional contingentism, such as its implications for semantics, have been discussed at length, there have been surprisingly few investigations into the seemingly more basic issue of developing a systematic theory of what propositions there are and what propositions there could have been. One exception is Fine (1980), who develops such a theory on the assumption that propositions are individuated relatively finely. As far as I am aware, there are only two such investigations which assume a more coarse-grained theory of propositions according to which propositions are identical if they are strictly equivalent (i.e., according to which $p$ is $q$ if necessarily, $p$ if and only if $q$). These are Fine (1977b) and Stalnaker (2012, Appendix A).

Both Fine and Stalnaker proceed model-theoretically, constructing classes of possible worlds models in which propositions are identified with sets of possible worlds. (Or rather, these are models in which sets of representatives of worlds represent propositions. As usual in possible worlds model theory, the entities representing worlds and propositions will be spoken of as if they were in fact worlds and propositions, although this is of course not required.) In principle, such a model theory is straightforward to define, following the variable domain possible worlds model theory of Kripke (1963), by associating each world with a domain of propositions. The model theory might therefore simply be the class of tuples $\langle W, D\rangle$ such that $W$ is a set and $D : W \to \mathcal{P}(\mathcal{P}(W))$. However, this model theory does not limit contingency in what propositions there are in any interesting way. E.g., it does not enforce the natural constraint that necessarily, the propositions there are are closed under negation.

Both the model theories of Fine and Stalnaker are more informative, and it turns out that they are closely related in both philosophical and formal respects. However, neither of them is easily accessible, although for very different reasons. Fine's model theory not only represents propositions but also individuals and relations in a complex hierarchy of intensional and extensional relations, and the development is dense and technical. Stalnaker's model theory is only sketched in a very short appendix, and the formal definitions are not related in any detail to the preceding philosophical discussion. Furthermore, Stalnaker gives two variants of his model theory which he claims to be equivalent; however, as will be shown below, one of his definitions must be corrected to establish the equivalence.

The aim of this paper is to provide an accessible but rigorous development of model theories for propositional contingentism along the lines of Stalnaker and Fine. In the interest of clarity, they are introduced on their own terms, without references to the literature. An appendix states Stalnaker's original definitions, shows how they differ from the ones proposed here, and argues for the latter. The models developed here are related to Fine's models in Fritz (unpublished a, here ch. 3), based on the work in Fritz and Goodman (unpublished c, here ch. 1).

The remainder of this paper is structured as follows: In section 2.2, a possible worlds model theory is developed whose models, called *equivalence systems*, associate with every world an equivalence relation of indistinguishability between worlds; they are interpreted as models of contingency in what propositions there are by taking the propositions at a world $w$ to be the sets of worlds which contain either both or neither of two worlds indistinguishable at $w$. In section 2.3, a second possible worlds model theory is developed whose models, called *permutation systems*, associate with every world a set of permutations representing the symmetries of modal space from the perspective of this world; they are interpreted as models of contingency in what propositions there are by mapping them to equivalence systems, associating each world $w$ with the equivalence relation which holds between two worlds if one is mapped to the other by some symmetry of $w$.

In sections 2.2 and 2.3, a restriction of coherence is imposed on each class of models. In section 2.4, it is shown that the two classes of coherent models represent the same patterns of contingency in what propositions there are, by showing that an equivalence system is coherent if and only if it is determined by a coherent permutation system. Section 2.5 shows that the two kinds of models nevertheless differ in what they represent, as different coherent permutation systems may determine the same coherent equivalence system. That this is in line with our philosophical interpretation of the systems is shown using a simple example. Section 2.6 delves deeper into the structural relations between the two kinds of systems, investigating both the structures formed by the two classes of coherent systems under natural orders and some relations between these two structures.

This paper is part of a larger body of work by Jeremy Goodman and myself; connections to related papers are discussed in the concluding section 2.7. Appendix 2.8 discusses Stalnaker's models, and shows that the present definition of coherent permutation systems matches Stalnaker's corresponding definition, whereas the present definition of coherent equivalence systems is more restrictive than Stalnaker's corresponding definition. Using a simple example, it is shown that Stalnaker's philosophical considerations support the present definition rather than his own. Since much of the following is formulated in terms of possible worlds, appendix 2.9 considers how such talk may be understood. It is argued that the version of propositional contingentism discussed here is incompatible with taking talk of worlds at face value. A well-known strategy for understanding such talk in terms of propositions is adapted to fit propositional contingentism, but it is noted that the strategy is limited in generality. Whether this lack of generality is a serious problem for the theory is left open.

## 2.2 Equivalence Systems

Consider again the second premise of the above argument for propositional contingentism: Why should there not have been the proposition that I am me, had I not been? An answer which motivates both classes of models to be explored is that without me, there would not have been the resources required to draw the distinction drawn by the proposition that I am me. This idea is best elaborated using a simpler, albeit more artificial, example: Consider the possibility of there being two fundamental particles $a$ and $b$ which actually are nothing. Assume that for both particles $a$ and $b$, there is a world in which this particle makes up the only matter in an otherwise completely homogenous space-time continuum. Let $w_a$ and $w_b$ be such a pair of worlds. Had there been $a$ and $b$, then $w_a$ and $w_b$ could be distinguished in terms of $a$ and $b$, but since actually there are neither $a$ nor $b$ and $w_a$ and $w_b$ differ only in which individual they contain, $w_a$ and $w_b$ can actually not be distinguished. Thus in particular, they cannot be distinguished by any proposition, so all propositions are either true in both or neither of $w_a$ and $w_b$. Of course, if there had been $a$ and $b$, $w_a$ and $w_b$ could be distinguished, and so there would be propositions true in only one of them.

This line of thought motivates the idea that what propositions there are at a given world depends on which distinctions among worlds can be drawn at it. Both classes of models to be explored take up this idea and model what distinctions among worlds can be drawn at a given world, from which what propositions there at that world is derived. Both classes of models identify propositions with sets of worlds, taking such a set to be true at a world if it contains it, and to be necessary if it is the set of all worlds.

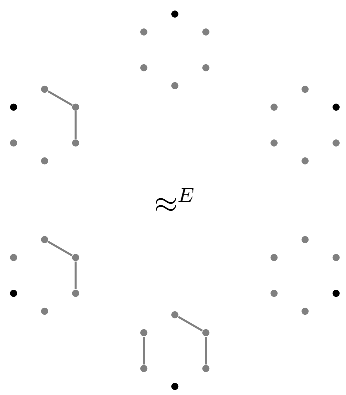
The first class of models represents what distinctions among worlds can be drawn at a given world in the most straightforward manner: such a model associates with each world $w$ a relation $\approx_w$, which relates two worlds if and only if they cannot be distinguished at $w$. Since the relevant notion of indistinguishability is plausibly reflexive, symmetric and transitive, $\approx_w$ will be assumed to be an equivalence relation. This determines what propositions there are as follows: At $w$, there are those propositions $P$ such that for all worlds $v$ and $u$ related by $\approx_w$, $P$ is true in $v$ if and only if $P$ is true in $u$. Equivalently, the propositions at $w$ are the unions of sets of equivalence classes of $\approx_w$; this is the unique complete atomic field of sets whose atoms are the equivalence classes of $\approx_w$. Formally, define:

**Definition 2.2.1.** *For every set $W$, an* equivalence system on $W$ *is a function $\approx$ mapping every $w \in W$ to an equivalence relation $\approx_w$ on $W$.*

As an example of an equivalence system, consider the function $\approx^E$ on $\{1, \ldots, 6\}$ which maps 1, 2 and 3 to the identity relation, which will be called id (letting the context determine its domain); which maps 4 to the equivalence relation on $\{1, \ldots, 6\}$ which relates two elements just in case they are both strictly less than 4, both identical to 4 or both strictly greater than 4; and which maps 5 and 6 to the equivalence relation on $\{1, \ldots, 6\}$ which relates two elements just in case they are identical or both strictly less than 4.

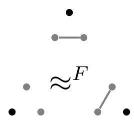
It will be helpful to represent such systems pictorially. Here is a natural way of drawing any equivalence system $\approx$ based on a set of worlds $\{1, \ldots, n\}$ for some natural number $n$: Draw representations of the worlds in a circle, starting with 1 at the top and turning clockwise. In this circle, each world $i$ is represented by a smaller circle of dots, each of which represents a world; again, start with 1 at the top and turn clockwise. In this smaller circle representing $i$, indicate which worlds are related by $\approx_i$ by drawing a line connecting dots which represent worlds related by $\approx_i$. There is no need to indicate a direction since $\approx_i$ is symmetric; dots don't have to be connected to themselves as $\approx_i$ is reflexive; and two dots need not be connected if they are already connected by a path since $\approx_i$ is transitive.

As an example, the following is one way of drawing $\approx^E$:




Here, for any $i \leq 6$, the dot representing $i$ in the circle representing $i$ has been distinguished by drawing it black instead of gray; this is simply to make it easier to see which dots correspond to which circles. The center of the big circle is used to label the system.

Not all equivalence systems are plausible models of which worlds can be distinguished at a given world. Consider the following equivalence system:




According to this system, 2 and 3 are indistinguishable at 1. Yet, 2 and 3 differ structurally in what can be distinguished at them: At 2, the other two worlds *cannot* be distinguished, although at 3 the other two worlds *can* be distinguished. Thus 2 and 3 can be distinguished purely in terms of what can be distinguished at them, and therefore 2 and 3 can't be indistinguishable at 1.

Here is another version of the same argument: According to $\approx^F$, there are four propositions in 1 and 2, and eight propositions in 3. So the proposition that there are exactly four propositions is the set $\{1, 2\}$. Delineating the worlds in which there are exactly four propositions seems not to depend on any special resources, so in every world, there should be the proposition that there are exactly four propositions. In particular, there should be this proposition in 1, so there should be the proposition $\{1, 2\}$ in 1. This conflicts with the fact that $2 \approx^F_1 3$, and therefore shows that $\approx^F$ is not a plausible model. This line of thought could also be spelled out more formally using so-called *comprehension principles*; see Fritz and Goodman (unpublished c, here ch. 1) and Fritz (unpublished c, here ch. 5) for further discussion.

The upshot is that in a coherent equivalence system, worlds indistinguishable at a given world must in particular be indistinguishable in terms of indistinguishability. In general, worlds indistinguishable at a given world $w$ must be indistinguishable in terms of all resources available in $w$, including the notion of indistinguishability. To turn this into a formal criterion, two questions must be answered. First, what are the resources available at a world, as represented by an equivalence system? And second, how can indistinguishability given those resources be understood?

Concerning the first question, three resources can be identified as being represented by equivalence systems: First, each world contains a set of propositions, given by its relation of indistinguishability. Second, as noted above, the notion of indistinguishability is a resource available at any world. Finally, it is natural to count each world as being one of the resources available at itself.

The natural answer to the second question is that $v$ and $u$ are indistinguishable given certain resources if $v$ and $u$ are symmetric with respect to them; that is, if there is a way of reconfiguring worlds which maps $v$ to $u$ but respects the given resources, in the sense of these resources being invariant under this reconfiguration. Formally, such a reconfiguration is a *permutation* – a bijection from worlds to worlds, i.e., a function from worlds to worlds which is both surjective (onto) and injective (one-to-one). It only remains to specify what it takes for a permutation of worlds to respect the three resources identified above. This is obvious in the case of the world itself: $w$ is invariant under $f$ just in case $f$ maps $w$ to itself. The other two resources require a bit more thought.

For propositions, note that it is straightforward to extend a permutation of worlds to a permutation of propositions, by letting the image of a proposition $P$ under a permutation $f$ be the set of the images of members of $P$ under $f$: $f.P = \{f(w) : w \in P\}$. (The notation $f.P$ indicates that from a group-theoretic perspective, the extension of $f$ from worlds to sets of worlds can be understood as an action.) Thus, a permutation $f$ respects the propositions at $w$ just in case it maps every union of a set of equivalence classes of $\approx_w$ to itself. It is easy to see that this is equivalent to requiring $f$ to map each world $v$ to one $\approx_w$-related to $v$. Taking $f$ to be the set of pairs $\langle v, u\rangle$ such that $f(v) = u$ and $\approx_w$ as the set of pairs $\langle v, u\rangle$ such that $v \approx_w u$, this is most concisely written as $f \subseteq \approx_w$.

For a permutation $f$ to respect the notion of indistinguishability, facts about which worlds can distinguished at a given world must be invariant under permuting the worlds using $f$. That is, $v$ and $u$ must be indistinguishable at $w$ just in case $f(v)$ and $f(u)$ are indistinguishable at $f(w)$; i.e., $v \approx_w u$ if and only if $f(v) \approx_{f(w)} f(u)$. In this case, $f$ is called an *automorphism of* $\approx$.

The coherence constraint can now be stated formally; it requires that if $v$ and $u$ are indistinguishable at $w$, then there is a permutation $f$ mapping $v$ to $u$ which (i) is a subset of $\approx_w$, (ii) is an automorphism of $\approx$, and (iii) maps $w$ to itself. To state this more concisely, let $\mathrm{aut}(\approx)$ be the set of automorphisms of $\approx$; this is a group, a fact which will be useful later. Further, let $\mathrm{aut}(\approx)_w$ be the set of elements of $\mathrm{aut}(\approx)_w$ which map $w$ to itself; this is called the *stabilizer of* $w$. With this, the condition can be formulated as follows:

**Definition 2.2.2.** *An equivalence system $\approx$ on a set $W$* coheres *if for all $w, v, u \in W$ such that $v \approx_w u$, there is an $f \in \mathrm{aut}(\approx)_w$ such that $f(v) = u$ and $f \subseteq \approx_w$.*

Note that since every element of $\mathrm{aut}(\approx)_w$ maps $w$ to itself, $w$ can only be $\approx_w$-related to itself. Consequently, $w$'s equivalence class under $\approx_w$ is its singleton: Every world contains its singleton proposition.

To illustrate how coherence is applied, it is helpful to introduce *cycle-notation* of permutations, which is best explained by examples: Considering permutations on $W = \{1, \ldots, 6\}$, the permutation which maps 1 to 2, 2 to 3, 3 to 1 and all other elements of $W$ to themselves can be written (123); the permutation which maps each $i < 5$ to itself and 5 and 6 to each other can be written (56).

Consider again the systems $\approx^F$ and $\approx^E$ depicted above. $\approx^F$ is easily seen to be incoherent: Since $2 \approx^F_1 3$, coherence requires there to be an $f \in \mathrm{aut}(\approx^F)_1$ such that $f(2) = 3$ and $f \subseteq \approx^F_1$. The only permutation of $\{1, 2, 3\}$ mapping 1 to itself and 2 to 3 is $f = (23)$; however, $f \notin \mathrm{aut}(\approx^F)$: $3 \approx^F_2 1$ holds, but $f(3) \approx^F_{f(2)} f(1)$, i.e., $2 \approx^F_3 1$ does not. Thus $\approx^F$ is incoherent.

In contrast, $\approx^E$ is coherent. This follows from the fact that for all $w, v, u \in \{1, \ldots, 6\}$, if $v \approx^E_w u$ then $(vu) \in \mathrm{aut}(\approx)_w$. (While this is evidently a sufficient condition for coherence, $\approx^B$ in the proof of Proposition 2.6.8 shows that it is not a necessary condition.) Although somewhat laborious, checking that this claim holds is a straightforward matter using the above definition of automorphisms and stabilizers.

## 2.3 Permutation Systems

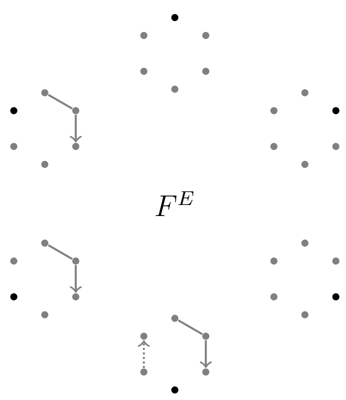
Models of the second class represent, for each world, which permutations of worlds respect *all* distinctions among worlds which can be drawn at that world. Call a permutation which does so a *symmetry* of the world. A model of the second class is therefore a function mapping each world to the set of its symmetries. Clearly, the identity permutation is a symmetry of every world. Further, if a permutation $f$ respects certain distinctions, then so does its inverse $f^{-1}$, and if two permutations $f$ and $g$ respect these distinctions, then so does their composition $fg$. Imposing these three constraints on the symmetries of each world is equivalent to requiring that the symmetries of each world form a *permutation group* on the set of worlds. Thus, define formally:

**Definition 2.3.1.** *For every set $W$, a* permutation system *on $W$ is a function $F$ mapping every $w \in W$ to a permutation group $F_w$ on $W$.*

To give an example of a permutation system, write, for a set of permutations $G$, $\langle G \rangle$ for the permutation group generated by $G$, i.e., the set of permutations which can be obtained by finite combinations of elements of $G$ by inverses and composition. E.g, $\langle \{(123)\} \rangle = \{(123), (321), \mathrm{id}\}$. Here is an example of a permutation system: Let $F^E$ be the function on $\{0,\ldots,6\}$ which maps 1, 2 and 3 to $\{\mathrm{id}\}$, 4 to $\langle\{(123),(56)\}\rangle$ and 5 and 6 to $\langle\{(123)\}\rangle$.

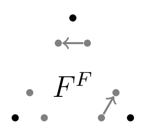
It will also be helpful to be able to draw any permutation system $F$ on $\{1,\ldots,n\}$, for some natural number $n$. Again, draw a circle of $n$ circles of $n$ dots. For each $i \leq n$, choose a set of permutations $G$ generating $F_i$, and assign to each member of $G$ a different style of arrow, such as solid versus dotted lines. Draw each permutation in $G$ in the circle of dots representing $i$ by arrows indicating which worlds are mapped to which worlds. Such arrows will be allowed to go through several worlds, leaving it implicit that the dot at the tip of an arrow represents the world mapped to the world represented by the dot at the start of the arrow; arrows from a world to itself will be omitted. E.g., to draw (123), draw a single arrow starting at 1, going through 2 and pointing to 3.

As an example, the following represents $F^E$, choosing $\{(123),(56)\}$ and $\{(123)\}$ to generate $F_4$ and $F_5/F_6$, respectively:




Note that for a given permutation system $F$, there might be several choices of associating each world $i$ with a set of permutations $G$ which generates $F_i$. Consequently, it is not always straightforward to tell whether two drawings represent the same permutation system.

Again, not every such system is plausible:



According to $F^F$, switching 2 and 3 is a symmetry of 1. But 2 and 3 can be distinguished in terms of their symmetries, since 2 does but 3 does not have a symmetry other than the trivial identity permutation.

A coherence constraint must be imposed which ensures that the permutations associated with a given world respects its resources, as represented by the permutation system. These resources are only the notion of a symmetry and the world itself. Note in particular that the symmetries of a world need not be resources available at that world: the symmetries associated with a world only *describe* what distinctions can drawn using the resources available at this world; they need not themselves be resources available at the world. In the model theory of Fritz and Goodman (unpublished c, here ch. 1), which extends the present treatment of contingency in what propositions there are to a type hierarchy of relations, this can be demonstrated more concretely by treating permutations of worlds as a special kind of binary relations among nullary relations. A related issue is discussed in appendix 2.9.

What does it take for a permutation of worlds $f$ to respect the notion of a symmetry? To spell this out, $f$ must first be extended to a permutation of permutations of worlds. The result of applying $f$ to $g$ should behave on the permuted elements as $f$ behaved on the original elements. So if $g$ maps $w$ to $v$, then the result of applying $f$ to $g$ should to map $f(w)$ to $f(v)$. Writing $f.g$ for the result of applying $f$ to $g$, it should thus be the case that $(f.g)f(w) = f(v)$, and since $v = g(w)$, $(f.g)f(w) = fg(w)$. Let $u = f(w)$; since $f$ is a bijection, $w = f^{-1}(u)$. So $(f.g)ff^{-1}(u) = fgf^{-1}(u)$, and therefore $f.g(u) = fgf^{-1}(u)$. So define $f.g$ to be $fgf^{-1}$, which is called the *conjugation of $g$ by $f$*.

This definition is naturally extended to sets of permutations, letting $f.G = \{f.g : g \in G\}$. The required constraint on $f$ is now formulated straightforwardly by requiring $f$ to map $F_w$, for any world $w$, to $F_{f(w)}$, i.e., $f.F_w = F_{f(w)}$. As above, call such permutations *automorphisms of $F$*, and the set of such permutations $\text{aut}(F)$. Again, the set of members of $\text{aut}(F)$ mapping $w$ to itself is the *stabilizer of $w$*, written $\text{aut}(F)_w$. Thus the members of $\text{aut}(F)_w$ are exactly the permutations which satisfy the constraint of respecting the resources available at $w$, as represented by the permutation system. Consequently, coherence can be defined as follows:

**Definition 2.3.2.** *A permutation system $F$ on a set $W$* coheres *if for all $w \in W$, $F_w \subseteq \mathrm{aut}(F)_w$.*

Consider again the systems $F^F$ and $F^E$ depicted above: $F^F$ is incoherent, since $f = (23)$ is a member of $F^F_1$ but not an automorphism of $F^F$: $(31) \in F^F_2$ holds, while $f.(31) \in F_{f(2)}$, i.e., $(21) \in F_3$, does not hold. And although somewhat laborious, it is routine to show that $F^E$ is coherent using the above definitions.

How do permutation systems model contingency in what propositions there are? They do so by determining an equivalence system, which itself can be seen as a model of contingency in what propositions there are, as described above. To see how a permutation systems $F$ determines an equivalence system, note first that if there is an $f \in F_w$ which maps $v$ to $u$, then there is a symmetry of $w$ which maps $v$ to $u$; consequently, $v$ and $u$ must be indistinguishable. Conversely, it was noted above that if $v$ and $u$ are indistinguishable at $w$, then they must be indistinguishable in terms of all resources available at $w$, so there must be a symmetry of $w$ mapping $v$ to $u$, i.e. an $f \in F_w$ such that $f(v) = u$. Thus the equivalence system determined by $F$ counts $v$ and $u$ as indistinguishable at $w$ if and only if there is an $f \in F_w$ such that $f(v) = u$. Taking relations and functions to be sets of pairs as noted above, this can be summed up as follows:

**Definition 2.3.3.** *For every permutation system $F$ on a set $W$, the* equivalence system determined by $F$*, written $\varepsilon(F)$, is such that for all $w \in W$:*

$$\varepsilon(F)_w = \bigcup F_w.$$

It is straightforward to check that this is well-defined, i.e., that $\varepsilon(F)_w$ is an equivalence relation for every $w \in W$. To illustrate the definition, note that $\varepsilon(F^F) = \approx^F$ and $\varepsilon(F^E) = \approx^E$.

## 2.4 Equivalence

Do coherent equivalence systems and coherent permutation systems encode the same theory of propositional contingency, in the sense of admitting the same patterns of contingency in what propositions there are? This section shows that this is so, by showing that an equivalence system is coherent if and only if it is determined by a coherent permutation system. That every coherent permutation system determines a coherent equivalence system is easy to show using the following lemma:

**Lemma 2.4.1.** *For any permutation system $F$,* $\mathrm{aut}(F) \subseteq \mathrm{aut}(\varepsilon(F))$.

*Proof.* Let $f \in \mathrm{aut}(F)$, and consider any $w, v, u \in W$ such that $v\varepsilon(F)_w u$. Then there is a $g \in F_w$ such that $g(v) = u$. Since $f \in \mathrm{aut}(F)$, $f.g \in F_{f(w)}$, so $f(v)\varepsilon(F)_{f(w)} f.g(f(v))$. As $f.g(f(v)) = fg(v) = f(u)$, it follows that $f(v)\varepsilon(F)_{f(w)} f(u)$, as required. The converse direction follows by a symmetric argument for $f^{-1}$. □

**Theorem 2.4.2.** *Every coherent permutation system determines a coherent equivalence system.*

*Proof.* Let $F$ be a coherent permutation system on a set $W$, and consider any $w, v, u \in W$ such that $v\varepsilon(F)_w u$. Then there is an $f \in F_w$ such that $f(v) = u$. Since $F$ is coherent, $f \in \mathrm{aut}(F)_w$, and so by Lemma 2.4.1, $f \in \mathrm{aut}(\varepsilon(F))_w$. By construction of $\varepsilon(F)$, $f \subseteq \varepsilon(F)_w$. □

To show that every coherent equivalence system is determined by a coherent permutation system, a mapping from equivalence systems to permutation systems will be used. The idea behind this mapping is to associate each world $w$ with the set of automorphisms which respect the propositions at $w$, using the above extension of permutations of worlds to propositions. It is easy to see that these are exactly the automorphisms which respect the equivalence classes at $w$. Formally, write $W/{\approx_w}$ for the set of equivalence classes under $W$, which is called the *quotient set of $W$ by $\approx_w$*, and write $\mathrm{aut}(\approx)_{(W/\approx_w)}$ for the set of automorphisms of $\approx$ which map each member of $W/{\approx_w}$ to itself, which is called the *point-wise stabilizer of* $W/{\approx_w}$. It is easy to see that $\mathrm{aut}(\approx)_{(W/\approx_w)} = \{f \in \mathrm{aut}(\approx) : f \subseteq {\approx_w}\}$. Thus, define formally:

**Definition 2.4.3.** *For every equivalence system* $\approx$ *on a set* $W$*, the* permutation system determined by $\approx$*, written* $\pi(\approx)$*, is such that for all* $w \in W$*:*

$$\pi(\approx)_w = \mathrm{aut}(\approx)_{(W/\approx_w)}.$$

The desired result can now be obtained from two lemmas. The first shows that every coherent equivalence system determines a coherent permutation system. The second shows that every coherent equivalence system is determined by the permutation system it determines. To prove the first, permutations of worlds, already extended to propositions, are analogously extended once more to sets of propositions.

**Lemma 2.4.4.** *Every coherent equivalence system determines a coherent permutation system.*

*Proof.* Let $\approx$ be a coherent equivalence system on a set $W$, and consider any $w \in W$ and $f \in \pi(\approx)_w$. To prove that $f \in \mathrm{aut}(\pi(\approx))$, consider any $v \in W$; we prove that $f.\pi(\approx)_v = \pi(\approx)_{f(v)}$. $f.\pi(\approx)_v = f.\mathrm{aut}(\approx)_{(W/\approx_v)}$; by a general principle for stabilizers, this is $\mathrm{aut}(\approx)_{(f.W/\approx_v)}$, which, as $f \in \mathrm{aut}(\approx)$, is $\mathrm{aut}(\approx)_{(W/\approx_{f(v)})}$, i.e., $\pi(\approx)_{f(v)}$. As noted above, $\{w\} \in W/\approx_w$, so $f(w) = w$, and thus $f \in \mathrm{aut}(\pi(\approx))_w$, as required. □

**Lemma 2.4.5.** *For every coherent equivalence system* $\approx$*,* $\approx = \varepsilon(\pi(\approx))$*.*

*Proof.* If $v \approx_w u$, then there is an $f \in \mathrm{aut}(\approx)_w$ such that $f(v) = u$ and $f \subseteq \approx_w$. So $f \in \pi(\approx)_w$, and hence there is an $f \in \pi(\approx)_w$ such that $f(v) = u$. Therefore $v\varepsilon(\pi(\approx))_w u$. If $v\varepsilon(\pi(\approx))_w u$, then there is an $f \in \pi(\approx)_w$ such that $f(v) = u$. Since $f \in \pi(\approx)_w$, $f \subseteq \approx_w$; in particular $v \approx_w f(v)$, so $v \approx_w u$. □

**Theorem 2.4.6.** *Every coherent equivalence system is determined by a coherent permutation system.*

*Proof.* If $\approx$ is a coherent equivalence system, then by Lemma 2.4.4, $\pi(\approx)$ is a coherent permutation system, and by Lemma 2.4.5, $\approx$ is the equivalence system determined by it. □

Together, Theorems 2.4.2 and 2.4.6 show that as models of contingency in what propositions there are, coherent equivalence systems and coherent permutation systems are equivalent.

## 2.5 Relating Coherent Systems

Given Theorems 2.4.2 and 2.4.6, one might conjecture that the two kinds of coherent systems are equivalent in a stronger sense, namely that the two determination relations are bijections and mutual inverses. This, however, is not the case; the relations between the two kinds of coherent systems are more interesting. To explore them in more detail in the following, fix an arbitrary set $W$ as the set of worlds, and consider $\varepsilon$ as a function from coherent permutation systems to coherent equivalence systems, and $\pi$ as a function from coherent equivalence systems to coherent permutation systems, all of them on $W$ – this will be left tacit in this section and the next.

Theorem 2.4.6 shows that every coherent equivalence system is determined by a coherent permutation system, so $\varepsilon$ is surjective. But as the following result shows, $\varepsilon$ is not injective, at least not for every choice of set of worlds $W$:

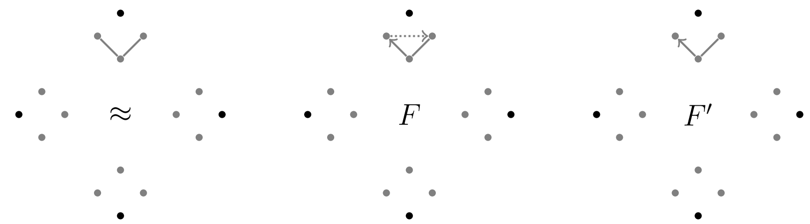
**Theorem 2.5.1.** *For some set $W$, there are distinct coherent permutation systems on $W$ which determine the same equivalence system.*

*Proof.* Let $F$ and $F^\omega$ be the permutation systems on an infinite set $W$ such that for every $w \in W$, $F_w$ is the set of permutations of $W$ which map $w$ to itself, and $F^\omega_w$ is the set of such permutations which also map only finitely many worlds not to themselves. It is routine to check that $F$ and $F^\omega$ are coherent. For both of these permutation systems, there is a permutation in the set associated with $w$ which maps $v$ to $u$ if and only if $w = v = u$ or $w \notin \{v, u\}$, so they determine the same equivalence system. □

This also shows that $\pi$ is not surjective: If it were, there would be coherent equivalence systems $\approx$ and $\approx'$ determining permutation systems $F$ and $F'$ which determine the same equivalence system, contradicting Lemma 2.4.5. However, Lemma 2.4.5 shows that $\pi$ is injective, and that the inverse of $\pi$ is the restriction of $\varepsilon$ to coherent permutation system which are determined by coherent equivalence systems.

Theorem 2.5.1 should not come as a surprise: given the interpretation of equivalence and permutation system, there was no reason to expect distinct coherent permutation systems to determine distinct equivalence systems. For recall that equivalence systems encode *which* worlds can be distinguished at a given world, whereas permutation systems encode *how* worlds can be permuted in ways which constitute a symmetry of the given world. While it is natural to think that the former information is contained in the latter information, there is no reason to expect the converse.

To illustrate this more concretely, it is helpful to consider a witness of the existential claim of Theorem 2.5.1 which is perhaps less elegant than the one used in the above proof. Let $F$ and $F'$ be the permutation systems on $W = \{1, 2, 3, 4\}$ such that $F_1$ is the set of permutations of $W$ which map 1 to itself, which can be represented as $\langle\{(234), (42)\}\rangle$, $F'_1 = \langle\{(234)\}\rangle$, and for all $w \in \{2, 3, 4\}$, $F_w = F'_w = \{\mathrm{id}\}$. It is routine to check that these are coherent and determine the same equivalence system $\approx$:




$F'_1 = \{(234), (432), \mathrm{id}\}$ does not contain (23), (24) or (34). So $F'$ might seem like a curious permutation system: How could it be that at 1, any two of 2, 3 and 4 are indistinguishable, yet not every way of permuting these is a symmetry of 1? How such a permutation system might arise can be motivated by considering individuals and their relations. To facilitate the comparison between $F$ and $F'$, the following story does so for both of these permutation systems. Assume that there are three possible electrons $a$, $b$ and $c$ which are nothing in 1, and some qualitative relation $R$ in which electrons can stand. Assume further that worlds 2, 3 and 4 are very simple, such that in some sense, all there is to be said about them is that in 2, there are $a$ and $b$, that in 3, there are $b$ and $c$, that in 4, there are $c$ and $a$, and for each world, which pairs of individuals instantiate $R$. Now consider two such configurations: according to the left, in each world, both individuals there are at this world stand in $R$ to each other; according to the right, $Rab$ in 2, $Rbc$ in 3 and $Rca$ in 4:

$$\begin{array}{llll} 1: & & 1: & \\ 2: & a \longleftrightarrow b & 2: & a \longrightarrow b \\ 3: & \qquad b \longleftrightarrow c & 3: & \qquad b \longrightarrow c \\ 4: & a \longleftarrow\!\!\!\longrightarrow c & 4: & a \longleftarrow\!\!\!\!\!\!\!\!\longrightarrow\!\!\!\!\!\!\!\!\!\!\!\!\phantom{\longrightarrow} c \end{array}$$

The symmetries of a given world $w$ were introduced above as the permutations of worlds which respect *all* distinctions among worlds which can be drawn using the resources at $w$. Clearly these resources include the individuals there are at $w$, and plausibly, they include the qualitative relation $R$ as well as the notion of being, i.e., which individuals there are at a world. As above, a symmetry is understood to respect these resources if it maps them to themselves. Permutations of worlds are not obviously extended to permutations of, e.g., individuals, in the way they were extended to propositions above. Rather, once individuals are considered, a reconfiguration of modal reality should be understood as consisting of a permutation of worlds $f$ as well as a permutation of individuals $g$. Such a reconfiguration can be understood to respect an individual if $g$ maps the individual to itself; it can be understood to respect $R$ if at any world $w$, individuals $x$ and $y$ stand in $R$ if and only if $g(x)$ and $g(y)$ stand in $R$ in $f(w)$, and it can be understood to respect the notion of being if at any world $w$, there is an individual $x$ if and only if there is $g(x)$ at $f(w)$. A permutation of worlds being a symmetry of a world can now be understood as it being part of a more comprehensive reconfiguration of modal space.

Much more in this direction can be found in Fritz and Goodman (unpublished c, here ch. 1), but what has been said so far suffices to indicate how $F$ is derived from the left configuration and how $F'$ is derived from the right configuration. E.g., the permutation of worlds (23) is a symmetry of 1 in $F$, since it can be extended by a permutation of individuals, namely $(ac)$, to form a reconfiguration of modal space which respects all resources at 1. That it respects all individuals at 1 is trivial, since there are none. That it respects $R$ is easily verified: since $Rab$ at 2, $Rcb$ must be the case at 3, which is the case; similarly for the other instances of the condition. Finally, the notion of being is respected: $a$ and $b$ (the individuals at 2) are mapped to $c$ and $b$ (the individuals at 3), and similarly for the other instances of the condition. In contrast, (23) is not a symmetry of 1 in $F'$: For this to be the case, there would have to be a permutation $g$ of individuals with which (23) forms a reconfiguration of modal space which respects all resources at 1. In particular, this would have to respect the notion of being, and so, since 2 is mapped to 3, map each of $a$ and $b$ to one of $b$ and $c$. Since $R$ must be respected as well and $Rab$ holds in 2, $Rg(a)g(b)$ must hold in 3, so $g$ must map $a$ to $b$ and $b$ to $c$. Since $g$ is a permutation it must map $c$ to $a$. But this means that another instance of the reconfiguration respecting $R$ is not satisfied: 3 is mapped to 2 and $Rbc$ holds in 3, but $Rg(b)g(c)$, i.e., $Rca$, does not hold in 2. So (23) is not a symmetry of 1 in $F'$.

To summarize, *how* worlds can be permuted in ways which constitute a symmetry of a world goes beyond *which* worlds can be distinguished at it. But then, what does $\pi$, the function which maps every coherent equivalence system to a coherent permutation system, do? Lemma 2.4.5 says that the permutation system determined by a coherent equivalence system $\approx$ is a coherent permutation system which determines $\approx$. But as the above examples witness, there might be more than one such permutation system, so the question is: which one does $\pi$ take us to? The definition of the permutation system determined by a given equivalence system works by associating with each world the set of *all* automorphisms of the equivalence system which respect the propositions at that world according to the equivalence system. The natural guess is therefore that $\pi(\approx)$ is the most inclusive among the coherent permutation systems which determine $\approx$, in the sense that for any world $w$, $\pi(\approx)_w$ contains all permutations in $F_w$ for any such permutation system $F$. In the next section, it is shown that this conjecture is correct. To do so, the idea of ordering permutation systems according to how inclusive they are is first made precise. It is clear that equivalence systems can be ordered analogously, which provides a new perspective on equivalence systems and permutation systems, namely as two ordered sets connected by two functions.

## 2.6 Ordering Coherent Systems

To start, the order among coherent permutation systems is defined formally:

**Definition 2.6.1.** $\sqsubseteq$ *is the binary relation on the set of coherent permutation systems such that* $F \sqsubseteq F'$ *just in case for all* $w \in W$, $F_w \subseteq F'_w$.

It is easy to see that $\sqsubseteq$ is a partial order, i.e., that it is reflexive, transitive and anti-symmetric. In such an order, an element $x$ which is greater than or equal to all elements of a subset $C$ of the ordered set is called an *upper bound* of $C$. An upper bound of $C$ which is an element of $C$ is called the *greatest* element of $C$. There need not always be such an element, but if there is one, it is unique. The conjecture ventured above is that for every coherent equivalence system $\approx$, $\pi(\approx)$ is the greatest element of the set of coherent permutation systems which determine $\approx$. To state this more concisely, define $\varepsilon^{-1}(\approx)$ to be the preimage of $\approx$ under $\varepsilon$, i.e., the set of permutation systems $F$ such that $\varepsilon(F) = \approx$. As the next lemma shows, $\pi(\approx)$ is an upper bound of $\varepsilon^{-1}(\approx)$:

**Lemma 2.6.2.** *For every coherent equivalence system* $\approx$ *and* $F \in \varepsilon^{-1}(\approx)$, $F \sqsubseteq \pi(\approx)$.

*Proof.* Consider any $w \in W$ and $f \in F_w$. As $\pi(\approx)_w = \mathrm{aut}(\approx)_{(W/\approx_w)}$, it suffices to show that $f \in \mathrm{aut}(\approx)_{(W/\approx_w)}$. Since $F$ is coherent, $f \in \mathrm{aut}(F)$, and so by Lemma 2.4.1, $f \in \mathrm{aut}(\varepsilon(F))$. $\varepsilon(F) = \approx$ by assumption, so $f \in \mathrm{aut}(\approx)$. By definition of $\varepsilon$, $f \subseteq \varepsilon(F)_w$, so $f \subseteq \approx_w$, and therefore $f \in \mathrm{aut}(\approx)_{(W/\approx_w)}$. □

For every coherent equivalence system $\approx$, $\pi(\approx) \in \varepsilon^{-1}(\approx)$ by Lemma 2.4.5, so the conjecture follows immediately from this and the previous lemma:

**Theorem 2.6.3.** *For every coherent equivalence system* $\approx$, $\pi(\approx)$ *is the greatest element of* $\varepsilon^{-1}(\approx)$ *under* $\sqsubseteq$.

While this shows that every preimage of a coherent equivalence system has a greatest element, this is clearly not the case for sets of coherent permutation systems in general: a set of two permutation systems $F$ and $F'$ such that for some $w \in W$, neither $F_w \subseteq F'_w$ nor $F'_w \subseteq F_w$, has no greatest element. But something closely related holds: Every set $C$ of coherent permutation systems has a *least upper bound*, written $\bigvee C$, i.e., an upper bound of $C$ which is less than or equal to all upper bounds of $C$. (Least upper bounds are also unique, and in general, there need not be one. If a set has a greatest element, this is its least upper bound.) $C$ also has a greatest lower bound, written $\bigwedge C$, i.e., a lower bound of $C$ which is greater than or equal to all lower bounds of $C$, where a lower bound of $C$ is of course an element which is less than or equal to all elements of $C$. A partial order in which every set $C$ has both a least upper bound and a greatest lower bound is called a *complete lattice*. In such an order, there are in particular the least upper bound and the greatest lower bound of the set of all elements; these can be thought of as the greatest and least elements overall and are written $\top$ and $\bot$. The following proposition shows that coherent permutation systems form a complete lattice, specifying greatest lower bounds, $\top$ and $\bot$. To state it, write $(S_W)_w$ for the set of permutations of $W$ which map $w$ to itself – again, this is the stabilizer of $w$, now with respect to $S_W$, the set of permutations of $W$, which is called the *symmetric group on* $W$.

**Proposition 2.6.4.** *The set of coherent permutation systems ordered by $\sqsubseteq$ is a complete lattice, where for any subset $C$, $\bigwedge C$, $\top$ and $\bot$ are the permutation systems such that for all $w \in W$:*

$$\bigwedge C_w = (S_W)_w \cap \bigcap_{F \in C} F_w$$

$$\top_x = (S_W)_w$$

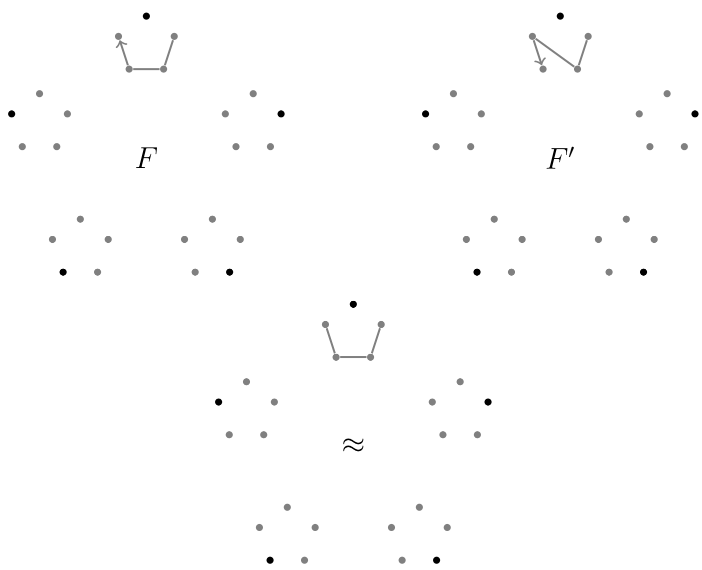
$$\bot_w = \{\mathrm{id}\}$$

*Proof.* It is routine to show that the function which maps every $w \in W$ to $(S_W)_w \cap \bigcap_{F \in C} F_w$ is a coherent permutation system and the greatest lower bound of $C$. Any partial order in which every set has a greatest lower bound (i.e., any complete meet-semilattice) is also a complete lattice. It is again routine to show that the functions which map every $w \in W$ to $(S_W)_w$ and $\{\mathrm{id}\}$ are $\top$ and $\bot$, respectively. □

Since sets of coherent permutation systems have both least upper bounds and greatest lower bounds and any preimage of a coherent equivalence system contains its least upper bound, one might wonder whether any such set also contains its greatest lower bound. It turns out that this is not the case, at least in the sense that for some sets $W$, the claim does not hold:

**Proposition 2.6.5.** *For some set $W$ and coherent equivalence system $\approx$ on $W$, $\bigwedge \varepsilon^{-1}(\approx) \notin \varepsilon^{-1}(\approx)$.*

*Proof.* $F$ and $F'$ are coherent and $\varepsilon(F) = \varepsilon(F') = \approx$, but $F \wedge F' = \bot \notin \varepsilon^{-1}(\approx)$:




□

This shows that there is in general no function analogous to $\pi$ which maps every coherent equivalence system to the least element of its preimage. To be more precise, $x$ is the *least* element of a set $C$ partially ordered by $\leq$ just in case $x \in C$ and $x \leq y$ for all $y \in C$. Such a least element is the greatest lower bound, so by Proposition 2.6.5, the preimages of some coherent equivalence systems do not have least elements.

Consider now the corresponding order on coherent equivalence systems:

**Definition 2.6.6.** *$\preccurlyeq$ is the binary relation on the set of coherent equivalence systems such that $\approx\ \preccurlyeq\ \approx'$ just in case for all $w \in W$, $\approx_w\ \subseteq\ \approx'_w$.*

Again, it is easy to see that this is a partial order. What else can be said about it? A natural conjecture is that it is isomorphic to the image of $\pi$ (the set of permutation systems determined by coherent equivalence systems) ordered by $\sqsubseteq$. More specifically, one might conjecture that for any coherent equivalence systems $\approx$ and $\approx'$, $\approx\ \preccurlyeq\ \approx'$ if and only if $\pi(\approx) \sqsubseteq \pi(\approx')$. This turns out not to be the case; while the 'if' direction holds, the 'only if' direction does not. To establish the former claim, it will be shown that $\varepsilon$ is order-preserving, in the sense that for all coherent permutation systems $F$ and $F'$, if $F \sqsubseteq F'$ then $\varepsilon(F) \preccurlyeq \varepsilon(F')$:

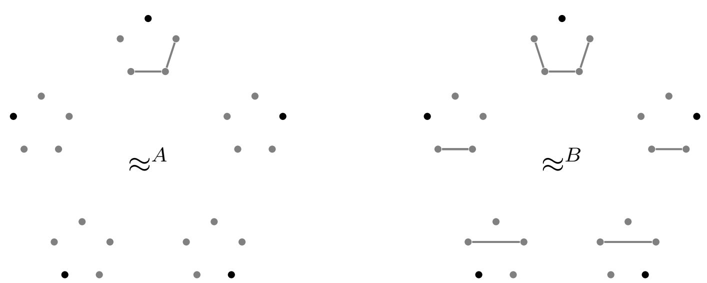
**Proposition 2.6.7.** *$\varepsilon$ is order-preserving.*

*Proof.* Assume that $F$ and $F'$ are coherent permutation systems such that $F \sqsubseteq F'$. If $v\varepsilon(F)_w u$, then there is an $f \in F_w$ such that $f(v) = u$, so given $F \sqsubseteq F'$, $f \in F'_w$ and therefore $v\varepsilon(F')_w u$. □

In particular, for every coherent equivalence systems $\approx$ and $\approx'$, if $\pi(\approx) \sqsubseteq \pi(\approx')$ then $\varepsilon(\pi(\approx)) \preccurlyeq \varepsilon(\pi(\approx'))$, and so by Lemma 2.4.5, $\approx\ \preccurlyeq\ \approx'$. To show that the other direction of the conjecture does not hold, it will be shown that $\pi$ is not guaranteed to be order-preserving, in the sense that for some coherent equivalence systems $\approx$ and $\approx'$, $\approx\ \preccurlyeq\ \approx'$ while $\pi(\approx) \not\sqsubseteq \pi(\approx')$:

**Proposition 2.6.8.** *For some set $W$, $\pi$ is not order-preserving.*

*Proof.* $\approx^A$ and $\approx^B$ are coherent and $\approx^A\ \preccurlyeq\ \approx^B$, but $\pi(\approx^A) \not\sqsubseteq \pi(\approx^B)$ since $(234) \in \pi(\approx^A)_1$ and $(234) \notin \pi(\approx^B)_1$:



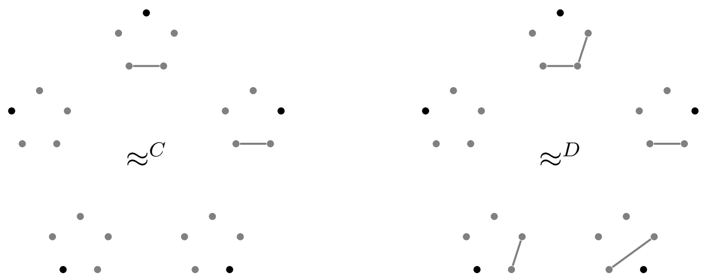
□

So while $\pi$ is a bijection from coherent equivalence systems to its image, it is not an isomorphism between the two sets ordered by $\preccurlyeq$ and $\sqsubseteq$.

What structure do coherent equivalence systems ordered by $\preccurlyeq$ form? It turns out that in contrast to coherent permutation systems, they do not form a complete lattice. It can even be shown that they neither form a join- nor a meet-semilattice, i.e., that two coherent equivalence systems are neither guaranteed to have a least upper bound nor guaranteed to have a greatest lower bound:[1]

**Proposition 2.6.9.** *For some set $W$, coherent equivalence systems on $W$ ordered by $\preccurlyeq$ form neither a join- nor a meet-semilattice.*

*Proof.* Let $\approx^A$ and $\approx^B$ be as in the proof of Proposition 2.6.8, and consider in addition the following coherent equivalence systems:




We argue (i) that $\approx^A$ and $\approx^C$ have no least upper bound, and (ii) that $\approx^B$ and $\approx^D$ have no greatest lower bound.

For (i), note that $\approx^B$ and $\approx^D$ are both upper bounds of $\approx^A$ and $\approx^C$. $\approx^D$ is the only upper bound $\approx$ of $\approx^A$ and $\approx^C$ such that $\approx \preccurlyeq \approx^D$, as none of the other three candidate equivalence systems is coherent. Since $\approx^D \not\preccurlyeq \approx^B$, it follows that there is no upper bound $\approx$ of $\approx^A$ and $\approx^C$ such that $\approx \preccurlyeq \approx^B$ and $\approx \preccurlyeq \approx^D$, and therefore no least upper bound of $\approx^A$ and $\approx^C$.

For (ii), note that $\approx^A$ is the only lower bound $\approx$ of $\approx^B$ and $\approx^D$ such that $\approx^A \preccurlyeq \approx$: Any such $\approx$ maps 3, 4 and 5 to the identity relation; since also $2 \approx_1 3$, $\approx_2$ is the identity relation; finally, since $\approx_1^A = \approx_1^D$, $\approx_1 = \approx_1^A$. Since

[1] The results established in this section also immediately show that $\pi$ and $\varepsilon$ do not form a Galois connection, as one might have conjectured.

$\approx^C \not\preccurlyeq \approx^A$, it follows that there is no lower bound $\approx$ of $\approx^B$ and $\approx^D$ such that $\approx^A \preccurlyeq \approx$ and $\approx^C \preccurlyeq \approx$, and therefore no greatest lower bound of $\approx^B$ and $\approx^D$. □

## 2.7 Conclusion

Two kinds of models for propositional contingentism were developed above, which were shown to be equivalent as models of contingency in what propositions there are, but not equivalent overall. Permutation systems were shown to draw finer distinctions than equivalence systems, and a philosophically motivated example was given for this difference using individuals and their relations. The details of the example suggest that the present treatment of contingency in what propositions there are can be expanded into a more comprehensive theory of higher-order contingency, i.e., contingency in what propositions, properties and relations there are. An investigation of this kind was already carried out in great detail in Fine (1977b). Fritz and Goodman (unpublished c, here ch. 1) explore variants of Fine's proposal which take up some further ideas from Stalnaker (2012), and argue that Fine's proposal must be revised to take contingency in what relations there are seriously. The discussion in Fritz and Goodman (unpublished c, here ch. 1) shows that there are a number of choice points in how to develop a theory of higher-order contingency. On the one hand, (Fritz, unpublished a, here ch. 3) shows that many but not all of them agree on the patterns of contingency in what propositions there are, which exactly correspond to the kinds of models developed here, i.e., coherent equivalence systems. Some similar results for patterns of symmetries are established there as well, showing that it depends on the particular details of the theory whether they exactly correspond to the class of coherent permutation systems.

Two aspects of the present model theory for propositional contingentism are explored elsewhere. First, (Fritz, unpublished c, here ch. 5) interprets two extensions of propositional modal logic on coherent equivalence systems. The first is an extension by propositional quantifiers, which are naturally interpreted at a given world as ranging over the propositions which there are at the world according to the system, i.e., the unions of the sets of equivalence classes of the equivalence relation associated with the world. This logic is shown not to be recursively axiomatizable, since it is recursively isomorphic to second-order logic. The second extension adds a modality which expresses that there is the proposition expressed by the formula it operates on, which can be seen as a fragment of the first extension. An axiomatization is proposed, but questions of completeness are left open.

Second, the ramifications of propositional contingentism on the semantics of counterfactuals are explored in (Fritz and Goodman, unpublished a, here ch. 4). It is argued that the present models of propositional contingentism are straightforwardly combined with the theory of counterfactuals of Lewis (1973), but that they are in tension with the theory of counterfactuals of Stalnaker (1968). The main point of tension arises from the principle of conditional excluded middle, which turns out to hold only for propositions there are at a world, not for propositions there could be.

## 2.8 Appendix on Stalnaker's Models

Stalnaker (2012, Appendix A) presents two classes of models on which the above development of coherent equivalence and permutation systems is based. The present appendix describes the differences between the above definitions and Stalnaker's definitions, and argues for the former. In inessential respects, Stalnaker's notation is modified to simplify the comparison.

Before considering the formal definitions, one merely terminological difference between Stalnaker (2012) and the present article must be mentioned: What is called a "world" here is called a "point (of logical space)" by Stalnaker. Stalnaker uses "world" for maximally strong non-trivial propositions at a world, which in equivalence systems are represented by equivalence classes.

Stalnaker defines the class of models corresponding to coherent permutation systems as follows: For each member $w$ of a set $W$, let $F_w$ be a set of permutations on $W$ such that:

(1') If $f \in F_w$, then $f(w) = w$.

(2') $F_w$ is closed under inverse and composition.

(3') If $f \in F_w$ and $g \in F_v$, then $f.g \in F_{f(v)}$.

It is routine to show that all of these constraints are satisfied if $F$ is a coherent permutation system; in particular, (2') follows from permutation systems mapping worlds to permutation groups, and (3') follows from the fact that the coherence constraint requires $F_w$ to be a set of automorphisms of $F$. As Stalnaker formulates his condition, the converse cannot be established, as his condition does not rule out $F_w$ being empty. This is ruled out for permutation systems, since permutation groups must contain the identity function. But Stalnaker seems committed to the stronger requirement as well, since he implicitly relies on it on p. 138, where he argues that the relation of a world being mapped to another by a member of $F_w$ is an equivalence relation. If the condition of $F_w$ not being empty is added to (2'), it is routine to show that any $F$ satisfying Stalnaker's constraints is a coherent permutation system: That $F$ is a permutation system follows from the fact that it is a function from $W$ to sets of permutations on $W$ satisfying the strengthened version of (2'). That each member of $F_w$ maps $w$ to itself is required by (1'), so it only remains to show that each $f \in F_w$ is an automorphism of $F$. So let $v \in W$, and consider any $g \in F_v$. Then by (3'), $f.g \in F_{f(v)}$, so $f.F_v \subseteq F_{f(v)}$. As by (2'), $f^{-1} \in F_w$, an analogous argument establishes that $f^{-1}.F_{f(v)} \subseteq F_{f^{-1}f(v)}$ and thus $F_{f(v)} \subseteq f.F_v$. Hence $F_{f(v)} = f.F_v$, as required.

Stalnaker defines the class of models corresponding to coherent equivalence systems as follows: For each member $w$ of a set $W$, let $\approx_w$ be an equivalence relation on $W$ such that:

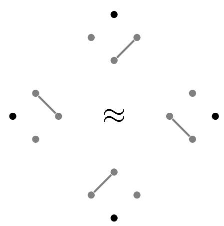
(1) If $w \approx_w v$ then $w = v$.

(2) If $v \approx_w u$, then there exists a permutation function $f$ from $W$ onto $W$ meeting these two constraints:

  (2a) $f(v) = u$

  (2b) for any $x$, $y$, and $z$, $y \approx_x z$ if and only if $f(y) \approx_{f(x)} f(z)$

It is routine to show that all of these constraints are satisfied if $\approx$ is a coherent equivalence system; in particular, (1) follows from the fact that if $w \approx_w v$ then there must be an automorphism mapping $w$ to itself as well as $w$ to $v$, and (2) follows from the fact that (2b) is equivalent to the condition of $f$ being an automorphism. Considering the converse direction, note that the permutation $f$ whose existence is required in (2) is not required to be a subset of $\approx_w$, as in the definition of coherent equivalence systems. This suggests the possibility of an equivalence system satisfying Stalnaker's constraints without being coherent. The following example shows that there are such equivalence systems:




This evidently satisfies condition (1) of Stalnaker's constraints; for condition (2), note that either (1234) or (4321) witnesses the existential claim in any non-trivial case. To see that $\approx$ is not coherent, consider the fact that $2 \approx_1 3$. The only non-trivial permutation $f$ which is a subset of $\approx_1$ is $f = (23)$, but this is not an automorphism: $3 \approx_2 4$ holds, but $f(3) \approx_{f(2)} f(4)$, i.e., $2 \approx_3 4$ does not.

Thus, Stalnaker's constraints on equivalence systems are strictly weaker than being coherent. With Theorem 2.4.2, it follows that not every equivalence system satisfying Stalnaker's constraints is determined by a coherent permutation system. This can also be shown directly by noting that the only permutation system which determines $\approx$ is incoherent. Thus Stalnaker's claim that his two models of propositional contingency are equivalent is incorrect; to reinstate it, the stronger condition of coherence for equivalence systems must be imposed.

Stalnaker does provide a formal argument to show that his constraints on $F$ and $\approx$ are equivalent. However, Stalnaker only shows that if a permutation system $F$ satisfies his constraint, then so does $\varepsilon(F)$, and if an equivalence system $\approx$ satisfies his constraint then so does $\pi(\approx)$. While the first of these results is to the point (cf. Theorem 2.4.2), the second results is strictly speaking irrelevant – what is required is that every equivalence system $\approx$ which satisfies his constraint is determined by some permutation system $F$ which satisfies his constraint (cf. Theorem 2.4.6), and this turns out not to be the case.

The philosophical discussion in section 2.2 already motivates coherence as defined here. It can be supported further by considering the above example of $\approx$ in more detail. According to $\approx$, 2 and 3 are indistinguishable at 1. But at 1, there is a proposition, namely $\{1\}$, which there is at 2 but not at 3. Thus 2 and 3 *can* be distinguished in terms of resources available at 1, and so cannot be indistinguishable at 1. As in the case of $\approx^F$, this line of thought can also be put in terms of what propositions there are: at 1, there should be the proposition that there is the proposition $\{1\}$, which is $\{1, 2\}$, contradicting the fact that $2 \approx_1 3$.

## 2.9 Appendix on Worlds

Each equivalence or permutation system is based on a set $W$, whose members where called "worlds" in incautious formulations, and said to "represent worlds" in more cautious formulations. Both of these formulations seem to presuppose that there are these entities – worlds. And this seems to be in conflict with the talk of indistinguishability between worlds engaged in above: Two worlds were said to be indistinguishable if they cannot be distinguished in terms of the resources available, i.e., in terms of what there is. Thus any two worlds should be distinguished in terms of themselves, and so there should not be any indistinguishable worlds. This conclusion might seem like a *reductio* of the whole project. But it is not a *reductio*, for even if there *are* no indistinguishable worlds, this does not rule out that there *could be* worlds which actually are indistinguishable. This suggests that talk of there being a certain world should be understood as it being possible that there is such a world.

To illustrate this understanding of world-talk in more detail, the following takes worlds to be maximally strong non-trivial propositions, as suggested, e.g., in Stalnaker (1976). For brevity, call such propositions “maximal”. Furthermore, quantification over propositions will be understood as quantification into sentence position, so the proposal will be spelled out in a formal language with propositional variables $p, q, \ldots$, the usual Boolean operators, $\Box$ for necessity and quantifiers $\exists$ and $\forall$ binding propositional variables. As the above models did not include an accessibility relation, the correctness of the modal logic **S5** will be assumed.

In such a setting, a proposition $p$ can be understood to be maximal if it is possible and strictly entails each proposition or its negation: $\Diamond p \wedge \forall q(\Box(p \rightarrow q) \vee \Box(p \rightarrow \neg q))$ (cf. Fine (1970) and Kaplan (1970)). If there is contingency in what propositions there are, being maximal might not suffice to count as a world, since a proposition might be maximal without being necessarily maximal. Given the models of propositional contingency developed here, being necessarily maximal suffices for being counted as a world, so define the following syntactic abbreviation:

$$world(p) := \Diamond p \wedge \Box\forall q(\Box(p \rightarrow q) \vee \Box(p \rightarrow \neg q))$$

Correspondingly, a claim being true at a world can be understood as being strictly entailed by it:

$$@w\varphi := \Box(w \rightarrow \varphi)$$

Talk of there being a world satisfying a certain condition is to be understood as talk of it being possible that there is a proposition which is necessarily maximal and actually satisfies the condition. Adapting a strategy found in Fine (1977a), the following spells this out formally in a way which makes sure that “actually” is interpreted in such a way that the construction can be embedded in modal operators:

$$\exists v(world(v) \wedge v \wedge \Diamond\exists w(world(w) \wedge @v \ldots))$$

This provides a way of understanding talk of there being a world satisfying a certain condition in terms of propositional quantification without begging the question of propositional contingentism. Some other forms of quantification such as universal quantification can of course be treated similarly. But the strategy is limited; it is neither obvious how to treat a generalized quantificational claim like "there are uncountably many worlds such that ..." nor higher-order quantificational claim like "there is a binary relation among worlds such that ...".

The latter is especially worrying, since on the present proposal of understanding of world-talk, quantification over relations and permutations among worlds, which the above model-theoretic discussion freely engaged in, is naturally understood in terms of higher-order quantification. This issue is most naturally investigated in the richer setting of higher-order modal logic, in which there are quantifiers over individuals, as well quantifiers over relations in a type hierarchy of relations, treating propositions as nullary relations. As mentioned in section 2.5, a fuller treatment of higher-order contingentism in such a setting is developed in Fritz and Goodman (unpublished c, here ch. 1). There, it is shown that the present worries about being able to make sense of quantification over relations among worlds are well-founded, since reformulating a theory of higher-order contingency in such a way that the relations used to formulate these theories are required to satisfy the constraints of the relevant theory themselves restricts which structures the theory admits. Whether this also restricts the patterns of indistinguishability and symmetries determined by these structures, and so puts additional restrictions on the coherence constraints developed here, is a difficult issue, which is only partly settled by the results in Fritz (unpublished a, here ch. 3).

The formal results are more conclusive in case of the generalized quantifier "there are uncountably many worlds such that ...". The results obtained in Fritz (unpublished b, here ch. 7) show that claims of this form cannot be expressed even in a higher-order modal language in which quantifiers are available for all types of relations, all generalized quantifiers are available as primitive expressions for all types, and which is infinitary in the sense of allowing conjunctions of arbitrary sets of formulas and universal and existential quantifiers binding sets of variables of arbitrary cardinality, and containing variables of all types.

Whether these limitative results tell against the present theory of propositional contingency is not obvious, and may depend on the use to which the models developed here are put. Fritz and Goodman (unpublished b, here ch. 6) use analogous limitations of talk of possible individuals to argue against higher-order contingentism in general, and so in particular against propositional contingentism.

# Chapter 3

# Higher-Order Contingentism, Part 2: Patterns of Indistinguishability

**Abstract.** We use the notions of equivalence systems and permutation system developed by Robert Stalnaker in his formal models of contingently existing propositions to investigate two kinds of patterns of indistinguishability of the structures for higher-order modal logic discussed in Part 1. We find that the three classes of structures which formalize the most plausible theories of higher-order contingentism all agree as far as the contingent existence of propositions is concerned.

This paper is a continuation of Part 1 (Fritz and Goodman, unpublished c, here ch. 1), familiarity with which is assumed. Stalnaker (2012, Appendix A) models the contingent existence of propositions using formal structures which have much in common with the models investigated in Part 1. In the terminology of Fritz (unpublished d, here ch. 2), he studies permutation systems, which map every member of a set of worlds to a group of permutations of the set of worlds, and equivalence systems, which map every member of a set of worlds to a set of equivalence relations on the set of worlds. We can understand the former as representing how the worlds can be permuted in ways indistinguishable from the perspective of a given world, and the latter as representing which pairs of worlds are indistinguishable from the perspective of a given world; I will call the former patterns of *global* indistinguishability and the latter patterns of *local* indistinguishability. In both cases, Stalnaker imposes coherence constraints, which he claims to be equivalent. Fritz (unpublished d, here ch. 2) shows that one of Stalnaker's constraints is too weak, but also shows how to strengthen it in a philosophically motivated way to obtain two matching notions of coherence.

From the structures studies in Part 1, we can straightforwardly derive these two kinds of patterns of indistinguishability, in the sense that from any structure, we can straightforwardly derive a permutation system, which in turn straightforwardly determines an equivalence system. We can thus use permutation and equivalence systems as formal tools to study the patterns of (global as well as local) indistinguishability admitted by various classes of structures. This is the aim of the present paper. In particular, I will focus on the question whether a given class of structures determines all and only *coherent* systems, for the notions of coherence defined in Fritz (unpublished d, here ch. 2). For both permutation and equivalence systems, we will obtain a number of positive results on this question. A particular reason for being interested in the results on equivalence systems is that they also establish that the propositional fragments of the relevant classes of structures exactly match the model of contingently existing propositions developed in Stalnaker (2012, Appendix A).

Before starting with permutation systems and facts about global indistinguishability, let me dispel a source of potential terminological confusion: Stalnaker (2012) uses the label 'point (of logical space)' for what is called a 'world' in Part 1 (as well as in Fine (1977b) and Williamson (2013)). Stalnaker uses the term 'world' for atomic propositions. (Note that since it may be a contingent matter which propositions are atomic, on Stalnaker's terminology, it may be a contingent matter which propositions are worlds.) To preserve continuity in terminology with Part 1, I here use 'world' as in Part 1 for what Stalnaker calls a 'point'.

## 3.1 Permutation Systems for Global Indistinguishability

We start with Stalnaker's definition of a permutation system. A *permutation group on a set* $X$ is a set of permutations of $X$ which forms a group, i.e., which is nonempty and closed under composition and inverses.

**Definition 3.1.1.** *For every set* $W$*, a* permutation system on $W$ *is a function* $F$ *mapping every* $w \in W$ *to a permutation group* $F_w$ *on* $W$*.*

For every structure $\mathfrak{S}$ based on a set of worlds $W$ and world $w \in W$, the model theory of Part 1 defines a set $\text{fix}(\mathfrak{S}, w)$ of automorphisms of the structure $\mathfrak{S}$, which represents the ways modal space can be permuted in ways indistinguishable from the perspective of $w$ in $\mathfrak{S}$. Automorphisms consist of a permutation of worlds and a permutation of possible individuals, so given the automorphisms at the worlds of the structure, we can derive how worlds can be permuted in ways indistinguishable from the perspective of a given world $w$ by taking the world-permutation components of the automorphisms in $\text{fix}(\mathfrak{S}, w)$. This motivates the following natural way of deriving a permutation system from a structure:

**Definition 3.1.2.** *For any structure* $\mathfrak{S} = \langle W, I, D\rangle$*,* the permutation system determined by $\mathfrak{S}$*, written* $\pi(\mathfrak{S})$*, is the function* $\pi(\mathfrak{S})$ *which maps every* $w \in W$ *to*

$$\pi(\mathfrak{S})_w = \{f : \textit{for some } g,\ \langle f, g\rangle \in \text{fix}(\mathfrak{S}, w)\}.$$

That this is well-defined is easily seen from the fact that $\text{fix}(\mathfrak{S}, w)$ is a group. Note that in Fritz (unpublished d, here ch. 2), another function $\pi$ is defined, which maps every equivalence system $\approx$ to a permutation system $\pi(\approx)$; we will let context disambiguate between these two functions $\pi$.

Our main result of this section is to show that a permutation system is coherent if and only of it is determined by some structure. We start by reproducing the definition of coherence of Fritz (unpublished d, here ch. 2), based on Stalnaker (2012, Appendix A). To define it, we let a permutation $f$ of a set $W$ be an *automorphism of a permutation system* $F$ if $f.F_w = F_{f(w)}$ for all $w \in W$, writing $\text{aut}(F)$ for the set of automorphisms of $F$ and $\text{aut}(F)_w$ for the stabilizer of $w$, i.e., the subset of automorphisms which map $w$ to itself:

**Definition 3.1.3.** *A permutation system $F$ on a set $W$* coheres *if for all $w \in W$, $F_w \subseteq \text{aut}(F)_w$.*

We start with the easier direction of the result:

**Proposition 3.1.4.** *Every structure determines a coherent permutation system.*

*Proof.* Let $\mathfrak{S} = \langle W, I, D \rangle$ be a structure. For any $w \in W$, $\text{fix}(\mathfrak{S}, w)$ is a subgroup of $\text{aut}(\mathfrak{S})$, so $\pi(\mathfrak{S})_w$ is a subgroup of $(S_W)_w$. Thus $\pi(\mathfrak{S})$ is a permutation system in which each member of $\pi(\mathfrak{S})_w$ maps $w$ to itself. To show that all such permutations are automorphisms, it suffices to consider any $w, v \in W$, $f \in \pi(\mathfrak{S})_w$ and $g \in \pi(\mathfrak{S})_v$, and prove that $f.g \in \pi(\mathfrak{S})_{f(v)}$. Since $f \in \pi(\mathfrak{S})_w$ and $g \in \pi(\mathfrak{S})_v$, there are $f', g' \in S_I$ such that $\langle f, f' \rangle \in \text{fix}(\mathfrak{S}, w)$ and $\langle g, g' \rangle \in \text{fix}(\mathfrak{S}, v)$. By Lemma 1.4.11 (ii) of Part 1, $\langle f, f' \rangle . \langle g, g' \rangle \in \text{fix}(\mathfrak{S}, f(v))$. $\langle f, f' \rangle . \langle g, g' \rangle = \langle f.g, f'.g' \rangle$, so $f.g \in \pi(\mathfrak{S})_{f(v)}$. □

To prove the harder direction, we need a few tools. First, note that a *strict total order of a set* $W$ is a binary relation $<$ which is *irreflexive* (not $w < w$), *transitive* (if $w < v$ and $v < u$ then $w < u$) and *total* (if $w \neq v$ then $w < v$ or $v < w$). Such a relation is a *strict well-ordering* if it is also *well-founded* (every non-empty set $W' \subseteq W$ has a least element, i.e., there is a $w \in W'$ such that for all $v \in W'$, not $v < w$). Note that by our conventions on extending functions to sets and sequences, we can apply a permutation $f$ of $W$ to a strict well-order $<$ of $W$: for all $w, v \in W$, $f(w) f(<) f(v)$ if and only if $w < v$. Similarly, for any set $\Theta$ of strict well-orderings, $f(\Theta) = \{f(<) : < \in \Theta\}$. Also, for any permutation group $G$ on $W$ and strict well-order $<$ of $W$, we write $G(<)$ for $\{f(<) : f \in G\}$; this is called the *orbit of* $<$. If $f$ is a permutation of $W$, then $fG$ is defined as $\{fg : g \in G\}$ and called a *left coset of* $G$.

Our proof strategy is to construct, for a given coherent permutation system $F$, a structure which determines it. On the one hand, this requires including enough in the domain of the structure at a given world $w$ to rule out all automorphisms whose world-permutation is not in $F_w$. To do so, we let the individual domain be empty and include in the higher-order domains, for some strict well-order $<$ of the worlds, the representation of the property of being a well-order to which $<$ is mapped by some member of $F_w$ – this representation is mapped to itself by an automorphism if and only if its world-permutation is in $F_w$. (From a group-theoretic perspective, this follows from the fact that if we let the permutations of a set $W$ act in the natural way on orders on $W$, then for every strict well-order $<$ of $W$, every permutation group on $W$ is the set-wise stabilizer of its own orbit of $<$.) On the other hand, we have to make sure that the condition of being an automorphism of the overall structure does not rule out too many permutations for the structure to determine $F$. It turns out that this can be ensured by including in the domain of a given world $w$ the representations of properties of being a well-order to which $<$ is mapped by some member of $F_w$ for *all* strict well-orders of $W$.

We start implementing this strategy by defining representations of properties of relations among worlds; as for other representations, we adopt the convention to drop the upper type index in representations where it is clear from context.

**Definition 3.1.5.** *Let $\mathfrak{F} = \langle W, I \rangle$ be a frame. For any set $\Theta$ of binary relations on $W$, define $\Theta_{\mathfrak{F}}^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle} \in \iota_{\mathfrak{F}}^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}$ such that for all $w \in W$:*

$$\Theta_{\mathfrak{F}}^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}(w) = \left\{ R_{\mathfrak{F}}^{\langle\langle\rangle,\langle\rangle\rangle} : R \in \Theta \right\}$$

**Lemma 3.1.6.** *Let $\mathfrak{F} = \langle W, I\rangle$ be a frame, $\langle f, g\rangle \in \mathrm{aut}(\mathfrak{F})$, and $\Theta$ a set of strict well-orderings of $W$. Then $\langle f, g\rangle.\Theta_{\mathfrak{F}} = (f(\Theta))_{\mathfrak{F}}$.*

*Proof.* Routine, using Lemma 1.5.13 (iv) of Part 1. □

**Theorem 3.1.7.** *Every coherent permutation system is determined by a structure.*

*Proof.* Consider any coherent permutation system $F$ on a set $W$. To specify the domain of our structure $\mathfrak{S}$ which we prove to determine $F$, let $[<]_G = \{f(<) : f \in G\}^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}_{\langle W,\emptyset\rangle}$ for any permutation group $G$ on $W$ and strict well-ordering $<$ of $W$. Let $\mathfrak{S}$ be the structure $\langle W, \emptyset, D\rangle$ such that for all types $t \neq \langle\langle\langle\rangle,\langle\rangle\rangle\rangle$, $D^t_W = \emptyset$, and for all $w \in W$,

$$D^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}_w = \{[<]_{F_w} : < \text{ is a strict well-ordering of } W\}.$$

We prove that $F = \pi(\mathfrak{S})$. Consider any $f \in S_W$ and $w \in W$. To prove that $f \in F_w$ iff $f \in \pi(\mathfrak{S})_w$, we make use of the following consequence of Lemma 3.1.6: For any strict well-order $<$ of $W$, $f^\emptyset.[<]_{F_w} = [<]_{fF_w}$ (where as in Part 1, $f^\emptyset = \langle f, \emptyset\rangle$). Call this observation $(*)$.

Assume first that $f \in F_w$. To show that $f \in \pi(\mathfrak{S})_w$, we have to show that $f^\emptyset \in \mathrm{fix}(\mathfrak{S}, w)$, which is the case if (i) $f^\emptyset \in \mathrm{aut}(\mathfrak{S})$, (ii) $f(w) = w$, and (iii) $f^\emptyset.[<]_{F_w} = [<]_{F_w}$ for any strict well-order $<$ of $W$. For (i), consider any $v \in W$; we show that $f^\emptyset.D^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}_v = D^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}_{f(v)}$. For $\subseteq$, consider any strict well-order $<$ of $W$. By $(*)$, $f^\emptyset.[<]_{F_v} = [<]_{fF_v} = \{fg(<) : g \in F_v\}_{\langle W,\emptyset\rangle} = \{f.g(f(<)) : g \in F_v\}_{\langle W,\emptyset\rangle} = \{gf(<) : g \in F_{f(v)}\}_{\langle W,\emptyset\rangle} = [f(<)]_{F_{f(v)}} \in D^{\langle\langle\langle\rangle,\langle\rangle\rangle\rangle}_{f(v)}$. (For the last identity, note that $f$ is an automorphism of $F$.) As usual, we can derive $\supseteq$ using inverses. (ii) follows from the coherence of $F$. For (iii), consider any strict well-order $<$ of $W$. As observed in $(*)$, $f^\emptyset.[<]_{F_w} = [<]_{fF_w} = [<]_{F_w}$ (for the last identity, note that since $f \in F_w$, $fF_w = F_w$).

Assume now that $f \in \pi(\mathfrak{S})_w$. Then $f^\emptyset \in \mathrm{fix}(\mathfrak{S}, w)$. Let $<$ be any strict well-ordering of $W$. Then $f^\emptyset.[<]_{F_w} = [<]_{F_w}$, so by $(*)$, $[<]_{fF_w} = [<]_{F_w}$, and therefore $\{fg(<) : g \in F_w\}_{\langle W,\emptyset\rangle} = \{g(<) : g \in F_w\}_{\langle W,\emptyset\rangle}$. Since the mapping of sets of strict well-orders of $W$ to their representations is injective, $\{fg(<) : g \in F_w\} = \{g(<) : g \in F_w\}$. So there is a $g \in F_w$ such that $f(<) = g(<)$. As $<$ is a strict well-ordering of $W$, $f = g$, and therefore $f \in F_w$. □

Once this result is in place, it is natural to go on asking whether it extends to $-/+$closed structures:

**Question 3.1.8.** *Is every coherent permutation system determined by a $-/+$closed structure?*

This is not obvious, and I will not settle the question here. We could easily derive a positive answer from the previous theorem if we had the result that the symmetries of modal space don't change under generation; i.e., if we could prove that for all signs $\times$, structures $\mathfrak{S}$ and worlds $w$: $\text{fix}(\mathfrak{S}, w) = \text{fix}(\otimes\mathfrak{S}, w)$. We proved that $\text{fix}(\mathfrak{S}, w) \subseteq \text{fix}(\otimes\mathfrak{S}, w)$ in Part 1, Lemma 1.4.13 (ii), but it is unclear whether these sets are in fact identical. To see how this may fail, note that while $\mathfrak{S}$ need not be a $-/+$structure, $-/+$generating imposes the being constraint or its positive weakening; therefore, the domain of $\otimes\mathfrak{S}$ need not contain the domain of $\mathfrak{S}$, and consequently, there might be an automorphism in $\text{fix}(\otimes\mathfrak{S}, w)$ not included in $\text{fix}(\mathfrak{S}, w)$.

This observation shows that even if not all coherent permutation systems are determined by a $-/+$closed structure, then this is in a sense a relatively fragile matter. This is because the observation in the last paragraph shows that the match established in Theorem 3.1.7 *can* be extended to closed structure in a setting in which neither the being constraint nor its positive weakening are imposed. Similarly, the match could also be established if we were dealing with $+$closed structures for a type hierarchy including types of relations of arbitrary ordinal arity, an extension we will come back to in Part 3, section 7.4.2. In such a setting, we could adapt the above proof by representing a property of orders of worlds not as above, but using the corresponding $|W|$-ary relation among world-propositions – the important difference is that such a relation would not be ruled out by the positive weakening of the being constraint.

While Question 3.1.8 seems difficult to answer, the further strengthening of the hypothesis to internally $+$closed structures can easily be refuted:

**Theorem 3.1.9.** *Not every coherent permutation system $F$ is determined by an internally $+$closed structure.*

*Proof.* Let $W = \{1, 2, 3, 4\}$ and $F$ the permutation system on $W$ such that for all $w \in W$, $F_w = (S_W)_w$. $F$ is coherent (in fact, $F$ is the top element of the lattice of coherent permutation systems on $W$; see Fritz (unpublished d, here ch. 2)). If $\mathfrak{S}$ is a structure such that $F = \pi(F)$, then for all $w \in W$ and $f \in (S_W)_w$, there is a $g$ such that $\langle f, g\rangle \in \text{fix}(\mathfrak{S}, w)$. So in the sense of Definition 1.6.18 of Part 1, for all $w \in W$, the members of $W\backslash\{w\}$ are collectively indistinguishable, and so by Proposition 1.6.21 of Part 1, $\mathfrak{S}$ is not internally +closed. □

Thus, while the patterns of global indistinguishability among worlds – how worlds can be permuted in ways indistinguishable from the perspective of a given world – admitted by all structures exactly correspond to coherent permutation systems, the ones admitted by internally +closed structures correspond to a more restricted class.

## 3.2 Equivalence Systems for Local Indistinguishability

We now turn to equivalence systems, which record which pairs of worlds can be distinguished from the perspective of a given world:

**Definition 3.2.1.** *For every set $W$, an* equivalence system on $W$ *is a function $\approx$ mapping every $w \in W$ to an equivalence relation $\approx_w$ on $W$.*

We can derive an equivalence system from a structure by associating with each world the equivalence relation which holds between two worlds if and only if there is a way of permuting modal space which maps one to the other. Since permutation systems naturally determine equivalence system in this way, as defined in Fritz (unpublished d, here ch. 2), we can define the determination of equivalence systems by chaining the two notions of determination we already have. We first reproduce the definition of the determination of equivalence systems from permutation systems:

**Definition 3.2.2.** *For every permutation system $F$ on a set $W$, the* equivalence system determined by $F$, *written $\varepsilon(F)$, is such that for all $w \in W$:*

$\varepsilon(F)_w = \bigcup F_w$.

We can now straightforwardly define the determination of an equivalence system from a structure, re-using, as in the case of $\pi$, the symbol $\varepsilon$ used for the function mapping permutation systems to equivalence systems also for the function mapping structures to equivalence systems:

**Definition 3.2.3.** *For any structure* $\mathfrak{S}$*,* the equivalence system determined by $\mathfrak{S}$*, written* $\varepsilon(\mathfrak{S})$*, is* $\varepsilon(\pi(\mathfrak{S}))$*.*

As in the case of permutation systems, we will focus on proving that structures determine all and only coherent equivalence systems. In fact, we will establish various stronger versions of this result, by showing that all structures determine coherent equivalence systems, while various interesting subclasses of structures determine all coherent equivalence systems. To do so, we first define coherence as in Fritz (unpublished d, here ch. 2), where an automorphism of an equivalence system $\approx$ is a permutation $f$ such that $v \approx_w u$ iff $f(v) \approx_{f(w)} f(u)$, and $\mathrm{aut}(\approx)$ and $\mathrm{aut}(\approx)_w$ are derived as above:

**Definition 3.2.4.** *An equivalence system* $\approx$ *on a set* $W$ coheres *if for all* $w, v, u \in W$ *such that* $v \approx_w u$*, there is an* $f \in \mathrm{aut}(\approx)_w$ *such that* $f(v) = u$ *and* $f \subseteq \approx_w$*.*

That structures only determine coherent equivalence systems follows straightforwardly from results we already have at our disposal:

**Proposition 3.2.5.** *For any structure* $\mathfrak{S}$*,* $\varepsilon(\mathfrak{S})$ *is a coherent equivalence system.*

*Proof.* By Proposition 3.1.4, $\pi(\mathfrak{S})$ is a coherent permutation system, so as proven in Fritz (unpublished d, here ch. 2), $\varepsilon(\pi(\mathfrak{S}))$ is a coherent equivalence system. □

We start by proving that every coherent equivalence systems is determined by a $\times$closed structure, for any sign $\times$, and then extend this to the analogous claim for Finely generated structures. Finally, we consider internally $-$closed structures and internally $+$closed structures, proving that the analogous claim does not hold in the former case, and leaving the issue for the latter case open.

Before delving into these details, let me mention a couple of interesting consequences we get from these results. The first is based on the observation that we can straightforwardly derive a propositional domain function from an equivalence system: with each world $w$, we associate the unions of equivalence classes of the equivalence relation associated with $w$. It is easy to see that for any (+ or −)closed structure $\mathfrak{S}$, the propositional domain function derived in this way from the equivalence system determined by $\mathfrak{S}$ is the propositional domain function of $\mathfrak{S}$ itself. Thus a corollary of the results to be established is that +closed, −closed, and Finely generated structures all agree on which propositional domain functions they admit.

From this we can derive our second observation, namely that these classes of structures agree on which sentences they count as valid if we limit ourselves to the propositional fragment of the language, i.e., the sentences in which all occurring variables and non-logical constants are of type $\langle\rangle$. Furthermore, we can interpret this propositional language on equivalence systems, and obtain that its validities on any of the three classes of structures agree with those on coherent equivalence systems. In particular, this is so for the finitary fragment of this propositional language, in which quantification and conjunction is only allowed over finite sets, which is especially interesting since we know that the validities in this fragment over coherent equivalence systems are recursively isomorphic to second-order logic, as proven in Fritz (unpublished c, here ch. 5).

### 3.2.1 Closed Structures

**Proposition 3.2.6.** *For any sign $\times$, every coherent equivalence system is determined by a $\times$closed structure.*

To keep it readable, I split up the proof into several lemmas. Let $\times$ be a sign and $\approx$ a coherent equivalence system on a set $W$. Let $\mathfrak{F}$ be the frame $\langle W, \emptyset\rangle$. For any $w, v \in W$, define $P^v_w = ([v]_{\approx_w})_{\mathfrak{F}}$; this can thought of as the propositional intension corresponding to the equivalence class of $v$ under the indistinguishability relation at $w$. The basic idea of this proof is that we can recover $\approx$ from these equivalence classes, so we can use the structure $\times$generated from the structure which contains only the corresponding propositional intensions in its higher-order domains. So we first define $\mathfrak{S} = \langle W, \emptyset, B\rangle$ to be the structure such that $B^t_W = \emptyset$ for all types $t \neq \langle\rangle$ and $B^{\langle\rangle}_w = \{P^v_w : v \in W\}$ for all $w \in W$. Let $\otimes\mathfrak{S} = \langle W, \emptyset, D\rangle$. We prove that $\approx = \varepsilon(\otimes\mathfrak{S})$. As above, we write $f^\emptyset$ for $\langle f, \emptyset\rangle$ for all $f \in S_W$.

**Lemma 3.2.7.** *Let $w, v \in W$ and $f \in S_W$ such that $f \subseteq\approx_w$. Then $f^\emptyset.P^v_w = P^v_w$.*

*Proof.* Let $u \in X$. $(f^\emptyset.P^v_w)(f(u)) = f^\emptyset.(P^v_w(u)) = f^\emptyset(\{\langle\rangle : u \in [v]_{\approx_w}\}) = \{\langle\rangle : u \in [v]_{\approx_w}\} = \{\langle\rangle : f(u) \in [v]_{\approx_w}\} = P^v_w(f(u))$. So $f^\emptyset.P^v_w = P^v_w$. □

**Lemma 3.2.8.** *Let $f$ be an automorphism of $\approx$. Then for any $w, v \in W$, $f^\emptyset.P^v_w = P^{f(v)}_{f(w)}$.*

*Proof.* Consider any $u \in W$. $(f^\emptyset.P^v_w)(f(u)) = f^\emptyset.(P^v_w(u)) = f^\emptyset.\{\langle\rangle : u \in [v]_{\approx_w}\} = \{\langle\rangle : v \approx_w u\} = \{\langle\rangle : f(v) \approx_{f(w)} f(u)\} = \{\langle\rangle : f(u) \in [f(v)]_{\approx_{f(w)}}\} = P^{f(v)}_{f(w)}(f(u))$. □

**Lemma 3.2.9.** *$B \sqsubseteq D$.*

*Proof.* Consider any $w \in W$ and $p \in B^T_w$. Then $p \in B^{\langle\rangle}_w$, so $D \boxtimes p$. Also, for any $\xi \in \mathrm{fix}(\mathfrak{S}, w)$, $\xi.p = p$, so $p \in D^{\langle\rangle}_w$. □

**Lemma 3.2.10.** *For any $w, v, u \in W$, $v \approx_w u$ iff there is an $f \in S_W$ such that $f^\emptyset \in \mathrm{fix}(\otimes\mathfrak{S}, w)$ and $f(v) = u$.*

*Proof.* Assume first that $v \approx_w u$. By coherence, there is an automorphism $f$ of $\approx$ such that $f(v) = u$ and $f \subseteq\approx_w$. By Lemma 1.4.13 (ii) of Part 1, $\mathrm{fix}(\mathfrak{S}, w) \subseteq \mathrm{fix}(\otimes\mathfrak{S}, w)$, so it suffices to show that $f^\emptyset \in \mathrm{fix}(\mathfrak{S}, w)$. To show that $f^\emptyset \in \mathrm{aut}(\mathfrak{S})$, we have to show that for all $x \in W$ and types $t$, $f^\emptyset.B^t_x = B^t_{f(x)}$. The claim is trivial for all types $t \neq \langle\rangle$. $f^\emptyset.B^{\langle\rangle}_x = f^\emptyset.\{P^y_x : y \in W\} = \{f^\emptyset.P^y_x : y \in W\}$. By Lemma 3.2.8, this is the set $\left\{P^{f(y)}_{f(x)} : y \in W\right\} = \left\{P^y_{f(x)} : y \in W\right\} = B^{\langle\rangle}_{f(x)}$. Since $f \subseteq\approx_w$, $f(w) = w$. Consider any $p \in B^T_w$. Then $p = P^x_w$ for some $x \in W$. By Lemma 3.2.7, $f^\emptyset.p = p$. So $f^\emptyset \in \mathrm{fix}(\mathfrak{S}, w)$, as required.

Assume now that there is an $f \in S_W$ such that $f^{\emptyset} \in \text{fix}(\otimes\mathfrak{S}, w)$ and $f(v) = u$. By Lemma 3.2.9, $P^v_w \in D^{\langle\rangle}_w$, so $f^{\emptyset}.P^v_w = P^v_w$. Firstly, $(f^{\emptyset}.P^v_w)(u) = f^{\emptyset}.(P^v_w(v)) = f^{\emptyset}.\{\langle\rangle : v \in [v]_{\approx_w}\} = \{\langle\rangle\}$. Secondly, $P^v_w(u) = \{\langle\rangle : u \in [v]_{\approx_w}\}$. Hence $u \in [v]_{\approx_w}$ and therefore $v \approx_w u$. □

*Proof of Proposition 3.2.6.* Let $w, v, u \in W$. $v\varepsilon(\otimes\mathfrak{S})_w u$ iff there is an $f \in \pi(\otimes\mathfrak{S})_w$ such that $f(v) = u$. The latter is the case iff there is a $\langle f, g\rangle \in \text{fix}(\otimes\mathfrak{S}, w)$ such that $f(v) = u$, which in turn is the case iff there is an $f \in S_W$ such that $f^{\emptyset} \in \text{fix}(\otimes\mathfrak{S}, w)$ and $f(v) = u$. By Lemma 3.2.10, this is the case iff $v \approx_w u$. So $\approx = \varepsilon(\otimes\mathfrak{S})$. □

Thus, the patterns of local indistinguishability among world – which pairs of worlds are indistinguishable from the perspective of a given world – admitted by all structures, by −closed structures, and by +closed structures all exactly correspond to coherent equivalence systems.

### 3.2.2 Finely Generated Structures

We can extend the result of the previous section to the class of structures satisfying the constraints on generation imposed in Fine (1977b):

**Proposition 3.2.11.** *Every coherent equivalence system is determined by a Finely generated structure.*

*Proof.* Consider any a coherent equivalence system $\approx$ on a set $W$. Let $\mathfrak{S} = \langle W, I, B\rangle$, where $I = \left\{([v]_{\approx_w})^{\langle\rangle}_{\langle W,I\rangle} : w, v \in W\right\}$, $B^t_W = \emptyset$ for all types $t \notin \{e, \langle e\rangle\}$, and for all $w \in W$, $B^e_w = \left\{([v]_{\approx_w})^{\langle\rangle}_{\langle W,I\rangle} : v \in W\right\}$ and $B^{\langle e\rangle}_w = \{E\}$, where $E(v) = \{P \in I : \langle\rangle \in P(v)\}$ for all $v \in W$. $\oplus\mathfrak{S}$ is Finely generated, and analogous to the proof of Proposition 3.2.6, we can show that it determines $\approx$. □

In Part 1, we noted that not every +closed structure is Finely generated; the present result shows that these differences do not show up at the propositional level.

### 3.2.3 Internally Generated Structures

We can easily see that the previous results do not extend to the case of internally −closed structures:

**Proposition 3.2.12.** *Some coherent equivalence system is not determined by any internally −closed structure.*

*Proof.* Let $\approx$ be the equivalence system on $W = \{1, 2, 3\}$ such that for all $w, v, u \in W$, $v \approx_w u$ iff ($v = w$ iff $u = w$). Assume for $\lightning$ that there is an internally −closed structure $\mathfrak{S} = \langle W, I, D\rangle$ such that $\varepsilon(\mathfrak{S}) = \approx$. Since $2 \approx_1 3$, $2^{\langle\rangle}_{\langle W,I\rangle} \notin D^{\langle\rangle}_1$, so there must be a $\xi \in \text{fix}(\mathfrak{S}, w)|_c\mathfrak{S}$ such that $\xi.1 = 1$ and $\xi.2 \neq 2$, and therefore $\xi.2 = 3$. So there must be a cumulative representation of (23) in $D^{\langle\langle\rangle,\langle\rangle\rangle}_W$, and hence by negativity of $\mathfrak{S}$ a $w \in W$ such that $\left\{2^{\langle\rangle}_{\langle W,I\rangle}, 3^{\langle\rangle}_{\langle W,I\rangle}\right\} \subseteq D^{\langle\rangle}_w$. $\lightning$. □

The analogous question concerning internally +closed structures seems quite difficult, and we leave it open here.

## 3.3 Global Versus Local Indistinguishability

In Fritz (unpublished d, here ch. 2), it is shown that permutation systems draw finer distinctions than equivalence systems. Give the results in the last two sections, we can use this to conclude that in the class of all structures, the patterns of global indistinguishability outrun the patterns of local indistinguishability: some structures share the same pattern of local indistinguishability while displaying different patterns of global indistinguishability. Since we have not established which permutation systems are determined by any of the other three classes of structures, we cannot straightforwardly extend this insight to these classes. Especially for the highly restrictive class of internally +closed structures, it seems possible that the difference between the patterns of global and local indistinguishability disappears. The following result shows that this is not the case:

**Proposition 3.3.1.** *There are internally +closed structures which determine the same equivalence systems but different permutation systems.*

*Proof.* Define $\mathfrak{S} = \langle W, I, B\rangle$ and $\mathfrak{S}' = \langle W, I, B'\rangle$ such that $W = \{1, 2, 3, 4, 5\}$, $I = \{2, 3, 4\}$, $B^t_W = B'^t_W = \emptyset$ for all types $t \notin \{e, \langle e, e\rangle\}$ and

$$B^e_1 = B'^e_1 = \emptyset$$

$$B^e_i = B'^e_i = I\backslash\{i\} \text{ for all } i \in \{2, 3, 4\}$$

$$B^e_5 = B'^e_5 = I$$

$$B^{\langle e,e\rangle}_i = \emptyset \text{ and } B'^{\langle e,e\rangle}_i = \{<\} \text{ for all } i \in W$$

where $<\in \iota^{\langle e,e\rangle}_{\langle W,I\rangle}$ such that for all $i \in W$,

$$< (i) = \{\langle 2, 3\rangle, \langle 3, 4\rangle, \langle 4, 2\rangle\} \cap (B'^e_i)^2$$

We show that $\oplus\mathfrak{S}$ and $\oplus\mathfrak{S}'$ witness the claim to be established. Since the representations of all permutations of both $W$ and $I$ are in the relevant domain of 5 in both $\oplus\mathfrak{S}$ and $\oplus\mathfrak{S}'$, they are internally +closed. To see that $\pi(\oplus\mathfrak{S}) \neq \pi(\oplus\mathfrak{S}')$, note that $(23) \in \pi(\oplus\mathfrak{S})_1$ but $(23) \notin \pi(\mathfrak{S}')_1$. To see that $\varepsilon(\oplus\mathfrak{S}) = \varepsilon(\oplus\mathfrak{S}')$, note that both of them are identical to the equivalence system $\approx$ such that $\approx_i = \mathrm{id}_W$ for all $i \in \{2, 3, 4, 5\}$, and $\approx_1$ is the equivalence relation corresponding to the partition $\{\{1\}, \{2, 3, 4\}, \{5\}\}$. □

## 3.4 Possible Symmetries

The existence condition for higher-order entities in internally +closed structures is formulated in terms of possibly existing permutations of worlds and possible individuals. One might therefore try to also use permutation and equivalence systems to represent the patterns of *possible* global and local indistinguishability of a structure using permutation and equivalence systems. It turns out that they are not an appropriate framework for doing so; exemplarily, I will show that even for the highly restrictive class of internally +closed structures, the patterns of possible global indistinguishability need not form a permutation system (coherent or otherwise):

**Definition 3.4.1.** *For any structure $\mathfrak{S} = \langle W, I, D\rangle$, define $\rho(\mathfrak{S})$ to be the function on $W$ mapping every $w \in W$ to*

$\rho(\mathfrak{S})_w = \{f : \textit{for some } g, \langle f, g\rangle \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}\}$.

**Proposition 3.4.2.** *There is an internally +closed structure $\mathfrak{S}$ such that $\rho(\mathfrak{S})$ is not a permutation system.*

*Proof.* Let $\mathfrak{S} = \langle W, I, D\rangle$, where $W = \{0, 1, 2, 3\}$, $I = \emptyset$, and $D$ is the domain assignment such that for all $w \in W$, types $t \neq e$ and $o \in \iota^t_{\langle W,I\rangle}$, $o \in D^t_w$ iff $D \boxplus o$ and $f^\emptyset.o = o$ for all $f \in \Xi_w$, where $\Xi = \{\text{id}, (01), (12), (20)\}$. (As above, we write $f^\emptyset$ for $\langle f, \emptyset\rangle$.)

We start by proving that for all $w \in W$, $\text{fix}(\mathfrak{S}, w)|\mathfrak{S} = \{f^\emptyset : f \in \Xi_w\}$.

Consider any $w \in W$ and $f \in \Xi_w$. Then $f(w) = w$, and by construction, $f^\emptyset.o = o$ for all $o \in D^T_w$. So $f^\emptyset \in \text{fix}(\mathfrak{S}, w)$. It remains to show that $f$ is possible in $\mathfrak{S}$. Since $v_{\langle W,I\rangle} \in D^{\langle\rangle}_v$ for all $v \in W$, $D \boxplus f_{\langle W,I\rangle}$. If $f = \text{id}$, then for all $v \in W$, trivially $f^\emptyset.o = o$ for all $f \in \Xi_v$, so $f_{\langle W,I\rangle} \in D^{\langle\langle\rangle,\langle\rangle\rangle}_v$. So assume $v \in \{0, 1, 2\}$ and $f = (v\ v + 1)$; we show that $f_{\langle W,I\rangle} \in D^{\langle\langle\rangle,\langle\rangle\rangle}_{v+2 \bmod 3}$. $\Xi_{v+2 \bmod 3} = \{f\}$, so it suffices to show that $f^\emptyset.f_{\langle W,I\rangle} = f_{\langle W,I\rangle}$, which follows from the fact that $(v\ v + 1).(v\ v + 1) = (v\ v + 1)$ via Lemma 1.5.13 (iv) of Part 1.

Now consider any $w \in W$ and $\xi \in \text{fix}(\mathfrak{S}, w)|\mathfrak{S}$. Then there is an $f \in (S_W)_w$ such that $\xi = f^\emptyset$. *Case 1:* $w = 3$. It suffices to show that $f \notin \{(012), (210)\}$, which we can do by showing that these permutations are not possible in $\mathfrak{S}$. This in turn follows from the fact that for any $v \in \{0, 1, 2\}$, $(v\ v + 1 \bmod 3).(012) \neq (012)$ and $(v\ v + 1 \bmod 3).(210) \neq (210)$ via Lemma 1.5.13 (iv) of Part 1. *Case 2:* $w \neq 3$. As in case 1, we can show that $f \notin \{(123), (321)\}$; it remains to show that $f \notin \{(w\ \ w + 1 \bmod 3), (w\ \ w + 2 \bmod 3)\}$. This follows from the fact that $(w + 1 \bmod 3\ \ w + 2 \bmod 3).(w\ \ w + 1 \bmod 3) \neq (w\ \ w + 1 \bmod 3)$ and $(w + 1 \bmod 3\ \ w + 2 \bmod 3).(w\ \ w + 2 \bmod 3) \neq (w\ \ w + 2 \bmod 3)$, again via Lemma 1.5.13 (iv) of Part 1.

From the claim proven, it follows immediately that $\mathfrak{S}$ is internally +closed. Further, it follows that $\{(01), (12)\} \subseteq \rho(\mathfrak{S})_3$ and $(210) \notin \rho(\mathfrak{S})_3$. Since $(01)(12) = (210)$, $\rho(\mathfrak{S})_3$ is not closed under composition, and thus $\rho(\mathfrak{S})$ is not a permutation system. □

# Chapter 4

# Counterfactuals and Propositional Contingentism

*with Jeremy Goodman*

**Abstract.** Robert Stalnaker has recently defended propositional contingentism, the claim that it is contingent what propositions there are. We consider the implications of this claim for Lewis's and Stalnaker's semantics of counterfactuals, showing how to adapt Lewis's and Stalnaker's semantics of counterfactuals to a formal setting of propositional contingentism. We show that in contrast to Lewis's semantics, adapting Stalnaker's semantics of counterfactuals to propositional contingentism produces two surprising consequences: on the one hand, it imposes strong restrictions on the patterns of contingency in what propositions there are; on the other hand, it weakens Stalnaker's logic of counterfactuals, in particular concerning the principle of conditional excluded middle. We argue that both of these consequences cast doubt on the viability of Stalnaker's semantics for counterfactuals given propositional contingentism.

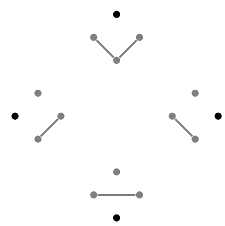
## 4.1 Propositional Contingentism

Stalnaker (2012) argues for propositional contingentism, the claim that it is contingent what propositions there are. E.g., according to him, if Saul Kripke had a daughter, there would be the proposition that she is identical to herself, but since (according to Stalnaker) there is in fact nothing which could be Saul Kripke's daughter, there is in fact no such proposition. In Stalnaker (2012, Appendix A), he develops a way of formally modeling contingency in what propositions there are. In Fritz (unpublished d, here ch. 2), it is shown that there is a problem with Stalnaker's formal development, and how to correct it. This corrected way of formally modeling contingency in what propositions there are takes the form of a class of possible-worlds structures, which will be presented in this section. For a more detailed presentation, see Fritz (unpublished d, here ch. 2).

As familiar from possible worlds structures, the formal models to be defined are based on a set $W$ representing possible worlds, and propositions are represented by subsets of $W$. There is no accessibility relation, so a proposition will be understood to be necessary just in case it is identical to $W$. Contingency in what propositions there are will be modeled by mapping each world $w$ to an equivalence relation $\approx_w$; the idea is that the propositions at $w$ are the subsets of $W$ which contain either both or neither of two worlds related by $\approx_w$. Equivalently, the propositions at $w$ can be taken to be the unions of sets of equivalence classes under $\approx_w$; thus the equivalence classes of $\approx_w$ can be understood as the atomic propositions at $w$. Formally, define the class of structures as follows:

**Definition 4.1.1.** *An* equivalence system on a set $W$ *is a function* $\approx$ *mapping every* $w \in W$ *to an equivalence relation* $\approx_w$ *on* $W$*.*

To visually represent equivalence systems, draw any finite equivalence system based on a set of worlds $\{1, \ldots, n\}$ as follows: Draw the worlds in a circle, starting with 1 at the top and going clockwise, drawing each world $i$ as a smaller circle of dots, again starting with 1 at the top and going clockwise. In the small circle representing $i$, represent $\approx_i$ by connecting dots with lines such that the dots representing $j$ and $k$ are connected by a finite sequence of lines just in case $j \approx_i k$. E.g., the following is a way of drawing the equivalence system based on $\{1, \ldots, 4\}$ which maps 1 to the equivalence relation whose equivalence classes are $\{1\}$ and $\{2, 3, 4\}$, and which maps each $i \in \{2, 3, 4\}$ to the equivalence relation whose equivalence classes are $\{1\}$, $\{i\}$ and $\{2, 3, 4\} \setminus \{i\}$:



Worlds $v$ and $u$ being related by the equivalence relation associated with world $w$ can be understood as $v$ and $u$ being indistinguishable at $w$, in the sense of being indistinguishable using the resources at $w$. This motivates a coherence constraint on equivalence systems: If $v$ and $u$ are indistinguishable using the resources at $w$, then they must in particular be indistinguishable using the resources at $w$ *as represented by the equivalence system*. So they must in particular be indistinguishable in terms of the propositions there are at $w$, the notion of indistinguishability, and finally $w$ itself. Consequently, $v$ and $u$ should be symmetric with respect to these resources, in the sense that there is a permutation of worlds $f$ which maps $v$ to $u$ such that all three resources are invariant under $f$. Since the propositions at $w$ are the unions of sets of equivalence classes of $\approx_w$, they are invariant under $f$ just in case $f$ maps each world to one to which it is related by $\approx_w$. On the usual identification of functions and relations as sets of pairs, this can be written as $f \subseteq \approx_w$. The notion of indistinguishability is invariant under $f$ if $f$ is an automorphism of $f$, in the sense that for all $x, y, z \in W$, $y \approx_x z$ if and only if $f(y) \approx_{f(x)} f(z)$. Of course, $w$ is invariant under $f$ just in case $f(w) = w$. Writing $\mathrm{aut}(\approx)$ for the set of automorphisms of $\approx$ and $\mathrm{aut}(\approx)_w$ for the subset of automorphisms which map $w$ to itself, the formal definition of coherence is therefore as follows:

**Definition 4.1.2.** *An equivalence system* $\approx$ *on a set* $W$ coheres *if for all* $w, v, u \in W$ *such that* $v \approx_w u$*, there is an* $f \in \mathrm{aut}(\approx)_w$ *such that* $f(v) = u$ *and* $f \subseteq \approx_w$*.*

The class of coherent equivalence systems forms the model theory of propositional contingentism used here. While the particular construction of equivalence systems is specific to Stalnaker, the main existing theories of propositional contingentism which assume a coarse-grained theory of propositions agree with this model of contingency in what propositions there are. In particular, this is the case for the development in Fine (1977b) as well as the classes of structures called $-$*closed* and $+$*closed* in Fritz and Goodman (unpublished c, here ch. 1). That these constructions all agree on what patterns of contingency in propositions they allow is proven in Fritz (unpublished a, here ch. 3).

## 4.2 Orderings for Counterfactuals

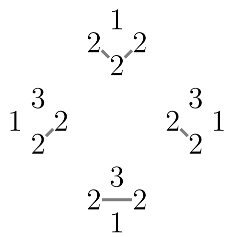
An influential semantics of counterfactuals is due to Lewis (1973), and can be formulated in terms of possible worlds models in which each possible world is associated with an ordering relation among worlds. Dispensing with further complications introduced by an accessibility relation among possible worlds, the semantics associates with each world $w$ a total preorder $\lesssim_w$ on the set of worlds, where $v \lesssim_w u$ represents that $v$ differs no more from $w$ than $u$, or, in other worlds, that $v$ is at least as similar to $w$ as $u$. *Preorders* are reflexive and transitive relations; *totality* requires that for any worlds $v$ and $u$, $v \lesssim_w u$ or $u \lesssim_w v$.

There are two natural ways of adopting this semantics of counterfactuals to the present setting of propositional contingentism as formalized using equivalence systems: each world $w$ of an equivalence system $\approx$ can either be associated with a total preorder of worlds or associated with a total preorder of equivalence classes under $\approx_w$. In the following, the second option will be adopted; the first option comes with interesting complexities which are beyond the scope of this paper and which we hope to explore in the future. The first step in implementing the second option is to extend the definition of an equivalence system:

**Definition 4.2.1.** *An* ordering equivalence system on a set $W$ *is a tuple* $\langle \approx, \lesssim \rangle$ *such that* $\approx$ *is an equivalence system on* $W$ *and* $\lesssim$ *is a function mapping each* $w \in W$ *to a total preorder* $\lesssim_w$ *on* $W/{\approx_w}$.

Here, $W/{\approx_w}$ is the quotient set of $W$ under $\approx_w$, i.e., the set of equivalence classes of $W$ under $\approx_w$. There are a few more items of notation which will be needed in the following: write $[v]_{\approx_w}$ for the equivalence class of $v$ under $\approx_w$, i.e., the set of $u \in W$ such that $v \approx_w u$. If $\lesssim$ is a preorder, write $w < v$ or $v > w$ if $w \lesssim v$ and not $v \lesssim w$, and $w \sim v$ if $w \lesssim v$ and $v \lesssim w$. Note that if $\lesssim$ is a *total* preorder, $w < v$ is the case if not $v \lesssim w$.

To draw an ordering equivalence system on $\{1, \ldots, n\}$, start by drawing the underlying equivalence system. For each $i \leq n$, replace the dots in the small circle representing $i$ by numbers, corresponding to how similar the worlds they represent are to $i$ according to $\approx_i$. I.e., the worlds most similar to $i$ are labeled 1, the worlds next similar to $i$ are labeled 2, etc. E.g., consider the example of an equivalence system presented in section 4.1, expanded to an ordering equivalence system by the ordering function $\lesssim$ such that $\{1\} <_1 \{2, 3, 4\}$, and for all $i \in \{2, 3, 4\}$, $\{i\} <_i \{2, 3, 4\} \backslash \{i\} <_i \{1\}$. This can be drawn as follows:



To adapt the notion of coherence to ordering equivalence systems, it must be ensured that if $v$ and $u$ are indistinguishable from $w$, then they are indistinguishable using all resources at $w$ as represented by the relevant ordering equivalence system. In addition to the resources represented by the underlying equivalence system, this now also includes the notion of similarity. Thus the required permutation mapping $v$ to $w$ must also be an automorphism of the ordering function $\lesssim$, in the sense that for all $w, v, u \in W$,

$[f(v)]_{\approx_{f(w)}} \lesssim_{f(w)} [f(u)]_{\approx_{f(w)}}$ if and only if $[v]_{\approx_w} \lesssim_w [u]_{\approx_w}$. Letting $f$ be an automorphism of an ordering equivalence system $\langle\approx, \lesssim\rangle$ just in case it is an automorphism of $\approx$ and $\lesssim$, and extending the notation of $\mathrm{aut}(\langle\approx, \lesssim\rangle)$ and $\mathrm{aut}(\langle\approx, \lesssim\rangle)_w$ accordingly, coherence for ordering equivalence systems is defined as follows:

**Definition 4.2.2.** *An ordering equivalence system $\langle\approx, \lesssim\rangle$ on a set $W$* coheres *if for all $w, v, u \in W$ such that $v \approx_w u$, there is an $f \in \mathrm{aut}(\langle\approx, \lesssim\rangle)_w$ such that $f(v) = u$ and $f \subseteq \approx_w$.*

In Lewis's discussion of his semantics, a number of additional constraints on the ordering function are important. These can straightforwardly be adapted to the present discussion. To state them, note that an order $\lesssim$ is *antisymmetric* if $x \sim y$ entails $x = y$, and that $\lesssim$ is *well-founded* if it has no infinite descending sequences, i.e., if there is no infinite sequence $x_1, x_2, x_3, \ldots$ such that $x_1 > x_2 > x_3, \ldots$.

**Definition 4.2.3.** *An ordering equivalence system $\langle\approx, \lesssim\rangle$ on a set $W$ satisfies*

weak centering *if for all $w, v \in W$, $[w]_{\approx_w} \lesssim_w [v]_{\approx_w}$,*

strong centering *if for all $w, v \in W$, if $w \neq v$ then $[w]_{\approx_w} <_w [v]_{\approx_w}$,*

*the* limit assumption *if for all $w \in W$, $\lesssim_w$ is well-founded, and*

Stalnaker's assumption *if it satisfies weak centering and the limit assumption, and for all $w \in W$, $\lesssim_w$ is anti-symmetric.*

Let $\Gamma$ be the set of these four conditions. It is easy to see that strong centering entails weak centering and that Stalnaker's assumption entails strong centering.

In Lewis's original discussion, Stalnaker's assumption is so-called because any ordering function satisfying it can be turned into one of the selection functions of the semantics of counterfactuals in Stalnaker (1968), and *vice versa*. This establishes the equivalence of Stalnaker's semantics and Lewis's semantics with Stalnaker's assumption; see Lewis (1973, section 3.4). In exactly the same way, an ordering equivalence system satisfying Stalnaker's assumption can be turned into a corresponding system using a selection function and *vice versa*; the details are routine, and therefore omitted.

## 4.3 Coherent Systems and Stalnaker's Assumption

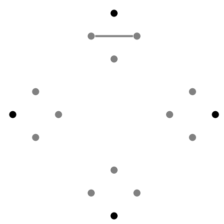
How do coherent ordering equivalence systems relate to coherent equivalence systems? It is clear that every coherent ordering equivalence system is based on a coherent equivalence system. But can every coherent equivalence system $\approx$ be extended to a coherent ordering equivalence system, in the sense that there is a function $\lesssim$ such that $\langle \approx, \lesssim \rangle$ is a coherent ordering equivalence system? In general, this is easily seen to be the case. In fact, there is always such a system which also satisfies strong centering and the limit assumption. But surprisingly, this cannot be extended to Stalnaker's assumption.

**Theorem 4.3.1.** *Every coherent equivalence system can be extended to a coherent ordering equivalence system which satisfies strong centering as well as the limit assumption.*

*Proof.* Consider any coherent equivalence system $\approx$ on a set $W$. Let $\lesssim$ be the function mapping each $w \in W$ to the binary relation $\lesssim_w$ on $W/\approx_w$ such that for all $A, B \in W/\approx_w$, $A \lesssim_w B$ iff $A = \{w\}$ or $B \neq \{w\}$. It is routine to show that $\langle \approx, \lesssim \rangle$ is a coherent ordering equivalence system which satisfies strong centering as well as the limit assumption. □

**Theorem 4.3.2.** *Some coherent equivalence system cannot be extended to a coherent ordering equivalence system satisfying Stalnaker's assumption.*

*Proof.* Let $\approx$ be the equivalence system on $\{1, 2, 3, 4\}$ which maps 1 to the equivalence relation whose equivalence classes are $\{1\}$, $\{2, 4\}$ and $\{3\}$, and maps 2, 3 and 4 to the identity relation:



It is routine to show that $\approx$ is coherent. Assume for contradiction that there is a function $\lesssim$ such that $\langle \approx, \lesssim \rangle$ is a coherent ordering equivalence system satisfying Stalnaker's assumption. Then by Stalnaker's assumption, $\{2\} <_3 \{4\}$ or $\{4\} <_3 \{2\}$. Either way, this contradicts the assumption of coherence, since by $2 \approx_1 4$ and $[3]_{\approx_1} = \{3\}$, this requires there to be an automorphism of $\langle \approx, \lesssim \rangle$ mapping 2 to 4 and 3 to itself. $\lightning$, so there is no such function. $\square$

A way of understanding this proof is as follows: Without considering counterfactuals, there is no problem with 1 being able to single out 3, but not being able to distinguish between 2 and 4, even though 3 is able to distinguish between 2 and 4, since 3 may make the distinction between 2 and 4 in terms which are not available in 1. But assuming that any world must distinguish between its atomic propositions in terms of similarity, then this situation is no longer available, since 3 is required to distinguish between 2 and 4 in terms of similarity, and the notion of similarity is assumed to be a resources available at 1.

Formally, the result opens up the following question, which will be left open:

**Question 4.3.3.** *How can the equivalence systems which can be extended to a coherent ordering equivalence system satisfying Stalnaker's assumption be characterized?*

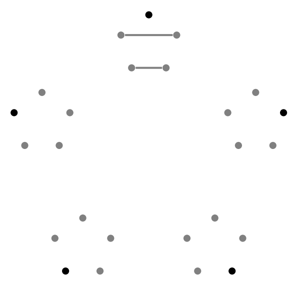
Theorem 4.3.2 shows that if Stalnaker's semantics for counterfactuals is adopted, then Stalnaker's theory of contingency in what propositions there are as encoded in the class of coherent equivalence systems is incomplete: the coherence constraint on equivalence systems formulated above is too weak.

It is instructive to flesh out the case of the equivalence system used in the proof of Theorem 4.3.2 in more detail. Assume that the universe of worlds is constructed by freely recombining two elementary particles $a$ and $b$, taking only into account what individuals there are. That is, there are four worlds, one for every subset $X$ of $\{a, b\}$, at which there are exactly the members of $X$. This naturally induces an equivalence system if one postulates that from the perspective of any world, worlds can only be distinguished using what elementary particles there are (a generalization of this process is central to Fine (1977b) and Fritz and Goodman (unpublished c, here ch. 1)). Thus one can make the following identifications: $1 = \emptyset$, $2 = \{a\}$, $3 = \{a, b\}$, and $4 = \{b\}$.

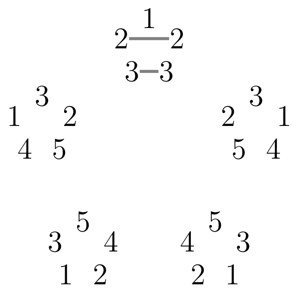
According to Stalnaker's theory of counterfactuals, this configuration of modal space is ruled out. How can this be? How can the structure of similarity between worlds impose such a constraint on how possible worlds might be obtained from freely recombining elementary particles? The following quote from Stalnaker (1984, p. 169) suggests an answer:

> [T]he introduction of counterfactuals allows for finer discriminations between possible worlds than could be made without counterfactuals. The selection functions relative to which counterfactuals are interpreted do not simply select on the basis of facts and criteria of similarity that are intelligible independently of counterfactuals. Rather, the claim is, the fact of selection gives rise to new ways of cutting up the space of possibilities, and so to a richer conception of the way the world is.

One way of reading this suggests that in some sense, counterfactual truths need not supervene on non-modal facts. (Goodman (unpublished) argues that Stalnaker is committed to such a claim independently of propositional contingency.) Thus, one might say that the four-world equivalence system discussed above is ruled out because counterfactuals draw distinctions between worlds which this system identifies. More concretely, one might propose that 3, the world in which there are both possible elementary particles, must be replaced by two worlds: the structure of counterfactuals (or more generally, similarity between worlds) enforces that there are two worlds in which there are two particles, which differ only in which one of the two worlds containing a single elementary particle is more similar. This leads to the following coherent equivalence system:



Along the lines suggested, this can in fact be extended to a coherent ordering equivalence system satisfying Stalnaker's assumption:




This line of response shows that it may be possible to accommodate the surprising result of Theorem 4.3.2 in Stalnaker's overall view. But on the one hand, it is not clear whether this procedure can be generalized; i.e., whether every coherent equivalence system which cannot be extended to a coherent ordering equivalence system satisfying Stalnaker's assumption can be modified analogously to one which can be so extended. On the other hand, although the response may be plausible on the view that the counterfactual facts do in some sense not supervene on non-modal facts, this is a surprising view which must be motivated independently.

## 4.4 Semantics

Lewis and Stalnaker use their semantics to interpret a propositional modal language with a unary connective $\Box$ for necessity and a binary connective $\Box\!\!\rightarrow$

for counterfactuals. For any class of models, a logic is naturally derived, by counting a formula as valid if it is true in every world of every model under every assignment of sets of worlds to propositional variables. In the present setting of propositional contingentism, there is no longer one but at least three natural notions of validity. E.g., a formula could be counted as valid on a class of coherent ordering equivalence systems if if it is true in every world $w$ of every model under every assignment of sets of worlds to propositional variables; alternatively, the universal quantification over assignments could be restricted to those which map every propositional variable to a set of worlds which is one of the propositions in some world, or to those which map every propositional variable to a set of worlds which is one of the propositions in $w$. In order to sidestep questions about whether there is a correct choice in this matter, and if so, which one it is, the following discussion will introduce propositional quantifiers which bind propositional variables, and define validity only for closed formulas. Since the truth of a closed formula is independent of the variable assignment, there is a single natural definition of validity of closed formulas on a class of structures, and this will be employed in the following.

So consider a formal language built up from a countably infinite set of propositional variables $p$, $q$, $r$, …, using a unary negation operator $\neg$, a binary conjunction operator $\wedge$, a unary necessity operator $\Box$, a binary operator $\Box\!\!\rightarrow$ for the counterfactual conditional, and a universal quantifier $\forall$ binding propositional variables. Let $\langle\approx,\lesssim\rangle$ be an ordering equivalence system on a set $W$. An assignment $a$ for this system is a function from proposition letters to subsets of $W$; for any proposition letter $p$ and $P \subseteq W$, write $a[P/p]$ for the assignment which maps every propositional variable $q \neq p$ to $a(q)$ and $p$ to $P$. Adapting the truth-conditions for $\Box\!\!\rightarrow$ from Lewis (1973, p. 49) and letting propositional quantifiers range over the domain of propositions at a world as specified by its relation of indistinguishability, truth is defined inductively as follows:

$\langle\approx,\lesssim\rangle, w, a \vDash p$ iff $w \in a(p)$

$\langle\approx,\lesssim\rangle, w, a \vDash \neg\varphi$ iff not $\langle\approx,\lesssim\rangle, w, a \vDash \varphi$

$\langle \approx, \lesssim \rangle, w, a \models \varphi \wedge \psi$ iff $\langle \approx, \lesssim \rangle, w, a \models \varphi$ and $\langle \approx, \lesssim \rangle, w, a \models \psi$

$\langle \approx, \lesssim \rangle, w, a \models \Box\varphi$ iff $\langle \approx, \lesssim \rangle, v, a \models \varphi$ for all $v \in W$

$\langle \approx, \lesssim \rangle, w, a \models \varphi \mathrel{\Box\!\!\rightarrow} \psi$ iff one of the following is the case:

(i) There is no $v \in W$ such that $\langle \approx, \lesssim \rangle, v, a \models \varphi$.

(ii) There is a $v \in W$ such that $\langle \approx, \lesssim \rangle, v, a \models \varphi$ and for all $u \in W$, if $[u]_w \lesssim_w [v]_w$, then $\langle \approx, \lesssim \rangle, u, a \models \varphi \rightarrow \psi$.

$\langle \approx, \lesssim \rangle, w, a \models \forall p\varphi$ iff $\langle \approx, \lesssim \rangle, w, a[P/p] \models \varphi$ for all $P \subseteq W$ such that for all $v, u \in W$, if $v \approx_w u$ then $v \in P$ iff $u \in P$

As usual, other Boolean operators as well as $\Diamond$ and $\exists$ will be treated as syntactic abbreviations.

It is well-known that in Lewis's and Stalnaker's semantics, necessity is definable in terms of counterfactuals, since $\neg\varphi \mathrel{\Box\!\!\rightarrow} \varphi$ is always equivalent to $\Box\varphi$; see Lewis (1973, chapter 1.5). This is the case as well for ordering equivalence systems; for any formula $\varphi \in L_{\Box\!\rightarrow}$, $\neg\varphi \mathrel{\Box\!\!\rightarrow} \varphi$ is true relative to an ordering equivalence system, point and assignment function if and only if $\Box\varphi$ is true relative to them.

A closed formula $\varphi$ is defined to be *true at* a world of an ordering equivalence system if it is true relative to this system, this world and all assignment functions (or equivalently, some assignment function). $\varphi$ is defined to be *valid on* a class of ordering equivalence systems $C$ just in case $\varphi$ is true at every world of every ordering equivalence system in $C$. For each set $\Delta$ of conditions on ordering equivalence systems defined in Definition 4.2.3 (i.e., $\Delta \subseteq \Gamma$), let $\Lambda_\Delta$ be the set of closed formulas valid on the class of coherent ordering equivalence systems satisfying all conditions in $\Delta$.

## 4.5 Logics

How can $\Lambda_\Delta$ for $\Delta \subseteq \Gamma$ be characterized? A result proven in Fritz (unpublished c, here ch. 5) shows for all except one of these that they cannot be recursively axiomatized, as they are recursively isomorphic to second-order logic. It is shown there that the propositionally quantified logic of the class of coherent equivalence systems is recursively isomorphic to second-order logic. By Theorem 4.3.1, for any $\Delta \subseteq \Gamma$ not including Stalnaker's assumption, $\Lambda_\Delta$ is a conservative extension of the propositionally quantified logic of coherent equivalence systems, and therefore also recursively isomorphic to second-order logic. Whether $\Lambda_\Gamma$ is recursively axiomatizable will be left open here.

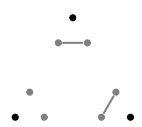
An important schema whose instances are valid on the class of all coherent ordering equivalence systems is a restricted comprehension schema which says, for any formulas $\varphi$, that there is the proposition expressed by $\varphi$ if all propositional variables in $\varphi$ are interpreted as propositions which there are. (An analogous comprehension principle for unary properties is discussed in Williamson (2013, p. 284).) Recall that a coarse-grained view of propositions is assumed on which necessarily equivalent propositions are identical. Thus, the comprehension principle can be formulated as the schema of closed formulas of the following form:

$\mathrm{Comp}_C$ $\forall p_1 \ldots \forall p_n \exists p \Box(p \leftrightarrow \varphi)$ (where $p$ is not free in $\varphi$)

**Theorem 4.5.1.** *All instances of* $\mathrm{Comp}_C$ *are valid on the class of coherent ordering equivalence systems.*

*Proof.* Analogous to the proof of the corresponding result in Fritz and Goodman (unpublished c, here ch. 1). □

Essential for this result is the condition of coherence. To illustrate this, consider the following incoherent equivalence system:

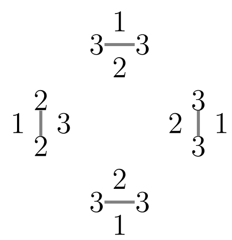


In any ordering equivalence system based on this equivalence system, the following instance of $\mathrm{Comp}_C$ is false at 1:

$\forall p_1 \exists p \Box(p \leftrightarrow \exists q \Box(p_1 \leftrightarrow q))$

To see that it is false at 1, interpret $p_1$ as $\{1\}$; on this interpretation, $\exists q\Box(p_1 \leftrightarrow q)$ expresses the proposition $\{1, 3\}$, which there is not in 1.

For an example essentially involving counterfactuals, consider the following incoherent ordering equivalence system based on a coherent equivalence system:



The following instance of $\mathrm{Comp}_C$ is false at 1 in this system:

$$\forall p_1 \forall p_2 \exists p \Box(p \leftrightarrow (p_1 \mathrel{\Box\!\!\rightarrow} p_2))$$

To see that it is false at 1, interpret $p_1$ as $\{2, 3, 4\}$ and $p_2$ as $\{3\}$; on this interpretation, $p_1 \mathrel{\Box\!\!\rightarrow} p_2$ expresses the proposition $\{1, 3, 4\}$, which there is not in 1.

A natural way of comparing the logics $\Lambda_\Delta$ to familiar logics of counterfactuals is by considering their fragments of closed formulas of the form $\forall p_1 \dots \forall p_n \varphi$ where $\varphi$ is quantifier-free. It turns out that if $\Delta \subseteq \Gamma$ does not contain Stalnaker's assumption, then these are exactly the formulas such that $\varphi$ is valid on the class of Lewis's standard ordering models satisfying the conditions corresponding to those in $\Delta$. Since this result is routine and unsurprising, the details will be skipped. Instead, the next section will show that this observation cannot be extended to coherent ordering equivalence systems satisfying Stalnaker's assumption, using the principle of conditional excluded middle.

## 4.6 Conditional Excluded Middle

A focal point of the debate between Lewis's and Stalnaker's semantics for counterfactuals is the principle of conditional excluded middle, which is often formulated as the following formula:

$(p \mathrel{\Box\!\!\rightarrow} q) \vee (p \mathrel{\Box\!\!\rightarrow} \neg q)$

To see that this is valid on Stalnaker's (standard) models but not on Lewis's, note that Stalnaker's assumption guarantees that if if $p$ is possible, then there is a unique $p$-world most similar to the world of evaluation, and that without Stalnaker's assumption, there might be several such $p$-worlds.

Given the observation in the previous section, it follows that prefixing this formula which universal quantifiers binding $p$ and $q$ yields a formula which is invalid on the class of coherent ordering equivalence systems satisfying strong centering and the limit assumption (and so in particular on all wider classes as well). As one might expect, the resulting closed formula is valid once Stalnaker's assumption is imposed. In fact, it remains valid even if it is strengthened by replacing the universal quantification into antecedent position by a modalized quantification such as $\Box\forall p\Box$. However, the principle cannot be analogously strengthened in the case of the universal quantification into consequent position. To state these observations precisely, define:

$(CEM^-)$ $\Box\forall p\Box\forall q((p \mathrel{\Box\!\!\rightarrow} q) \vee (p \mathrel{\Box\!\!\rightarrow} \neg q))$

$(CEM)$ $\Box\forall p\Box\forall q\Box((p \mathrel{\Box\!\!\rightarrow} q) \vee (p \mathrel{\Box\!\!\rightarrow} \neg q))$

Note that since $\mathrm{Comp}_C$ is valid on the class of coherent ordering equivalence systems, the antecedent $p$ could equivalently be replaced by a complex formula $\varphi$ and the universal quantifier binding $p$ by a sequence of universal quantifiers binding the propositional variables occurring free in $\varphi$, and similarly for the consequent $q$, given that the relevant sequences of propositional variables are disjoint.

**Theorem 4.6.1.** *On the class of coherent ordering equivalence systems satisfying Stalnaker's assumption,*

*(1) $CEM^-$ is valid, but*

*(2) $CEM$ is not valid.*

*Proof.* (1) follows from Theorem 4.6.2 below. For (2), let $\langle \approx, \lesssim \rangle$ be the following ordering equivalence system:

1<br>2–2

3      3<br>1 2    2 1

It is routine to show that $\langle \approx, \lesssim \rangle$ is coherent and satisfies Stalnaker's assumption. There are both propositions $\{2\}$ and $\{2,3\}$ in 2, but for any assignment $a$ such that $a(p) = \{2,3\}$ and $a(q) = \{2\}$, it is not the case that $\langle \approx, \lesssim \rangle, 1, a \vDash (p \mathrel{\Box\!\!\rightarrow} q) \vee (p \mathrel{\Box\!\!\rightarrow} \neg q)$. □

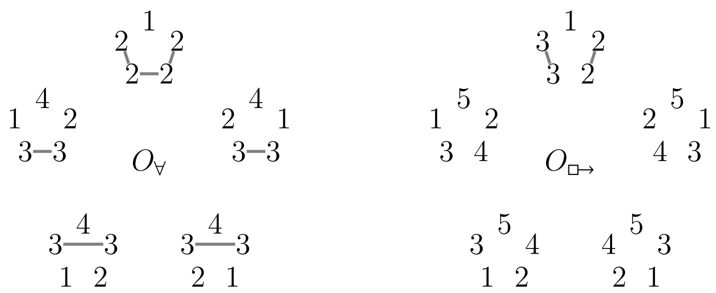
The next section considers this result from a philosophical perspective. To conclude this section, it is shown that a principle slightly stronger than $CEM^-$ is valid on coherent ordering equivalence system satisfying Stalnaker's assumption. The principle is the schema of closed formulas of the following form:

$(CEM^B)$ $\Box \forall p \Box \forall q_1 \ldots \forall q_n((p \mathrel{\Box\!\!\rightarrow} \varphi) \vee (p \mathrel{\Box\!\!\rightarrow} \neg\varphi))$ ($\varphi$ free of $\forall$ and $\Box\!\!\rightarrow$)

**Theorem 4.6.2.** *All instances of $(CEM^B)$ are valid on the class of coherent ordering equivalence systems satisfying Stalnaker's assumption.*

*Proof.* Let $\langle \approx, \lesssim \rangle$ be a coherent ordering equivalence system on a set $W$ which satisfies Stalnaker's assumption, and let $w \in W$. Let $a$ be an assignment function such that for all $i \leq n$, $a(p_i)$ is a union of a set of equivalence classes of $\approx_w$. It suffices to prove that $\langle \approx, \lesssim \rangle, w, a \vDash (p \mathrel{\Box\!\!\rightarrow} \varphi) \vee (p \mathrel{\Box\!\!\rightarrow} \neg\varphi)$. The claim is trivially true if there is no $v \in W$ such that $\langle \approx, \lesssim \rangle, v, a \vDash p$, so assume that there is such a world. By Stalnaker's assumption, there is a $v \in W$ such that $\langle \approx, \lesssim \rangle, v, a \vDash p$ and for all $[u]_w <_w [v]_w$, $\langle \approx, \lesssim \rangle, u, a \nvDash p$. So it suffices to prove that either for all $u \in [v]_w$, $\langle \approx, \lesssim \rangle, u, a \vDash p \rightarrow \varphi$, or for all $u \in [v]_w$, $\langle \approx, \lesssim \rangle, u, a \vDash p \rightarrow \neg\varphi$. Since on a given assignment, a formula of the form of $\Box\psi$ is either true at all or no worlds, any subformula of $\varphi$ of this form can be replaced by either $\top$ or $\bot$ without changing whether $\varphi$ is true in a given world, so for present purposes, $\varphi$ can be taken to be a Boolean combination of $p_1, \ldots, p_n$ and $p$; since $a$ maps each of $p_1, \ldots, p_n$ to a union of sets of equivalence classes of $\approx_w$, the claim follows. □

The exclusion of $\forall$ and $\Box\!\!\rightarrow$ in the schema $(CEM^B)$ raises the question whether allowing one of these operators to occur in $\varphi$ produces instances which are not valid on the class of coherent ordering equivalence systems satisfying Stalnaker's assumption. That this is the case can be shown using the following systems:




For the case of universal quantifiers, consider $O_\forall$ and let $a$ be an assignment function such that $a(p) = \{2, 3, 4\}$ (note that at 2 and 5, there is the proposition $\{2, 3, 4\}$). Then $\forall q \Box(p \rightarrow \exists r \Box((p \wedge q) \leftrightarrow r))$ is true in 2 but not in 3 or 4, so:

$$O_\forall, 1, a \not\models \ (p \mathrel{\Box\!\!\rightarrow} \forall q \Box(p \rightarrow \exists r \Box(p \rightarrow (q \leftrightarrow r)))) \vee$$
$$(p \mathrel{\Box\!\!\rightarrow} \neg\forall q \Box(p \rightarrow \exists r \Box(p \rightarrow (q \leftrightarrow r))))$$

For the case of counterfactuals, consider $O_{\Box\!\!\rightarrow}$ and let $b$ be an assignment function such that $b(p) = \{2, 3, 4\}$ and $b(q) = \{4, 5\}$ (note that 1, there is the proposition $\{4, 5\}$). Then:

$$O_{\Box\!\!\rightarrow}, 1, b \not\models \ (p \mathrel{\Box\!\!\rightarrow} (q \mathrel{\Box\!\!\rightarrow} p)) \vee (p \mathrel{\Box\!\!\rightarrow} \neg(q \mathrel{\Box\!\!\rightarrow} p))$$

## 4.7 Arguments for Conditional Excluded Middle

Given the informal explanation of Lewis's and Stalnaker's semantics of counterfactuals as one associating worlds with a *similarity* order, Stalnaker's assumption is *prima facie* counterintuitive – why shouldn't there be ties in similarity? Arguments are therefore needed to motivate Stalnaker's semantics over Lewis's. The arguments which have been discussed in the literature have mainly taken the form of ruling out Lewis's semantics by arguing for conditional excluded middle. The fact that in the context of propositional contingentism, only certain unmodalized forms of conditional excluded middle are validated by Stalnaker's semantics poses the threat of a dilemma for Stalnaker: if the main arguments for conditional excluded middle extend to the modalized principle (*CEM*), then they cannot be used to motivate his semantics.

One important motivation for conditional excluded middle are pre-theoretic judgements of the truth of its instances independently of which sentences are chosen as antecedents and consequents. These pre-theoretic judgements are even acknowledged by David Lewis, the principal critic of conditional excluded middle, in Lewis (1973, p. 80). A defense of conditional excluded middle based on them is given in Stalnaker (1981). An example of the kind of judgements this defense relies on is the following, letting $c$ be some particular coin which is in fact not tossed:

(1) $c$ would come up heads were it tossed or
$c$ would not come up heads were it tossed.

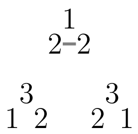
If this is judged to be true, then the following modalized example should also be judged as true, assuming the possibility of $c_1$ being a coin:

(2) Necessarily,
$c_1$ would come up heads were it tossed or
$c_1$ would not come up heads were it tossed.

However, because of the initial "necessarily", (*CEM*), not just ($CEM^B$) is required to explain this judgement. Thus the pre-theoretic judgements which Stalnaker uses to motivate his theory of counterfactuals also tell against the combination of this theory with propositional contingentism.

Might there be another general principle which is valid on coherent ordering equivalence systems satisfying Stalnaker's assumption which can be used to explain our judgement that (2) is true? This can be ruled out by exhibiting a natural system in which (2) is false. To do so, let 2 be a world in which coins $c_1$ and $c_2$ are flipped, $c_1$ comes up heads and $c_2$ comes up tails.

Let 3 be a world which is exactly like 2 in qualitative respects, although $c_1$ and $c_2$ have witched places, so that $c_1$ comes up tails and $c_2$ comes up heads. Let 1 be a world in which there are neither $c_1$ nor $c_2$. It is natural to model these three worlds using the following system, already discussed in the proof of Theorem 4.6.1:



Interpret $p$ as the proposition that $c_1$ is tossed, i.e., $\{2, 3\}$, and $q$ as the proposition that $c_1$ comes up heads, i.e., $\{2\}$. Then $\Box((p \mathrel{\Box\!\!\rightarrow} q) \vee (p \mathrel{\Box\!\!\rightarrow} \neg q))$ is false in all worlds, but this is just a formalization of (2).

More sophisticated arguments for conditional excluded middle have been discussed recently, e.g., in Williams (2010) and Klinedinst (2011). In a number of cases, these can be extended to arguments for (*CEM*), and so to arguments against the combination of Stalnaker's views on counterfactuals and propositional contingentism. Although a fuller discussion of these issues is left for a future occasion, the fact that the basic argument for conditional excluded middle using pre-theoretical judgements generalizes to (*CEM*) in such a straightforward way already indicates that the worry for the combination of Stalnaker's views presented in this section must be taken very seriously.

## 4.8 Indeterminacy

In the last section, it was mentioned that even in the standard case in which there is no contingency in what propositions there are, the assumption of Stalnaker's semantics that there cannot be ties in similarity between worlds is implausible. Similarly, one might note that if the disjunction (1) is true, one of its disjuncts should be true, which is also implausible. Stalnaker (1981, pp. 89–91) responds to these worries by appealing to semantic indeterminacy in counterfactuals. The idea is that there is not one but many similarity relations associated with a given world, each of which corresponds to one way of resolving the indeterminacy in counterfactuals. Adopting a supervaluational treatment of indeterminacy, Stalnaker takes a sentence to be true (*simpliciter*) if it is true on all ways of resolving the indeterminacy, false (*simpliciter*) if it is false on all of them, and indeterminate otherwise. This allows each similarity relation to admit no ties while allowing for cases in which it depends on the how the indeterminacy in counterfactuals is resolved which worlds are counted as more similar to a given world than another. Similarly, this allows (1) to be true without either of its disjuncts being true.

Can such an appeal to indeterminacy be used to respond to the worries for Stalnaker's theory discussed above? In the case of conditional excluded middle, indeterminacy does not seem to help. As Stalnaker remarks, introducing indeterminacy into his semantics of counterfactuals does not change which formulas are valid; consequently (*CEM*) will still not be valid once a supervaluational treatment of indeterminacy of counterfactuals is added to the present model theory. More concretely, there seems to be no way of applying the indeterminacy of counterfactuals to the counterexample to (2) in the previous section: The ordering function used in this example is the only natural one on the toy model of modal reality the example uses.

Appealing to indeterminacy in counterfactuals is more promising as a way of replying to the worries discussed in section 4.3. It was shown there that a coherent equivalence system which is a natural toy model of modal space cannot be extended to a coherent ordering equivalence system satisfying Stalnaker's assumption. If, following the idea of indeterminacy, each world is not associated with a single ordering relation but several, there might be a way of providing a coherent extension, by associating world 3 with the following two orderings $\lesssim_3^1$ and $\lesssim_3^2$: $\{3\} <_3^1 \{2\} <_3^1 \{4\} <_3^1 \{1\}$ and $\{3\} <_3^2 \{4\} <_3^2 \{2\} <_3^2 \{1\}$. Setting up such a model theory rigorously and investigating whether every coherent equivalence system can be so extended will be left open.

## 4.9 Conclusion

Propositional contingentism as formalized using coherent equivalence systems is straightforwardly combined with Lewis's semantics for counterfactuals, leading to a natural and interesting model theory. Two surprising consequences emerge from restricting the model theory to systems which satisfy the constraints of Stalnaker's semantics for counterfactuals. The first is that this restriction rules out natural patterns of contingency in what propositions there are. While this is a surprising consequence, it is one that defenders of Stalnaker's semantics of counterfactuals might simply accept; alternatively, it may be possible to argue that no natural patterns of contingency in what propositions there are are ruled out once indeterminacy in counterfactuals is taken into account. The second surprising consequence is that on the restricted model theory of systems satisfying the constraints of Stalnaker's semantics for counterfactuals, only an unmodalized version of conditional excluded middle is valid. This is a serious problem for the combination of propositional contingentism with Stalnaker's theory of counterfactuals, since the basic argument for Stalnaker's semantics of counterfactuals is an argument for conditional excluded middle which turns out to support the stronger modalized version of conditional excluded middle. This makes it doubtful that Stalnaker's theory of counterfactuals is compatible with his propositional contingentism.

# Chapter 5

# Logics for Propositional Contingentism

**Abstract.** Robert Stalnaker has recently advocated propositional contingentism, the claim that it is contingent what propositions there are. He has proposed a philosophical theory of contingency in what propositions there are and sketched a possible worlds model theory for it, on which a propositional modal language with propositional quantifiers can straightforwardly be interpreted. In this note, it is proven that the propositionally quantified modal logic of Stalnaker's model theory is recursively isomorphic to second-order logic and so has no complete recursive axiomatization. A more restrictive language in which propositional quantifiers are replaced by a modality of existence is briefly discussed.

In many cases, adding propositional quantifiers to modal logics significantly increases their complexity. E.g., Kaminski and Tiomkin (1996) show that for any normal modal logic $\Lambda$ contained in $\mathbf{S4.2}$, the propositionally quantified modal logic of the class of $\Lambda$-frames is recursively isomorphic to second-order logic. (Here, propositional quantifiers range over all subsets of the set of worlds of the frame on which they are interpreted.) This is in sharp contrast to the case of the propositionally quantified modal logic of **S5**-frames (or equivalently, frames with a universal accessibility relation), which was axiomatized in Bull (1969), Kaplan (1970) and Fine (1970), and shown to to be decidable in Kaplan (1970) and Fine (1970).

These results are especially noteworthy since many applications of modal logic deal with modalities whose propositional logic is **S5**. However, it turns out that the tractability of propositional quantification in this case is not only highly specific to the use of equivalence or universal relations in interpreting the modality, but fragile in another dimensions as well: If a variable propositional domain function is added to frames with a universal accessibility relation according to the construction sketched in (Stalnaker, 2012, Appendix A), the same sharp increase in complexity occurs as in the case of wider classes of frames, as the logic of this class of structures is also recursively isomorphic to second-order logic.

## 5.1 Propositional Contingentism

In the literature on propositional quantification in modal logic, quantifiers are usually interpreted as ranging over a fixed domain of subsets of the set of worlds; in many cases, this domain is simply the power set of the set of worlds. So according to these models, the existence of propositions is a non-contingent matter. However, many philosophers are propositional contingentists; they think that there is contingency in what propositions there are; see, e.g., Stalnaker (2012) and (Williamson, 2013, chapter 6) for recent discussion. The main motivation for this is the idea that it is contingent what individuals there are (e.g., I might not have been, maybe if my parents had never met), and that some propositions about such individuals existentially depend on them, in the sense that they wouldn't have been had the relevant individuals not been (e.g., the proposition that I am me might not have been had I not been). Taking this kind of aboutness of propositions seriously might motivate one to adopt a finer-grained understanding of propositions than the one assumed in possible world semantics, where propositions are understood as sets of worlds. E.g., this approach is pursued in Fine (1980) in the form of a number of first-order modal theories of propositions. But it is also possible to stick to the coarse-grained conception of propositions as sets of possible worlds; this is what Stalnaker and Williamson do, and I will follow them here.

Formally, it is straightforward to adapt the standard model theory for propositional quantification based on Kripke frames to contingency in what propositions there are, understanding propositions as sets of possible worlds: Simply add a domain function which maps every world to a set of sets of worlds – the set of propositions which there are at that world. Adding such a domain function was in fact already briefly suggested in (Fine, 1970, section 2.4), but it seems that this suggestion has not been taken up in the literature on propositional quantification. However, Robert Stalnaker develops a variant of such models in (Stalnaker, 2012, Appendix A), which makes two assumptions that allow him to simplify the models. The first assumption is that the accessibility relation among worlds is the universal relation, and so can simply be dropped. The second assumption is that the propositional domain function maps every world to a complete atomic field of sets; since complete atomic fields of sets bijectively correspond to equivalence relations, a function can be used which maps every world to an equivalence relation among worlds (taking the corresponding propositional domain to be the set of unions of equivalence classes of this relation). So formally, define the structures Stalnaker is concerned with as follows:

**Definition 5.1.1.** *For every set $W$, an* equivalence system on $W$ *is a function $\approx$ mapping every $W \in W$ to an equivalence relation $\approx_w$ on $W$.*

Restricting propositional quantification at a world $w$ of an equivalence system $\approx$ on $W$ to the unions of equivalence classes of $\approx_w$, it is clear how to evaluate truth of a propositionally quantified modal formula relative to $\approx$, one of its worlds $w$, and an assignment function which maps every proposition letter to a subset of $W$. In (Stalnaker, 2012, Appendix A), Stalnaker works out what amounts to two constraints on equivalence systems, which he claims to be equivalent. I show in Fritz (unpublished d, here ch. 2) that they are in fact distinct, and that only the one which I call *coherence* is philosophically motivated. To define coherence, let an automorphism of an equivalence system $\approx$ on a set $W$ be a permutation $f$ of $W$ such that for all $w, v, u \in X$, $v \approx_w u$ iff $f(v) \approx_{f(w)} f(u)$, and write $\text{aut}(\approx)$ for the set of automorphisms of $\approx$ and $\text{aut}(\approx)_w$ for the subset of automorphisms which map $w$ to itself. With this, coherence is defined as follows:

**Definition 5.1.2.** *An equivalence system $\approx$ on a set $W$* coheres *if for all $w, v, u \in W$ such that $v \approx_w u$, there is an $f \in \text{aut}(\approx)_w$ such that $f(v) = u$ and $f \subseteq \approx_w$.*

In brief, the idea behind this condition is that if $v \approx_w u$, then $v$ and $u$ cannot be distinguished at $w$, and so there must be a permutation $f$ of worlds which maps $v$ to $u$ and respects all resources at $w$. That $f$ respect the propositions at $w$ is enforced by requiring $f \subseteq \approx_w$; that $f$ respect the notion of indistinguishability is enforced by requiring $f$ to be an automorphism of $\approx$; and that $f$ respect $w$ itself is enforced by requiring $f$ to map $w$ to itself. A much more detailed discussion of coherence can be found in Fritz (unpublished d, here ch. 2).

## 5.2 Propositional Quantifiers

In this section, it will be shown that the logic of coherent equivalence systems is recursively isomorphic to second-order logic. Essential to the proof is the fact that second-order logic is recursively isomorphic to second-order logic with second-order quantifiers restricted to binding binary variables and ranging only over symmetric and irreflexive binary relations. Therefore, it suffices to simulate second-order quantification over symmetric and irreflexive binary relations in the propositional modal language. The central idea of the proof is that in certain equivalence systems, this restricted form of second-order quantification can be simulated over a subset of worlds $i$: Simulate first-order quantification as propositional quantification over world-propositions whose worlds are contained in $i$, and this restricted form of second-order quantification as quantification over all propositions which there are in some world – the idea is that a proposition $P$ corresponds to the relation which relates elements $w$ and $v$ of $i$ just in case $P$ contains a world at which the world-proposition of a world $u$ in $i$ exists just in case $u$ is identical to $w$ or $v$. For this to work, it must be guaranteed (a) that for any two worlds in $i$, there is a corresponding 'doubleton-world'. But for quantification over possible propositions to have the desired effect, it also has to be ensure that every set of worlds is a proposition at some world. This can be done by requiring (b) that some world $w$ contains every possible proposition: Since every world contains its own world-proposition, $w$ contains every world-proposition, so since the propositional domain of $w$ forms is a complete atomic field of sets, it must contain all sets of worlds.

To make this strategy precise, syntactic abbreviations of formulas are introduced which express that $p$ is a world-proposition; that $p$ and $q$ are identical; that there is the proposition $p$; that $\varphi$ is true at the world of a world-proposition $w$; that $\varphi$ is true when $w$ is bound to the world-proposition of the world of evaluation; and that world-propositions $w$ and $v$ are such that the only world-propositions there are in whose worlds $i$ is true are $w$ and $v$:

$$\begin{aligned}
world(p) &:= \Diamond p \wedge \Box \forall q(\Box(p \to q) \vee \Box(p \to \neg q)) \\
p = q &:= \Box(p \leftrightarrow q) \\
Ep &:= \exists q(q = p) \\
@w\varphi &:= \Box(w \to \varphi) \\
\downarrow w\varphi &:= \exists w(world(w) \wedge w \wedge \varphi) \\
D(w,v) &:= \downarrow u\Box(i \to \downarrow x@u(Ex \leftrightarrow (x = w \vee x = v)))
\end{aligned}$$

With this, a theory $T$ can be specified which imposes constraints (a) and (b):

$$T := \Box(i \rightarrow \downarrow w \Box((i \wedge \neg w) \rightarrow \downarrow v \Diamond D(w, v))) \wedge \Diamond \downarrow w \Box \forall p @ w E p$$

Define a recursive mapping $\cdot^*$ from the formulas of the language of pure second-order logic without identity using only binary second-order variables ('pure' meaning without non-logical constants) to the formulas of the propositionally quantified modal language, with the only non-trivial clauses as follows:

$$\begin{aligned} (\exists X \varphi)^* &:= \Diamond \exists p_X \varphi^* \\ (\exists x \varphi)^* &:= \Diamond(i \wedge \downarrow p_x \varphi^*) \\ (Xxy)^* &:= \Diamond(p_X \wedge D(p_x, p_y)) \end{aligned}$$

Here, it is assumed that the first- and second-order variables are injectively associated with proposition letters, mapping $x/X$ to $p_x/p_X$. Define a second translation $\cdot^\dagger$ between the *sentences* of same languages as follows:

$$\varphi^\dagger := \Box \forall i (T \rightarrow \varphi^*)$$

With this mapping, the theorem can be proven:

**Theorem 5.2.1.** *The propositionally quantified modal logic of coherent equivalence systems is recursively isomorphic to second-order logic.*

*Proof.* We first show that the mapping $\cdot^\dagger$ embeds pure second-order logic without identity using only binary second-order variables, with second-order quantifiers interpreted as ranging over symmetric irreflexive binary relations, in the propositionally quantified logic of coherent equivalence systems. Let $\varphi$ be a sentence of this second-order language, $\approx$ a coherent equivalence system, and $\Gamma$ the set of sets of worlds $I$ such that $T$ is true in $\approx$ when $i$ is interpreted as $I$. (Note that $T$ is necessary if true, so we can simply speak of its truth in an equivalence system; similarly for any formula $\varphi^\dagger$.) Then $\varphi^\dagger$ is true in $\approx$ iff $\varphi$ is true on all $I \in \Gamma$. Thus it suffices to show that for every cardinality

$\kappa$, there is a coherent equivalence system $\approx$ and a set of its worlds $I$ of cardinality $\kappa$ such that $T$ is true in $\approx$ when $i$ is interpreted as $I$.

So let $\kappa$ be a cardinal and $I$ a set of cardinality $\kappa$. Define $X = \{I\} \cup \{J \subseteq I : |J| \in \{1, 2\}\}$ and $\approx$ as the equivalence system on $X$ such that for all $w, v, u \in X$, $v \approx_w u$ iff there is a permutation $f$ of $I$ such that $\{f(i) : i \in v\} = u$ and $f(i) = i$ for all $i \in w$. $\approx$ is a coherent equivalence system which satisfies $T$ when $i$ is interpreted as $\{\{i\} : i \in I\}$: $I$ is a world in there is every proposition, and for any distinct $i, j \in I$, $\{i, j\}$ is a world in which the only world-propositions there is whose worlds are members of $I$ are the singletons of $\{i\}$ and $\{j\}$. This shows that $\cdot^\dagger$ embeds second-order logic as claimed.

We now show how the restrictions on second-order logic in this embedding can be lifted. The assumption of purity and the omission of identity are easily dealt with: An impure sentence is valid if and only if the result of replacing all constants by variables and binding them by a prefix of universal quantifiers is valid, and identity is straightforwardly definable using (our restricted form of) binary second-order quantifiers. That the restriction to binary variables and quantifiers ranging over symmetric irreflexive binary relation can likewise be lifted was proven by Scott and Rabin; see the presentation of their proofs in Nerode and Shore (1980) or the variant construction in (Kremer, 1997, Appendix). This gives us a recursive embedding of full standard second-order logic in the propositionally quantified modal logic of coherent equivalence systems, and it is routine to derive from this that the two logics are recursively isomorphic (see Kremer (1993) for details). □

## 5.3 Existence as a Modality

Given the result just proven, an axiomatic investigation of propositional contingentism must be formulated in a more restrictive language than propositional modal logic with propositional quantifiers. A natural option is to replace propositional quantifiers by an existential modality $E$, where $E\varphi$ is interpreted as saying that there is the proposition expressed by $\varphi$. This can be viewed as a fragment of the language with propositional quantifiers, by taking $E$ to be defined as above. If $E$ is taken as primitive, one can view equivalence systems as neighborhood structures, taking the neighborhood function interpreting $\Box$ to be the function mapping each world to the set of all sets of worlds, and the neighborhood function interpreting $E$ to be the function mapping each world $w$ to the set of worlds there are at $w$, i.e., the set of unions of sets of equivalence classes of the equivalence relation associated with $w$.

Let $\mathbf{S5}E$ be the set of formulas in this restricted language which are valid on the class of coherent equivalence systems. Since $\mathbf{S5}E$ is characterized by a class of neighborhood structures, it is a classical modal logic in the sense of Segerberg (1971). That is, $\mathbf{S5}E$ contains all propositional tautologies and is closed under modus ponens, uniform substitution and the congruence rules (if $\varphi \leftrightarrow \psi$ is a theorem of $\mathbf{S5}E$ then so are $\Box\varphi \leftrightarrow \Box\psi$ and $E\varphi \leftrightarrow E\psi$). In addition, the following are easily seen to be theorems of $\mathbf{S5}E$, where $\mathrm{pl}(\varphi)$ is the set of proposition letters in $\varphi$:

$N\Box : \Box\top$

$K_\Box : \Box(p \to q) \to (\Box p \to \Box q)$

$T_\Box : \Box p \to p$

$5_\Box : \Diamond p \to \Box\Diamond p$

$NE : E\top$

$\Box E : \Box(p \leftrightarrow q) \to (Ep \leftrightarrow Eq)$

$C : \bigwedge_{p \in \mathrm{pl}(\varphi)} Ep \to E\varphi$

Let $\Lambda$ be the classical modal logic axiomatized by these axioms; i.e., let $\Lambda$ be the smalles classical modal logic containing all of these axioms. Whether $\Lambda$ is complete (weakly or strongly) with respect to the class of coherent equivalence systems is left open here. I close with a few remarks on the axioms:

First, $\mathbf{S5}E$ is closed under necessitation and its analog for $E$: if $\varphi \in \mathbf{S5}E$ then $\Box\varphi \in \mathbf{S5}E$ and $E\varphi \in \mathbf{S5}E$. With $N\Box$ and $NE$, it follows that $\Lambda$ is closed under these rules as well. Second, $NE$ is an instance of $C$ if $\top$ is a primitive logical operator, but not if it is taken as a syntactic abbreviation of a classical tautology. Third, the $E$-free fragment of $\Lambda$ is identical to **S5**, as is, of course, the $E$-free fragment of $\mathbf{S5}E$. Fourth, the axiom schema $C$ can be understood as the natural comprehension schema saying that there are all the propositions expressed by formulas with parameters which there are; an analog for properties is discussed under the label $\mathrm{Comp}_{\mathrm{MC}}$ in Williamson (2013, p. 284). Fifth, instead of using the infinite axiom schema $C$, one could use the following finite list of its instances:

$C\neg :\ Ep \rightarrow E\neg p$

$C\wedge :\ (Ep \wedge Eq) \rightarrow E(p \wedge q)$

$C\Box :\ Ep \rightarrow E\Box p$

$CE :\ Ep \rightarrow EEp$

# Part III

# Contingentism and Paraphrase

# Chapter 6

# Counting Incompossibles

*with Jeremy Goodman*

**Abstract.** We often speak as if there are merely possible people – for example, when we make such claims as that most possible people are never going to be born. Yet most metaphysicians deny that anything is both possibly a person and never born. Since our unreflective talk of merely possible people serves to draw non-trivial distinctions, these metaphysicians owe us some paraphrase by which we can draw those distinctions without committing ourselves to there being merely possible people. We show that such paraphrases are unavailable if we limit ourselves to the expressive resources of even highly infinitary first-order modal languages. We then argue that such paraphrases are available in higher-order modal languages only given the assumption that it is a necessary matter what properties there are. We canvas a range of alternative paraphrase strategies and argue that none of them are tenable. We conclude with a dilemma. If talk of merely possible people can be paraphrased, then it is a necessary matter what properties there are. If such talk cannot be paraphrased, then it must be taken at face value, in which case it is a necessary matter not only what properties there are but also what individuals there are.

We often find ourselves speaking of things there could have been, as much in everyday life as when doing philosophy. We talk about buildings that could have been built but never will be, and of children who could have been born but never will be. Yet, in our more reflective moods, we are inclined to deny that there are any such things. Surely there are no merely possible buildings or merely possible people: nothing that could be a building fails to ever be one, and nothing that could be a person fails to ever be one. So our opinions appear inconsistent: We appear committed to both there being and to there not being things that could be people but never are. Which is it?

Philosophical orthodoxy sides with non-being: There are no merely possible people. If this is right, then the distinctions that we cognize and communicate when we unreflectively speak of merely possible people must be understood in other terms. For such distinctions are clearly intelligible. Consider the opening lines of Richard Dawkins' *Unweaving the Rainbow*:

> We are going to die, and that makes us the lucky ones. Most people are never going to die because they are never going to be born. The potential people who could have been here in my place but who will in fact never see the light of day outnumber the sand grains of Arabia. Certainly those unborn ghosts include greater poets than Keats, scientists greater than Newton. We know this because the set of possible people allowed by our DNA so massively exceeds the set of actual people. In the teeth of these stupefying odds it is you and I, in our ordinariness, that are here.[1]

Dawkins is clearly not saying something obviously false. Yet it *is* obviously false that most people are never going to be born: Every person has been born. So if we wish to speak perspicuously while communicating what he is communicating, we will have to paraphrase him. His talk of 'potential people' and 'possible people' suggests that we paraphrase him as saying that most *possible* people are never going to be born. But according to philosophical

[1]Dawkins (1998, p. 1). Merely possible people also figure prominently in ethical debates about the so-called 'non-identity' problem. See Parfit (1984, chapter 16), Hare (1988), and Hare (2007).

orthodoxy, this paraphrase is little improvement, since all possible people have been (or will be) people, and so have been (or will be) born. The orthodox position therefore faces a challenge of expressive power: to say what substantive claim Dawkins is making without in so doing saying that there are merely possible people (things which could have been people but never are). The challenge is an urgent one. If it cannot be met, then we will have no choice but to take talk of merely possible people at face value, and, more generally, accept that it is a necessary matter which things there are.[2]

This paper explores strategies for how the expressive power challenge might be met. In section 6.1 we argue that it cannot be met using the resources afforded by even highly infinitary first-order modal languages. In section 6.2 we show that it can arguably be met by moving to a higher-order modal language, but only given the controversial principle that every possible individual is singled out by some actual property. In section 6.3 we argue that anyone engaged in the project of paraphrase must accept this principle, since all other paraphrase strategies are problematic. In section 6.4 we examine a familiar but ultimately inconclusive line of argument from this principle to the stronger principle that it is a necessary matter what properties there are. In section 6.5 we propose a framework for understanding our expressive power challenge, reconsider the arguments of the previous sections in light of it, and argue that the project of paraphrase does commit one to the stronger principle. In section 6.6 we conclude with a dilemma: If our paraphrase challenge can be met, then it is a necessary matter what properties there are. If it cannot be met, then we have no choice but to conclude that it is a nec-

[2]We have slightly oversimplified the orthodox position, which ought to be compatible with the claim that some future machines will count as people and so be people that are never born. What the orthodox position is committed to is that anything that could be a person is at some time a *person*. (Our ascription of orthodoxy is based on the sociological observation that Williamson's (2013) view that there actually are things that could have been people yet are never born (e.g., Wittgenstein's possible children) has mostly been met with incredulity.)

One might think that human zygotes that fail to develop are merely possible people, on the grounds that people are identical to human organisms that were zygotes before they were people (and hence people are not essentially people). Those who accept such views should replace our talk of merely possible people with talk of merely possible organisms. In footnote 21 we consider views according to which there are merely possible organisms because organisms are identical to their underlying matter.

essary matter what individuals there are, and hence also a necessary matter what properties there are. An appendix establishes the technical results appealed to in section 6.1; the technical results appealed to in section 6.2 are proved elsewhere. Section 6.3 can be skipped without loss of continuity and does not presuppose familiarity with section 6.2; most of section 6.5 does not presuppose familiarity with section 6.4.

## 6.1 The challenge of incompossibility

Before considering potential paraphrases of Dawkins' passage, let us consider some simpler examples.[3]

### 6.1.1 A possible knife

In a knife factory there is a spare handle and a spare blade. No knife will ever be made using either of them. Nevertheless, we might naturally say that there is a possible knife that could have been made from this blade and handle had the two been joined.[4] But if, in keeping with metaphysical orthodoxy, we wish to deny that there are any such possible knives, then this claim, like Dawkins', is in need of paraphrase. Luckily, the following paraphrase is available: It is possible for there to be a knife made by joining this handle and this blade.

[3]Throughout our paper we will engage in the very talk of merely possible things that we are exploring strategies for eliminating. Since such talk will not figure in any of the paraphrases we will be considering, this is unobjectionable.

In what follows we deliberately refrain from theorizing in terms of 'possibilia' that 'exist' only in 'non-actual possible worlds'. We agree with Williamson (2013, chapter 1) that such formulations at best obscure the most productive questions in their vicinity – namely, questions that can be formulated in the language of quantified modal logic, on an unrestricted interpretation of its quantifiers and a metaphysical interpretation of its modalties (see section 6.1.6). In particular, 'possible worlds'-talk prejudices the issues of higher-order contingency discussed in section 6.2.3; see Williamson (2013, chapter 6) and Stalnaker (2012, chapter 2).

[4]Where it is clear from context, we will use 'possible $\varphi$s' as shorthand for 'possible things that are possibly $\varphi$' and 'merely possible $\varphi$s' as shorthand for 'possible things that are possibly $\varphi$ and never $\varphi$'. On this usage, Saul Kripke is a merely possible circus performer.

### 6.1.2 Two possible knives

In the knife factory there is a spare handle and two spare blades. No knife will ever be made using any of these materials. Nevertheless, we might naturally say that there are two possible knives that could be made by joining the handle with one of the blades.

The most flatfooted generalization of the paraphrase strategy from the previous section will not work: It is not possible for there to be two knives made by putting together these materials, since the handle would have to be part of both knives at once, which is impossible. A different strategy is needed. Luckily, the following paraphrase is available: It is possible for there to be a knife made by joining this handle with one of these blades such that it is possible for there to be a different knife made by joining this handle with one of these blades.

This strategy generalizes to cases with any finite number of spare blades. In general, we can paraphrase the claim that there are $n$ possible knives that could be made from these materials as follows: It is possible for there to be a knife $x_1$ made from some of these materials such that it is possible for there to be a knife $x_2$ distinct from $x_1$ made from some of these materials such that $\ldots$ it is possible for there to be a knife $x_n$ distinct from each of $x_1 \ldots x_{n-1}$ made from some of these materials. Unlike the flatfooted generalization of the paraphrase strategy from the previous section, we need not assume that it is possible for there to be more than one object that is possibly a knife made from some of these materials. In other words, the strategy is compatible with the *pairwise incompossibility* of the $n$ merely possible knives. We achieve this compatibility by interleaving $n$ possibility operators and $n$ existential quantifiers.

Notice that the claim that these merely possible knives are pairwise incompossible is stronger than the claim that, necessarily, no two of them are knives. One might think that, necessarily, everything that is possibly a knife is at some time a knife, without thinking that, necessarily, everything that is possibly a knife *is* a knife. On such a view, it might be possible for there to be all of the merely possible knives after all, were the handle joined with each of the blades one after another. One might think that it is therefore unnecessary that our paraphrase strategy be compatible with pairwise incompossibility. Why not instead use the simpler paraphrase: It is possible for there to be $n$ things each of which is at some time a knife made from some of these materials? The problem is that the proposal doesn't generalize: there are two possible dishes – an omelette and a soufflé – that I could make from these eggs, but it is not possible that at some time I make an omelette from these eggs and at another time I make a soufflé from these same eggs. So we need paraphrases compatible with pairwise incompossibility.

### 6.1.3 Infinitely many possible knives

In Hilbert's knife factory there is a spare handle and infinitely many spare blades. No knife will ever be made using any of these materials. Nevertheless, we might naturally say that there are infinitely many possible knives that could be made by joining the handle with one of the blades.

This claim is no less in need of paraphrase than the claim that there are two possible knives that could be made by joining the handle with one of the blades. Yet unlike the latter claim, this claim cannot be paraphrased in a standard first-order modal language. This is an instance of a more general limitative result. In order to state it, we need a sharper characterization of the paraphrastic project we are trying to carry out.

The existential and universal quantifiers familiar from first-order logic are instances of a well-studied family of variable-binding operators known as *generalized quantifiers*, other examples of which include 'there are $n$ things such that ...', 'there are infinitely many things such that ...', 'there are uncountably many things such that ...' and 'most things such that ... are such that ...'.[5] The claims we can make about knives, buildings, people, omelettes, and soufflés using generalized quantifiers correspond to intelligible modalized claims. Corresponding to the claim that there are $n$ knives is the claim that there are $n$ possible knives. Corresponding to the claim that most people will never be born is Dawkins' claim that most possible people will

[5] See Westerståhl (2011) for a concise introduction to generalized quantifiers.

never be born. This raises a general expressive power challenge: to be able, for any generalized quantifier and claim involving it, to express the corresponding modalized claim without contradicting the philosophical orthodoxy according to which knives, buildings, people, omelettes, and soufflés have contingent being. Think of generalized quantifiers as expressing structural conditions on the pattern of satisfaction of the conditions expressed by the formulas whose variables they bind. Our challenge is to be able to express the same structural conditions on the pattern of satisfaction of those conditions by all *possible* individuals.[6]

Here is our first limitative result: In standard first-order modal logic enriched with all generalized quantifiers, our expressive power challenge can be met for only those generalized quantifiers that can be defined in standard first-order logic.[7] Since 'there are infinitely many ... ' cannot be defined in standard first-order logic, it follows that not all modalized claims of the form 'there are infinitely many possible ... ' can be expressed in a standard first-order modal language, no matter which (unmodalized) generalized quantifiers the language includes. If we hope to paraphrase such claims, we must move to a language whose expressive resources go beyond those of standard first-order modal logic.

Luckily, there are languages in which we can paraphrase such claims. If we enrich our standard first-order modal language with a device for forming infinite conjunctions, we can then paraphrase claims of the form 'there are infinitely many possible ... ' as infinite conjunctions of the corresponding claims of the form 'there are $n$ possible ... ', for all natural numbers $n$. We see nothing illegitimate about paraphrase strategies that employ conjunctions of

[6]Fine (2003), Leuenberger (2006) and Williamson (2010, 2013) raise similar challenges. Lewis (2004), Sider (2006) and Szabó (2006) raise analogous challenges for tensed quantifiers. But note that the way we conceive of our paraphrase challenge differs substantially from the ways in which most of the aforementioned authors conceive of their respective challenges; see section 6.5.

[7]See Theorem 6.7.1 of Appendix 6.7.2. We discuss the philosophical interpretation of these model theoretic results in more detail in section 6.5. Throughout we will be considering modal languages that include devices for 'undoing' or 'scoping out of' modal operators. Such devices are well studied and correspond to certain uses of 'actually' and 'in fact' in English; for technical details, see Fine (1977a, section 6), Hodes (1984), and Correia (2007).

arbitrary sets of formulas, or that employ infinitary analogues of standard existential and universal quantifiers capable of binding arbitrary sets of variables.[8] The important question is whether such strategies are sufficiently general to answer our expressive power challenge.

### 6.1.4 Uncountably many possible knives

In Cantor's knife factory there is a spare handle and uncountably many spare blades. No knife will ever be made using any of these materials. Nevertheless, we might naturally say that there are uncountably many possible knives that could be made by putting together some of these materials.

This claim is no less in need of paraphrase than the claim that there are infinitely many possible knives that could be made by joining the handle with one of the blades. Yet unlike the latter claim, this claim cannot be paraphrased in a standard first-order modal language in which we can form conjunctions of arbitrary sets of formulas and existential quantifications that bind arbitrary sets of variables, and which also includes all generalized quantifiers. This is an instance of a more general limitative result: In such a language, our expressive power challenge can be met for only those generalized quantifiers that are definable in first-order logic with infinitary conjunction (but only finitary existential and universal quantification).[9] Since 'there are uncountably many ...' is not definable in first-order logic with infinitary conjunction but only finitary quantification, it follows that not all modalized claims of the form 'there are uncountably many possible ...' can be expressed in the aforementioned modal language. If we hope to paraphrase such claims, we must move to a yet more expressive language.

This result might come as a surprise to those familiar with infinitary languages, since the unmodalized 'there are uncountably many ...' *can* be paraphrased in standard infinitary first-order logic as 'some $x_1, x_2, \ldots$ are pairwise distinct and are each ...', where '$x_1, x_2, \ldots$' stands for an uncountable set of variables and 'pairwise distinct' and 'are each ...' stand for the

[8] See Dickmann (1985) for discussion of such languages.
[9] See Theorem 6.7.2 of Appendix 6.7.2.

corresponding uncountable conjunctions. The key observation is that in our language we cannot bind all uncountably many free variables in '$x_1, x_2, \ldots$ are pairwise distinct and are each ...' without binding uncountably many of them at once with a single existential quantifier, forcing them to have the same scope. But as we saw in section 6.1.2, *modalized* existential quantification requires all variables so-bound to have different scopes, since for any two of them there must be a possibility operator with intermediate scope to accommodate pairwise incompossibility.

This suggests a new paraphrase strategy. Suppose we move to a language in which infinitary quantification is achieved not by letting existential and universal quantifiers bind infinitely many variables at once, but instead by allowing for strings of infinitely many existential and universal quantifiers. Perhaps in such a language we can paraphrase claims of the form 'there are uncountably many possible ...' as claims of the form 'possibly, for some $x_1$, possibly, for some $x_2$, *etc.*: $x_1, x_2, \ldots$ are pairwise distinct and are each possibly ...', where 'possibly, for some $x_1$, possibly, for some $x_2$, *etc.*' stands for an uncountable string of interwoven possibility operators and existential quantifiers.[10]

In Appendix 6.7.3 we give a formal syntactic characterization of quantified modal languages allowing for such infinite embeddings, and provide a model-theoretic semantics relative to which our expressive power challenge can indeed be met for 'there are uncountably many possible ...' in the above way. For the sake of argument we will grant that languages allowing for infinitely embedded quantifiers and modal operators are legitimate for the purpose of paraphrasing modalized quantification.[11] Again, we will focus on the question of whether such languages are sufficiently expressive to answer our expressive power challenge.

[10] Fine (2003, section 4) makes a similar proposal.

[11] Williamson (2013, chapter 7) questions their legitimacy in a related context; Fine (unpublished) replies.

### 6.1.5 Most possible people

Let us finally return to Dawkins' claim that most possible people will never be born. It turns out that it cannot be paraphrased even in the highly infinitiary language considered in the previous section. Again, this follows from a more general result: In a first-order modal language enriched with all generalized quantifiers and in which we can infinitely nest conjunctions, disjunctions, existential and universal quantifiers, and modal operators, our expressive power challenge can be met for only those generalized quantifiers that are definable in a non-modal first-order language in which we can infinitely nest conjunctions, disjunctions, and existential and universal quantifiers.[12] Since the unmodalized 'most ... are ... ' is not definable in the latter non-modal language, it follows that the modalized 'most possible ... are ... ' cannot be paraphrased in the former modal language.

One might hope to answer our expressive power challenge by moving to an even more infinitary language in which we can form conjunctions of proper classes of formulas and have strings of quantifiers and modal operators that bind proper classes of variables. The general idea would be to embed some enormous conjunction within a string of interleaved possibility operators and existential quantifiers that bind as many variables as there are possible individuals.[13] But we see no consistent way of implementing this strategy. Presumably, for any formula of any language and any variables that occur in it, there should be some formula of some language such that those variables are all and only the variables that occur in it. It then follows by a plural version of Cantor's diagonal argument that no formula contains as many variables as there are possible individuals, since there would have to be more formulas than possible individuals, which is inconsistent, since formulas are individuals. To be clear: the problem is not that a certain specified formula fails to be an adequate paraphrase but rather that it is impossible for there to be a formula satisfying the given specification.[14]

[12] See Theorem 6.7.5 of Appendix 6.7.4. This result vindicates the skepticism about the expressivity of such languages voiced in Leuenberger (2006, p. 157) and Williamson (2013, p. 354).

[13] Fine (1977a, section 6) advocates such a proposal.

[14] Fine might resist the diagonal argument by rejecting standard plural logic in favor of a

We conclude that the expressive power challenge cannot be met in even highly infinitary first-order modal languages. In retrospect, this should be somewhat unsurprising. After all, the way we actually express generalized quantifiers like 'there are uncountably many ...' is not using infinite conjunction and quantification, but through higher-order resources. In the next section we explore the possibility of using such resources to answer our expressive power challenge. But first, a brief interlude on context sensitivity.

### 6.1.6 Interlude on Context Sensitivity

Both quantifiers and modal adverbs breed context-sensitivity, in that sentences containing them can be used literally to express different propositions on different occasions of use. So perhaps we speak truly both when in unreflective moods we utter 'Most possible people are never going to be born' and when in more metaphysically minded moods we utter 'There are no merely possible people'. To be clear: when *we* raise the challenge to say what non-trivial claim Dawkins is making without in so doing saying that there are merely possible people, we are (i) resolving the context-sensitivity of 'there are' and 'possible' uniformly with the way in which orthodox metaphysicians resolve it when they utter 'There are no merely possible people', and (ii) challenging those metaphysicians to say what non-trivial claim Dawkins is making while resolving the context-sensitivity of their quantifiers and modal adverbs in only that way.

Why is this an urgent challenge? Because on any *uniform* resolution of their context-sensitivity, these two sentences express incompatible propositions. And it is widely agreed that there is a particular way in which we

logic akin to the non-standard class theory of Fine (2005), which allows for a universal class. One might also deny that we should think of formulas as individuals (i.e., as sequences of expressions), but instead treat them as well-orderings of expressions, understood in higher-order terms. But the problem then immediately reoccurs if we think that there is contingency in what conditions (e.g., well-orderings) there are, as we discuss in section 6.2. Generalized quantifiers and their modalizations make just as much sense concerning well-orderings as they do concerning individuals. But using the present strategy to paraphrase modalized generalized quantification over well-orderings would require as many second-order variables as there are well-orderings, and hence well-orderings of variables, which leads to another cardinality crash.

should uniformly resolve the context-sensitivity of our quantifiers and modal adverbs when we do metaphysics.[15] It follows that when engaging in metaphysical theorizing we cannot accept both of the above two sentences as literally true. Since orthodox metaphysicians accept 'There are no merely possible people' as literally true, they cannot likewise accept 'Most possible people are never going to be born'. So they must provide some other way of making sense of the claim that we would normally communicate by uttering the latter sentence, and do so holding fixed the meaning of their quantifiers and modal adverbs.

## 6.2 Singling out merely possible objects

### 6.2.1 Singling out merely possible knives

Why are we so confident that there are uncountably many possible knives that could be made from the materials in Cantor's knife factory? Because the factory contains uncountably many blades, each of which could, together with the handle, compose a knife were the two joined, and any two of which are such that, necessarily, any knife made with one of them could not possibly be made with the other. In other words, we think that there are uncountably many possible knives that could be made using those materials because we know that there are uncountably many ways of making knives from these

[15] On quantifiers, see Williamson (2003), Dorr (2005), and Sider (2009); on modal adverbs, see Kripke (1980 [1972]) and Plantinga (1974). For general discussion of such metametaphysical issues, see the papers in Chalmers et al. (2009).

Although we will not question it here, the methodological assumption about quantifiers has recently been questioned under the heading of 'ontological pluralism'; see McDaniel (2009) and Turner (2010). Someone sympathetic to ontological pluralism might, in the present setting, suggest that quantifiers come in two fundamental families – 'actualist' and 'possibilist' – such that, for example, $\text{most}_p$ things that are possibly people are never going to be born, yet $\text{everything}_a$ that is possibly a person is at some time born. The strategy rejects the paraphrase challenge by positing an ambiguity: modalized quantification just is possibilist quantification. Whatever one thinks of ontological pluralism in general, this particular version of it faces the following difficulty. The view accepts that $\text{everything}_p$ is possibly $\text{something}_a$, and that it is contingent what things there $\text{are}_a$. If these commitments are extended to one's higher-order quantifiers, as seems natural, then the technical results cited in section 6.2.4 show that one will not be able to make sense of the possibilist generalized quantifier 'there are uncountable $\text{many}_p$' in higher-order possibilist terms.

materials and that each of these ways would yield a different knife. (Dawkins offers an analogous justification of his claim that most possible people are never going to be born, noting that the possibly exemplified human genetic profiles outnumber all the people there will ever be.)

This suggests a general strategy for answering our expressive power challenge. First, we find some ways of singling out the possible things we are interested in. In the case of the possible knives, for example, we appeal to the conditions under which they would be constructed. We then quantify over these conditions, letting them go proxy for the possible objects themselves.[16] Of course, we have only worked a simple example involving knives, and even in that case one might reject some of our underlying assumptions. We need to examine more systematically the conditions under which such strategies can be used to paraphrase modalized quantification.

### 6.2.2 Haecceities in higher-order logic

First, we need a more precise statement of what it is for a condition to single out a possible object. Say that $F$ *singles out* $x$ just in case, necessarily, everything is $F$ just in case it is identical to $x$. It follows, by the principle that what is possibly necessary is necessarily necessary, that, if $F$ possibly singles out $x$, then $F$ necessarily singles out $x$, even if $x$ is a contingent being. (Note that it does not follow from the fact that being identical to $x$ singles out $x$ that being identical to $x$ is the *only* condition that singles out $x$, since we have not assumed that necessarily co-extensive conditions are identical.) A condition $F$ is a *haecceity* just in case it is possible that there be something that it singles out.

We can now present the general strategy. The idea is to replace talk of possible $F$s with talk of haecceities that possibly apply to something that is $F$. For example, instead of saying that there are uncountably many possible knives, we might say that there are uncountably many haecceities each of which is possibly instantiated by a knife and no two of which are possibly

[16] Related strategies were independently proposed by Plantinga (1976) and Fine (1977a, section 2).

instantiated by the same thing. Parallel paraphrases can be given for all generalized quantifiers. The strategy requires moving to a higher-order modal language in which we can quantify in to predicate positions; in what follows talk of conditions, properties, and the like should be understood as shorthand for such quantification.[17]

This paraphrase strategy relies on the following assumption if it is to apply generally:

*Haecceitistic Plenitude*: For every possible individual there is a condition that singles it out.

Yet most philosophers who have considered the issue reject this principle: they deny that there is a way to single out every possible object, usually invoking merely possible elementary particles as purported counterexamples.[18] The idea is that, just as there is contingency as regards what elementary particles there are, there is parallel contingency as regards which possible elementary particles can be singled out: Had there not been this electron, there would have been no haecceities of it either. On these grounds metaphysicians usually dismiss attempts to answer the expressive power challenge by using haecceities as surrogates for possible objects.[19]

But such dismissals are too quick, since there are more sophisticated paraphrase strategies that can get by with weaker assumptions. Suppose we assume only that haecceitistic plenitude *could* be true, i.e.:

*Haecceitistic Compossibilism*: It is possible that for every possible individual there be a condition that singles it out.

We can now paraphrase the claim that there are uncountably many possible $F$s as the claim that, possibly, there are uncountably many haecceities each

[17]Fritz (unpublished b, here ch. 7) formalizes this paraphrase strategy. Jager (1982) also takes himself to be formalizing something in the vicinity of Plantinga's informal remarks about haecceities, but his project is quite different from ours: we are offering a strategy for paraphrasing modalized first-order quantification in a higher-order language, whereas Jager takes himself to be offering a more metaphysically hygienic alternative to Kripke's (1963) model theory for first-order modal languages.

[18]See, e.g., Fine (2003) and Stalnaker (2012); Plantinga (1976, 1983) is a notable exception.

[19]See, e.g., Fine (1977a, section 4), Fine (1985), and Williamson (2013, sections 6.2 and 7.6).

of which possibly applies to an $F$ and no two of which possibly apply to the same thing. We can in fact prove that, if haecceitistic compossibilism is true, then there is a general paraphrase strategy that answers our expressive power challenge[20] – at least insofar as such questions can be answered using standard model-theoretic methods (an issue we take up in section 6.5).

The modified strategy is of interest because incompossible possible objects can have compossible haecceities. For example, our incompossible merely possible knives seem to have compossible haecceities, since we seem to be able to single out each of them in its absence by appealing to the conditions under which it would be manufactured. And while perhaps merely possible elementary particles (and other merely possible matter more generally) cannot be singled out in their absence, this is not obviously a problem, since bits of matter seem not to be incompossible in the manner of other possible enmattered objects like possible knives and possible people. Those who think that it is contingent what objects can be singled out might therefore reasonably hope that haecceitistic compossibilism is the solution to their problems of expressive power.

### 6.2.3 Incompossible haecceities

Unfortunately, such optimism is misplaced. Anyone who rejects haecceitistic plenitude should also reject haecceitistic compossibilism, both with respect to the haecceities of possible enmattered objects (e.g., possible people) and with respect to the haecceities of possible bits of matter (e.g., possible electrons). We will consider the two cases in turn.

Consider a healthy egg and two healthy sperm neither of which will ever fertilize it. Assume that eggs can be fertilized by at most one sperm and that people have their biological origins essentially. It follows that the possible people who could be born from these biological materials are incompossible. But unlike the case of possible knives, we do not seem to be able to single out the possible people who could be born from these biological materials. This is because, for each sperm, there is *more than one* possible person who could

[20]See Fritz (unpublished b, here ch. 7).

have had it and the egg as its biological origins, since such biological origins could have produced monozygotic twins. We seem to have no way of singling out any such merely possible person from its merely possible twins. We would be able to single out the individual possible twins were they born, but then we wouldn't be able to single out the merely possible people who could be born from the egg and the unused sperm. So we have a counterexample to haecceitistic compossibilism.

The case of matter is somewhat different, since bits of matter are arguably never *pairwise* incompossible. Nevertheless, it seems impossible that there be *every possible* elementary particle, since, whatever elementary particles there ever are, it is possible that there be an elementary particle that is not one of them. Assuming that necessarily there are no haecceities of merely possible elementary particles, and that elementary particles are essentially elementary particles, it follows that the possible haecceities of possible elementary particles are incompossible.[21]

[21] The incompossibility of all possible bits of matter is compatible with the weaker hypothesis that, for any cardinal $\kappa$, any $\kappa$ possible bits of matter are compossible. This hypothesis is somewhat plausible, and would allow us to paraphrase modalized cardinality quantifiers restricted to possible matter in a finitary language. The further hypothesis that any possible individuals are either equinumerous with all pure sets or equinumerous with the members of some pure set would even allow us to paraphrase 'most possible bits of matter such that $\varphi$ are such that $\psi$'. (Sloppily: either it is possible that there be all possible bits of matter that are actual $\varphi$s and most of them be actual $\psi$s, or it is impossible that there be all possible bits of matter that are actual $\varphi$-and-$\psi$s.)

The existence of such piecemeal paraphrases raises the question of whether those who think that necessarily all contingent beings are matter – or are at least similarly modally recombinable (e.g., Cartesian egos) – can evade the limitative results of section 1. (Such a view might be held out of a commitment to the necessary truth of classical extensional mereology or mereological nihilism, although one would also need to reject incompossible possible facts, events, etc.) The relevant technical question is whether, in a first-order modal language with all generalized quantifiers, can we express all claims involving modalized generalized quantifiers *relative to the class of models that respect the aforementioned 'set-sized-compossiblism' and 'modalized limitation of size' constraints*. We leave this as an open question, both because investigating it would require developing a model theory in which the domain of a model can be a proper class and generalizing the notion of a generalized quantifier accordingly, and because paraphrase strategies that build in these sorts of substantive metaphysical assumptions are philosophically problematic, as we argue in section 6.5.3.

That said, we doubt that the above question has a positive answer, since we see no way to generalize the trick by which we paraphrased 'most possible ... are ...' to other generalized quantifiers whose behavior on a proper-class sized domain is not determined

Since the purported counterexamples to haecceitistic plenitude are, if genuine, also counterexamples to mere haecceitistic compossibilism, the two claims stand or fall together.

### 6.2.4 The challenge of incompossible haecceities

One might hope that there is some yet more sophisticated paraphrase that can get by without relying on the assumption of haecceitistic compossibilism. After all, we are allowing ourselves the expressive resources of an infinitary higher-order modal language. Unfortunately, there are very strong model-theoretic indications that there is no such paraphrase. In particular, we can show that in a higher-order modal language enriched with infinitary conjunction and quantification, there are pairs of models that agree on all formulas of the language but disagree on whether there are uncountably many objects in the extension of the open formula '$x$ is a possible being'.[22] Of course, this result is only as probative as the background model theory is well motivated. And developing variable-domain model theories for higher-order modal languages is a non-trivial technical project with a number of important choice points. Here we can only report that our models are neutral with respect to

---

by their behavior on all set-sized sub-domains. For example, consider the generalized quantifier binding two variables stating that the relation expressed by the open formula whose variables it binds is a *complete partial order* on the domain, in the sense that the relation partially orders the domain and any sub-class of the domain with an upper bound in the domain has a least upper bound in the domain. Now consider the open formula 'either $x$ and $y$ are sets and $x \subseteq y$, or $x$ and $y$ are natural numbers and $x > y$, or $x$ is a set and $y$ is a natural number'. The binary relation expressed by this formula is a complete order on the class of pure sets, but it is not a complete order on the class of pure sets and natural numbers, since the latter has a subclass (the class of pure sets) which has an upper bound (any natural number) but no least upper bound. Yet the relation is a complete order on every set-sized subclass of both the pure sets and the pure sets with the natural numbers. We conjecture that the corresponding modalized quantifier cannot be defined in a first-order modal language relative to the class of models satisfying the above two constraints. (This example was inspired by the discussion in Williamson (2013, section 6.4).)

[22] This result is the main theorem of Fritz (unpublished b, here ch. 7); it can be seen as a generalization of the results of Leuenberger (2006), Williamson (2010) and Fritz (2013) to a higher-order setting. It would be prohibitively difficult to generalize this limitative result to a higher-order language that allows for infinite embeddings, but the considerations of section 6.1.5 give us every reason to expect that 'most possible ... are ... ' is undefinable in such languages.

these choice points and are in keeping with the best motivated and developed extant theories of contingency in what objects can be singled out.[23]

Here is an intuitive characterization of the models. In one model the domain of all possible individuals is countable while in the other model the domain of all possible individuals is uncountable. In both models every finite subset of possible individuals is the domain of some world, all worlds have finite domains, and no two worlds have the same domain; the actual worlds of both models have the same domain. We generate higher-order domains from the first-order domains in accordance with the idea that the distinctions that there are at a world are exactly those that can be drawn using the materials that there are at that world. These two models clearly differ as regards whether they ought to validate 'there are uncountably many possible beings' on the reading we have been trying to paraphrase. But we can prove that the two models validate exactly the same formulas. In this formal sense, there is no paraphrase of 'there are uncountably many possible . . . ' in our infinitary higher-order modal language.

We might interpret these models as follows. Suppose that any finite number of people could be born from a particular sperm and egg, that any finite collection of $n$ possible people who could be born from the sperm and egg could be born together as monozygotic $n$-tuplets, and that it is impossible that infinitely many people be born from the sperm and egg. Perhaps this isn't the most plausible view about what sperms and eggs can do, but it is clearly a view that we should be able to express. We should also be able to distinguish, on the one hand, the conjunction of this view with the claim that there are only countably many possible people who could be born from the sperm and egg, and, on the other hand, the conjunction of this view with the claim that there are uncountably many possible people who could be born from the sperm and egg. The two hypotheses are naturally modeled by the pair of models described above. The inability of our language to distinguish the models indicates that it cannot distinguish these two hypotheses.

[23]We develop various model theories for infinitary higher-order modal languages in Fritz and Goodman (unpublished c, here ch. 1), drawing on the work of Stalnaker (2012), Williamson (2013, chapter 6), and, especially, Fine (1977b). In section 6.5 we take up the general question of the role of model theory in assessing candidate paraphrase strategies.

## 6.3 Other expressive resources

We have seen that there is a natural and general strategy for answering our expressive power challenge using higher-order resources, but that it relies on the controversial assumption of haecceitistic plenitude. Those who wish to reject haecceitistic plenitude must therefore find other expressive resources with which to answer the expressive power challenge, beyond those available in infinitary higher-order modal logic. In this section we survey a number of paraphrase strategies that employ such resources, and argue that none of them succeed.

### 6.3.1 Fictionalism

Consider first the proposal that we should paraphrase a claim involving modalized generalized quantifiers with the claim that its umodalized counterpart is true according to a certain fiction, one according to which it is necessary what things there are. Let '$T$' abbreviate 'the conjunction of all non-modal truths and the claim that it is necessary what things there are'. The proposal is that we paraphrase, e.g., the claim that most possible people are never going to be born as the claim that, *according to* $T$, most things that are possibly people will never be born.

This proposal faces a number of problems. One is that it requires distinguishing the 'non-modal' truths (which get built into the fiction) from modal truths, such as the purported fact that it is contingent what individuals there are (which had better not get built into the fiction). Yet it is far from clear that the distinction between the modal and the non-modal is in good standing.[24]

A more serious problem is the following. Proponents of fictionalist paraphrase think that it is contingent what things there are. *A fortiori*, they think that all possible people are such that it is contingent what things there are. But the present proposal ends up paraphrasing this claim with the clearly false claim that, according to $T$, everything that is possibly a person is such

[24] See Stalnaker (1984, chapter 8) and Williamson (2013, chapter 8.4) for skepticism.

that it is contingent what things there are. The source of the problem is that sentences containing modalized quantifiers can also contain unmodalized quantifiers. (This point tends to be overlooked because the fictionalist project is usually conceived as that of translating the sentences of a language containing only modalized quantifiers into a distinct language containing only unmodalized quantifiers.[25])

One might try to get around this problem by modifying the paraphrase strategy so that all originally unmodalized quantifiers get restricted by an 'existence' predicate. This strategy will only work if, according to $T$, not everything exists. But it is unclear what 'existence' is supposed to be such that it plays this role. Most contemporary metaphysicians, and especially those who think it is contingent what things there are, deliberately use the word 'exist' as synonymous with 'is identical to something'.[26] If fictionalist paraphrase is committed to a notion of existence that comes apart from being identical to something – at least according to $T$ – that fact significantly undercuts the strategy's appeal.[27]

Perhaps the most serious problem for the proposed strategy is that fictions are incomplete. There are many questions not settled by $T$. Let $\varphi$ be some such claim not containing modalized quantifiers. It is neither the case that, according to $T$, $\varphi$, nor the case that, according to $T$, not-$\varphi$. The proposed paraphrase strategy then entails that it is neither the case that $\varphi^M$ nor the case that not-$\varphi^M$, where $\varphi^M$ is the result of replacing all quantifiers in $\varphi$ with modalized quantifiers. But this is a contradiction. The source of the problem is that 'according to $T$ ...' fails to commute with negation.[28]

[25]Compare Sider (2002) and Fine (2003, section 5).

[26]This is no coincidence: Goodman (in preparation a) argues that one of the strongest reasons to think that it is contingent what things there are is that it allows one to avoid recognizing any primitive notion of 'existence' according to which not everything exist.

[27]Perhaps 'according to $T$' is an opaque context such that the differential behavior of 'exists' and 'is identical to something' in that context is compatible with existence in fact being the very same thing as being identical to something. But if so, then the paraphrase strategy is subject to a more serious problem – namely, that it wrongly predicts that Leibniz's law is invalid in the scope of modalized quantifiers.

[28]One might try to mitigate this problem by adding some general modal principles to the fiction, akin to the way Rosen (1990) builds into his fiction about possible worlds various *dicta* from Lewis (1986). But we see no non-*ad hoc* way of enriching $T$ so as to

### 6.3.2 Counterfactuals

One might hope to solve all of these problems by moving from a fictional paraphrase to a counterfactual one, by paraphrasing claims involving modalized quantifiers as claims about what *would* have been the case had it been necessary what individuals there are.[29] For example, we might paraphrase Dawkins' claim that most possible people are never born as the claim that, had it been necessary what things there are, most things that actually could have been people would actually never be born. We don't have to specify the actual non-modal truths or restrict our originally non-modalized quantifiers: the needed effect is achieved by the 'actually' operator. As for incompleteness, the problem will not arise if we assume that 'had it been necessary what things there are . . . ' commutes with negation.

This proposal faces two immediate but surmountable problems. The first is that it delivers the wrong paraphrase of 'All possible things are possibly identical to something'. Assuming it is contingent what things there are, it is presumably *not* the case that, had it been necessary what things there are, then everything would have actually been possibly identical to something. For presumably the unrestricted theory of ordered pairs would still have been true had it been necessary what things there are. So, had it been necessary what things there are, there would have been ordered pairs of things that are actually pairwise incompossible (assuming, as the strategy requires, that there would have been things which are actually pairwise incompossible). Such pairs would actually not even *possibly* be identical to anything. But this problem can be solved: we simply restrict our originally modalized quantifiers to things that are actually possibly identical to something.

A second problem for the proposal is that it is founded on the unmotivated presupposition that every possible being is something there would have been had it been necessary what things there are. This claim is far from obvious. Maybe, if it had been necessary what things there are, it would have

---

settle, for example, the question of section 6.2.4 of whether the number of things that are possibly people who could be born from a particular sperm and egg is countable.

[29] This strategy is inspired by Dorr (2005, section 3) and Dorr (2008, section 2); see also Woodward (2012).

been because Pythagoreanism was true and everything was a mathematical object. If so, *you* are a counterexample to the aforementioned presupposition (assuming that, even if it had been necessary what things there are, you would not have been a mathematical object). In general, if it is in fact contingent what things there are, then it is not at all clear what things there would have been had it been necessary what things there are.

This problem can be solved as follows. Instead of the claim that it is necessary what things there are, we take as the antecendent of our counterfactual the claim that there are all actually possible individuals. Making this claim precise requires scope-indicating devices of the sort mentioned in footnote 7. We introduce an 'in fact$_i$' operator that has the effect of letting the subformula it embeds be evaluated at a wider scope indicated by a co-indexed $\uparrow_i$. We can now understand claims of the form 'Had there been all actually possible individuals, ...' as shorthand for claims of the form 'Had it been $\uparrow_1$ actually necessary that everything is in fact$_1$ identical to something, ...'.[30]

Unfortunately, the present proposal faces more serious challenges. The first is that it is unclear whether 'had there been all actually possible things ...' does commute with negation. One reason for doubt is that counterfactuals with impossible antecedents are widely thought to be vacuously true, which immediately rules out such commutativity (assuming it is in fact impossible that it be necessary what things there are).[31] Another reason for doubt is that, even if we grant that not all counterfactuals with impossible antecedents are vacuously true, commutativity with negation requires the validity of the schema: Either, had it been necessary what things there are, it would have been the case that $\psi$, or, had it been necessary what things there are, it would have been the case that not-$\psi$. It would be *ad hoc* to accept this schema without accepting the more general principle of *conditional excluded middle* – that, either, had it been the case that $\varphi$, it would have been the

[30] Note that although 'everything' superficially has narrow scope with respect to the counterfactual, it is effectively quantifying *into* the scope of the counterfactual (indicated by $\uparrow_1$) from the *outside* (i.e., 'actual possibility').

[31] See Lewis (1973) and, more recently, Williamson (2007, pp. 171–175).

case that $\psi$, or, had it been the case that $\varphi$, it would have been the case that not-$\psi$ – which is highly controversial.

Furthermore, since the modalized quantifiers that are to be paraphrased might occur in modal contexts, the paraphrase in fact requires a necessitated version of conditional excluded middle according to which, necessarily for all $p$ and $q$, it is *necessary* that, had it been the case that $p$, it would have been the case that $q$, or, had it been the case that $p$, it would have been the case that not-$q$. Surprisingly, this principle entails that haecceitistic plenitude is necessarily true.[32] Since the whole point of the present strategy was to meet the expressive power challenge without having to accept haecceitistic plenitude, this result significantly undercuts the strategy's appeal.[33]

A final worry for the proposal is that it requires an actuality operator that enables us to scope out of the relevant counterfactual environments and a family of 'in fact' operators that allow us to scope back into them. Such operators are well studied and well behaved when we consider embeddings under necessity and possibility operators. But it is not at all clear that we can make good sense of operators that achieve the same effect with respect to these hyperintensional environments.

### 6.3.3 Hyper-possibility

Rather than explore these epicycles further, let us take a step back. We saw in section 6.1 that incompossibility is the barrier to paraphrasing modalized

[32]See Fritz and Goodman (unpublished a, here ch. 4).

[33]We could avoid this problem by adopting the following non-compositional paraphrase strategy: We paraphrase any sentence by making it the consequent of a counterfactual whose antecedent is the claim that there are all actually possible things, and replace all modalized quantifiers by unmodalized quantifiers evaluated at the scope of the counterfactual's consequent and restricted to actually possible things. That is, we paraphrase a sentence $\varphi$ containing modalized quantifiers as 'Had there been all actually possible things, then $\uparrow_1$ actually $\varphi^{\dagger}$', where $\varphi^{\dagger}$ is the result of replacing all modalized quantifiers in $\varphi$ as described above. For example, 'There are uncountably many possible $F$s'$^{\dagger}$ will be '$\uparrow_2$ in fact$_1$ uncountably many actually possible things are in fact$_2$ $F$'. This paraphrase has the feature that even sentences containing modalized quantifiers in modal scope do not get paraphrased using counterfactuals in modal scope, and therefore only the unnecessitated version of conditional excluded middle is required, which is compatible with the falsity of haecceitistic plenitude.

quantifiers in terms of ordinary quantifiers and modal operators. If we could only somehow bring together all possible individuals in a way that would let us talk about the actual distribution of properties and relations among them, that would be enough. In other words, we could answer the expressive power challenge if only there were *some* notion of possibility – call it *hyper-possibility* – such that it is hyper-possible that there be all actually possible individuals, and such that we can use actuality operators to 'scope out' of hyper-possibility operators and thereby, within the scope of the hyper-possible, draw distinctions among all actually possible things with respect to their actual properties and relations. Retrospectively, we can think of the preceding two sections as exploring two particular proposals for defining such a notion of hyper-possibility, the first appealing to possibilities afforded by fiction and the second to counterfactual possibilities. This section explores the more abstract question of whether there is *any* modality that satisfies this theoretical role, and if so, whether it can be used to answer the expressive power challenge.

In order to get clearer on what this theoretical role is, let us return to Dawkins' claim that most possible people are never born. The proposal is to paraphrase the claim as follows: it is hyper-possible that both (i) there be all actually possible individuals and (ii) most things that are actually (possibly people and possibly identical to something) be actually (never born and possibly identical to something). This paraphrase strategy can be straightforwardly generalized to all modalized quantifiers.

We know that any notion of hyper-possibility for which the above paraphrase strategy succeeds must be a weaker notion of possibility than the notion of 'metaphysical' possibility we have been operating with thus far. But hyper-possibility must also be sufficiently 'metaphysically robust' to permit 'quantifying in' and 'scoping out'. The question is whether any notion of possibility meets these two requirements.

Suppose, for example, that we can make sense of a notion of 'logical possibility' such that, for any formula $\varphi$ of a first-order modal language, $\ulcorner$it is logically possible that $\varphi\urcorner$ is true just in case $\varphi$ is satisfiable relative to the standard Kripke semantics. Note first that Leibniz's law fails to hold within the scope of logical possibility, since it is logically possible that Hesperus is distinct from Phosphorus but not logically possible that Hesperus is distinct from Hesperus. Given such opacity, one might worry whether one can unambiguously quantify into the scope of this operator. Second, and more importantly, we cannot use an actuality operator to scope out of 'it is logically possible that . . . ', since, although there are no flying pigs, it is logically possible that there are actually flying pigs. So this notion of logical possibility is not metaphysically robust, and is therefore not suitable to play the role of hyper-possibility.[34]

Why think that any metaphysically robust notion of possibility supports the hyper-possibility paraphrase? The only positive conception of such a modality we can imagine is one according to which hyper-possibility involves relaxing certain 'laws of metaphysics' in something like the way that metaphysical possibility is sometimes thought to involve relaxing the laws of physics. In particular, the idea would be that metaphysically necessary 'essentialist' principles that entail the incompossibility of the possible people who could be born from a particular sperm and egg do not hold with hyper-necessity. In this way, we might hope that hyper-possibility allows us to get around incompossibility.

It is far from clear that there is a metaphysically robust modality that behaves in the way just sketched. But for present purposes we can sidestep that question. Even if some such modality can be used to evade essentialism-induced incompossibility, there is no reason to think it will evade the second sort of incompossibility discussed in section 2.3 – namely, the incompossibility that results from the fact that, necessarily, whatever possible elementary particles there are, it is possible that there be all of them and one more possible elementary particle. This argument does not seem to rely on any 'hyper-contingent laws of metaphysics'. The same considerations that lead us to think that all metaphysically possible elementary particles are incompossible should lead us to think that all metaphysically possible elementary particles are *hyper*-incompossible – namely, that it is *hyper*-necessary that, whatever *actually metaphysically* possible elementary particles there are, it

[34] Not all notions of logical possibility fail to be metaphysically robust; see Fine (1990).

is *hyper*-possible that there be all of them and one more actually metaphysically possible elementary particle. Compare: it is metaphysically necessary that whatever *actually nomologically* possible elementary particles there are, it is metaphysically possible that there be all of them and one more actually nomologically possible elementary particle.

We therefore see no reason to think that there is any metaphysically robust notion of hyper-possibility according to which it is hyper-possible that there be all actually metaphysically possible elementary particles. And *even if* there were such a notion, it would not allow us to answer the expressive power challenge in full generality, because we could still run a different analogue of the argument from section 2.3. Presumably it is hyper-necessary that, whatever hyper-possible elementary particles there are, it is hyper-possible that there be all of them and one more. It follows that the hyper-possible elementary particles are hyper-incompossible. This raises such questions as: Are *most* hyper-possible elementary particles *metaphysically* possible? The proponent of the hyper-possibility paraphrase now faces an expressive power challenge exactly analogous to the one with which we began. It is no help to appeal to an infinite hierarchy of notions of hyper-possibility, since the disjunction of all such modalities will itself be a notion of hyper-possibility to which the revenge argument applies.

## 6.4 From haecceitistic plenitude to higher-order necessitism

We began with a dilemma: one must either provide some paraphrase of modalized quantification, or take such quantification at face value and accept that there are merely possible people. In section 6.2 we argued that, limiting ourselves to the expressive resources of infinitary higher-order modal logic, taking the first horn of the dilemma carries a commitment to haecceitistic plenitude. In section 6.3 we canvassed a range of proposals to avoid this commitment by appealing to expressive resources beyond those available in standard infinitary higher-order modal logic and found them all deeply problematic. So taking the first horn of the dilemma carries a commitment to haecceitistic plenitude. And taking the second horn of the dilemma also carries a commitment to haecceitistic plenitude, since taking modalized quantification at face value in general commits one to the view that it is necessary what things there are, from which haecceitistic plenitude straightforwardly follows (since every individual has a haecceity). We therefore conclude that haecceitistic plenitude is true: every possible object actually has a haecceity. The otherwise well motivated and well developed theories to the contrary must be rejected. This is a surprising result; let us explore its ramifications.

### 6.4.1 Aboutness

Since haecceitistic plenitude is presumably not a contingent truth, we may conclude that, *necessarily*, every possible object has a haecceity. That is, we ought to accept the following principle:

*Weak Haecceitistic Necessitism*: Necessarily everything necessarily has a haecceity.

Together with the controversial assumption that necessarily co-extensive conditions are identical, weak haecceitistic necessitism entails the following principle:

*Strong Haecceitistic Necessitism*: It is necessary what haecceities there are.[35]

But what about those who deny that all necessarily co-extensive conditions are identical? Might they resist the move from weak to strong haecceitistic necessitism?

Few have wanted to occupy such a position. For among those who deny that necessarily co-extensive conditions are identical, those who have also wanted to resist strong haecceitistic necessitism have usually been motivated by the thought that (i) necessarily equivalent conditions can be distinct by

[35]Strong haecceitistic necessitism entails weak haecceitistic necessitism given the assumption that necessarily everything has a haecceity, which should be uncontroversial given the operative abundant understanding of conditions; see Williamson (2013, chapter 6).

being *about* different individuals, (ii) which individuals a condition is about is essential to it, and (iii) for all $x$, necessarily, no $F$ is about $x$ unless something is identical to $x$. For example, the condition of being identical to you and the condition of possibly being your biological father are respectively haecceities of you and of your biological father that are about you, and therefore have contingent being on account of your continent being.

By itself this combination of commitments is perfectly consistent with weak haecceitistic necessitism, since it is consistent with the claim that, necessarily, every individual has a qualitative haecceity with necessary being.[36] But this claim is widely rejected, since it is equivalent to an implausibly strong version of the identity of indiscernibles, according to which, necessarily, every object has a qualitative property that is necessarily co-extensive with being identical to it.[37] Ordinary material objects, for example, seem to lack qualitative haecceities. Without qualitative haecceities, weak haecceitistic necessitism entails (given the aforementioned assumptions) that, for every material object $x$, necessarily, there are some individuals that *single out* $x$, where some individuals single out $x$ just in case there is some condition which singles out $x$ which is about all and only them. And this claim is also thought to be wildly implausible, presumably on the grounds that, e.g., had spacetime been empty but for a lone elementary particle, there would not have been any individuals that singled you out.[38]

So it seems that the package of commitments that normally motivates hyperintensional contingentist higher-order metaphysics, and thereby opens up logical space for accepting weak but not strong haecceitistic necessitism, supports an independent argument against weak haecceitistic necessitism, rendering such a split decision ultimately unsustainable. (Furthermore, given that haecceities are generally considered the best candidates for contingent

[36]A condition is *qualitative* just in case it is not about any possible individual.

[37]See McMichael (1983) for discussion.

[38]Fine (1985, p. 190) gives an inchoate version of this argument, writing that 'even though a property [of being identical to some merely possible individual] has no [...] counterpart [...] involving actual individuals alone, there may, in each world, exist individuals that suffice to specify its application conditions. This, though, would be a kind of modal freak.'

higher-order beings, it would be *ad hoc* having accepted strong haecceitistic necessitism to not accept necessary being at higher-orders across the board.)

### 6.4.2 Plenitude

This argument is too quick. It turns out that a popular and independently motivated view about the metaphysics of material objects entails that, even in a one-particle universe, there would be material objects in terms of which you could be singled out. Say that two material objects *coincide* just in case they mereologically overlap the same things. It is widely believed that distinct material objects can coincide – for example, a statue and the lump of clay that composes it. According to one popular version of this view, coincidence is ubiquitous in the following sense: for any function from possibilities to possible objects that would have been material objects had the relevant possibility been realized, there is a possible material object whose modal profile of mereological coincidence is given by that function.[39] It follows that, in any possible one-particle universe, there is an object coincident with the particle in that situation, coincident with you in all possible situations in which you have being, and without parts in all other possibilities. We might then conjecture that, in a one-particle universe, you could be singled out in terms of your possible mereological relations to such objects. In fact, this conjecture can be established from the plausible assumption that, necessarily, any two distinct yet necessarily coincident material objects will differ in some qualitative respect: It then follows that for every possible material object has there is a haecceity of it that is about only individuals that there actually are (and this argument for haecceitistic plenitude generalizes to weak haecceitistic necessitism if we assume that necessarily there is at least one material object).[40]

[39]A closely related thesis is defended by Hawthorne (2006) under the heading of 'plenitude' and is formally explored by Hovda (2013).

[40]We argue as follows. For convenience we will indulge in 'possible worlds'-jargon, although such talk can in principle be eliminated in a higher-order language. Consider a possible material object $a$ and world $w$ which does not contain $a$. We will argue that in $w$ there is a property that singles out $a$. We may assume that in $w$ there is at least one material object $b$. Plenitude entails that in $w$ there is an object $c$ which coincides with

Although the above argument establishes weak haecceitistic necessitism only for possible material objects, we can establishes full weak haecceitistic necessitism if we assume that all possible objects can be singled out in terms of qualitative relations to possible material objects and to abstract objects which have necessary being. This assumption is not implausible – it seems to hold for impure sets, for example, which are normally taken to be neither material objects nor necessary beings – although Cartesian egos and the like, if possible, might be counterexamples. These considerations show that independently motivated assumptions – most centrally, a form of mereological

---

$b$ in $w$, coincides with $a$ in all worlds containing $a$, and is nothing in all other worlds. It also entails that in $w$ there is an object $d$ that coincides with $b$ in $w$ and is nothing in all other worlds. We now define the condition $F$ as follows: $Fx =_{df}$ necessarily: $x$ is something if and only if $c$ is something but $d$ isn't, and, if $x$ is something, it coincides with $c$. By construction, $F$ is necessarily co-extensive with being necessarily coincident with $a$. And $w$ contains $F$, since it was specified only in terms of individuals and conditions that there are in $w$ – namely in terms of $c$, $d$, and modal and mereological conditions (which have necessary being).

Now consider some world $v$ containing $a$. Assuming it is necessary that necessarily coincident objects are qualitatively discernible, it follows that, for every possible object $x$ distinct from $a$ that is $F$ at $v$, there is a qualitative condition $Q_x$ that distinguishes $x$ and $a$ at $v$. Since the negation of any qualitative condition is itself qualitative, we may choose $Q_x$ such that $a$ has $Q_x$ at $v$. Choose some such $Q_x$ for each $x$ distinct from $a$ that is $F$ in $v$. Let $Q$ be the conjunction of these conditions. By construction, $a$ is unique among the objects that are $F$ in $v$ in having $Q$ in $v$. And $w$ contains $Q$, since it is the conjunction of qualitative conditions, qualitativeness is closed under conjunction, and qualitative conditions have necessary being.

By plenitude, there is a material object $e$ in $w$ that coincides with $b$ in $w$, coincides with $a$ in $v$, and is nothing in all other worlds. Now let $G$ be the condition of being $F$ and being, necessarily, $Q$ if coincident with $e$. As can be straightforwardly verified, $G$ singles out $a$: necessarily, everything is $G$ if and only if it is identical to $a$. And $w$ contains $G$, since it was specified only in terms of individuals and conditions that there are in $w$ – namely, in terms of $e$, $F$, $Q$, and modal and mereological conditions which have necessary being. So there is a condition in $w$ which singles out $a$, even though $w$ does not contain $a$. Since $w$ and $a$ are arbitrary, we may conclude that, necessarily, every material object necessarily has a haecceity. (Here we assume that aboutness is the only source of higher-order contingency; see Fritz and Goodman (unpublished c, here ch. 1) for discussion.)

Note that our assumption that, necessarily, distinct necessarily coincident material objects are qualitatively discernible should be distinguished from the stronger claim that, necessarily, distinct necessarily coincident material objects are of different 'sorts'. In particular, our assumption is not threatened by the putative counterexample to the latter principle given in Fine (2000). Note also that our other assumption – that necessarily there is at least one material object – is in fact dispensable, since, as discussed in section 6.2.2, haecceitistic *compossibilism* is all one really needs to meet our expressive power challenge.

plenitude – support an argument for weak haecceitistic necessitism that is both independent of our expressive power challenge and consistent with the claim that haecceities have contingent being in virtue of being about possible individuals with contingent being. While these assumptions are admittedly tendentious, they show that the widely accepted inference from weak to strong haecceitistic necessitism is too quick.[41]

These considerations raise the question of whether one could answer the expressive power challenge using the higher-order paraphrase discussed in section 6.2 by accepting plenitude and so weak haecceitistic necessitism but denying strong haecceitistic necessitism. We will ultimately argue for a negative answer to this question, but first we need to get a clearer understanding of the nature of our expressive power challenge.

## 6.5 Expressive Power Revisited

In section 6.1 we posed the challenge: for any generalized quantifier and claim involving it, to express the corresponding modalized claim without contradicting the philosophical orthodoxy according to which ordinary things are contingent beings. In this section we advocate a particular understanding of this challenge and reconsider the arguments of the previous sections in light of it.

### 6.5.1 The Challenge

We understand our expressive power challenge as follows: we use modalized quantifiers to convey particular *propositions* about modal reality, and the paraphrase challenge is to provide a *systematic* way of expressing these propositions without using modalized quantifiers. For example, when Dawkins says that most possible people are never going to be born, there is a non-trivial proposition he succeeds in conveying (whether or not it counts as the 'literal semantic content' of his utterance). The challenge requires that one produce

[41] The argument is of course specific to the particular form of mereological plenitude discussed here; it is likely that there are variants which do not entail weak haecceitistic necessitism.

a sentence not containing modalized quantifiers that (literally) expresses this proposition. Furthermore, the paraphrase cannot be *ad hoc*. There must be a general algorithm for producing paraphrases, one that cares only about the form of the sentences being paraphrased.

It is important to distinguish this proposed understanding of the expressive power challenge from an alternative understanding that might otherwise be suggested by our technical appendices and by much of the related literature: namely, to produce a sentence that is *logically equivalent* to the one in need of paraphrase. We don't have a firm enough grip on the notion of logical equivalence to say precisely how such a proposal might come apart from ours. But our feeling is that, were someone to meet our challenge but fail to produce a 'logically' equivalent paraphrase (on some given notion of logical equivalence), that fact would not constitute a particularly pressing worry for their position.[42]

The advantage of the alternative proposal – of requiring paraphrases to be logically equivalent to what they paraphrase – is that it makes clear the relevance of the model-theoretic results we appealed to above. On our proposed understanding of the expressive power challenge, the role of model theory in evaluating candidate paraphrase strategies is less familiar. Here is how we understand it. Call a class of models *probative* just in case, for any two sentences of the formal language interpreted over those models, the two sentences express the same proposition only if they define the same class of those models.[43] From the assumption that a given class of models is probative, we can argue that a proposed paraphrase is inadequate by showing that some sentence in need of paraphrase and its proposed paraphrase define

[42]One could also strengthen our systematicity condition by requiring the paraphrase to be *compositional*: by demanding, for example, that the paraphrase of a sentence $\varphi$ occur as a subformula of the paraphrase of any sentence $\psi$ of which $\varphi$ is subformula. Such demands strike us as unduly restrictive: Russell's celebrated theory of descriptions, after all, fails to be compositional in this sense.

[43]More precisely: any interpretation of the non-logical constants of the language determines a function from sentences of the language to the propositions expressed by those sentences, and a class of models is probative just in case, for any interpretation of the language, two sentences of the language express the same proposition on this interpretation only if they define the same class of models. (The class of models defined by a sentence is the class of models in which it is true.)

different classes of models: by probativity, it follows that they express distinct propositions, and so the expressive power challenge has not been met. More generally, we can argue that a certain sentence has *no* paraphrase in a given formal language by showing that no sentence of that language defines the same class of models as the sentence in need of paraphrase defines. The model-theoretic results appealed to in sections 6.1 and 6.2 are of this form. In this way we can use intensional model theory to answer hyperintensional questions.[44]

Probative classes of models do not straightforwardly give us a way of arguing that a given paraphrase *is* adequate.[45] But we can still use probative classes of models as a heuristic in assessing the adequacy of candidate paraphrases, to be supplemented by independent judgments of propositional identity. Every probative class of models provides a necessary condition of adequacy on any candidate paraphrase strategy. Satisfying such conditions is non-trivial, and so provides substantial, though defeasible, evidence for a strategy's adequacy.

(Our understanding of the expressive power challenge in terms of propositional identity will likely be rejected by those who accept fine-grained theories of propositions that would render the challenge so-construed hopelessly demanding. It would be natural for friends of such views to instead understand the challenge in terms of some relation of 'metaphysical equivalence' between propositions, such as corresponding to the same 'state of affairs'. For example, proponents of structured propositions will distinguish the proposition that it is raining from the proposition that it is raining and it is raining, and Fregeans will distinguish the proposition that Hesperus is a planet from the proposition that Phosphorus is a planet, but everyone ought to recognize a sense in which these propositions make the same demands on reality.

[44] In Fritz and Goodman (unpublished c, here ch. 1, section 3.4) we address the question of how to understand higher-order quantifiers interpreted over intensional models without prejudging the issue of hyperintensionality.

[45] Such an argument would require an *individuating* class of models: one such that any two sentences that define the same class of models express the same proposition on any interpretation of their non-logical constants. But assuming our language contains classical Boolean connectives, there could only be such a class of models if the class of propositions forms a Boolean algebra, a hypothesis about which we are skeptical.

All of what follows could be reframed in terms of such a notion, although we will continue to frame the paraphrase challenge in terms of propositional identity.)

### 6.5.2 Negative Arguments Reconsidered

Having clarified the role of model theory in assessing candidate paraphrase strategies, we can now make explicit the structure of the negative arguments of sections 6.1 and 6.2.

The tacit assumption of the arguments of sections 6.1.3–6.1.5 is that the class of Kripke models for first-order modal languages deployed in the appendices is probative. Given this premise, the aforementioned technical results establish that the expressive power challenge, on our understanding thereof, cannot be met in a first-order modal language, since some sentences containing modalized quantifiers define classes of models not defined by any even highly-infinitary first-order sentence not containing modalized quantifiers.

The tacit assumption of the argument of section 6.2.4 is that, *if haecceitistic plenitude is false*, then the class of 'internally closed' models described in Fritz and Goodman (unpublished c, here ch. 1) is probative. Given this premise, the results in Fritz (unpublished b, here ch. 7) establish that claims involving 'there are uncountably many possible . . . ' cannot be paraphrased in an infinitary higher-order modal language, and hence the expressive power challenge cannot be met using higher-order resources if haecceitistic plenitude is false.

These premises are not undeniable. They will be rejected, for example, by those who think that necessarily equivalent propositions are identical. This is because two sentences of a first-order modal language can define different classes of Kripke models while both expressing necessary falsehoods (e.g., concerning the cardinality of possible individuals). Since they express necessarily equivalent propositions, the coarse-grained view of propositions under consideration entails that they express the *same* proposition, in which case the class of Kripke models is not probative and our deductive arguments do not go through. But resisting an argument that the challenge *cannot* be met using certain expressive resources does nothing to suggest that it *can* be met using those resources. Accepting a coarse grained theory of propositions does not by itself suggest any such strategy. To appreciate this point, consider the extensionalist view according to which there are only two propositions: the True and the False. On this view, the propositions expressed using modalized quantifiers can clearly be expressed in all sorts of other terms. But answering the expressive power challenge requires a *systematic* paraphrase: an *algorithm* that cares only about the form of the sentences being paraphrased. No algorithm suggests itself.[46]

### 6.5.3 Positive Arguments Reconsidered

Our claim that the class of internally closed models was probative was conditional on the falsity of haecceitistic plenitude. What if haecceitistic plenitude is true? Then, for all we have said, the haecceitistic paraphrase strategy outlined in section 6.2.2 meets the expressive power challenge. But what positive considerations might be offered in its favor?

One thing to note is that the class of Kripke models with constant higher-order domains is arguably probative regardless of whether haecceitistic plenitude is true.[47] And the proposed haecceitistic paraphrase strategy succeeds as judged by this restricted class of models, in the sense that it maps every formula containing modalized quantifiers to a formula not containing modalized quantifiers that is true in exactly the same models in this restricted class.

[46]It is worth noting that certain eccentric metaphysical views would support systematic paraphrases given the assumption that propositions are individuated intensionally. For example, if one thought that there were only finitely many possible contingent beings, then one should be able to systematically paraphrase all modalized generalized quantifiers in a first-order modal language with infinitary conjunction. But this sort of modal finitism is wildly implausible. More interestingly, the intensionalist might be able to cook up a systematic first-order paraphrase were they to accept the plenitudinous modal mereology described in section 6.4.2. But even if this can be done, the resulting view entails weak haecceitistic necessitism, which, assuming the hypothetical advocate of this strategy rejects hyperintensional distinctions for properties and propositions alike, leads to strong haecceitistic necessitism. So the availability of such a strategy does not threaten the main conclusion of this paper.

[47]They are probative if haecceitistic plenitude is false because they are a subclass of the internally closed models, and any subclass of a probative class of models is itself probative.

Clearing such a hurdle is no small feat, so we ought to take the proposed paraphrase seriously. But the viability of the paraphrase must ultimately be judged on more general theoretical grounds. We can only assess it in conjunction with a more general picture of modal reality relative to which the paraphrase is offered.

In section 6.4 we considered two such pictures. The first accepts strong haecceitistic necessitism. The second accepts weak but not strong haecceitisitic necessitism on the basis of the following combination of commitments: principles (i)–(iii) from section 6.4.1 (i.e., the aboutness-theoretic hyperintensional contingentist theory of properties), a plenitudinous modal mereology, and a few other backgrounds assumptions that, together with the modal mereology, entailed weak haecceitistic necessitism. Let us consider these views in turn.

If strong haecceitistic necessitism is true, then it is not implausible that the haecceitistic paraphrase yields sentences that express the same propositions as those expressed by the sentences they paraphrase (or, if you prefer, 'metaphysically equivalent' propositions, as per section 6.5.1). We are not endorsing this claim, but it is not without its attractions and is at any rate not obviously false.

If, on the other hand, strong haecceitistic plenitude is false on hyperintensional aboutness-theoretic grounds, and yet weak haecceitistic plenitude manages to come out true as the result of a fortuitous modal mereology, then it is not plausible that the haecceitistic paraphrase yields sentences that express the same propositions as those expressed by the sentences they paraphrase. Given the conspiratorial air surrounding weak haecceitistic necessitism on this view, and given the fact that the view already involves drawing substantive hyperintensional distinctions, it is not credible that the proposition that most possible people are never born is identical to the proposition that most haecceities that possibly single out people possibly single out things that are actually never born. (The paraphrase will actually need to be even more complicated owing to the fact that the view in question entails that every possible object has many haecceities.) Reflection on these two cases suggests that the haecceitistic paraphrase strategy is successful only if strong haecceitistic plenitude is true – and, more generally, only if what properties and relations there are is a necessary matter.

## 6.6 Conclusion

We have argued as follows. If it is *not* contingent what individuals there are, then is it not contingent what properties and relations there are either. If it *is* continent what individuals there are, then we face a challenge of making sense of modalized quantification. This challenge can be met only if it is not contingent what properties and relations there are. So, reasoning by cases, we conclude that it is not contingent what properties and relations there are (where 'property'- and 'relation'-talk is shorthand for what would be properly expressed in higher-order terms). One might think that, once we have gone this far, we should go further and embrace the conclusion that it is necessary what individuals there are.[48] But while we are ultimately sympathetic to the view that it is necessary what individuals there are, we do not think it follows in any straightforward way from what we have established here.[49]

## 6.7 Appendix: Definability of modalized generalized quantifiers in infinitary first-order modal logics

In this appendix, we prove and discuss the results on the definability of modalizations of generalized quantifiers appealed to in section 6.1. In the

[48]Fine (1985, section 2) and Williamson (2013, section 6.2) defend this conditional claim.

[49]We do not mean to suggest that orthodox metaphysicians who think it is contingent what individuals there are can rest content so long as they are willing to accept that it is necessary what properties and relations there are. There are other expressive power challenges they must still answer: for example, the challenge to make sense of modalized *plural* quantification. Goodman (in preparation a) argues that this further challenge cannot be met *even assuming that it is necessary what properties and relations there are*, and on these grounds argues that it is in fact necessary what individuals there are.

Note that while the technical results of Williamson (2010, 2013, chapter 7) are similar to those of our appendix, his mode of argument is radically different from ours. For criticism of his argument, see Goodman (in preparation b).

interest of brevity and readability, we will keep the definitions somewhat informal, and only sketch proofs of the results.

### 6.7.1 Well-founded languages

We start with the more standard languages appealed to in sections 6.1.3 and 6.1.4, which require the subformula relation to be well-founded. For simplicity, we assume that we are working in a relational signature, so we don't allow individual constants or function symbols. As usual, we let $\mathcal{L}_{\omega\omega}$ be the language of first-order logic, i.e., the set of formulas built out of atomic predications (including identity statements, $=$ being treated as a logical connective) using negation ($\neg$), binary disjunction and conjunction ($\vee$ and $\wedge$), and existential and universal quantification ($\exists$ and $\forall$). Likewise, we use $\mathcal{L}_{\infty\omega}$ for the extension of this language obtained by allowing the disjunction and conjunction operators to apply to sets of formulas of arbitrary cardinality (writing them $\bigvee$ and $\bigwedge$), and $\mathcal{L}_{\infty\infty}$ for the further extension obtained by also allowing existential and universal quantification binding sets of variables of arbitrary cardinality. These languages are interpreted as usual over a model $\mathfrak{A}$ given by a set $|\mathfrak{A}|$ (the domain of quantification) and a relation $R^{\mathfrak{A}}$ of the appropriate arity on $|\mathfrak{A}|$ for each relational symbol $R$ in our signature. We write $\mathfrak{A}, a \vDash \varphi$ for $\varphi$ being true in $\mathfrak{A}$ relative to an assignment function $a$; note that $a$ may be partial, as long as it is defined on all free variables in $\varphi$.

From $\mathcal{L}_{\omega\omega}$ and $\mathcal{L}_{\infty\infty}$, we derive first-order modal languages $\mathcal{L}^{\mathrm{mq}}_{\omega\omega}$ and $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ by adding four resources: (i) all generalized quantifiers; (ii) a new kind of variables, called world-variables, which can be used as atomic formulas; (iii) existential and universal quantifiers binding a single world variable; and (iv) an additional operator @ operating on a world variable and a formula. We will mark the difference between the two kinds of variables by using Roman letters for individual variables and Greek letters for world variables.

We interpret the modal languages over the usual Kripke models, understood as tuples $\langle W, D, i, w\rangle$, where $W$ is a set (the 'set of worlds'), $D$ is a function mapping each $v \in W$ to a set $D_v$ (the 'domain of $w$'), and $i$ is a function mapping each relation symbol of our signature to a function which maps every world $v$ to a relation of the appropriate arity on $D_v$. We evaluate formulas relative to a world $v$ and an assignment function $a$, writing $\mathfrak{M}, v, a \vDash \varphi$ for $\varphi$ being true in $\mathfrak{M}$ and $v$ relative to $a$. Assignment functions may again be partial, and in addition to mapping individual variables to individuals, now also map world variables to worlds. We let $D_v$ be the domain of individual quantifiers – including generalized quantifiers – evaluated at $v$. A world variable is true in just the world assigned to it; world quantifiers range over the set of worlds; and the effect of $@\xi$ is to let the world of evaluation be the world assigned to $\xi$.

More precise definitions of such modal languages can be found in Williamson (2013, section 7.9) and Fritz (2013), although both use a kind of generalized actuality operators instead of explicit quantification over worlds. We choose explicit quantification over worlds here since this generalizes more naturally to the case of non-well-founded languages; see Appendix 6.7.5 for further discussion of this difference. For now, we only note that it is straightforward to use the resources available here to define the familiar operators for possibility and necessity as follows, where $\xi$ is some world variable not free in $\varphi$:

$$\Diamond\varphi =_{\mathrm{df}} \exists\xi @\xi\varphi$$

$$\Box\varphi =_{\mathrm{df}} \forall\xi @\xi\varphi$$

For more on quantifiers binding world variables, see the literature on hybrid logic, e.g., Areces and ten Cate (2007). See Westerståhl (2011) for a definition of generalized quantifiers and their interpretation, and Fritz (2013, section 1) for generalized quantifiers in modal logic.

We say that a generalized quantifier $\mathcal{Q}$ is *definable* in a given language just in case adding it to the language does not increase its expressivity, in the sense that every sentence of the expanded language is equivalent to some sentence in the original language (where two sentences are equivalent just in case they are true in the same models and false in the same models – the non-well-founded languages discussed below introduce gaps in the truth-value assignment, so agreement on truth and falsity have to be imposed separately). A helpful reformulation of this notion uses what we might call the *canonical sentence* $\gamma_{\mathcal{Q}}$ of a generalized quantifier $\mathcal{Q}$, which is simply the result of applying it to a sequence of atomic formulas using relation symbols of appropriate arities, keeping all variables and relation symbols distinct. We can show that in any non-modal language considered here (including the non-well-founded languages to be introduced below), $\mathcal{Q}$ is definable if and only if the language contains a sentence equivalent to $\gamma_{\mathcal{Q}}$.

For any generalized quantifier $\mathcal{Q}$, define the modalization of $\mathcal{Q}$ to be an operator $\mathcal{Q}^O$ which when added to any modal language considered here has the following truth-conditions:

$$\mathfrak{M}, v, a \vDash \mathcal{Q}^O \bar{\bar{x}}\bar{\varphi} \text{ iff } \mathcal{Q}_{\bigcup_{u \in W} D_u} \left( \varphi_0(\bar{x}_0)_O^{\mathfrak{M},v,a}, \ldots, \varphi_{l-1}(\bar{x}_{l-1})_O^{\mathfrak{M},v,a} \right)$$

Here, $\mathfrak{M} = \langle W, D, i, w \rangle$ and

$$\varphi(\bar{x})_O^{\mathfrak{M},v,a} = \left\{ \bar{o} \in \left( \bigcup_{u \in W} D_u \right)^n : \mathfrak{M}, v, a[\bar{o}/\bar{x}] \vDash \varphi \right\}.$$

(In general, we indicate a sequence of elements by putting a bar over the relevant term, leaving the length of the sequence to be determined by the context. We also adopt the usual convention of writing $a[o/x]$ for the $x$-variant of the assignment $a$ which maps $x$ to $o$, and extend this to the case of tuples in the obvious way.) Definability of modalized generalized quantifiers is understood as before, except that we take two sentences to be equivalent if they are true/false in the same Kripke models in the same worlds. We also extend the definition of the canonical sentence of a modalized generalized quantifier in the obvious way, but note that in the Kripke models used here, the extension of a relation at a world is confined to the individuals at that world, so it is no longer clear that being able to express the canonical sentence of a modalized generalized quantifier implies being able to define it, although of course the converse still holds.

### 6.7.2 Definability in well-founded languages

**Theorem 6.7.1.** *The modalization of a generalized quantifier $\mathcal{Q}$ is definable in $\mathcal{L}_{\omega\omega}^{\mathrm{mq}}$ if and only if $\mathcal{Q}$ is definable in $\mathcal{L}_{\omega\omega}$.*

*Proof.* It is straightforward to adapt the proof of Fritz (2013, p. 656, Theorem 1) to the present setting. □

**Theorem 6.7.2.** *The modalization of a generalized quantifier $\mathcal{Q}$ is definable in $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ if and only if $\mathcal{Q}$ is definable in $\mathcal{L}_{\infty\omega}$.*

*Proof.* If $\mathcal{Q}$ is definable in $\mathcal{L}_{\infty\omega}$ then there is a sentence $\varphi$ of $\mathcal{L}_{\infty\omega}$ equivalent to $\gamma_{\mathcal{Q}}$. Define $\varphi^O$ to be the result of replacing quantifiers by complex constructions in $\varphi$, replacing

$\exists x\psi$ by $\exists\xi(\xi \wedge \Diamond\exists x@\xi\psi)$, and

$\forall x\psi$ by $\exists\xi(\xi \wedge \Box\forall x@\xi\psi)$.

For any $\chi$ in $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ enriched by $\mathcal{Q}^O$, replacing every occurrence of $\mathcal{Q}^O$ by $\varphi^O$ (in which atomic predications are in turn replaced by the relevant subformulas of $\chi$) results in a sentence of $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ which we can prove to be equivalent to $\chi$.

For the other direction, we adapt the proof of Fritz (2013, p. 664, Proposition 10). Let $\mathcal{Q}$ be a generalized quantifier not definable in $\mathcal{L}_{\infty\omega}$. Consider any infinite ordinal $\alpha$ and let $\mathcal{L}^{\alpha}_{\infty\omega}$ be the class of $\mathcal{L}_{\infty\omega}$ sentences of quantifier rank up to $\alpha$. As noted in Väänänen (2004, p. 46, Lemma 9), $\mathcal{L}^{\alpha}_{\infty\omega}$ has at most $\beth_\alpha$ sentences up to equivalence, so the class of models defined by $\gamma_{\mathcal{Q}}$ is not the union of equivalence classes of the relation of satisfying the same sentences of $\mathcal{L}^{\alpha}_{\infty\omega}$. Hence there are models $\mathfrak{A}$ and $\mathfrak{B}$ which satisfy the same sentences of $\mathcal{L}^{\alpha}_{\infty\omega}$, while $\mathfrak{A}$ but not $\mathfrak{B}$ satisfies $\gamma_{\mathcal{Q}}$. By Karp (1965, p. 410, Theorem 1), there is a back-and-forth system of length $\alpha$ relating $\mathfrak{A}$ and $\mathfrak{B}$. As in Fritz (2013, p. 664, Lemma 9), we can extend this to a kind of back-and-forth system which holds between Kripke models $\mathfrak{A}^n$ and $\mathfrak{B}^n$ based on $\mathfrak{A}$ and $\mathfrak{B}$, and conclude from this that $\mathfrak{A}^n$ and $\mathfrak{B}^n$ satisfy the same sentences of $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ up to modal depth $\alpha$ (which can now be understood as world quantifier depth). From the fact that $\mathfrak{A}$ but not $\mathfrak{B}$ satisfies $\gamma_{\mathcal{Q}}$, we can conclude that $\mathfrak{A}^n$ but not $\mathfrak{B}^n$ satisfies $\gamma_{\mathcal{Q}^O}$. Since every sentence of $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ has some ordinal modal depth, we conclude that no such sentence defines $\mathcal{Q}^O$. □

### 6.7.3 Non-well-founded languages

We now turn to the languages appealed to in section 6.1.5, which allow for infinitely deep nestings of operators. The literature on such languages goes back to Hintikka and Rantala (1976), whose approach we follow here. According to it, formulas are given by labeled trees – allowing infinite branching as well as infinitely long branches – in which leaf nodes are labeled by atomic formulas and non-leaf nodes are labeled by operators. Here, we will define these languages in a somewhat unorthodox way, in a sense combining the traditional recursive definition of formulas with Hintikka and Rantala's tree-based approach. We motivate setting up the languages in this way at the end of this section; now, we start by defining the syntax and semantics.

Define a *tree* to be a partial order $\langle T, \leq \rangle$ such that for all $t, t' \in T$:

(a) $\{s : s \leq t\}$ is well-ordered by $\leq$,

(b) there is a leaf (maximal element) $s$ such that $t \leq s$, and

(c) if $\{s : s \leq t\} = \{s : s \leq t'\}$ and the order type of this set is a limit ordinal, $t = t'$.

With this, we define the formulas of $\mathcal{N}_{\infty\infty}$ using the standard method of (finitary) structural recursion, but with the following (partly non-standard) clauses:

- For any relation symbol $R$ and variables $\bar{x}$, $R\bar{x}$ is a formula.
- $\langle T, \leq, g \rangle$ is a formula, given that
  - $\langle T, \leq \rangle$ is a tree (as defined above),
  - $g$ is a function on $T$ mapping any leaf to a formula and any non-leaf to $\neg$, $\vee$, $\wedge$, $\exists x$ or $\forall x$ (for some individual variable $x$), and
  - if $g(t) \notin \{\vee, \wedge\}$, then $t$ has a unique successor.

To interpret formulas of $\mathcal{N}_{\infty\infty}$, we adapt Hintikka and Rantala's game-theoretic semantics. Since we have a hybrid syntax, we also proceed with a hybrid semantics, using games only for the step of trees in the usual recursion of truth-conditions relative to a model and an assignment function. As before, we allow assignment functions to be partial, and in this case, don't even require them to be defined on all free variables. As infinite embeddings may introduce truth-value gaps, we define separate properties of truth and falsity, noting that the semantics will never assign both truth and falsity, but sometimes neither.

An atomic formula $R\bar{x}$ is undefined if $a$ is undefined on some variable in $\bar{x}$, in all other cases, truth and falsity are defined in the usual bivalent manner. For a formula based on a tree, we construct a game between two players, $V$ ('verifier') and $F$ ('falsifier'), defining the formula to be true if and only if $V$ has a winning strategy and false if and only if $F$ has a winning strategy. We define this game as follows:

Let $\varphi = \langle T, \leq, g \rangle$ be a formula of $\mathcal{N}_{\infty\infty}$, $\mathfrak{A}$ a model and $a$ an assignment function $\varphi$. Plays of the game determined by these three items consist of a (possibly transfinite) sequence of stages. Each such stage is given by a node of the tree (i.e., an element of $T$) and an assignment function. $V$ wins the game if the node of the last stage is labeled by a formula which is true in $\mathfrak{A}$ and the assignment function of the last stage; $F$ wins the game if this formula is false relative to these parameters; the game is a draw otherwise. Note that the game can be a draw because the formula labeling the node of the last stage is neither true nor false relative to the relevant parameters, or because there is no last stage. We define the plays of the game determined by $\varphi$, $\mathfrak{A}$ and $a$ by transfinite induction (where the play concludes once the node of the current stage is a leaf of the tree):

- Stage 0 is given by the root node of $\varphi$ and $a$.
- If stage $\alpha$ is given by node $t$ and assignment $b$, then:
  - If $t$ is labeled by $\neg$, then $V$ and $F$ switch roles; stage $n+1$ is given by the successor of $t$ and $b$.
  - If $t$ is labeled by $\vee$, then $V$ chooses one of its successors $t'$; stage $n+1$ is given by $t'$ and $b$.
  - If $t$ is labeled by $\wedge$, then $F$ chooses one of its successors $t'$; stage $n+1$ is given by $t'$ and $b$.
  - If $t$ is labeled by $\exists x$, then $V$ chooses an element $o$ of the domain of $\mathfrak{A}$; stage $n+1$ is given by the successor of $t$ and $b[o/x]$.

  – If $t$ is labeled by $\forall x$, then $F$ chooses an element $o$ of the domain of $\mathfrak{A}$; stage $n+1$ is given by the successor of $t$ and $b[o/x]$.

- Stage $\lambda$, for limit ordinal $\lambda$, is the limit of the stages $< \lambda$, defined as follows: Let $\langle t_\alpha : \alpha < \lambda \rangle$ and $\langle b_\alpha : \alpha < \lambda \rangle$ be the sequences of nodes and assignment functions of the stages $< \lambda$. Then we define the node of the limit stage to be the first node after the elements of $\langle t_\alpha : \alpha < \lambda \rangle$ (this is guaranteed to be unique by constraint (c) of the definition of trees above). We define the assignment of the limit stage to map every variable $x$ to the element $o$ such that for some $\alpha < \lambda$, $b_\beta(x) = o$ for all $\beta < \lambda$ such that $\alpha < \beta$, and to be undefined on $x$ of there is no such element.

This concludes the definition of $\mathcal{N}_{\infty\infty}$. As noted in Rantala (1979, p. 122), we can turn every sentence of $\mathcal{L}_{\infty\infty}$ into an equivalent sentence of $\mathcal{N}_{\infty\infty}$ by replacing every quantification over a set of variables of size $\kappa$ by an infinite sequence of existential quantifiers of length $\kappa$. However, there are sentences of $\mathcal{N}_{\infty\infty}$ which have no equivalent in $\mathcal{L}_{\infty\infty}$; this follows from the fact that the so-called *game quantifier*, which can be thought of as an $\omega$ sequence of alternating existential and universal quantifiers, is clearly definable in $\mathcal{N}_{\infty\infty}$, but as noted in Kolaitis (1985, p. 370), it can be shown not to be definable in $\mathcal{L}_{\infty\infty}$ (see Väänänen (2011, p. 244, Proposition 9.38) for a proof).

Formulas of the modal extension $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ are defined as in the case of well-founded languages by adding clauses for the resources (i) – (iv) specified above. More precisely, we define a formula of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ using the following recursion:

- For any relation symbol $R$ and individual variables $\bar{x}$, $R\bar{x}$ is a formula.

- $\xi$ is a formula, for any world variable $\xi$.

- $\mathcal{Q}\bar{\bar{x}}\bar{\varphi}$ is a formula, given that $\mathcal{Q}$ is a generalized quantifier, each $\bar{x}_i$ is a sequence of variables and $\varphi_i$ is a formula.

- $\langle T, \leq, g \rangle$ is a formula, given that

- $\langle T, \leq \rangle$ is a tree (as defined above),
- $g$ is a function on $T$ mapping any leaf to a formula and any non-leaf to $\neg$, $\vee$, $\wedge$, $\exists x$, $\forall x$, $\exists \xi$, $\forall \xi$ or $@\xi$ (for some individual variable $x$ and world variable $\xi$), and
- if $g(t) \notin \{\vee, \wedge\}$, then $t$ has a unique successor.

We define truth and falsity of a formula $\varphi$ of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ in a Kripke model $\mathfrak{M}$ and a world $w$ relative to an assignment function $a$. The conditions for atomic formulas are as above. A world variable $\xi$ is true in $w$ if and only if $w$ is assigned to it; it is false in $w$ if and only if another world is assigned to it; and so it is undefined if the assignment function is not defined on it. Similarly, the truth-conditions for a generalized quantifiers are standard, except that $\mathcal{Q}\bar{\bar{x}}\bar{\varphi}$ is undefined if for some $i$ and sequence $\bar{o}$ of elements in $D_w$, $\varphi_i$ is undefined in $\mathfrak{M}$, $w$ and $a[\bar{o}/\bar{x}_i]$. The definition of truth of a tree-based formula differs only in the construction of the relevant game. Given a formula $\varphi = \langle T, \leq, g \rangle$ of $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$, a Kripke model $\mathfrak{M}$, a world $w$ of $\mathfrak{M}$ and an assignment function $a$, this is defined as follows: A stage of the game is given by an element of $T$, a world (which might be undefined) and an assignment function. Plays are defined as above, with obvious minor amendments, as well as the following new rules:

- Stage 0 is given by the root node of $\varphi$, $w$ and $a$.
- If stage $\alpha$ is given by node $t$, world $v$ and assignment $b$, then:
  - If $t$ is labeled by $\exists x$, then $V$ chooses an element $o \in D_v$; stage $n+1$ is given by the successor of $t$, $v$ and $b[o/x]$.
  - If $t$ is labeled by $\forall x$, then $F$ chooses an element $o \in D_v$; stage $n+1$ is given by the successor of $t$, $v$ and $b[o/x]$.
  - If $t$ is labeled by $\exists \xi$, then $V$ chooses a world $u$; stage $n+1$ is given by the successor of $t$, $v$ and $b[u/\xi]$.
  - If $t$ is labeled by $\forall \xi$, then $F$ chooses a world $u$; stage $n+1$ is given by the successor of $t$, $v$ and $b[u/\xi]$.

  - If $t$ is labeled by $@\xi$, then stage $n+1$ is given by the successor of $t$, $b(\xi)$ and $b$.

- Stage $\lambda$ is given as above, except that the world of the limit stage is the world $v$ such that for some $\alpha < \lambda$, the world of stage $\beta$ is $v$ for all $\beta < \lambda$ such that $\alpha < \beta$, and is undefined if there is no such element.

This concludes the definition of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$. As before, we note that every sentence of $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ can be turned into an equivalent sentence of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$. As in the case of well-founded languages, we use $\vDash$ to express truth of a formula relative to the appropriate parameters, and now use $\Dashv$ similarly for falsity.

Our unusual way of setting up $\mathcal{N}_{\infty\infty}$ and $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ is mainly motivated by the inclusion of all generalized quantifiers in the latter language. While a number of generalized quantifiers can be given a natural game-theoretic semantics (see Pietarinen (2007)), to our knowledge, there is no general game-theoretic semantics for arbitrary generalized quantifiers. (Engström (2012) makes steps in this direction, but his constructions only apply to a limited class of generalized quantifiers and also require us to move from a two-player game to a game between teams of players.) We therefore have to combine a truth-conditional semantics for generalized quantifiers with a game-theoretic semantics for infinitary embeddings; the above hybrid provides a natural and general way of doing so.

Another feature of our presentation to note is the definition of limit stages of a game, which explicitly allows for variables not to be assigned an element – this can happen if there is an infinite sequence of quantifiers binding the same variables and the players keep choosing different elements for it. Such cases are usually implicitly assumed to be ruled out syntactically; Oikkonen (1979, p. 104, (iv)) does so explicitly. Ruling out these cases syntactically is unnatural in the setting of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ since the same problem occurs in the case of the world of evaluation for formulas with branches containing an infinite sequence of $\Diamond$s – of course, the formulas we are most interested in are exactly such formulas, so we cannot rule these out syntactically.

### 6.7.4 Definability in non-well-founded languages

To characterize which generalized quantifiers are modalizable in $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$, we adapt a number of ideas from Williamson (2010, Appendix 3). We start with some definitions, observations and lemmas. Define a function $\cdot^n$ for every natural number $n$ which maps every model $\mathfrak{A}$ to the Kripke model $\mathfrak{A}^n = \langle W, D, i, \emptyset\rangle$, where $W$ is the set of subsets of $|\mathfrak{A}|$ of cardinality $\leq n$, $D_w = w$ for all $w \in W$, and $i(R)(w) = R^{\mathfrak{A}} \cap w^n$ for all relation symbols $R$.

Note that for every generalized quantifier $\mathcal{Q}$ and $n < \omega$, there is a formula $\delta^n_{\mathcal{Q}}$ of $\mathcal{L}_{\omega\omega}$ which defines $\mathcal{Q}$ on models up to cardinality $n$; i.e., for every model $\mathfrak{A}$ of cardinality $\leq n$, $\mathfrak{A} \vDash \delta^n_{\mathcal{Q}}$ iff $\mathfrak{A} \vDash \gamma_{\mathcal{Q}}$. We write $\delta^n_{\mathcal{Q}}(\bar{\varphi})$ for the result of replacing the atomic formulas in $\delta^n_{\mathcal{Q}}$ by the formulas of $\bar{\varphi}$, leaving the appropriate replacements of variables implicit. Define the *relativization* of a generalized quantifier $Q$ of type $\bar{n}$ to be the generalized quantifier $Q^{\mathrm{rel}}$ of type $\langle 1, \bar{n}\rangle$ such that for all sets $D$ and $D' \subseteq D$ and sequence of relations $\bar{R}$ on $D$, $Q^{\mathrm{rel}}_D(D', \bar{R})$ if and only if $Q'_D(\bar{R}')$, where $\bar{R}'$ is the sequence of relations in $\bar{R}$ restricted to $D'$. See Westerståhl (2011, section 7) for a more precise definition.

For every $n < \omega$, we define a mapping $[\cdot]_n$ from $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ to $\mathcal{N}_{\infty\infty}$ by a number of replacements. For present purposes, we assume that all individual variables of $\mathcal{N}_{\infty\infty}$ are individual variables of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$, and that in addition, $\mathcal{N}_{\infty\infty}$ contains distinct individual variables $\xi_0, \ldots, \xi_{n-1}$ for each world variable $\xi$ of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$, as well as new individual variables $w_0, \ldots, w_{n-1}$. We replace

- $R\bar{x}$ by $R\bar{x} \wedge \bigwedge_{x\in\bar{x}} \bigvee_{i<n} x = w_i$
- $\xi$ by $\forall x \left(\bigvee_{i<n} x = \xi_i \leftrightarrow \bigvee_{i<n} x = w_i\right)$
- $\mathcal{Q}\bar{\bar{x}}\bar{\varphi}$ by $\delta^n_{\mathcal{Q}^{\mathrm{rel}}}\left(\bigvee_{i<n} x = w_i, \bar{\varphi}\right)$, where $x$ is not free in any $\bar{\varphi}$
- $\exists x\varphi$ by $\exists x \left(\bigvee_{i<n} x = w_i \wedge \varphi\right)$
- $\forall x\varphi$ by $\forall x \left(\bigvee_{i<n} x = w_i \rightarrow \varphi\right)$
- $\exists\xi\varphi$ by $\exists\bar{\xi}\varphi$
- $\forall\xi\varphi$ by $\forall\bar{\xi}\varphi$

- $@\xi\varphi$ by $\forall\bar{w}\left(\bigwedge_{i<n}\xi_i = w_i \rightarrow \varphi\right)$

For any sentence sentence $\varphi$ of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$, we define $[\varphi]_n$ to be $\forall\bar{w}\varphi'$, where $\varphi'$ is the result of carrying out the above replacements on $\varphi$. Note that we write $\mathfrak{M} \vDash \varphi/\mathfrak{M} \mathrel{=\!\!|} \varphi$ for $\varphi$ being true/false in every world of the Kripke model $\mathfrak{M}$.

**Lemma 6.7.3.** *For any $n < \omega$, sentence $\varphi$ of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ and model $\mathfrak{A}$, if all relation symbols occurring in $\varphi$ are of arity $\leq n$, then $\mathfrak{A}^n \vDash \varphi$ iff $\mathfrak{A} \vDash [\varphi]_n$, and $\mathfrak{A}^n \mathrel{=\!\!|} \varphi$ iff $\mathfrak{A} \mathrel{=\!\!|} [\varphi]_n$.*

*Proof.* By induction on the complexity of $\varphi$. □

**Lemma 6.7.4.** *Let $\mathfrak{A}$ be a model and $\mathcal{Q}$ a generalized quantifier of type $\bar{n}$. Then $\mathfrak{A} \vDash \gamma_{\mathcal{Q}}$ iff $\mathfrak{A}^{\max(\bar{n})} \vDash \gamma_{\mathcal{Q}^O}$, and $\mathfrak{A} \mathrel{=\!\!|} \gamma_{\mathcal{Q}}$ iff $\mathfrak{A}^{\max(\bar{n})} \mathrel{=\!\!|} \gamma_{\mathcal{Q}^O}$.*

*Proof.* By the construction of $\mathfrak{A}^{\max(\bar{n})}$. □

**Theorem 6.7.5.** *The modalization of a generalized quantifier $\mathcal{Q}$ is definable in $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ if and only if $\mathcal{Q}$ is definable in $\mathcal{N}_{\infty\infty}$.*

*Proof.* The right-to-left direction can be established as in the proof of Theorem 6.7.2. For the left-to-right direction, let $\mathcal{Q}$ be a generalized quantifier of type $\bar{n}$ whose modalization $\mathcal{Q}^O$ is definable in $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$. Then there is a formula $\varphi$ of $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$ such that for all models $\mathfrak{A}$, $\mathfrak{A}^{\max(\bar{n})} \vDash \gamma_{\mathcal{Q}^O}$ iff $\mathfrak{A}^{\max(\bar{n})} \vDash \varphi$, and $\mathfrak{A}^{\max(\bar{n})} \mathrel{=\!\!|} \gamma_{\mathcal{Q}^O}$ iff $\mathfrak{A}^{\max(\bar{n})} \mathrel{=\!\!|} \varphi$. By Lemmas 6.7.3 and 6.7.4, it follows that for all models $\mathfrak{A}$, $\mathfrak{A} \vDash \gamma_{\mathcal{Q}}$ iff $\mathfrak{A} \vDash [\varphi]_{\max(\bar{n})}$, and $\mathfrak{A} \mathrel{=\!\!|} \gamma_{\mathcal{Q}}$ iff $\mathfrak{A} \mathrel{=\!\!|} [\varphi]_{\max(\bar{n})}$. So $\mathcal{Q}$ is definable in $\mathcal{N}_{\infty\infty}$. □

### 6.7.5 Conclusions and Remarks

Using Theorems 6.7.1, 6.7.2 and 6.7.5, we can use standard results on the undefinability of generalized quantifiers in $\mathcal{L}_{\omega\omega}$, $\mathcal{L}_{\infty\omega}$ and $\mathcal{N}_{\infty\infty}$ to deduce the undefinability of the corresponding modalized generalized quantifiers in $\mathcal{L}^{\mathrm{mq}}_{\omega\omega}$, $\mathcal{L}^{\mathrm{mq}}_{\infty\infty}$ and $\mathcal{N}^{\mathrm{mq}}_{\infty\infty}$. In particular, we note the following, using open English sentences to denote the relevant generalized quantifiers:

- 'there are infinitely many . . . ' is not definable in $\mathcal{L}_{\omega\omega}$.

- 'there are uncountably many ...' is not definable in $\mathcal{L}_{\infty\omega}$.
- 'most ... are ...' is not definable in $\mathcal{N}_{\infty\infty}$.

The first follows from the compactness of $\mathcal{L}_{\omega\omega}$, the second from Fact 1.1.1 of Dickmann (1985, p. 318), and the last from a version of the downward Löwenheim-Skolem theorem for $\mathcal{N}_{\infty\infty}$ (Karttunen, 1983, p. 228, Theorem 2.1).

Let us also add a few remarks on the robustness of the above results: First, we impose a negative free logic in the model theory for our modal languages, requiring the interpretation of a relation at a world in a Kripke model to be restricted to the domain of that world. All of the results presented above go through as well if this assumption is dropped. Second, we used the relatively rich resources of world quantification rather than the more modest generalized actuality operators used in Williamson (2013, chapter 7). Since the latter are easily definable using the former, it is clear that our choice does not weaken the direction of our results which deduces the definability of $\mathcal{Q}$ from the definability of $\mathcal{Q}^O$ in the relevant languages. And in the mapping which establishes the reverse direction, we could have used the operators Williamson uses, so all of the results presented here are independent of this choice.

Finally, note that the preceding results are established using two distinct proof strategies; the proofs for well-founded languages employ the use of back-and-forth systems as in Fritz (2013), while the proofs for non-well-founded languages proceed roughly along the lines of Williamson (2010, Appendix 3). Call the latter the 'direct method' and the former the 'indirect method'. The direct method is applicable to well-founded languages as well, and might even provide simpler proofs. However, the indirect method gives us a way of constructing Kripke models which are equivalent up to some modal depth but disagree on the canonical sentence of a modalized quantifier; this makes it easy to see how the result can be strengthened by enriching the language under consideration. E.g., it is clear from this strategy that the result is not affected by expanding the language to include plural quantifiers. Plural quantifiers can also be accommodated on the direct method, but how to do so is is less obvious. Beside simplicity, the direct method has the advantage of not requiring us to characterize the equivalence of models up to a given quantifier rank in terms of a back-and-forth system. This need not be an obstacle to applying it to $\mathcal{N}_{\infty\infty}$, as back-and-forth systems have been developed for non-well-founded languages; see Rantala (1979) and Karttunen (1979).

# Chapter 7

# Higher-Order Contingentism, Part 3: Expressive Limitations

**Abstract.** Two expressive limitations of an infinitary higher-order modal language interpreted on models for higher-order contingentism are established: First, the inexpressibility of certain relations, which leads to the fact that certain model-theoretic existence conditions for relations cannot equivalently be reformulated in terms of being expressible in such a language. Second, the inexpressibility of certain modalized cardinality claims, which shows that in such a language, higher-order contingentists cannot express what we communicate by saying that there are uncountably many possible stars.

## 7.1 Introduction

This paper is a continuation of Part 1 (Fritz and Goodman, unpublished c, here ch. 1), familiarity with which is assumed. Familiarity with Part 2 (Fritz, unpublished a, here ch. 3) is not required. In Part 1, several versions of the Fine-Stalnaker view of higher-order contingentism are explored. Specifically, these are understood to be theories of the contingent existence of relations which are hereditarily intensional, a qualification which will mostly be left tacit in the following. In this third part, we are concerned with two questions which arise from expressive limitations of the infinitary higher-order modal language introduced in Part 1. The first is the question how to state the Fine-Stalnaker view; here we continue the discussion of this issue started in Part 1. The second is the question what claims about possible individuals can be expressed if the Fine-Stalnaker view is correct; the formal results on this issue we will establish here serve as central premises in a philosophical critique of higher-order contingentism in Fritz and Goodman (unpublished b, here ch. 6).

### 7.1.1 Stating the Views

The investigation of theories of higher-order contingency in Part 1 starts with two variants of the Fine-Stalnaker view, called the *higher-order closure view* and the *qualitative generation view*. Formally, the class of closed models $C\times$ is developed, where $\times$ is a parameter which indicates whether a positive or negative semantics is used. It is argued that this does not capture the views expressed in the philosophical writings that serve as its motivation. In the case of closure, this is spelled out formally by showing that the comprehension principle $\times\text{Comp}_{FS}$ is not valid on $C\times$. The model theory is restricted accordingly, but the resulting classes of models turn out to be highly restrictive, ruling out systems that are paradigmatic instances of the informal picture-thinking that motivated the views.

In order to keep the original model theory, one might reject the way the basic idea underlying the Fine-Stalnaker view was spelled out in terms of automorphisms. Instead, one might propose to spell it out by formulating the existence condition for relations linguistically. The higher-order closure view would then be cashed out as saying that necessarily, a relation exists if it is expressible in principle using only existing parameters, and the qualitative generation view as saying that necessarily, a relation exists if and only if it is expressible in principle using only generating parameters. (Here, an existing/generating parameter is an expression which is interpreted as a relation which, in the world in question, exists/is among the choice of relations from which the higher-order domains are generated.)

While the notion of expressibility in principle – given certain parameters – is somewhat unclear, it is not uncommon to find philosophers appeal to it. E.g., as observed in Stalnaker (2012, p. 61), it is used in explications of the notion of qualitativeness of relations in Adams (1979, p. 7) and Lewis (1986, p. 221). Concerning the present issue of formulating a theory of higher-order contingentism, the idea is discussed in Fine (1977b, section V). Adams, Lewis and Fine all seem to suggest that expressibility in principle can be understood as expressibility in a sufficiently rich language, but only Fine is more specific about what such a language might look like. Fine in fact specifies an infinitary language very similar to the one we have been working with in Part 1, and *proves* that his semantic criterion of the existence of relations in terms of automorphisms coincides with the linguistic criterion of being expressible in his infinitary language using only generating parameters. Fine's result is therefore a promising sign for formulating the higher-order closure and qualitative generation views linguistically.

However, as will be shown here, Fine's result essentially depends on a questionable resource to which he avails himself, namely a primitive infinitary "outer" first-order quantifier, which he writes $\exists\Diamond$. Understood primitively, such a quantifier is highly suspect from the point of view of a contingentist. Fine's explanation of it on p. 161 suggests that it can be understood as an infinite sequence of possibility operators and existential quantifiers, and this understanding is supported further by the fact that in other writings on the subject, Fine explicitly appeals to such embeddings; see Fine (1977a) and Fine (2003). Infinitary embeddings of this kind are not allowed in the infinitary language we are working with presently. The first main result to be established here is that without these resources of infinitary outer quantification or infinitary embeddings, the analogs to Fine's result fail: In the case of closure, we can show that there are models in which necessarily, every relation expressible in our language using existing parameters exists but which are not closed. We establish this in section 7.2. It would be relatively straightforward to adapt this result to the case of qualitative generation, using the extensions of syntax and semantics sketched in Part 1, section 1.6.5. For simplicity, we restrict ourselves to the case of closure in the following. We return to a discussion of languages with infinitary embeddings in section 7.4.1.

The condition on models that necessarily, every relation expressible using existing parameters exists is equivalent to validating $\times\text{Comp}_C$, as we prove in Proposition 7.2.2. Since we know that $\times\text{Comp}_C$ is valid on $C\times$, we can restate the result to be proven as saying that $\times\text{Comp}_C$ does not define the class $C\times$, in the sense that it is not the case for every model that it validates $\times\text{Comp}_C$ if and only if it is in $C\times$; this results also holds when we restrict ourselves to world-selective models. In fact, the way the result is proven establishes something stronger, namely that *no* class of sentences defines $C\times$ (Corollary 7.2.15).

These results show that in formulating the higher-order closure and qualitative generation views of higher-order contingency, one cannot simply assume that being expressible in principle can be cashed out as being expressible in a particular language which provides the required infinitary resources; this depends on subtle issues concerning which infinitary resources are available. Of course, this does not mean that $\times\text{Comp}_{FS}$ (or its analog for generation), which commits one to the more restrictive model theory of internally closed or internally generated models, is the only way of formulating a theory of higher-order contingency in the vicinity of the Fine-Stalnaker view. E.g., one might hold the view that necessarily, a relation exists just in case all possible individuals it is about exist.[1] However, it should be noted that such a view is a significant departure from the guiding ideas behind the Fine-

[1] This was suggested by Kit Fine (pc). As he noted, this proposal also brings the existence condition for relations more in line with that of the extensional entities treated in Fine (1977b), but omitted in the present type hierarchy.

Stalnaker view, and in particular from the reductive ambitions which seem to lie at the heart of Fine (1977b).

### 7.1.2 Paraphrase

In Fritz and Goodman (unpublished b, here ch. 6), it is argued that we can make sense of claims in which we seem to be quantifying over merely possible individuals, such as the claim that there are possible buildings which have never and will never be built. Such claims are trivially false according to the contingentist's metaphysics, so it is argued that they must provide a paraphrase. In the case just mentioned, this is easily done by saying that there could have been possible buildings which actually have never and will never be built. It is shown there that using even highly infinitary first-order resources, analogous modalized cardinality claims such as the claim that there are uncountably many possible stars cannot be paraphrased. Whether such claims can be paraphrased using infinitary *higher-order* resources is the second issue of expressivity of this part, which is the topic of section 7.3.

We will prove a positive and a negative expressivity result. The positive result is that assuming first-order contingentism but higher-order necessitism, we can paraphrase any claim formulated with modalizations of generalized quantifiers using the corresponding unmodalized quantifiers. The negative result is that using either the class of closed or internally closed models (whether positive or negative), the claim that there are at least $\kappa$ many possible individuals, for a given uncountable cardinality $\kappa$, is inexpressible. Again, the results can easily be adapted to the cases of generation and internal generation, but we focus on closure for simplicity. On the basis of these results, we argue against higher-order contingentism in Fritz and Goodman (unpublished b, here ch. 6).

Before proceeding to proving these limitative results in sections 7.2 and 7.3, we define the main tool for doing so, namely back and forth systems. In section 7.4, we consider some possible extensions of the formal object language with which we are working and discuss how likely they are to overcome the expressive limitations discussed here.

### 7.1.3 Back and Forth Systems

The central tool in proving the limitative results of this paper are back and forth systems. Although their definition is somewhat complex, they are straightforward extensions of well-known definitions; see Fritz (2013, section 2.2) for references. It may be of historical interest that extensions of such systems to higher-order logics were in fact defined relatively early; they go back at least to Fraïssé (1958).

Recall that for a function $f$ from a set $A$ to a set $B$, we write $\mathrm{dom}(f)$ for the domain of $f$. We now also write $\mathrm{im}(f)$ for the *image of* $f$, the set $\{y \in B : f(x) = y \text{ for some } x \in A\}$. When convenient, we consider functions as functional relations. For the rest of the paper, we will tacitly assume a choice of a signature $\sigma$ and sign $\times$.

**Definition 7.1.1.** *Let* $\mathfrak{M} = \langle W, I, D, V, w\rangle$ *and* $\mathfrak{M}' = \langle W', I', D', V', w'\rangle$ *be models. A* partial isomorphism from $\mathfrak{M}$ to $\mathfrak{M}'$ *is a tuple* $\langle \tau, \rho\rangle$ *such that*

- $\tau$ *is a partial injection from* $W$ *to* $W'$
- $\rho$ *is a function on types mapping each type* $t$ *to a bijection* $\rho^t$ *from* $D^t_{\mathrm{dom}(\tau)}$ *to* $D'^t_{\mathrm{im}(\tau)}$ *such that for all* $v \in \mathrm{dom}(\tau)$*:*
  - *for all types* $t$, $\rho^t|D^t_v$ *is a bijection from* $D^t_v$ *to* $D'^t_{\tau(v)}$
  - *for all types* $\bar{t}$, *for all* $o \in D^t_{\mathrm{dom}(\tau)}$ *and* $\bar{o} \in \Pi_{i\leq n} D^{t_i}_{\mathrm{dom}(\tau)}$, $\bar{o} \in o(v)$ *iff* $\langle \rho^{t_i}(o_i) : i \leq n\rangle \in \rho^t(o)(\tau(v))$
- $\tau(w) = w'$
- *for all types* $t$ *and* $a \in \sigma(t)$, $\rho^t V(a) = V'(a)$

*Let a* back and forth system from $\mathfrak{M}$ to $\mathfrak{M}'$ *be a non-empty set* $J$ *of partial isomorphisms from* $\mathfrak{M}$ *to* $\mathfrak{M}'$ *such that for all* $\langle \tau, \rho\rangle \in J$*:*

- *For all* $v \in W$*, there is a* $\langle \tau', \rho'\rangle \in J$ *such that* $\tau \subseteq \tau'$ *and* $\rho^t \subseteq \rho'^t$ *for all types* $t$*, and* $v \in \mathrm{dom}(\tau')$*.*
- *For all* $v' \in W'$*, there is a* $\langle \tau', \rho'\rangle \in J$ *such that* $\tau \subseteq \tau'$ *and* $\rho^t \subseteq \rho'^t$ *for all types* $t$*, and* $v' \in \mathrm{im}(\tau')$*.*

*We write* $J : \mathfrak{M} \cong^\infty \mathfrak{M}'$ *for* $J$ *being a back and forth system from* $\mathfrak{M}$ *to* $\mathfrak{M}'$ *and* $\mathfrak{M} \cong^\infty \mathfrak{M}'$ *for there being a* $J$ *such that* $J : \mathfrak{M} \cong^\infty \mathfrak{M}'$.

**Proposition 7.1.2.** *For any models* $\mathfrak{M}$ *and* $\mathfrak{M}'$ *of the same signature,*

*if* $\mathfrak{M} \cong^\infty \mathfrak{M}'$ *then* $\mathfrak{M} \equiv \mathfrak{M}'$.

*Proof.* By induction on the complexity of formulas. □

## 7.2 Expressing Relations

We start by making the claim to be proven precise. Focussing on the case of closure, we want to show that there are models in which the semantic criterion of closure comes apart from the syntactic criterion of necessarily containing every relation expressible using existing parameters. We therefore formally define the latter criterion, calling it expressible closure, using the notion of a formula expressing a relation introduced in Part 1, Definition 1.5.3. (Although we won't consider them, there are alternative conceptions of expressibility; see Fine (1977b, pp. 162–163). The option used here is both natural from a conceptual point of view and useful from a technical point of view.)

**Definition 7.2.1.** *Let* $\mathfrak{M} = \langle W, I, D, V, w\rangle$ *a model and* $\bar{t}$ *a sequence of types.*

*Define any* $o \in \iota^{\bar{t}}_{\langle W,I\rangle}$ *to be* $\times$expressible in $v \in W$ *if there is a formula* $\varphi$ *of* $\mathcal{L}(\emptyset)$*, sequence of variables* $\bar{x}$ *of types* $\bar{t}$ *and assignment* $a$ *for* $\langle W, I\rangle$ *admissible for* $\varphi$ *such that* $\mathrm{im}(a) \subseteq D^T_v$ *and* $o = \varphi(\bar{x})^\times_{\mathfrak{M},a}$.

$\mathfrak{M}$ *is* expressibly $\times$closed *if for all* $v \in W$, $D^T_v$ *contains all* $o \in \iota^T_{\langle W,I\rangle}$ *which are* $\times$*expressible in* $v$.

As noted above, being expressibly closed is equivalent to verifying $\mathrm{Comp}_C$:

**Proposition 7.2.2.** *A model* $\mathfrak{M}$ *is expressibly* $\times$*closed if and only if* $\mathfrak{M} \vDash \times\mathrm{Comp}_C$.

*Proof.* Immediate. □

Since by Part 1, Proposition 1.5.6, every $\times$closed model verifies $\times\mathrm{Comp}_C$, it follows that every $\times$closed model is expressively $\times$closed. We will therefore have to show that some expressively $\times$closed model is not closed. To do so, we consider a model in which there is a class of worlds which share a certain individual, and a class of worlds which share a different individual. One of the two classes of worlds is countably infinite and the other uncountably infinite, so there is no automorphism which maps one to the other. Therefore closure forces the two propositions corresponding to the two classes of worlds to exist at the distinguished world. However, by letting the individual domain of this world be empty, we can construct the model in such a way that these propositions are not expressible at that world, which gives us an expressibly closed model which is not closed.

The difficult part of this proof is specifying the higher-order domains of the model in a way which guarantees that the model is expressibly closed. To do so, we first define a similar model in which the two classes of worlds – distinguished by the individuals the respective worlds share – are both countably infinite; to ensure that this model is closed, we construct it by generation. We then project the higher-order domains of this model onto the frame of the model to be constructed using a technique I call *projective generation*, which we now develop. While this is a highly specific technique, we will be able to apply it again in section 7.3. This construction gives us a back and forth system from the original model to the one which is projectively generated. Since the original model is generated, it is closed, and therefore verifies $\times\mathrm{Comp}_C$. So by the equivalence of models related by a back and forth system, the projectively generated model also verifies $\times\mathrm{Comp}_C$, and so by Proposition 7.2.2 is expressibly closed. That it is not closed can easily be established by showing that the two propositions indicated earlier are not in the higher-order domain of the distinguished world.

### 7.2.1 Projective Generation

Projective generation deals only with models which are determined by their worlds and the distribution of individuals at the worlds. To make this precise, we first single out the structures in which all higher-order domains are empty and in which worlds can be distinguished by the individuals they contain:

**Definition 7.2.3.** *Let an* individual structure*, in short* IS*, be a structure* $\mathfrak{S} = \langle W, I, D\rangle$ *such that*

*(i) for all types* $t \neq e$*,* $D^t_W = \emptyset$*, and*

*(ii) for all* $w \in W$*,* $w = D^e_w$*.*

In the context of ISs, we can reduce both automorphisms of a structure and partial isomorphisms, as used in back and forth systems between models, to partial injections between individuals. We single out the relevant partial injections in the following definition:

**Definition 7.2.4.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *and* $\mathfrak{S}' = \langle W', I', D'\rangle$ *be* IS*s and* $f$ *a partial injection from* $I$ *to* $I'$*.*

- $f$ respects worlds *if for all* $X \subseteq \mathrm{dom}(f)$*,* $X \in W$ *iff* $\{f(x) : x \in X\} \in W'$*.*
- *If* $f$ *respects worlds, let* $\dot{f}$ *be the partial function from* $W$ *to* $W'$ *mapping each* $w \in W$ *such that* $w \subseteq \mathrm{dom}(f)$ *to* $\{f(x) : x \in w\}$*, and* $\hat{f} = \langle \dot{f}, f\rangle$*. Note that* $\dot{f}$ *is injective.*

This definition is slightly sloppy as the relativity to the structures is not noted, but context will make this clear in all applications below. We can now show how instead of an automorphism of a structure consisting of a permutation of worlds and a permutation of individuals, we can use a permutation of individuals alone. To state this, recall that if a group $G$ acts on some set $X$ and $x \in X$, we write $G_x$ for the stabilizer subgroup of $x$, the set of $g \in G$ which map $x$ to itself. We now extend this notion to sets: for any $Y \subseteq X$, we write $G_{(Y)}$ for the *point-wise stabilizer* subgroup of $Y$, the set of $g \in G$ which map each $x \in Y$ to itself.

**Definition 7.2.5.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be an* IS *and* $w \in W$*. Define:*

- $\mathrm{aut}^i(\mathfrak{S}) = \{f \in S_I : f$ *respect worlds*$\}$

- $\text{fix}^i(\mathfrak{S}, w) = \text{aut}^i(\mathfrak{S})_{(w)}$

**Lemma 7.2.6.** *For any* IS $\mathfrak{S} = \langle W, I, D\rangle$ *and* $w \in W$, $\text{aut}(\mathfrak{S}) = \{\hat{f} : f \in \text{aut}^i(\mathfrak{S})\}$ *and* $\text{fix}(\mathfrak{S}, w) = \{\hat{f} : f \in \text{fix}^i(\mathfrak{S}, w)\}$.

*Proof.* Immediate. □

For the definition of projective generation, we further single out a class of special ISs, inspired by the notion of homogeneity in model theory, as defined, e.g., in Hodges (1997, p. 160). To define it, we understand a *partial permutation of a set* $X$ to be a partial injection from $X$ to $X$. It should be noted that the present notion of homogeneity is unrelated to the notion of homogeneity discussed in Fine (1977b, p. 150); a notion similar to Fine's notion of homogeneity is discussed in section 7.3.2 under the label of being "fully symmetric".

**Definition 7.2.7.** *An* IS $\mathfrak{S} = \langle W, I, D\rangle$ *is* homogeneous *if each finite partial permutation of* $I$ *which respects worlds can be extended to a permutation of* $I$ *which respects worlds, i.e., an element of* $\text{aut}^i(\mathfrak{S})$.

Projective generation will allow us to project the higher-order domains of a structure generated from a homogeneous IS onto an IS. This will be done relative to a construction connecting the first to the second IS; this construction we call a projection. Projections can be thought of as back and forth systems from the first to the second IS which satisfy certain additional conditions. In the following definition, condition (i) and (ii) correspond to the conditions on back and forth systems; condition (iii) encodes the idea that the projection maps every world of the first IS to a world of the second IS and condition (iv) to the idea that a projection must cohere with the automorphisms of the first structure. For condition (iv), we adopt the convention of composing partial functions and total functions just as relations in general; i.e., $fg = \{\langle x, z\rangle : \text{there is a } y \text{ such that } \langle x, y\rangle \in g \text{ and } \langle y, z\rangle \in f\}$.

**Definition 7.2.8.** *Let* $\mathfrak{S} = \langle W, I, B\rangle$ *be an homogeneous* IS, $\mathfrak{S}' = \langle W', I', B'\rangle$ *an* IS. *A* projection from $\mathfrak{S}$ to $\mathfrak{S}'$ *is a set* $P$ *of finite partial injections from* $I$ *to* $I'$ *which respect worlds such that for all* $p \in P$*:*

*(i) For all $w \in W$, there is an $q \in P$ such that $p \subseteq q$ and $w \subseteq \mathrm{dom}(q)$.*

*(ii) For all $w' \in W'$, there is an $q \in P$ such that $p \subseteq q$ and $w' \subseteq \mathrm{im}(q)$.*

*(iii) For all $w \in \mathrm{dom}(\dot{p})$, $p|w \in P$.*

*(iv) For all $f \in \mathrm{aut}^i(\mathfrak{S})$, $pf \in P$.*

*We call the members of $P$* perspectives.

For the rest of this section, assume that $\mathfrak{S}$, $\mathfrak{S}'$ and $P$ are given as in Definition 7.2.8. Let $\otimes\mathfrak{S} = \langle W, I, D\rangle$, $\mathfrak{F} = \langle W, I\rangle$ and $\mathfrak{F}' = \langle W', I'\rangle$. Using $P$, we will project a domain assignment from $\otimes\mathfrak{S}$ onto $\mathfrak{F}'$; this will coincide with $\mathfrak{S}'$ for individuals, so we can also think of $P$ as projecting the higher-order domains of $\otimes\mathfrak{S}$ onto $\mathfrak{S}'$.

For this definition, we simultaneously define, first, a relation $Z$, parametric to types and perspectives, from $D$ to intensions on $\mathfrak{F}'$ and, second, a domain assignment $D^P$ on $F'$. The idea behind this definition is that we can extend any perspective $p$ to a connection between intensions corresponding to relations among individuals: An intension $o$ in the domain of $\otimes\mathfrak{S}$ is connected to an intension $o'$ on $\mathfrak{F}'$ just in case for every perspective $q$ extending $p$, $o$ and $o'$ agree on individuals related by $q$ in worlds related by $q$. We can then use this extension of $P$ to define the domains of intensions corresponding to relations among individuals on $\mathfrak{F}'$, including in the domain of a given world $v$ all intensions in the image of each of the extensions of a perspective restricted to the domain of the world which it maps to $v$. Iterating this procedure along the type hierarchy gives us the definitions of $Z$ and $D^P$. To state it, we adopt the convention of writing, for a binary relation $R \subseteq X \times Y$ and $Z \subseteq X$, $R[Z]$ for the image of $Z$ under $R$, i.e., the set $\{y : \langle x, y\rangle \in R \text{ for some } x \in Z\}$.

**Definition 7.2.9.** *We define a relation $Z^t_p \subseteq D^t_{\mathrm{dom}(\dot{p})} \times \iota^t_{\mathfrak{F}'}$ for each type $t$ and $p \in P$ and a domain assignment $D^P$ for $\mathfrak{F}'$ by simultaneous induction on types:*

$t = e$: *For all $p \in P$, $Z^e_p = p|D^e_{\mathrm{dom}(\dot{p})}$.*

*For all $v \in W'$, ${D^P}^e_v = B'^e_v$.*

$t = \bar{t}$: *For all* $p \in P$, $o \in D^t_{\mathrm{dom}(\dot{p})}$ *and* $o' \in \iota^t_{\mathfrak{F}'}$, $oZ^t_p o'$ *iff*

(1) $D^P \boxtimes o'$ *and*

(2) *for all* $q \in P$ *such that* $p \subseteq q$, $w \in \mathrm{dom}(\dot{q})$ *and* $n$*-tuples* $\bar{o}, \bar{o}'$ *such that* $o_i Z^{t_i}_q o'_i$ *for all* $i \leq n$, $\bar{o} \in o(w)$ *iff* $\bar{o}' \in o'(\dot{q}(w))$.

*For all* $v \in W'$, $D^{P^t}_v = \bigcup\{Z^t_p[D^t_{\dot{p}^{-1}(v)}] : p \in P$ *and* $v \in \mathrm{im}(\dot{p})\}$.

*For brevity, we write* $\bar{o}Z_p\bar{o}'$ *for the claim that* $o_i Z^{t_i}_p o'_i$ *for all* $i \leq n$. *Let* the structure projectively $\times$generated by $P$ *be* $\mathfrak{S}^P = \langle W', I', D^P \rangle$.

This is well-defined since $D^P$ only needs to be defined for lower types to evaluate whether $D^P \boxtimes o'$.

We now need to show that $Z$ is well-behaved. The main claim to be established is that all relations $Z^t_p$ are bijections, which we prove in the next lemma. It turns out to be convenient to prove also, by simultaneous induction, that if two perspectives agree on the individuals of the generating structure, their extensions agree on the higher-order domains of that world as well.

**Lemma 7.2.10.** *For all types* $t$ *and* $p \in P$*:*

*(i)* $Z^t_p$ *is a bijection from* $D^t_{\mathrm{dom}(\dot{p})}$ *to* $(D^P)^t_{\mathrm{im}(\dot{p})}$.

*(ii) For all* $q \in P$, $w \in W$ *such that* $w \subseteq \mathrm{dom}(p \cap q)$ *and* $o \in D^t_w$, $Z^t_p(o) = Z^t_q(o)$.

Since the proof of this lemma is somewhat involved, it is given in Appendix 7.5. We draw out the following immediate consequence, which will be useful several times below:

**Lemma 7.2.11.** *For any type* $t$, $p \in P$ *and* $w \in \mathrm{dom}(\dot{p})$, $Z^t_p|D^t_w = Z^t_{p|w}$.

*Proof.* By condition (iii) of the definition of projections, $p|w \in P$. If $\langle o, o' \rangle \in Z^t_{p|w}$, then by Lemma 7.2.10 (i), there is a $v \in \mathrm{dom}(\dot{p})$ such that $v \in W$ and $o \in D^t_v$. Since $v \subseteq w$, $\mathrm{fix}(\mathfrak{S}, w) \subseteq \mathrm{fix}(\mathfrak{S}, v)$, so $D^t_v \subseteq D^t_w$, and thus $o \in D^t_w$. By Lemma 7.2.10 (ii), $Z^t_p(o) = Z^t_{p|w}(o)$, so $\langle o, o' \rangle \in Z^t_p|D^t_w$. If $\langle o, o' \rangle \in Z^t_p|D^t_w$, then $o \in D^t_w$, so by Lemma 7.2.10 (i), $o \in \mathrm{dom}(Z^t_{p|w})$, and so by Lemma 7.2.10 (ii), $Z^t_p(o) = Z^t_{p|w}(o)$, and therefore $\langle o, o' \rangle \in Z^t_{p|w}$. □

Finally, we extend our construction from structures to models. To do so, consider a model $\mathfrak{M} = \langle W, I, D, V, w\rangle$ on $\otimes\mathfrak{S}$. The following definition extends the generation of $\mathfrak{S}^P$ to a model, relative to a perspective, and defines the corresponding back and forth system between the two models:

**Definition 7.2.12.** *For any $p \in P$ such that $w \subseteq \mathrm{dom}(p)$, let* the model projectively $\times$generated from $\mathfrak{M}$ by $P$ and $p$ *be* $\mathfrak{M}_p^P = \langle W', I', D^P, V_p^P, \dot{p}(w)\rangle$, *where for all $t \in T$ and $a \in \tau(t)$, $V_p^P(a) = Z_p^t V(a)$.*

*For every $q \in P$, let $Z_q$ be the function on types mapping each type $t$ to $Z_q^t$. Let $J_p^P = \{\langle \dot{q}, Z_q\rangle : q \in P$ such that $p \subseteq q\}$.*

By construction, $\mathfrak{M}_p^P$ is a $\times$model.

**Theorem 7.2.13.** $J_p^P : \mathfrak{M} \cong^\infty \mathfrak{M}_p^P$.

*Proof.* Since $P$ is non-empty, so is $J_p^P$. Consider any $q \in P$; we first show that $\langle \dot{q}, Z_q\rangle$ is a partial isomorphism from $\mathfrak{M}$ to $\mathfrak{M}_p^P$. Since $q$ respects worlds, $\dot{q}$ is a partial injection from $W$ to $W'$. Let $t$ be a type. By Lemma 7.2.10, $Z_q^t$ is a bijection from $D^t_{\mathrm{dom}(\dot{q})}$ to $D'^t_{\mathrm{im}(\dot{q})}$. Consider any $v \in \mathrm{dom}(\dot{q})$. By Lemmas 7.2.10 and 7.2.11, $Z_q^t | D_v^t$ is a bijection from $D_v^t$ to $D'^t_{\dot{q}(v)}$. So let $t = \bar{t}$ be a type, $o \in D^t_{\mathrm{dom}(\dot{q})}$ and $\bar{o} \in \Pi_{i \leq n} D^{t_i}_{\mathrm{dom}(\dot{q})}$. Using $Z_q^t$ both as a relation and function, note that trivially, $\bar{o} Z_q \langle Z_q^{t_i}(o_i) : i \leq n\rangle$, so by construction of $Z$, $\bar{o} \in o(v)$ iff $\langle Z_q^{t_i}(o_i) : i \leq n\rangle \in Z_q^t(o)(\dot{q}(v))$, as required. The last two conditions required for $\langle \dot{q}, Z_q\rangle$ being a partial isomorphism from $\mathfrak{M}$ to $\mathfrak{M}_p^P$ likewise follow straightforwardly from the construction of $\mathfrak{M}_p^P$ and $J_p^P$. By conditions (i) and (ii) of the definition of projections and Lemma 7.2.10 (ii), we can establish the required closure conditions on $J_p^P$. □

### 7.2.2 Inexpressible Relations

**Theorem 7.2.14.** *There is a world-selective $\times$model which is expressibly $\times$closed but not $\times$closed.*

*Proof.* For any ordinal $\alpha > \omega$, let

$$W_2^\alpha = \{\emptyset, \{0, \beta\}, \{\omega, \gamma\} : 0 < \beta < \omega < \gamma < \alpha\},$$

and let $\mathfrak{S}_2^\alpha$ be the unique IS determined by $W_2^\alpha$ and $\alpha$. Let $P$ be the set of partial injections $p$ from $\omega 2$ to $\omega_1$ with finite domains satisfying the following two conditions:

- For all $\beta \in \mathrm{dom}(p)$, $\beta \in \{0, \omega\}$ iff $p(\beta) \in \{0, \omega\}$.
- For all $\beta, \gamma \in \mathrm{dom}(p)$, ($\beta < \omega$ iff $p(\beta) < \omega$) iff ($\gamma < \omega$ iff $p(\gamma) < \omega$).

It is routine to verify that $P$ is a projection from $\mathfrak{S}_2^{\omega 2}$ to $\mathfrak{S}_2^{\omega_1}$. Let $\mathfrak{M}$ be a model based on $\otimes\mathfrak{S}_2^{\omega 2}$, and $\mathfrak{M}_\emptyset^P$ the model projectively $\times$generated from $\mathfrak{M}$ by $P$ and $\emptyset$.

To see that $\mathfrak{M}_\emptyset^P$ is world-selective, consider any $v \in W_2^{\omega_1}$. By conditions (ii) and (iii) of the definition of projections, there is a $p \in P$ such that $\mathrm{im}(p) = v$. By the construction of $\mathfrak{S}^P$, it suffices to show that $\mathrm{dom}(p)^{\langle\rangle}_{\langle W_2^{\omega 2}, \omega 2\rangle} Z_p^{\langle\rangle} v^{\langle\rangle}_{\langle W_2^{\omega_1}, \omega_1\rangle}$, which is routine. As noted above, $\mathfrak{M}_\emptyset^P$ is a $\times$model by construction.

We show that $\mathfrak{M}_\emptyset^P$ is expressibly $\times$closed: Since $\mathfrak{M}$ is $\times$closed, $\mathfrak{M} \vDash \times\mathrm{Comp}_C$, so by Theorem 7.2.13 and Proposition 7.1.2, $\mathfrak{M}_\emptyset^P \vDash \times\mathrm{Comp}_C$. That $\mathfrak{M}_\emptyset^P$ is expressively $\times$closed follows with Proposition 7.2.2.

Finally, we show that $\mathfrak{M}_\emptyset^P$ is not $\times$closed. Let $D$ and $D^P$ be the domain assignments of $\mathfrak{M}$ and $\mathfrak{M}_\emptyset^P$, respectively. Assume for contradiction that $\mathfrak{M}_\emptyset^P$ is closed. Then $o' = \{0, \beta : 0 < \beta < \omega\}^{\langle\rangle}_{\langle W_2^{\omega_1}, \omega_1\rangle} \in D^P{}^{\langle\rangle}_\emptyset$. Hence by construction of $S^P$, there must be a $p \in P$ and $o \in D_\emptyset^{\langle\rangle}$ such that $oZ_p^{\langle\rangle}o'$. With condition (iii) of the definition of projections and Lemma 7.2.10, it follows that $oZ_\emptyset^{\langle\rangle}o'$. But this conflicts with condition (2) of the definition of $Z$. $\lightning$. □

In this proof, we make use of two models verifying $\times\mathrm{Comp}_C$, only one of which is $\times$closed. It follows that $\times\mathrm{Comp}_C$ does not define $C\times$. Moreover, the two models satisfy the same sentences, so *no* class of sentences defines $C\times$, and since both models are world-selective, we obtain the following corollary:

**Corollary 7.2.15.** *$C\times$ is undefinable relative to the class of world-selective models. I.e., there is no class of sentences $\Gamma$ such that a world-selective model $\mathfrak{M}$ is in $C\times$ if and only if $\mathfrak{M} \vDash \Gamma$.*

## 7.3 Expressing Modalized Cardinality Claims

We now consider what can be expressed given various theories of higher-order contingentism. As argued in some detail in Fritz and Goodman (unpublished b, here ch. 6, section 6.5), the relevant distinctions which are in need of being expressed relative to a certain theory of higher-order contingtism can be identified, at least for present purposes, with the classes of models in the model theory developed for the relevant theory. E.g., assuming that higher-order necessitists use the class of models in which necessarily, all relations exist which are compatible with the being constraint or its positive weakening, we can understand the claim that there are uncountably many possible individuals as the class of such models in which the union of individual domains of all worlds is uncountable. We understand this claim to be expressible, given higher-order contingentism, if there is a sentence which is true in such a model if and only if it belongs to this class.

### 7.3.1 Expressivity via Haecceities

We start by defining the class of models for higher-order necessitism, calling them *full*:

**Definition 7.3.1.** *A structure* $\langle W, I, D\rangle$ *is* $\times$full *if for all* $w \in W$*, types* $t \neq e$ *and* $o \in \iota^t_{\langle W,I\rangle}$*:*

$$o \in D^t_w \textit{ iff } D \boxtimes o.$$

*A model is* $\times$full *if it is based on a full structure.*

Recall the definition of existential and universal outer quantifiers in Part 1, section 1.5.3. Analogous to these defined outer quantifiers we can consider primitive generalized quantifiers which operate on the outer domain of models. We will show that on full models, such generalized quantifiers can always be eliminated in favour of the corresponding (inner) generalized quantifiers over properties, restricted to certain haecceities. Thus we first define – assuming fullness – what it is to be a haecceity, what it is for a property of haecceities to contain a unique haecceity for every possible individual (in which case we think of it as a choice of representation), and what it is for a relation among haecceities to relate exactly those haecceities of a given choice of representations which single out individuals which satisfy a given open formula.

$$\textsc{h}(X^{\langle e\rangle}, x^e) := \Box\forall y^e\Box(Xy \leftrightarrow (x = y \wedge \exists z^e(z = x)))$$

$$\textsc{ch}(X^{\langle\langle e\rangle\rangle}) := \forall Y^{\langle e\rangle}(XY \rightarrow \Sigma x^e\textsc{h}(Y, x)) \wedge \Pi x^e\exists Y^{\langle e\rangle}(XY \wedge \textsc{h}(Y, x))$$

$$\textsc{hmap}(X^{\langle\langle e\rangle\rangle}, Y^{\langle e\rangle^n}, \varphi, \bar{x}^e) := \forall\bar{Z}^{\langle e\rangle}\Pi\bar{x}^e(\textstyle\bigwedge_{i<n}(XZ_i \wedge \textsc{h}(Z_i, x_i)) \rightarrow (Y\bar{Z} \leftrightarrow \varphi))$$

Now consider the extension of $\mathcal{L}$ obtained by adding, for every generalized quantifier $Q$, the quantifier $Q^o$, operating on first-order variables, and the quantifier $Q^{\langle e\rangle}$, operating on variables of type $\langle e\rangle$. We interpret $Q^o$ on the outer domain (see Fritz (2013) for details) and $Q^{\langle e\rangle}$ on the properties of the world of evaluation. We can eliminate every occurrence of $Q^o$ in favour of an occurrence of $(Q^{\text{rel}})^{\langle e\rangle}$, where $Q^{\text{rel}}$ is the relativization of $Q$ (again, see Fritz (2013) for details) as follows:

$$\tau(Q^o\bar{\bar{x}}^e\bar{\varphi}) := \forall X^{\langle\langle e\rangle\rangle}Y_1^{\langle e\rangle^{k_1}}\ldots Y_n^{\langle e\rangle^{k_n}}((\textsc{ch}(X) \wedge \textstyle\bigwedge_{i<n}\textsc{hmap}(X, Y_i, \varphi_i, \bar{x}_i)) \rightarrow (Q^{\text{rel}})^{\langle e\rangle}Z^{\langle e\rangle}\bar{\bar{Z}}^{\langle e\rangle}(XZ, Y_1\bar{Z}_1, \ldots, Y_n\bar{Z}_n))$$

In particular, if $Q$ is definable in infinitary higher-order logic, we can completely eliminate the use of generalized quantifiers.

### 7.3.2 FFISs and Bi-Projections

To show that an analogous result cannot be proven on the classes of models for the various versions of the higher-order closure view of higher-order contingentism we have been exploring, we refine the technique of projective generation to allow us to construct back and forth systems between closed models. We start by imposing even stronger constraints on IS, forcing their worlds to contain only finitely many individuals and them being fully symmetric:

**Definition 7.3.2.** *Let $\mathfrak{S} = \langle W, I, D\rangle$ be an* IS.

- $\mathfrak{S}$ *is* finitary *if for all* $w \in W$, $D^e_w$ *is finite.*
- $\mathfrak{S}$ *is* fully symmetric *if* $\mathrm{aut}^i(\mathfrak{S}) = S_I$.

*We write* FFIS *to abbreviate* finitary fully symmetric individual structure.

Note that every fully symmetric IS is trivially homogeneous. One important feature of FFISs is that for purposes of generation, we can restrict our attention to permutations with finite support, as we now show. (Recall that the support of a permutation is the set of elements it does not map to itself.)

**Definition 7.3.3.** *Let* $\mathfrak{S} = \langle W, I, D\rangle$ *be an* IS *and* $w \in W$. *Define:*

- $\mathrm{aut}^i_\omega(\mathfrak{S}) = \{f \in \mathrm{aut}^i(\mathfrak{S}) : \mathrm{supp}(f)$ *is finite*$\}$
- $\mathrm{fix}^i_\omega(\mathfrak{S}, w) = \{f \in \mathrm{fix}^i(\mathfrak{S}, w) : \mathrm{supp}(f)$ *is finite*$\}$

**Lemma 7.3.4.** *Let* $\times$ *be a sign,* $\mathfrak{S}$ *a* FFIS *and* $\otimes\mathfrak{S} = \langle W, I, D\rangle$. *Then for all types* $t \neq e$, $w \in W$ *and* $o \in \iota^t_{\langle W,I\rangle}$, $o \in D^t_w$ *iff* $D \boxtimes o$ *and* $\hat{f}.o = o$ *for all* $f \in \mathrm{fix}^i_\omega(\mathfrak{S}, w)$.

*Proof.* The left-to-right direction is immediate. For the right-to-left direction, assume $o \notin D^t_w$ and $D \boxtimes o$. Then there is an $f \in \mathrm{fix}^i(\mathfrak{S}, w)$ and $\hat{f}.o \neq o$. So there are $v \in W$, $\bar{v} \in W^n$ and $\bar{o} \in \Pi_{i \leq n} D^{t_i}_{v_i}$ such that not $\hat{f}.\bar{o} \in o(\hat{f}.v)$ iff $\bar{o} \in o(v)$. Let $X = \bigcup\{v, v_0, \ldots, v_{n-1}, w\}$. Since $\mathfrak{S}$ is a FFIS, there is a $g \in \mathrm{fix}^i_\omega(\mathfrak{S}, w)$ such that $f|X = g|X$ and $\mathrm{supp}(g)$ is finite. Then $\hat{g}.v = \hat{f}.v$; similarly $\hat{g}.\bar{v} = \hat{f}.\bar{v}$ from which it is straightforward to derive that $\hat{g}.\bar{o} = \hat{f}.\bar{o}$. So not $\hat{g}.\bar{o} \in o(\hat{g}.v)$ iff $\bar{o} \in o(v)$, and thus $\hat{g}.o \neq o$. □

To be able to use projections to construct back and forth systems between closed models, we show that if a projection from one FFIS to another is such that its inverse (the set of inverses of its perspectives) is a projection as well, then the structure it projectively generates is the structure generated by the second FFIS:

**Definition 7.3.5.** *A projection* $P$ *from an* IS $\mathfrak{S}$ *to an* IS $\mathfrak{S}'$ *is a* bi-projection *if* $P^{-1} = \{p^{-1} : p \in P\}$ *is a projection from* $\mathfrak{S}'$ *to* $\mathfrak{S}$.

**Theorem 7.3.6.** *If $\mathfrak{S}$ and $\mathfrak{S}'$ are* FFIS*s and $P$ is a bi-projection from $\mathfrak{S}$ to $\mathfrak{S}'$, then $\otimes\mathfrak{S}'$ is the structure projectively generated by $P$.*

As it is somewhat complex, the proof is given in Appendix 7.6. The condition of being a FFIS used here is of course extremely strong. This is mainly since it is simple and the structures we will be looking at satisfy it, but the results of this section could very likely be proven for much weaker assumptions.

### 7.3.3 Inexpressibility of Cardinality Claims

To show that various claims concerning the cardinality of possible individuals are not expressible over closed and internally closed structures, we define, for every infinite cardinality $\kappa$, a highly symmetric structure with $\kappa$ individuals in the outer domain. To do so, we take a set of cardinality $\kappa$ and define an IS on it, taking its finite subsets as the worlds. These are clearly FFIS, and the structures they generated turn out to be internally closed. Moreover, any two such FFIS can be related by a bi-projection, and we can connect models based on the structures they generate by back and forth system, which allows us to show that no distinctions among infinite cardinalities of possible individuals can be drawn on internally closed, and so in particular also on closed, models. To define these models, we write $X \subseteq_\omega Y$ for $X$ being a finite subset of $Y$.

**Definition 7.3.7.** *For any set $X$, let $W^X_\omega = \{Y : Y \subseteq_\omega X\}$, $\mathfrak{F}^X_\omega = \langle W^X_\omega, X\rangle$ and $\mathfrak{S}^X_\omega$ the unique* IS *determined by $W^X_\omega$ and $X$. Let $\mathfrak{M}^X_\omega$ be the model for the empty signature based on $\otimes\mathfrak{S}^X_\omega$ with distinguished world $\emptyset$.*

**Lemma 7.3.8.** *For any infinite set $X$, $\otimes\mathfrak{S}^X_\omega$ is internally $\times$closed.*

*Proof.* Let $\mathfrak{S}^X_\omega = \langle W^X_\omega, X, B\rangle$ and $\otimes\mathfrak{S}^X_\omega = \langle W^X_\omega, X, D\rangle$. Consider any $w \in W^X_\omega$, type $t \neq e$ and $o \in \iota^t_{F^X_\omega}$. By Part 1, Proposition 1.6.11, it suffices to show that $o \in D^t_w$ iff $D \boxtimes o$ and $\xi.o = o$ for all $\xi \in \mathrm{fix}(\otimes\mathfrak{S}^X_\omega, w)|_c \otimes \mathfrak{S}^X_\omega$. Since $\otimes\mathfrak{S}^X_\omega$ is a $\times$structure, if $o \in D^t_w$ then $D \boxtimes o$, so the left to right direction is immediate. So assume that $o \notin D^t_w$ and $D \boxtimes o$. By Lemma 7.3.4, there is an $f \in \mathrm{fix}^i_\omega(\mathfrak{S}^X_\omega, w)$ such that $\hat{f}.o \neq o$. It only remains to show that $\hat{f} \in \mathrm{fix}(\otimes\mathfrak{S}^X_\omega, w)|_c \otimes \mathfrak{S}^X_\omega$.

Since $f \in \mathrm{fix}^i_\omega(\mathfrak{S}^X_\omega, w)$, $\hat{f} \in \mathrm{fix}(\mathfrak{S}^X_\omega, w)$, and so by Part 1, Lemma 1.4.13 (ii), $\hat{f} \in \mathrm{fix}(\otimes\mathfrak{S}^X_\omega, w)$. Also, since $f \in \mathrm{fix}^i_\omega(\mathfrak{S}^X_\omega, w)$, $\mathrm{supp}(f)$ is finite, and therefore $\mathrm{supp}(f) \in W^X_\omega$. Define $f_c \in \iota^{\langle e,e\rangle}_{\mathfrak{F}^X_\omega}$ and $\dot{f}_c \in \iota^{\langle\langle\rangle,\langle\rangle\rangle}_{\mathfrak{F}^X_\omega}$ such that for all $v \in W^X_\omega$:

$$f_c(v) = \begin{cases} \{\langle x, f(x)\rangle : x \in v\} & \text{if } \mathrm{supp}(f) \subseteq v \\ \emptyset & \text{otherwise} \end{cases}$$

$$\dot{f}_c(v) = \begin{cases} \left\{\left\langle u^{\langle\rangle}_{\mathfrak{F}^X_\omega}, \dot{f}(u)^{\langle\rangle}_{\mathfrak{F}^X_\omega}\right\rangle : u \subseteq v\right\} & \text{if } \mathrm{supp}(f) \subseteq v \\ \emptyset & \text{otherwise} \end{cases}$$

Using the fact that $g.f = f$ for any $g \in \mathrm{fix}^i_\omega(\mathfrak{S}^X_\omega, \mathrm{supp}(f))$, it is routine to show that $f_c \in D^{\langle e,e\rangle}_{\mathrm{supp}(f)}$ and $\dot{f}_c \in D^{\langle\langle\rangle,\langle\rangle\rangle}_{\mathrm{supp}(f)}$. □

**Theorem 7.3.9.** *For any infinite sets $X$ and $Y$, $\mathfrak{M}^X_\omega \cong^\omega \mathfrak{M}^Y_\omega$.*

*Proof.* Let $P$ be the set of partial injections from $X$ to $Y$ with finite domains. It is straightforward to see that $P$ is a bi-projection from $\mathfrak{S}^X_\omega$ to $\mathfrak{S}^Y_\omega$. So by Theorem 7.2.13, $J^P_\emptyset : \mathfrak{M}^X_\omega \cong^\omega \mathfrak{M}^P_\emptyset$, where $J^P_\emptyset$ and $\mathfrak{M}^P_\emptyset$ are defined as above. By Theorem 7.3.6, $\mathfrak{M}^P_\emptyset$ is based on $\otimes\mathfrak{S}^P_\omega$, so $\mathfrak{M}^P_\emptyset = \mathfrak{M}^Y_\omega$, hence $J^P_\emptyset : \mathfrak{M}^X_\omega \cong^\omega \mathfrak{M}^Y_\omega$. □

**Theorem 7.3.10.** *For any uncountable cardinality $\kappa$, there is no set of sentences $\Gamma$ such that for every internally $\times$closed model $\mathfrak{M} = \langle W, I, D, V, w\rangle$, $\mathfrak{M} \models \Gamma$ if and only if $\left|\bigcup_{w\in W} D_w\right| = \kappa$.*

*Proof.* Using Lemma 7.3.8, Theorem 7.3.9 and Proposition 7.1.2. □

## 7.4 Extending the Language

In response to the expressive limitations seen above, one might suggest to enrich the language with which we are working here. One natural suggestion would be to lift the restriction of higher-order quantifiers to hereditarily intensional relations. We won't consider this in the following, simply because the behaviour of higher-order modal logic is completely unclear without this restriction unless one makes the controversial assumption that the restriction is vacuous, in which case it is uninteresting to lift it. Similarly, we won't consider adding hyperintensional operators as logical constants. See Fritz and Goodman (unpublished b, here ch. 6, section 6.3) for arguments that such resources are unpromising to deal with the problem of expressing modalized cardinality claims. Instead, we consider two kinds of infinitary resources.

### 7.4.1 Non-well-founded Languages

Although the language $\mathcal{L}$ used here allows conjunctions of infinite sets of formulas and quantifiers binding infinite sets of variables, it is defined in the usual recursive manner. Consequently, although a node in a syntax tree of one of its formulas may have infinitely many immediate successors, this tree may not contain a path, following the successor relation, of infinite length. That is, such a tree may be infinitely branching, but it may not have infinite branches. This is equivalent to the condition that the subformula relation among formulas is well-founded, and we therefore call $\mathcal{L}$ a well-founded language.

Since in such languages, formulas cannot contain infinite branches, every subformula is in the scope of only a finite number of modal operators. Therefore, in evaluating a sentence, a subformula will only be evaluated relative to an assignment which maps its free variables to parameters from a finite number of worlds. In this sense, it is impossible to say anything in these languages which requires comparing parameters from an infinite number of worlds; it is exactly this feature which both of the limitative results proven here exploit. In fact, it would be possible to give an abstract characterization of the class of well-founded languages in which $\Box$ is the only non-extensional operator, and to prove that both of the limitative theorems hold for any such language.

It is therefore natural to consider non-well-founded languages. As we saw above, this is exactly what Fine does in appealing to languages in which infinite embeddings of operators are allowed. To evaluate the use of such languages, we have to consider two questions: first, whether they are in good standing, and second, whether they overcome the expressive limitations discussed here.

While Fine does not provide a formal syntax or a model-theoretic semantics for such a language, this is done in Leuenberger (2006) and Fritz and Goodman (unpublished b, here ch. 6), building on Hintikka and Rantala (1976). Formally, these languages are therefore in good standing. But they may still be philosophically problematic, and arguments to this effect are given in Williamson (2013, section 7).

We turn to the question whether non-well-founded languages overcome the expressive limitations discussed here. Concerning the first issue of expressing the higher-order closure and qualitative generation view, we know from the results in Fine (1977b) that allowing infinite embeddings of operators suffices to establish the equivalence of the semantic and the linguistic existence criteria in Fine's formal setting. Given the details of his proof, it is not to be expected that the formal differences between this and the present setting will affect the result, and it is therefore to be expected that the analog of Theorem 7.2.14 or a similar result for generation do not hold for a extension of the language by infinitary embeddings. However, such a positive result would crucially rely on the fact that since models are based on sets, there are guaranteed to be formulas containing as many variables as individuals, which is arguably an artifact of the model theory (see the discussion in Part 1, section 1.2.2). It is therefore doubtful whether the non-well-founded languages considered here address the first expressive limitation in a satisfactory way.

Consider now the case of paraphrasing claims formulated using modalized generalized quantifiers. As shown in Fritz and Goodman (unpublished b, here ch. 6), infinitary embeddings suffice to paraphrase any claim formulated using modalized cardinality quantifiers (i.e., "there are $\kappa$ many possible $\varphi$s", for any cardinality $\kappa$), using only first-order quantifiers. But it is also shown there that in a first-order language with such infinitary resources, claims of the form "most possible $\varphi$s are possible $\psi$s" cannot be paraphrased. Can this be done if we add higher-order quantifiers to such a language and interpret it over (internally) closed structures?

I conjecture that this is not the case: Even though for every cardinality

$\kappa$, there are sentences in such a language which "collect and compare" possible individuals from $\kappa$ worlds, every particular sentence of such a language is itself a member of the set-theoretic hierarchy, and so can only "collect and compare" possible individuals from collections of worlds up to some particular cardinality. For a generalized quantifier like "most", no particular cardinality suffices, and we can always find countermodels to a proposed paraphrase using sufficiently large models. While I won't attempt to do so here, I conjecture that this rough idea can be turned into a rigorous proof by combining the model-construction idea of section 7.3.3 with the syntactic approach of Fritz and Goodman (unpublished b, here ch. 6, appendix 6.7.4). If this is right, then even on set-sized models, which introduce the arguably unrealistic artifact that formulas can contain as many variables as there are individuals, infinitary embeddings don't suffice to paraphrase "most possible $\varphi$s are possible $\psi$s".

### 7.4.2 Transfinite Types

As noted in Fine (1977b, p. 144), there are two dimensions along which the finitary type hierarchy explored here can be extended to admit transfinite types. On the one hand, we might extend the recursive definition of types by allowing relational types of arbitrary arity. On the other hand, we might admit cumulative types by adding in the recursive definition of types that each set of types $T'$ is a type as well, understanding an entity to be of type $T'$ just in case it is of a type which is a member of $T'$. Conceptually, the former extension is much more natural, and we therefore concentrate on it in the following. It is important to note that by extending the type hierarchy, we are not only extending the formal language, but also the semantics structures on which they are interpreted.

Working in a positive setting, consider the following further extension of the language: For every sequence of types $\bar{t}$, we add a logical constant $\approx$ of type $\langle \bar{t}, \bar{t} \rangle$, which expresses hereditary intensional equivalence, and a $\lambda$ operator binding a sequence of variables of types $\bar{t}$, where $\lambda\bar{v}\varphi$ is an expression of type $\bar{t}$ interpreted as the intension expressed by $\varphi$, abstracted over $\bar{v}$.

With these additional resources we can define Fine's infinitary outer quantifier, since the universal outer quantification $\Pi\bar{v}\varphi$ is true if and only if $\varphi$ and $\top$, abstracted over $\bar{v}$, are hereditarily intensionally equivalent. So we can define:

$$\Pi\bar{v}\varphi := \lambda\bar{v}\varphi \approx \lambda\bar{v}\top$$

This gives us a way of expressing infinitary outer quantification without appealing to the potentially problematic resources of non-well-founded languages discussed above. But as pointed out there, it is not clear that this will overcome either of the expressive limitations discussed here, likely leaving the higher-order contingentist unable to express their own view by cashing out talk of expressibility in principle by appealing to a particular infinitary language, as well as unable to express claims about most possible individuals.

## 7.5 Appendix on Projective Generation

In this appendix, we prove Lemma 7.2.10. To do so, we first establish two subsidiary lemmas.

**Lemma 7.5.1.** *Let $\mathfrak{S} = \langle W, I, D\rangle$ be an homogeneous* IS*, $\mathfrak{S}' = \langle W', I', D'\rangle$ an* IS *and $P$ a projection from $\mathfrak{S}$ to $\mathfrak{S}'$. For all $p, q \in P$, there are $p' \in P$ and $f \in \mathrm{aut}^i(\mathfrak{S})_{(\mathrm{dom}(p\cap q))}$ such that $p \subseteq p'$ and $q \subseteq p'f$.*

*Proof.* By condition (ii) of projections, there is a $p' \in P$ such that $p \subseteq p'$ and $\mathrm{im}(q) \subseteq \mathrm{im}(p')$. Let $f$ be the partial function from $I$ to $I$ mapping every $x \in \mathrm{dom}(q)$ to $p'^{-1}q(x)$. It is routine to show that $f$ is a finite partial permutation of $I$ which respects worlds. Since $\mathfrak{S}$ is homogeneous, there is an $f' \in \mathrm{aut}^i(\mathfrak{S})$ such that $f \subseteq f'$. It is routine to show that $f' \in \mathrm{aut}^i(\mathfrak{S})_{(\mathrm{dom}(p\cap q))}$ and $q \subseteq p'f'$. □

**Lemma 7.5.2.** *For any type $t$, $o \in D^t_W$, $o' \in \iota^t_{F'}$, $p \in P$ and $f \in \mathrm{aut}^i(\mathfrak{S})$, if $oZ^t_p o'$ then $\hat{f}.oZ^t_{pf^{-1}}o'$.*

*Proof.* By induction on types. For $t = e$, note that if $oZ^e_p o'$, then $p(o) = o'$, so $pf^{-1}f(o) = o$, whence $\hat{f}.oZ^e_{pf^{-1}}o'$. Let $t = \bar{t}$. To show that $\hat{f}.oZ^t_{pf^{-1}}o'$, note that condition (1) follows from $oZ^t_p o'$. For condition (2), consider any $q \in P$ such that $pf^{-1} \subseteq q$, $v \in \mathrm{dom}(\dot{q})$, $\bar{v} \in \mathrm{dom}(\dot{q})^n$, $\bar{o} \in \Pi_{i\leq n} D^{t_i}_{v_i}$, and $\bar{o}' \in \Pi_{i\leq n} \iota^{t_i}_{\langle W',I'\rangle}$ such that $\bar{o}Z_q\bar{o}'$. We show that $\bar{o} \in \hat{f}.o(v)$ iff $\bar{o}' \in o'(\dot{q}(v))$. By induction hypothesis, it follows from $\bar{o}Z_q\bar{o}'$ that $\hat{f}^{-1}.\bar{o}Z_{qf}\bar{o}'$. Note also that $p \subseteq qf$ (for any $o \in \mathrm{dom}(p)$, $p(o) = pf^{-1}f(o) = qf(o)$). Thus it follows from $oZ^t_p o'$ that $\hat{f}^{-1}.\bar{o} \in o(\hat{f}^{-1}.v)$ iff $\bar{o}' \in o'(\dot{q}(v))$. Since $\bar{o} \in \hat{f}.o(v)$ iff $\hat{f}^{-1}.\bar{o} \in o(\hat{f}^{-1}.v)$, the desired equivalence follows. □

We are now able to prove Lemma 7.2.10, which claims that for all types $t$ and $p \in P$:

(i) $Z^t_p$ is a bijection from $D^t_{\mathrm{dom}(\dot{p})}$ to $(D^P)^t_{\mathrm{im}(\dot{p})}$.

(ii) For all $q \in P$, $w \in W$ such that $w \subseteq \mathrm{dom}(p \cap q)$ and $o \in D^t_w$, $Z^t_p(o) = Z^t_q(o)$.

*Proof of Lemma 7.2.10.* By induction on types. The case of $t = e$ is trivial for both (i) and (ii). So consider any type $t = \bar{t}$ and $p \in P$. We start with (i):

*Claim 1:* $Z^t_p$ is functional. *Proof.* Consider any $o \in D^t_{\mathrm{dom}(\dot{p})}$ and $o', o'' \in (D^P)^t_{\mathrm{im}(\dot{p})}$ such that $oZ^t_p o'$ and $oZ^t_p o''$. Let $v \in W'$; we show that $o'(v) \subseteq o''(v)$ (the other direction follows by symmetry). So consider any $\bar{o}' \in o'(v)$. Since $D^P \boxtimes o'$, there are $\bar{v} \in W'^n$ such that $\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_{v_i}$. By condition (ii) of the definition of projections, there is a $q \in P$ such that $p \subseteq q$ and $v, v_1, \ldots, v_n \in \mathrm{im}(\dot{q})$. By induction hypothesis (i), there are $\bar{o} \in \Pi_{i\leq n} D^{t_i}_{\mathrm{dom}(\dot{q})}$ such that $\bar{o}Z_q\bar{o}'$. So it follows from $oZ^t_p o'$ that $\bar{o} \in o(\dot{q}^{-1}(v))$, and therefore with $oZ^t_p o''$ that $\bar{o}' \in o''(v)$.✓

*Claim 2:* $Z^t_p$ is total. *Proof.* Consider any $o \in D^t_{\mathrm{dom}(\dot{p})}$. Define $o' \in \iota^t_{\mathfrak{F}'}$ such that for all $v \in W'$, $o'(v)$ is the set of $\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_{W'}$ (if $\times = +$) / $\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_v$ (if $\times = -$) such that there is a $q \in P$ such that $p \subseteq q$ and $v \subseteq \mathrm{im}(q)$, and $n$-tuple $\bar{o}$ such that $\bar{o}Z_q\bar{o}'$ and $\bar{o} \in o(\dot{q}^{-1}(v'))$. We show that $oZ^t_p o'$. $D^P \boxtimes o'$ is immediate by construction. So consider any $q \in P$ such that $p \subseteq q$, $w \in \mathrm{dom}(\dot{q})$ and $n$-tuples $\bar{o}, \bar{o}'$ such that $\bar{o}Z_q\bar{o}'$. We prove that $\bar{o} \in o(w)$ iff $\bar{o}' \in o'(\dot{q}(w))$. By construction of $o'$, the latter follows from the former. So assume that $\bar{o}' \in o'(\dot{q}(w))$. Then by construction of $o'$, there is an $r \in P$ such that $p \subseteq r$ and $\dot{q}(w) \subseteq \mathrm{im}(r)$, and $n$-tuple $\bar{o}^*$ such that $\bar{o}^*Z_r\bar{o}'$

and $\bar{o}^* \in o(\dot{r}^{-1}\dot{q}(w))$. By Lemma 7.5.1, there are $q', r' \in P$ such that $q \subseteq q'$, $r \subseteq r'$ and an $f \in \mathrm{aut}^i(\mathfrak{S})_{(\mathrm{dom}(p))}$ such that $q'f = r'$. By induction hypothesis (ii), $\bar{o}Z_{q'}\bar{o}'$ and $\bar{o}^*Z_{r'}\bar{o}'$. Thus $\bar{o}^*Z_{q'f}\bar{o}'$, and so with Lemma 7.5.2, $\hat{f}.\bar{o}^*Z_{q'}\bar{o}'$. By induction hypothesis (i), $Z^{t_i}_{q'}$ is a bijection, for each $i \leq n$, so $\hat{f}.\bar{o}^* = \bar{o}$. Thus from the fact that $\bar{o}^* \in o(\dot{r}^{-1}\dot{q}(w))$, we obtain $\bar{o} \in \hat{f}.o(\hat{f}.\dot{r}^{-1}\dot{q}(w))$. Since $f \in \mathrm{aut}^i(\mathfrak{S})_{(\mathrm{dom}(p))}$ and $o \in D^t_{\mathrm{dom}(\dot{p})}$, $\hat{f}.o = o$, and so $\bar{o} \in o(\hat{f}.\dot{r}^{-1}\dot{q}(w))$, from which $\bar{o} \in o(w)$ follows with $q'f = r'$ as required. ✓

*Claim 3:* $Z^t_p$ is injective. *Proof.* Consider any $o, o^* \in D^t_{\mathrm{dom}(\dot{p})}$ and $o' \in (D^P)^t_{\mathrm{im}(\dot{p})}$ such that $oZ^t_po'$ and $o^*Z^t_po'$. Let $w \in W$; we show that $o(w) \subseteq o^*(w)$ (the other direction follows by symmetry). So consider any $\bar{o} \in o(w)$. By condition (i) of the definition of projections, there is a $q \in P$ such that $p \subseteq q$, $w \in \mathrm{dom}(\dot{q})$ and $\bar{o} \in \Pi_{i\leq n}D^{t_i}_{\mathrm{dom}(\dot{q})}$. So by induction hypothesis (i), there are $\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_{\mathrm{im}(\dot{q})}$ such that $\bar{o}' \in o'(\dot{q}(w))$, and so $\bar{o} \in o^*(w)$. ✓

*Claim 4:* $Z^t_p$ is surjective. *Proof.* Consider any $o' \in (D^P)^t_{\mathrm{im}(\dot{p})}$. By construction of $D^P$, there is a $v \in \mathrm{im}(\dot{p})$ and $q \in P$ such that $v \subseteq \mathrm{im}(q)$ and $o' \in Z^t_q[D^t_{\dot{q}^{-1}(v)}]$. Hence there is an $o \in D^t_{\dot{q}^{-1}(v)}$ such that $oZ^t_qo'$. By Lemma 7.5.1, there are $p', q' \in P$ such that $p \subseteq p'$, $q \subseteq q'$, and an $f \in \mathrm{aut}^i(\mathfrak{S})$ such that $p'f = q'$. So by Lemma 7.5.2, $\hat{f}.oZ^t_{q'f^{-1}}o'$, hence $\hat{f}.oZ^t_{p'}o'$. Since $o \in D^t_{\dot{q}^{-1}(v)}$, $\hat{f}.o \in D^t_{\hat{f}.q^{-1}(v)} = D^t_{p^{-1}(v)}$. So $\hat{f}.o \in \mathrm{dom}(Z^t_p)$, and thus with induction hypothesis (ii), $\hat{f}.oZ^t_po'$. Hence $o' \in \mathrm{im}(Z^t_p)$ as required.

(ii): Since we have established (i), we can rely on the fact that $Z^t_p$ is a bijection from $D^t_{\mathrm{dom}(\dot{p})}$ to $(D^P)^t_{\mathrm{im}(\dot{p})}$, for all $o \in P$. Note first that it is routine to show that for all $p, q \in P$, if $p \subseteq q$ and $Z^t_p \subseteq Z^t_q$. Now consider any $p, q \in P$, $w \in W$ such that $w \subseteq \mathrm{dom}(p \cap q)$ and $o \in D^t_w$; we show that $Z^t_p(o) = Z^t_q(o)$. By Lemma 7.5.1, there are $p', q' \in P$ and $f \in \mathrm{aut}^i(\mathfrak{S})_{(w)}$ such that $p \subseteq p'$, $q \subseteq q'$ and $p'f = q'$. Since $Z_p \subseteq Z_{p'}$, $Z^t_p(o) = Z^t_{p'}(o)$, which by Lemma 7.5.2 is $Z^t_{p'f}(\hat{f}^{-1}.o)$. Since $o \in D^t_w$, $\hat{f}^{-1}.o = o$, so $Z^t_p(o) = Z^t_{p'f}(o)$, which is $Z^t_{q'}(o)$. With the fact that $Z^t_q \subseteq Z^t_{q'}$, it follows that $Z^t_p(o) = Z^t_q(o)$. □

## 7.6 Appendix on FFISs and Bi-Projections

In this appendix, we prove Theorem 7.3.6. In the following, we assume that $\mathfrak{S} = \langle W, I, B\rangle$ and $\mathfrak{S}' = \langle W', I', B'\rangle$ are FFIS, and that $P$ is a bi-projection from $\mathfrak{S}$ to $\mathfrak{S}'$. Let $D$ and $D'$ be the domain assignments of $\otimes\mathfrak{S}$ and $\otimes\mathfrak{S}'$, $Z$ the extension of $P$ as defined above, and $D^{P^{-1}}$ and $D^P$ the domain assignments of the structures projectively generated by $P^{-1}$ and $P$, respectively.

**Definition 7.6.1.** *For any $p \in P$, define a relation $Z^p_p \subseteq \mathrm{aut}^i_\omega(\mathfrak{S}) \times \mathrm{aut}^i_\omega(\mathfrak{S}')$ such that for all $f \in \mathrm{aut}^i_\omega(\mathfrak{S})$ and $g \in \mathrm{aut}^i_\omega(\mathfrak{S}')$, $fZ^p_p g$ iff*

$\mathrm{supp}(f) \subseteq \mathrm{dom}(p)$,

$\mathrm{supp}(g) \subseteq \mathrm{im}(p)$, *and*

$pf(o) = gp(o)$ *for all* $o \in \mathrm{supp}(f)$.

**Lemma 7.6.2.** *For any $p \in P$, $w \in \mathrm{dom}(\dot{p})$ and $f \in \mathrm{fix}^i_\omega(\mathfrak{S}', \dot{p}(w))$ such that $\mathrm{supp}(f) \subseteq \mathrm{im}(p)$, there is a $g \in \mathrm{fix}^i_\omega(\mathfrak{S}, w)$ such that $gZ^p_p f$.*

*Proof.* Define $g : I \to I$ such that for all $o \in I$,

$$g(o) = \begin{cases} p^{-1}fp(o) & \text{if } o \in \mathrm{dom}(p) \\ o & \text{otherwise} \end{cases}$$

It is routine to check that $g \in \mathrm{fix}^i_\omega(\mathfrak{S}, w)$ and $gZ^p_p f$. □

**Lemma 7.6.3.** *Let $p, q \in P$, $f \in \mathrm{aut}^i_\omega(\mathfrak{S})$ and $g \in \mathrm{aut}^i_\omega(\mathfrak{S}')$. If $fZ^p_p g$ and $p \subseteq q$ then $f^{-1}Z^p_q g^{-1}$.*

*Proof.* Routine. □

**Lemma 7.6.4.** *For all types $t$:*

*(i)* *For all $p \in P$, $f \in \mathrm{aut}^i_\omega(\mathfrak{S})$, $g \in \mathrm{aut}^i_\omega(\mathfrak{S}')$ such that $fZ^p_p g$ and $\langle o, o'\rangle \in Z^t_p$, $\hat{f}.oZ^t_p\hat{g}.o'$.*

*(ii)* *For all $p \in P$, $(Z^t_p)^{-1} = Z^t_{p^{-1}}$.*

*(iii)* *For all $v \in W'$, $D'^t_v = (D^P)^t_v$.*

*Proof.* By induction on types. Let $t = e$. (i): Let $p \in P$, $f \in \mathrm{aut}^i_\omega(\mathfrak{S})$, $g \in \mathrm{aut}^i_\omega(\mathfrak{S}')$ such that $fZ^p_p g$ and $\langle o, o'\rangle \in Z^e_p$. Then $o \in \mathrm{dom}(p)$, so $p(o) = o'$. Since $fZ^p_p g$, $pf(o) = gp(o)$, so $pf(o) = g(o')$, i.e., $\hat{f}.oZ^e_p\hat{g}.o'$. (ii) and (iii) are immediate. So let $t = \bar{t}$.

(i): Consider any $p \in P$, $f \in \mathrm{aut}^i_\omega(\mathfrak{S})$, $g \in \mathrm{aut}^i_\omega(\mathfrak{S}')$ such that $fZ^p_p g$ and $\langle o, o'\rangle \in Z^t_p$. Then $\hat{f}.o \in D^t_{\hat{f}.\mathrm{dom}(\dot{p})}$. Since $\mathrm{supp}(f) \subseteq \mathrm{dom}(p)$, $\hat{f}.\mathrm{dom}(\dot{p}) = \mathrm{dom}(\dot{p})$, so $\hat{f}.o \in D^t_{\mathrm{dom}(\dot{p})}$. We show $\hat{f}.oZ^t_p\hat{g}.o'$ by checking conditions (1) and (2) of the construction of $Z$.

(1): Consider any $v \in W'$ and $\bar{o}' \in \hat{g}.o'(v)$. Then $\hat{g}^{-1}.\bar{o}' \in o'(\hat{g}^{-1}.v)$. Since $oZ^t_p o'$, $D^P \boxtimes o'$, so there are $\bar{v} \in W'^n$ such that $\hat{g}^{-1}.\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_{v_i}$. $g \in \mathrm{aut}^i_\omega(\mathfrak{S}')$, so by Part 1, Lemma 1.4.13 (i), $\hat{g} \in \mathrm{aut}(\otimes\mathfrak{S}')$. Since by IH (iii), $(D^P)^{t_i}_{v_i} = D'^{t_i}_{v_i}$ for all $i \leq n$, it follows that $\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_{\hat{g}.v_i}$. If $\times = -$, we can assume $v_i = v$ for all $i < n$, therefore $\bar{o}' \in \Pi_{i\leq n}(D^P)^{t_i}_{\hat{g}.v}$. Thus $D^P \boxtimes \bar{o}'$.

(2): Consider any $q \in P$ such that $p \subseteq q$, $w \in \mathrm{dom}(\dot{q})$ and $n$-tuples $\bar{o}, \bar{o}'$ such that $\bar{o}Z_q\bar{o}'$. We prove that $\bar{o} \in \hat{f}.o(w)$ iff $\bar{o}' \in \hat{g}.o'(\dot{q}(w))$. By Lemma 7.6.3, $f^{-1}Z^p_q g^{-1}$, so by IH (i), $\hat{f}^{-1}.\bar{o}Z_q\hat{g}^{-1}.\bar{o}'$. Since $\mathrm{supp}(f) \subseteq \mathrm{dom}(\dot{q})$, $\hat{f}^{-1}.w \in \mathrm{dom}(\dot{q})$, so $\hat{f}^{-1}.\bar{o} \in o(\hat{f}^{-1}.w)$ iff $\hat{g}^{-1}.\bar{o}' \in o'(\dot{q}(\hat{f}^{-1}.w))$, and so $\bar{o} \in \hat{f}.o(w)$ iff $\bar{o}' \in \hat{g}.o'(\hat{g}.\dot{q}(\hat{f}^{-1}.w))$. $qf = gq$, so $\hat{g}.\dot{q}(\hat{f}^{-1}.w) = \dot{q}(w)$, from which the desired claim follows.

(ii): Let $p \in P$. By symmetry, it suffices to show that if $\langle o, o'\rangle \in Z^t_p$ then $\langle o', o\rangle \in Z^t_{p^{-1}}$. So assume $\langle o, o'\rangle \in Z^t_p$. Then there is a $w \in \mathrm{dom}(\dot{p})$ such that $o \in D^t_w$ and $o' \in (D^P)^t_{\dot{p}(w)}$. We first show that $o' \in D'^t_{\dot{p}(w)}$. By Lemma 7.3.4, it suffices to show that $D' \boxtimes o'$ and $\hat{f}.o' = o'$ for all $f \in \mathrm{fix}^i_\omega(\mathfrak{S}', \dot{p}(w))$. By Lemma 7.2.10 (ii), $o' \in (D^P)^t_{\mathrm{im}(\dot{p})}$, so with the fact that the structure projectively generated by $P$ is a $\times$structure, it follows that $D^P\boxtimes o'$. Therefore by IH (iii), $D'\boxtimes o'$. So consider any $f \in \mathrm{fix}^i_\omega(\mathfrak{S}', \dot{p}(w))$. Since $\mathrm{supp}(f)$ is finite, there is a $q \in P$ such that $p \subseteq q$ and $\mathrm{supp}(f) \subseteq \mathrm{im}(q)$. So by Lemma 7.6.2, there is a $g \in \mathrm{fix}^i_\omega(\mathfrak{S}, w)$ such that $gZ^p_q f$. By Lemma 7.2.10 (ii), $oZ^t_q o'$, so with claim (i) of the present lemma, $\hat{g}.oZ^t_q\hat{f}.o'$. As $o \in D^t_w$, $\hat{g}.o = o$, so $oZ^t_q\hat{f}.o'$. Hence by the functionality of $Z^t_q$, established in Lemma 7.2.10 (i), $\hat{f}.o' = o'$.

With $o' \in D'^t_{\dot{p}(w)}$ established, we can prove that $o'Z^t_{p^{-1}}o$ by checking conditions (1) and (2) of the definition of $Z$. Since $o \in D^t_w$, $D\boxtimes o$, so by IH, $D^{P^{-1}}\boxtimes o$. For condition (2), consider any $q \in P^{-1}$ such that $p^{-1} \subseteq q$, $v \in \mathrm{dom}(\dot{q})$ and

$n$-tuples $\bar{o}, \bar{o}'$ such that $\bar{o} Z_q \bar{o}'$. By IH, $\bar{o}' Z_{q^{-1}} \bar{o}$; also $q^{-1} \in P$, $p \subseteq q^{-1}$ and $\dot{q}(v) \in \mathrm{dom}(\dot{q}^{-1})$. So by $o Z^t_p o'$, $\bar{o}' \in o(\dot{q}(v))$ iff $\bar{o} \in o'(v)$, as required for condition (2). So $o' Z^t_{p^{-1}} o$.

(iii): Let $v \in W'$. By condition (iii) of the definition of projections, there is a $p \in P$ such that $v \in \mathrm{im}(\dot{p})$. By Lemma 7.2.11, $D'^t_v = \mathrm{dom}\left(Z^t_{p^{-1}|v}\right) = \mathrm{dom}\left(Z^t_{(p|p^{-1}(v))^{-1}}\right)$. By (ii), this is $\mathrm{im}\left(Z^t_{p|p^{-1}(v)}\right)$, which by Lemma 7.2.11 again is $\mathrm{im}\left(Z^t_p | D^t_{p^{-1}(v)}\right)$. By Lemma 7.2.10 (i), this is $(D^P)^t_v$. □

*Proof of Theorem 7.3.6.* Immediate by Lemma 7.6.4. □

# Bibliography

Robert Merrihew Adams. Primitive thisness and primitive identity. *The Journal of Philosophy*, 76:5–26, 1979.

Robert Merrihew Adams. Actualism and thisness. *Synthese*, 49:3–41, 1981.

Carlos Areces and Balder ten Cate. Hybrid logics. In Patrick Blackburn, Johan van Benthem, and Frank Wolter, editors, *Handbook of Modal Logic*, pages 821–868. Amsterdam: Elsevier, 2007.

George Bealer. Universals. *The Journal of Philosophy*, 90:5–32, 1993.

George Bealer. Propositions. *Mind*, 107:1–32, 1998.

Karen Bennett. Two axes of actualism. *The Philosophical Review*, 114:297–326, 2005.

Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*, volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge: Cambridge University Press, 2001.

R. A. Bull. On modal logics with propositional quantifiers. *Journal of Symbolic Logic*, 34:257–263, 1969.

David J. Chalmers, David Manley, and Ryan Wasserman. *Metametaphysics*. Oxford: Oxford University Press, 2009.

Fabrice Correia. Modality, quantification, and many Vlach-operators. *Journal of Philosophical Logic*, 36:473–488, 2007.

Marian David. Defending existentialism? In Maria Reicher, editor, *States of Affairs*, pages 167–208. Frankfurt: Ontos Verlag, 2009.

Richard Dawkins. *Unweaving the Rainbow: Science, Delusion and the Appetite for Wonder*. London: Penguin Books, 1998.

M. A. Dickmann. Larger infinitary languages. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 317–363. New York: Springer-Verlag, 1985.

John D. Dixon and Brian Mortimer. *Permutation Groups*. New York: Springer-Verlag, 1996.

Cian Dorr. What we disagree about when we disagree about ontology. In Mark Eli Kalderon, editor, *Fictionalism in Metaphysics*, pages 234–286. Oxford: Oxford University Press, 2005.

Cian Dorr. There are no abstract objects. In Theodore Sider, John Hawthorne, and Dean W. Zimmerman, editors, *Contemporary Debates in Metaphysics*, pages 32–64. Malden, MA: Blackwell, 2008.

Fredrik Engström. Generalized quantifiers in dependence logic. *Journal of Logic, Language and Information*, 21:299–324, 2012.

Kit Fine. Propositional quantifiers in modal logic. *Theoria*, 36:336–346, 1970.

Kit Fine. *Postscript* to *Worlds, Times and Selves* (with A. N. Prior). London: Duckworth, 1977a.

Kit Fine. Properties, propositions and sets. *Journal of Philosophical Logic*, 6:135–191, 1977b.

Kit Fine. First-order modal theories II – Propositions. *Studia Logica*, 39: 159–202, 1980.

Kit Fine. Plantinga on the reduction of possibilist discourse. In J. Tomberlin and Peter van Inwagen, editors, *Alvin Plantinga*, pages 145–86. D. Reidel, 1985.

Kit Fine. Quine on quantifying in. In C. A. Anderson and J. Owens, editors, *Proceedings of the Conference on Propositional Attitudes*, pages 1–26. Stanford: CSLI, 1990.

Kit Fine. A counter-example to Locke's thesis. *The Monist*, 83:357–361, 2000.

Kit Fine. The problem of possibilia. In Michael J. Loux and Dean W. Zimmerman, editors, *The Oxford Handbook of Metaphysics*, pages 161–179. Oxford: Oxford University Press, 2003.

Kit Fine. Class and membership. *The Journal of Philosophy*, 102:547–572, 2005.

Kit Fine. Williamson on Fine on Prior on the reduction of possibilist discourse. unpublished.

G. W. Fitch. In defense of Aristotelian actualism. *Philosophical Perspectives*, 10:53–71, 1996.

Roland Fraïssé. Sur une extension de la polyrelation et des parentés tirant son origine du calcul logiques du k-ème échelon. In *Le raisonnement en mathématiques et en sciences expérimentales*, volume 70 of *Colloques Internationaux du CNRS*, pages 45–50. Paris: Editions du Centre National de la Recherche Scientifique, 1958.

Peter Fritz. Modal ontology and generalized quantifiers. *Journal of Philosophical Logic*, 42:643–678, 2013.

Peter Fritz. Higher-order contingentism, part 2: Patterns of indistinguishability. unpublished a, here ch. 3.

Peter Fritz. Higher-order contingentism, part 3: Expressive limitations. unpublished b, here ch. 7.

Peter Fritz. Logics for propositional contingentism. unpublished c, here ch. 5.

Peter Fritz. Propositional contingentism. unpublished d, here ch. 2.

Peter Fritz and Jeremy Goodman. Counterfactuals and propositional contingentism. unpublished a, here ch. 4.

Peter Fritz and Jeremy Goodman. Counting incompossibles. unpublished b, here ch. 6.

Peter Fritz and Jeremy Goodman. Higher-order contingentism, part 1: Closure and generation. unpublished c, here ch. 1.

Daniel Gallin. *Intensional and Higher-Order Modal Logic*. Amsterdam: North-Holland, 1975.

Jeremy Goodman. Consequences of conditional excluded middle. unpublished.

Jeremy Goodman. The case for necessitism. in preparation a.

Jeremy Goodman. Williamson on necessitism and plural quantification. in preparation b.

Caspar Hare. Voices from another world: Must we respect the interests of people who do not, and will never, exist? *Ethics*, 117:498–523, 2007.

R. M. Hare. Possible people. *Bioethics*, 2:279–293, 1988.

John Hawthorne. *Metaphysical Essays*. Oxford: Oxford University Press, 2006.

Jaakko Hintikka and Veikko Rantala. A new approach to infinitary languages. *Annals of Mathematical Logic*, 10:95–115, 1976.

Harold Hodes. On modal logics which enrich first-order S5. *Journal of Philosophical Logic*, 13:423–454, 1984.

Wilfrid Hodges. *A Shorter Model Theory*. Cambridge: Cambridge University Press, 1997.

Paul Hovda. Tensed mereology. *Journal of Philosophical Logic*, 42:241–283, 2013.

Thomas Jager. An actualistic semantics for quantified modal logic. *Notre Dame Journal of Formal Logic*, 23:335–349, 1982.

Michael Kaminski and Michael Tiomkin. The expressive power of second-order propositional modal logic. *Notre Dame Journal of Formal Logic*, 37: 35–43, 1996.

David Kaplan. S5 with quantifiable propositional variables. *Journal of Symbolic Logic*, 35:355, 1970.

David Kaplan. A problem in possible-world semantics. In Walter Sinnott-Armstrong, Diana Raffman, and Nicholas Asher, editors, *Modality, Morality, and Belief*, pages 41–52. Cambridge: Cambridge University Press, 1995.

Carol R. Karp. *Languages with Expressions of Infinite Length*. Amsterdam: North-Holland, 1964.

Carol R. Karp. Finite-quantifier equivalence. In J. W. Addison, Leon Henkin, and Alfred Tarski, editors, *The Theory of Models: Proceedings of the 1963 International Symposium at Berkeley*, pages 407–412. Amsterdam: North-Holland Publishing Company, 1965.

Maaret Karttunen. Infinitary languages $N_{\infty\lambda}$ and generalized partial isomorphisms. In Jaakko Hintikka, Ilkka Niiniluoto, and Esa Saarinen, editors, *Essays on Mathematical and Philosophical Logic*, pages 153–168. Dordrecht: D. Reidel, 1979.

Maaret Karttunen. Model theoretic results for infinitely deep languages. *Studia Logica*, 42:223–241, 1983.

Nathan Klinedinst. Quantified conditionals and conditional excluded middle. *Journal of Semantics*, 28:149–170, 2011.

Ph. G. Kolaitis. Game quantification. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 365–421. New York: Springer-Verlag, 1985.

Philip Kremer. Quantifying over propositions in relevance logic: Nonaxiomatizability of primary interpretations of $\forall p$ and $\exists p$. *Journal of Symbolic Logic*, 58:334–349, 1993.

Philip Kremer. On the complexity of propositional quantification in intuitionistic logic. *Journal of Symbolic Logic*, 62:529–544, 1997.

Saul A. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14, 1959.

Saul A. Kripke. Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16:83–94, 1963.

Saul A. Kripke. *Naming and Necessity.* Cambridge, MA: Harvard University Press, 1980 [1972]. First published in *Semantics of Natural Language*, edited by Donald Davidson and Gilbert Harman, pages 253–355, 763–769, Dordrecht: D. Reidel, 1972.

Stephan Leuenberger. A new problem of descriptive power. *The Journal of Philosophy*, 103:145–162, 2006.

David Lewis. *Counterfactuals.* Oxford: Basil Blackwell, 1973.

David Lewis. *On the Plurality of Worlds.* Oxford: Basil Blackwell, 1986.

David Lewis. Tensed quantifiers. In Dean W. Zimmerman, editor, *Oxford Studies in Metaphysics*, volume 1, pages 3–14. Oxford: Oxford University Press, 2004.

Sten Lindström. Possible worlds semantics and the liar. In Joseph Almog and Paolo Leonardi, editors, *The Philosophy of David Kaplan*, pages 93–108. Oxford: Oxford University Press, 2009.

Kris McDaniel. Ways of being. In David J. Chalmers, David Manley, and Ryan Wasserman, editors, *Metametaphysics*, pages 290–319. Oxford: Oxford University Press, 2009.

Alan McMichael. A problem for actualism about possible worlds. *Philosophical Review*, 92:49–66, 1983.

Michael Nelson. Contingently existing propositions. *Canadian Journal of Philosophy*, 43:776–803, 2014.

Anil Nerode and Richard A. Shore. Second order logic and first-order theories of reducibility orderings. In Jon Barwise, H. Jerome Keisler, and Kenneth Kunen, editors, *The Kleene Symposium*, pages 181–200. Amsterdam: North Holland, 1980.

Juha Oikkonen. A generalization of the infinitely deep languages of Hintikka and Rantala. In Esa Saarinen, Risto Hilpinen, Ilkka Niiniluoto, and Merrill Provence Hintikka, editors, *Essays in Honour of Jaakko Hintikka*, pages 101–112. Dordrecht: D. Reidel, 1979.

Derek Parfit. *Reasons and Persons*. Oxford: Oxford University Press, 1984.

Ahti-Veikko Pietarinen. Semantic games and generalised quantifiers. In Ahti-Veikko Pietarinen, editor, *Game Theory and Linguistic Meaning*, pages 183–206. Oxford: Elsevier, 2007.

Alvin Plantinga. *The Nature of Necessity*. Oxford: Oxford University Press, 1974.

Alvin Plantinga. Actualism and possible worlds. *Theoria*, 42:139–160, 1976.

Alvin Plantinga. On existentialism. *Philosophical Studies*, 44:1–20, 1983.

A. N. Prior. Platonism and quantification. In *Objects of thought*, chapter 3, pages 31–47. Oxford: Oxford University Press, 1971.

Arthur Prior. *Past, Present and Future*. Oxford: Oxford University Press, 1967.

Veikko Rantala. Game-theoretical semantics and back-and-forth. In Jaakko Hintikka, Ilkka Niiniluoto, and Esa Saarinen, editors, *Essays on Mathematical and Philosophical Logic*, pages 119–151. Dordrecht: D. Reidel, 1979.

Gideon Rosen. Modal fictionalism. *Mind*, 99:327–354, 1990.

Joseph J. Rotman. *An Introduction to the Theory of Groups*. New York: Springer-Verlag, fourth edition, 1995.

Krister Segerberg. *An Essay in Classical Modal Logic*. Number 13 in Filosofiska Studier. Uppsala: Uppsala Universitet, 1971.

Theodore Sider. The ersatz pluriverse. *The Journal of Philosophy*, 99:279–315, 2002.

Theodore Sider. Quantifiers and temporal ontology. *Mind*, 115:75–97, 2006.

Theodore Sider. Ontological realism. In David Chalmers, David Manley, and Ryan Wasserman, editors, *Metametaphysics*, pages 384–423. Oxford: Oxford University Press, 2009.

Jeff Speaks. On possibly nonexistent propositions. *Philosophy and Phenomenological Research*, 85:528–562, 2012.

Robert Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Oxford: Blackwell, 1968.

Robert Stalnaker. Possible worlds. *Noûs*, 10:65–75, 1976.

Robert Stalnaker. A defense of conditional excluded middle. In William L. Harper, Robert Stalnaker, and Glenn Pearce, editors, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pages 87–104. D. Reidel, 1981.

Robert Stalnaker. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Robert Stalnaker. The interaction of modality with quantification and identity. In *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*, pages 144–161. Oxford: Oxford University Press, 2003 [1994]. First published in an earlier form in *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*, edited by Walter Sinnott-Armstrong, Diana Raffman and Nicholas Asher, pages 12–28, Cambridge: Cambridge University Press, 1994.

Robert Stalnaker. *Mere Possibilities.* Princeton: Princeton University Press, 2012.

Zoltán Gendler Szabó. Counting across times. *Philosophical Perspectives*, 20:399–426, 2006.

Jason Turner. Ontological pluralism. *The Journal of Philosophy*, 107:5–34, 2010.

Jouko Väänänen. Barwise: Abstract model theory and generalized quantifiers. *The Bulletin of Symbolic Logic*, 10:37–53, 2004.

Jouko Väänänen. *Models and Games.* Cambridge: Cambridge University Press, 2011.

Dag Westerståhl. Generalized quantifiers. In Edward N. Zalta, editor, *Stanford Encyclopedia of Philosophy.* Summer 2011 edition, 2011.

J. Robert G. Williams. Defending conditional excluded middle. *Noûs*, 44: 650–668, 2010.

Timothy Williamson. Necessary existents. In Anthony O'Hear, editor, *Logic, Thought and Language*, pages 233–51. Cambridge: Cambridge University Press, 2002.

Timothy Williamson. Everything. *Philosophical Perspectives*, 17:415–465, 2003.

Timothy Williamson. *The Philosophy of Philosophy.* Malden: Blackwell Publishing, 2007.

Timothy Williamson. Necessitism, contingentism, and plural quantification. *Mind*, 119:657–748, 2010.

Timothy Williamson. *Modal Logic as Metaphysics.* Oxford: Oxford University Press, 2013.

Richard Woodward. Fictionalism and incompleteness. *Noûs*, 46:781—90, 2012.