The logic of sequences

Cian Dorr and Matt Mandelkern*

November 22, 2024

Abstract

In the course of proving a tenability result about the probabilities of conditionals, van Fraassen (1976) introduced a semantics for conditionals based on ω-sequences of worlds, which amounts to a particularly simple special case of ordering semantics for conditionals. On that semantics, 'If p, then q' is true at an ω-sequence just in case q is true at the first tail of the sequence where p is true (if such a tail exists). This approach has become increasingly popular in recent years. However, its logic has never been explored. We axiomatize the logic of ω-sequence semantics, showing that it is the result of adding two new axioms to Stalnaker's logic C2: one, Flattening, which is prima facie attractive, and, and a second, Sequentiality, which is complex and difficult to assess, but, we argue, likely invalid. But we also show that when sequence semantics is generalized from ω-sequences to arbitrary (transfinite) ordinal sequences, the result is a more attractive logic that adds only Flattening to C2. We also explore the logics of a few other interesting restrictions of ordinal sequence semantics. Finally, we address the question of whether sequence semantics is motivated by probabilistic considerations, answering, pace van Fraassen, in the negative.

1 Introduction

Stalnaker's (1968) 'A theory of conditionals' launched the modern study of the conditional with a semantics for natural-language conditionals and a description of the corresponding logic C2, laying the groundwork for the rich subsequent literature on conditionals in philosophy, linguistics, and logic. One subject of lively debate in this literature concerns the probabilities of conditionals. Our topic is the logic corresponding to an intriguing semantics for conditionals, based on ω-sequences, developed by van Fraassen (1976) in the course of that debate.

*Thanks to Melissa Fusco, Wesley Holliday, Calum McNamara, Nick Ramsey, and especially Snow Zhang for invaluable and patient help; as well as audiences at UConn, Stanford and Berkeley.

Stalnaker (1970) developed a theory of probability—treated as a property of sentences, and equated with "degree of rational belief"—for a language with a binary conditional connective $>$ standing for 'if... then' (on an indicative interpretation). The theory includes the following characteristic principle:

> *Stalnaker's Thesis*: $Pr(p > q) = Pr(q \mid p)$ (provided the right-hand-side is defined).[1]

There is robust empirical motivation for thinking that there is some important truth in the vicinity of Stalnaker's Thesis (see Douven and Verbrugge 2013 and many citations therein). For instance, if David is holding a fair die, the probability that if he rolls even, then he'll roll two is intuitively 1/3—equal to the probability that he rolls a two, conditional on rolling even.

Lewis (1976) showed that Stalnaker's Thesis was in tension with the assumption that we update our credences by conditionalization: in particular, no non-trivial class of probability functions closed under conditionalization satisfies Stalnaker's Thesis for a single interpretation of $>$. This left it open, however, that there is some way of co-ordinating different interpretations of the conditional with different interpretations of 'probability'—e.g., whether we are talking about credences before or after some update—in such a way that Stalnaker's Thesis always holds *within a single context*. However, a striking result in Stalnaker 1974 showed that even this is impossible given the background logic C2, except in certain trivial cases.

At the same time, however, van Fraassen (1976) showed that models of C2 *can* be equipped with nontrivial probability functions satisfying a restricted form of Stalnaker's Thesis.[2] In van Fraassen's models, the probability function on non-conditional sentences can be freely specified, and Stalnaker's Thesis holds for conditionals whose antecedents do not themselves contain conditionals.[3]

Ours is not, however, a paper about the probabilities of conditionals (a topic we return to only briefly, in §7), but rather about the construction which van Fraassen developed in the course of modeling the restricted version of Stalnaker's Thesis. In that construction, van Fraassen used a semantics for conditionals with the following form. Start with a set W of "worlds" and a valuation that specifies which atomic sentences are true at elements of

[1] See Stalnaker 1970: p. 75. Like Popper (1959), Stalnaker himself sets things up in such a way that the right-hand-side is always defined—e.g., $Pr(q \mid p \wedge \neg p) = 1$ for all p and q.

[2] Publication dates in this literature are confusing. To our knowledge, Lewis's was the first triviality result. Stalnaker's 1974 letter was a response to a draft of van Fraassen's paper, which, in turn, was a response (in part) to Lewis's result. Van Fraassen's published 1976 paper appears to leave open whether his construction validates Stalnaker's Thesis for the whole language, a possibility which Stalnaker's letter rules out; our understanding is that the published version was in fact the one Stalnaker's letter was responding to, despite the later publication date. Thanks to Bas van Fraassen for correspondence about this.

[3] It also holds for some conditionals whose antecedents *do* contain conditionals: see §7 for details.

W. Now consider the set of *ω-sequences* over W: that is, functions from the natural numbers to W. These sequences will serve as indices in a model for a language containing the conditional connective $>$. In this model, an atom is true at a sequence $\sigma = \langle w_0, w_1, w_2, \ldots \rangle$ just in case it is true at w_0 according to the old valuation. The clauses for negation and conjunction are classical. Finally, a conditional $p > q$ is true at a sequence σ just in case either σ has a tail at which p is true, and q is true at the first such tail, or σ has no tail at which p is true (the *tails* of $\langle w_0, w_1, w_2, w_3, w_4 \ldots \rangle$ are $\langle w_0, w_1, w_2, w_3, w_4 \ldots \rangle, \langle w_1, w_2, w_3, w_4 \ldots \rangle, \langle w_2, w_3, w_4, \ldots \rangle$, and so on).

Sequence semantics has become increasingly popular in recent years.[4] But, surprisingly, some basic questions about the semantics have never been answered, including what its logic is. The goal of this paper is to axiomatize the logic of ω-sequence models, as well as some interesting generalizations and restrictions of that approach that base the same semantics on different classes of ordinal sequences (that is, functions from arbitrary ordinals, possibly larger or smaller than ω, to an underlying set).

We have a few motivations for this project. One is its intrinsic interest: ω-sequence semantics is an intriguing, and in some ways very simple, semantics for conditionals. So we should understand it, and part of understanding the semantics is knowing the logic it gives rise to. Besides being of intrinsic interest, this will help us assess the viability of sequence semantics for modeling conditionals in natural language.[5]

A final motivation comes from particularities of the logic that arises from sequence semantics. As we will show, sequence semantics can be viewed as a special case of Stalnaker's semantics for conditionals—a special case which, it turns out, strictly strengthens Stalnaker's logic C2. This is of special interest to both authors, who believe that all the principles of C2 are plausible as far as natural language conditionals go.

This is a controversial position. The commitment of C2 to the validity of Conditional Excluded Middle (CEM) has historically been rather unpopular, due to influential criticism by Lewis. In fact, essentially every commitment of C2 has been rejected somewhere in the subsequent literature. However, our commitment to the correctness of a logic at least as strong as C2 makes us particularly interested in strengthenings of C2. (We will not do anything here to defend C2, but see Dorr and Hawthorne 2022 for extensive discussion.)

To our knowledge, however, no logics which are stronger than C2 but weaker than Materialism have ever been explored. Materialism is the logic which collapses the natural language conditional $p > q$ to the material conditional $p \to q$, that is, the logic which simply adds to classical logic the principle $(p > q) \leftrightarrow (p \to q)$. There exist powerful arguments against this equivalence (Edgington 1995). Famously, however, Dale (1974), Dale (1979),

[4]See e.g. Kaufmann 2009; 2015; Bacon 2015; Schultheis 2022; Santorio 2021; Goldstein and Santorio 2021; Khoo 2022.

[5]See Holliday and Icard 2018 on the methodological importance of axiomatization in semantics.

Gibbard (1981), and McGee (1985) showed that the gap between C2 and Materialism is surprisingly small: in particular, it is fully closed by the Import-Export principle (which we discuss in §4.1). To our knowledge, no logics residing in the gap bewteen C2 and Materialism have ever been studied, perhaps because of these famous results. But will turn out that the logic of ω-sequences is strictly intermediate between C2 and Materialism. In fact, in the course of exposition, we will explore two such logics: we will show that the logic of ω-sequences is the logic we call C2.FS, comprising C2 plus every instance of the following two axiom schemas.

Flattening $\quad (p > ((p \wedge q) > r)) \leftrightarrow ((p \wedge q) > r)$

Sequentiality $\quad \Box(p \to (\neg p > r)) \wedge \Box(q \to (\neg q > r)) \to ((p \vee q) \to (\neg(p \vee q) > r))$

where $\Box p$ is defined as $\neg p > p$. We will argue that Flattening is at least prima facie appealing for conditionals in natural language, while Sequentiality is, at best, too complex to reasonably assess, and, at worst, invalid. This suggests that ω-sequence semantics is not a strong contender for a logic of the natural language conditional. But we will show that the logic of a semantics based on *ordinal sequences* in general, rather than just ω-sequences, is the more attractive logic comprising C2 together with just Flattening. We will explore the logics of a few other interesting restrictions of ordinal sequence semantics, and, finally, argue that, *pace* van Fraassen, sequence semantics cannot be motivated by indirect considerations about the probabilities of conditionals.

2 C2 and its semantics

We will begin with some important background, reviewing Stalnaker's (1968) conditional logic C2, and a class of models corresponding to that logic. (Cognoscenti may wish to skip to the next section.)

The language of C2 and all the logics we will be considering is a standard propositional language $\mathcal{L}$ equipped with a binary conditional connective $>$. So, where $At = \{p_0, p_1, \ldots\}$ is a countably infinite set of atomic sentences:[6]

$$p ::= p_k \in At \mid \neg p \mid (p \wedge p) \mid (p > p)$$

We use $\to$, $\leftrightarrow$, and $\vee$ as abbreviations for the material conditional, material biconditional, and disjunction defined as usual. For brevity, we sometimes use a compact notation for negation and conjunction applied to atoms and metavariables: $\overline{p}$ for $\neg p$ and pq for $(p \wedge q)$ (thus $\neg pq$ is $\neg(p \wedge q)$). We also sometimes omit parentheses: where we do, the order of operations is negation,

[6]Stalnaker and Thomason (1970) extend C2 to a language with quantifiers, but here we are concerned only with the propositional fragment from Stalnaker 1968.

then $>$, then $\wedge$ and $\vee$, and finally $\to$ and $\leftrightarrow$, so for instance $p > q \to \neg r > s \wedge t$ is to be read as $(p > q) \to ((\neg r > s) \wedge t)$.[7]

C2 is the closure of the following set of axiom schemas:[8]

PC	Every theorem of classical propositional logic
Identity	$p > p$
Reciprocity	$(p > q) \wedge (q > p) \wedge (p > r) \to q > r$
MP	$p > q \to (p \to q)$
CEM	$p > q \vee p > \neg q$

under the following two inference rules:

Detachment	$\vdash p \to q$ and $\vdash p$ together imply $\vdash q$
Normality	$\vdash (p \wedge q) \to r$ implies $\vdash ((s > p) \wedge (s > q)) \to s > r$

When p is a theorem of C2 we write $\vdash_{C2} p$.

Stalnaker's own axiomatization of C2 is somewhat different, and uses two further abbreviations, which will be useful in what follows and hence worth noting here:

$\Box$, defined by $\Box p := \neg p > p$; and

$\Diamond$, defined by $\Diamond p := \neg \Box \neg p$. [9]

Stalnaker then defines C2 with the axioms PC, MP, and Reciprocity plus four further axioms (the first of which is just the familiar K axiom for $\Box$):

$$\Box(p \to q) \to (\Box p \to \Box q)$$
$$\Box(p \to q) \to p > q$$
$$\Diamond p \to (p > q \to \neg(p > \neg q))$$
$$p > (q \vee r) \to (p > q \vee p > r)$$

and closing the result under Detachment together with the rule:

Necessitation	$\vdash p$ implies $\vdash \Box p$

It is a good exercise to show that these two axiomatizations of C2 are indeed equivalent. Deriving our axiomatization from Stalnaker's is easy, using the definition of $\Box$ via $>$. For the other direction, the key principle to derive is:

MOD	$\Box p \to q > p$

[7] The defined unary operators $\Box$ and $\Diamond$ which we introduce below also take highest priority.

[8] Reciprocity is often called CSO, but the source of that name is lost (to us, at least), so we will use the more mnemonic name.

[9] $\Diamond p$ could equivalently be defined as $\neg(p > \neg p)$.

Stalnaker's axiomatization brings out the fact that there is a sense in which C2 *contains* (though is not reducible to) a modal logic. The modal logic KT is the logic containing every instance of the PC and K schemas as above, along with the further axiom schema:

T $$\Box p \to p$$

closed under the rules of Necessitation and Detachment. We will show below that the theorems of C2 expressible using atoms, Boolean connectives, and $\Box$ (that is, the theorems in *the modal fragment* of $\mathcal{L}$) are exactly the theorems of KT.

While $\Box$ in the formal language is simply a shorthand, one might also think there are indeed close connections between necessity and conditionals, in particular between, on the one hand, epistemic modals and indicative conditionals; and, on the other, circumstantial modals and subjunctive conditionals.[10] Such connections would make the modal logics of various conditional logics especially interesting. Even in the absence of such connections, however, the modal fragments of our conditional logics are well worth studying on a purely logical basis.

2.1 Order models for C2

There are many model-theoretic semantics for conditional connectives in the literature, generalizing Kripke's possible-worlds semantics for modal logic to conditional languages. Our focus in this paper will be on *order models* for conditionals, which were introduced by Lewis (1973) as models for his logic (which is strictly weaker than C2). Order models turn out to be particularly intuitive for the study of C2 and its strengthenings. In particular, we can model C2 with order models where each world is associated with a *well-ordering* of worlds. As we will see, sequence models can be naturally viewed as a special case of well-order models, making it easy to see that their logic includes C2.

A Kripke model equips a set of possible worlds with a binary accessibility relation R, representing relative necessity and possibility: p is necessary at w just in case p is true throughout the worlds accessible from w, which we write $R(w)$, and p is possible at w just in case p is true somewhere in $R(w)$. An order model is like a Kripke model but with additional structure: in addition to a set of worlds $R(w)$, an order model associates each world w with an *ordering* $<_w$ of $R(w)$. We pronounce $u <_w v$ as 'u is closer to w than v'. In such a model, assuming there are some *closest* p-worlds to w, $p > q$ is true at w just in case all of these worlds are q-worlds. $p > q$ is also ('vacuously') true at w when there aren't any p-worlds in $R(w)$.[11]

[10]See Dorr and Hawthorne 2022 for one defense of such a position.

[11]In Lewis's generalization of order models, we also need to say something about the case where $R(w)$ contains some p-worlds, but for each one of them, there is another that is closer. However this case will conveniently not arise in the models we deal with.

Lewis conceived of closeness in terms of *similarity*: $x <_w y$ means that x is more similar to w than y is (in whatever respects turn out to be relevant). But using order models does not commit us to a similarity-based interpretation of the order functions, any more than Kripke semantics for the modal operators commits us to any particular theory of necessity and possibility. Thus, skeptics of similarity-based approaches to conditionals (a group in which we include ourselves) have no special reason to object to the use of order models. Indeed, in modelling strengthenings of C2, we will need to impose conditions on order models which would be completely implausible if closeness had to be interpreted as similarity, so insofar as the strengthenings are well-motivated, they will add to the already strong case against similarity-based approaches (see §4.1).

In the order models characteristic of C2, each ordering $<_w$ is a *well-order*: that is, a transitive, connected, asymmetric, well-founded relation on $R(w)$. This means that we can restate the order semantics for conditionals in terms of the *unique* closest antecedent world, if there is one: $p > q$ is true at w just in case q is true at the first p-world in $<_w$, or there are no p-worlds in $R(w)$. This uniqueness assumption guarantees that the controversial CEM axiom holds in the model; logics without CEM (like Lewis's) can be obtained by relaxing this assumption. However, we will not further discuss such models in this paper, so by 'order model' we will always mean a well-order model.[12]

Let us lay this all out more formally, and introduce some standard terminology which will be helpful:

Definition 2.1. An *order frame* is a pair $\langle W, < \rangle$, where W is a non-empty set and $<$ is a function which takes any $w \in W$ to a strict total well-order $<_w$ on some subset of W such that whenever $x <_w y$, $w = x$ or $w <_w x$.

[12]This order semantics for conditionals is equivalent to the semantics based on *selection functions* given in Stalnaker 1968. A selection function is a function f which takes a set of worlds φ and world w and returns a set $f(\varphi, w)$. We can use a selection function to evaluate conditionals, by defining $[\![p > q]\!]$ to be $\{w : f([\![p]\!], w) \subseteq [\![q]\!]\}$. In the case of interest for C2, f is required to obey constraints corresponding to MP, Reciprocity, Identity, and CEM:

1. If $w \in \varphi$, $w \in f(\varphi, w)$
2. If $f(\varphi, w) \subseteq \psi$ and $f(\psi, w) \subseteq \varphi$, $f(\varphi, w) = f(\psi, w)$.
3. $f(\varphi, w) \subseteq \varphi$
4. $f(\varphi, w)$ has cardinality at most 1

Given these constraints, we can move freely between selection functions and order functions. Given a selection function f, we define an isomorphic order function $<$ by saying that $x <_w y$ just in case $f(\{x, y\}, w) = \{x\}$ and $f(\{y\}, w) = \{y\}$. Conversely, given an order function $<$, we define an isomorphic selection function f by saying that $f(\varphi, w)$ is $\{w\}$ when $w \in \varphi$; otherwise the singleton of the first φ-world ordered by $<_w$, if there is one; and otherwise the empty set. (Reflection on the respective properties of order and selection functions show that these defined constructions are, indeed, order and selection functions, respectively.) So there is no deep difference between these two kinds of model. However, order models lend themselves more naturally to the study of sequence semantics, as we shall see.

We can read off an accessibility relation from an order frame: the worlds accessible from w are those that w strictly precedes in the ordering induced at w, together with w itself:[13]

$$R(w) = \{w\} \cup \{v : w <_w v\}$$

As we will see, R plays the same role with respect to the defined $\Box$ as accessibility relations usually do in Kripke models. We write $x \leq_w y$ whenever $x, y \in R(w)$ and $y \not<_w x$, i.e., whenever either $x <_w y$, or $x = y$ and $x \in R(w)$.

Definition 2.2. An *order model* is an order frame $\langle W, < \rangle$ together with a valuation function $V : At \to \mathcal{P}(W)$.

Definition 2.3. When $\langle W, <, V \rangle$ is an order model, its denotation function is the function $[\![\cdot]\!]^{\langle W,<,V \rangle} : \mathcal{L} \to \mathcal{P}(W)$ such that for any atom p_k and sentences p, q:

$$\begin{aligned}
[\![p_k]\!] &= V(p_k) \\
[\![\neg p]\!] &= W \setminus [\![p]\!] \\
[\![p \wedge q]\!] &= [\![p]\!] \cap [\![q]\!] \\
[\![p > q]\!] &= \{w \in W : R(w) \cap [\![p]\!] = \varnothing \vee \exists y \in R(w) \cap [\![p \wedge q]\!] : \forall x <_w y : x \notin [\![p]\!]\}
\end{aligned}$$

For readability, relativization of $[\![\cdot]\!]^{\cdot}$ to a model is usually left implicit. As usual, we can define a *pointed* order model as a pair of an order model with a world from its set of worlds, i.e. a pair $\langle w, \langle W, <, V \rangle \rangle$ such that $w \in W$ and $\langle W, <, V \rangle$ is an order model; when $\langle w, \langle W, <, V \rangle \rangle$ is a pointed order model, we say that it is *based on* $\langle W, <, V \rangle$. p is true at a pointed order model $\langle w, \mathcal{M} \rangle$ just in case $w \in [\![p]\!]^{\mathcal{M}}$; when the implicit model is clear, we write $[\![p]\!]^w = 1$ for $w \in [\![p]\!]$, and $[\![p]\!]^w = 0$ for $w \notin [\![p]\!]$. When $\Gamma \subseteq \mathcal{L}$, we can also speak of Γ being true at a pointed model to mean that all its elements are. We also write $w, \mathcal{M} \Vdash p$ when p is true $\langle w, \mathcal{M} \rangle$; when $\mathcal{M}$ is implicit from the context, we write simply $w \Vdash p$. For brevity we sometimes talk about p being true at every model in a given class; by this we mean true at every pointed model based on a model in that class. Two pointed models are *equivalent* just in case they verify exactly the same sentences of $\mathcal{L}$.

A standard induction on formulae shows:

Theorem 2.4. C2 is sound for order models: that is, $\vdash_{\mathsf{C2}} p$ implies that p is true in every pointed order model.

We also have a corresponding completeness result: every sentence that is true in every pointed order model is a theorem of C2. Equivalently, whenever p is *consistent* in C2 (that is, $\nvdash_{\mathsf{C2}} \neg p$) it is true in some pointed order model. In fact, we can show something stronger: C2 is complete with respect to the class of *finite* order models:

[13] Accessibility relations are sometimes specified as independent parameters, but as this shows, they needn't be.

Theorem 2.5. If p is true in every finite pointed order model, then $\vdash_{\mathsf{C2}} p$.

The proof of this result, together with all the other completeness theorems we will claim in this paper, is given in the Appendix.

One corollary is that C2 is *decidable*. Since every non-theorem is false in some finite pointed order model, and we can effectively enumerate all the finite pointed order models (up to isomorphism), we can test for non-theoremhood by searching through the finite pointed order models until we find a countermodels; this provides an effective decision-procedure when run in parallel with a proof search.

This soundness and completeness theorem also allows us to prove our earlier assertion about the modal fragment of C2:

Theorem 2.6. When p is a sentence in the modal fragment of $\mathcal{L}$, $\vdash_{\mathsf{C2}} p$ iff $\vdash_{\mathsf{KT}} p$.

Proof. We rely on the well-known fact that KT is sound and complete with respect to modal modals with a reflexive accessibility relation.

- $\Rightarrow$ Normality for $>$ gives the K axiom for $\Box$, and MP for $>$ gives T for $\Box$.
- $\Leftarrow$ Suppose we have a modal model W with a reflexive accessibility relation R; we can extend this to an order model that respects R by fixing a strict well-ordering $<$ on W, and for $x \neq y$, let $x <_w y$ iff wRx and wRy and either $x = w$, or $x \neq w$ and $y \neq w$ and $x < y$. If $\vdash_{\mathsf{C2}} p$ then by soundness p is true in every order model, and hence in every reflexive modal model, and so by completeness of KT, $\vdash_{\mathsf{KT}} p$.

$\Box$

2.2 The failure of strong completeness

Theorem 2.5 is about individual sentences: it is equivalent to the claim that whenever p is *consistent* in C2 (meaning that $\neg p$ is not a theorem of C2), p is true in some pointed order model. This kind of result is sometimes called a *weak* completeness theorem. By contrast, a *strong* completeness theorem would say that for every *set* of sentences that is consistent in a certain logic, there is a model based on a frame in the relevant class in which *every* sentence in that set is true. We define consistency for sets of sentences as usual: Γ is consistent relative to $\vdash$ when there is no finite conjunction p of elements of Γ such that $\vdash p \to \bot$, where $\bot$ abbreviates $p_0 \wedge \neg p_0$. Somewhat surprisingly, we do not have a strong completeness theorem analogous to Theorem 2.5.[14]

Theorem 2.7. There are C2-consistent sets $\Gamma \subseteq \mathcal{L}$ which are not true in any pointed order model.

[14] We do not know of any explicit discussion of this point in the literature. But see Kaufmann 2017 for related points.

Proof. One such Γ is the following:

$$\Gamma = \{\neg((p_i \vee p_{i+1}) > p_i) \mid i \in \mathbb{N}\}$$

Suppose all the members of Γ were true in some pointed order model $\langle w, \langle W, <, V\rangle\rangle$. Consider the set of worlds φ which verify some atom in this model, i.e., $\varphi = \bigcup_i V(p_i)$. We must have $\varphi \cap R(w) \neq \varnothing$, for otherwise the elements of Γ would all be false at w (e.g., if p_1 and p_2 are nowhere true in $R(w)$, then $(p_1 \vee p_2) > p_1$ is vacuously true at w). Since $<_w$ is a well-order, φ must have a least element x in $<_w$. Some atom p_k is true at x, by definition of φ. Now consider $(p_k \vee p_{k+1}) > p_k$. This is true at w, since the first world in $<_w$ where $p_k \vee p_{k+1}$ is true must be x (since x is the first world in $<_w$ where *any* atom is true), where p_k is true. Hence its negation is false at w, contrary to the assumption that w verifies all the elements of Γ.

Nevertheless, Γ is consistent in C2. If it were not, then by definition, it would have some inconsistent finite subset. But every non-empty finite subset $\Delta \subset \Gamma$ is true in a pointed order model, which with soundness shows that every finite subset of Γ is consistent. To see this, let p_k be the atom with the highest index of any atom which appears in Δ. Consider any set U with $k+1$ members, which we label $w_0, w_1, \ldots w_k$. Δ is true at the pointed order model $\langle w_k, \langle U, <, V\rangle\rangle$, where $<$ is any order function with $w_n <_{w_k} w_{n-1} : n \leq k$, and V any valuation such that $V(p_i) = \{w_i\}$ for $i \leq k$. □

The same reasoning shows that no extension of C2 in which Γ remains consistent is strongly complete for any class of order models.

It is possible, however, to formulate a notion of a "general" order frame, and hence order model, relative to which we do have strong completeness (cf. Segerberg 1989). The idea is to add to our frames an extra "propositional domain" parameter—a set of subsets of worlds, representing the allowable denotations for sentences—and only require that our orders are well-founded relative to the elements of that parameter, rather than relative to all subsets of

In more detail, let a generalized order frame be a triple $\langle W, \mathcal{B}, <\rangle$, where W is any non-empty set; $<$ is a function which takes any $w \in W$ to a total linear order $<_w$ on a subset of W, such that whenever $<_w$ orders any element of $\varphi \in \mathcal{B}$, φ has a first element in $<_w$; and $\mathcal{B}$ is a set of subsets of W, closed under the set-theoretic operations corresponding to $\wedge, \neg$, and $>$ in order semantics (relative to $<$). A generalized order model is a generalized order frame $\langle W, \mathcal{B}, <\rangle$ equipped with a valuation $V : At \to \mathcal{B}$. The definition of $[\![\cdot]\!]$ that worked for order models still works in generalized models, and yields a function from $\mathcal{L}$ to $\mathcal{B}$. C2 is sound and *strongly* complete with respect to generalized order models; completeness can be shown with a standard canonical model construction. (To model a set of sentences like Γ which is not true in any order model, we need only consider a generalized order model in which the set φ of worlds that verify some atom is not in the propositional domain, and does not have a minimal element in $<_w$.)

A generalized order frame is *full* if $\mathcal{B} = \wp(W)$; full generalized frames are equivalent to order frames.[15]

Presumably because of facts along these lines, Segerberg (1989) writes that 'in modal logic it seems quite natural to restrict one's interest—at least initially—to full frames. In conditional logic, studied in the present vein, this is not so.' Nevertheless, our interest in this paper will be primarily in full order frames, since our main goal is to identify the logics of various kinds of sequence models, which (as we will shortly see) can be viewed as special cases of (full) order models.

3 Omega-sequence Semantics

With this set-up in hand, we are now in a position to give a more rigorous presentation of van Fraassen's (1976) ω-sequence models. We will present these models in a way which brings out the fact that these are in fact a special case of order models. This will let us show immediately that the logic of ω-sequence models is at least as strong as $\mathsf{C2}$.

First, we will introduce some general terminology for talking about sequences. Although for van Fraassen's models the sequences of interest are ω-sequences, which can be understood as functions from the natural numbers to an underlying set, for the sake of later generalizations we will consider these as a special case of "ordinal sequences", whose domain can be any arbitrary ordinal.

Definition 3.1. Given a non-empty set P and an ordinal α, an *α-sequence* over P is a function $\sigma : \alpha \to P$. A function is an *ordinal sequence* just in case it is an α-sequence for some ordinal α.

When σ is an α-sequence and $\beta < \alpha$, we write σ^β for the value of σ at β, that is, $\sigma(\beta)$. When $\beta \leq \alpha$, we write $\sigma^{[\beta:]}$ for the *βth tail* of σ, i.e., the length $\alpha - \beta$ sequence $\langle \sigma^\beta, \sigma^{\beta+1}, \sigma^{\beta+2}, \ldots \rangle$.

When τ is a tail of σ, the *rank* of τ in σ is the least β such that $\tau = \sigma^{[\beta:]}$.

Any set of ordinal sequences can be endowed with an order function in a natural way:

[15]The selection function models described in Footnote 12 can be "generalized" in an analogous way to order models. In a generalized order model, the selection function f is only defined for on pairs w, φ where φ belongs to the propositional domain; as with order models, we also require the propositional domain to be closed under all the operations on sets corresponding to the semantic clauses.

In both generalized order models and generalized selection models, elements of the propositional domain that happen not to be denoted by any sentence in $\mathcal{L}$ are logically irrelevant: restricting the propositional domain of a model to the sets that are in fact denoted by sentences of $\mathcal{L}$ (in the original model) will not change the truth value of any sentence at any world. This is worth noting because it brings us back to the kind of models developed by Stalnaker and Thomason (Stalnaker 1968; Stalnaker and Thomason 1970) in which the selection functions are defined not on pairs of worlds and sets of worlds, but for pairs of worlds and sentences. Given the constraints they place on such selection functions, such models are equivalent to generalized order models as we have defined them here.

Definition 3.2. The *tail order function* $\prec^S$ on any set S of sequences has $\tau \prec^S_\sigma \rho$ iff τ and ρ are tails of σ and the rank of τ in σ is less than the rank of ρ in σ.

Definition 3.3. An order frame $\langle W, < \rangle$ is an *ω-sequence frame* iff W is a set of ω-sequences on some underlying set P (which we call the "protoworlds"); W is *closed under tailhood* (that is, if it contains σ then it contains every non-empty tail of σ); and $<$ is the tail order function $\prec^W$ as in Definition 3.2. A [pointed] *ω-sequence model* is a [pointed] order model based on an ω-sequence frame.

For brevity, we write $\langle \sigma, W, V \rangle$ for the pointed ω-sequence model $\langle \sigma, \langle W, \prec^W, V \rangle \rangle$, since the tail order function supervenes on W. For even more brevity we sometimes will simply specify a sequence and a valuation $\langle \sigma, V \rangle$; in that case, W is (implicitly) the set of all and only σ's non-empty tails.

It should be clear how this notion of model, together with the semantics given above for our language, is equivalent to the (more standard) presentation of van Fraassen's models we gave in the introduction: on this semantics, $p > q$ is true at a sequence σ just in case σ has a tail at which p is true and q is true at the first such tail, or else σ doesn't have any tails at which p is true.

Van Fraassen's models have two further particular properties: they are *full* and *categorical*:

Definition 3.4. An ω-sequence frame $\langle W, \prec^W \rangle$ is *full* when W is the set of *all and only* ω-sequences over some non-empty P.

Definition 3.5. An ω-sequence frame $\langle W, \prec^W \rangle$ is *categorical* iff $\sigma \in V(p_k)$ iff $\rho \in V(p_k)$ whenever $\sigma(0) = \rho(0)$.

As we will see, however, these two restrictions on models are logically immaterial.

Van Fraassen called the members of the underlying set P 'worlds'. We call them 'protoworlds' to avoid confusion—after all it is not elements of P, but sequences over P, that play the standard model-theoretic role of worlds in assigning truth values to sentences. The choice to call them "worlds" might go along with a metaphysically ambitious take on the significance of the models, on which the contrast between subsets of W that do not divide sequences with the same first element and the rest is taken to model a non-model-relative contrast between "factual"/"objective"/"heavyweight" questions on the one hand and "non-factual"/"subjective"/"lightweight" questions on the other, with categorical (conditional-free) propositions depending only on the answers to the former kind of question, and conditionals depending non-trivially on the latter kind of questions. The former questions might be supposed to be settled by *how things are*, whereas the latter are in some sense mere expressions of the way we think, or artifacts of the way we talk (see Khoo 2022 for a recent picture along these lines). A proponent of this metaphysical distinction might think of "worlds" as things that merely answer all *factual* questions; in that case, 'world' will seem a good name for elements of P, since, given van Fraassen's assumptions, a single element of

P is enough to determine the extension of any conditional-free sentences. But we will set aside these questions of metaphysical interpretation here, since they are irrelevant to our logical concerns: even if you hold the contrast between categorical and conditional to be metaphysically chimerical, you could still accept the logic of sequence models; conversely, you could reject that logic while still maintaining that there is an important metaphysical distinction between the categorical and conditional.

3.1 Some variations on ω-sequence models

In the previous section we introduced both ω-sequence models and the more specific class of full, categorical models. *Prima facie*, since not every ω-sequence model is a full and categorical model, one might expect that some sentences that hold in all full and categorical models do not hold in all ω-sequence models. But it turns out that this is not the case: the restriction to full and categorical models makes no difference to the logic. To see why this is the case, the following fact will be useful:

Fact 3.6. Pointed ω-sequence models $\langle \sigma, W, V \rangle$ and $\langle \tau, W', V' \rangle$ are equivalent (that is, verify the same sentences of $\mathcal{L}$) whenever $\sigma^{[j:]} \in V(p_k)$ iff $\tau^{[j:]} \in V'(p_k)$ for all $j, k \in \mathbb{N}$.

The proof is a routine induction on the length of formulae. The intuition is that all that matters in assessing the truth of a sentence at the distinguished sequence in a pointed ω-sequence model is how the tails of that sequence are valued; just as the actual identity of worlds doesn't matter in Kripke semantics, likewise the actual identity of sequences doesn't matter in sequence semantics.

As an immediate consequence of Fact 3.6, we have:

Fact 3.7. Any pointed ω-sequence model $\mathcal{M} = \langle \sigma, W, V \rangle$ is equivalent to the pointed ω-sequence model $\mathcal{M}_\omega = \langle \vec{\mathbb{N}}, \Omega, V' \rangle$ where $\vec{\mathbb{N}}$ is the sequence $\langle 0, 1, 2, 3 \ldots \rangle$ of the natural numbers in their standard order, Ω is the set of non-empty tails of $\vec{\mathbb{N}}$, and $\vec{\mathbb{N}}^{[j:]} \in V'(p_k)$ iff $\sigma^{[j:]} \in V(p_k)$.

We can think of $\mathcal{M}_\omega$ as a kind of minimal representation of $\mathcal{M}$. Its underlying frame is obviously isomorphic to the order frame with domain $\mathbb{N}$ in which $j <_i k$ iff $i \leq j < k$. Thus the logic of the class of all ω-sequence frames is the same as the logic of $\langle \Omega, <^\Omega \rangle$ which is also the same as the logic of the singleton of that order frame.

From Fact 3.7 a number of interesting invariance facts immediately follow. First, we can extend any pointed ω-sequence model to a *full* pointed ω-sequence model in which W includes all ω-sequences over P, extending the valuation to the new sequences however we please, without making any difference to what's true in the model. The logic of full ω-sequence frames is thus the same as the logic of all ω-sequence frames. From the other end, we can prune any pointed ω-sequence frame back to the *generated* frame in which W is just the set of all non-empty tails of the designated sequence without making a difference to what's true in the model. Thus the logic of generated ω-sequence frames is also the same as the logic of all ω-sequence frames.

We can also use Fact 3.7 to show that requiring categoricity makes no difference to the logic. For however a pointed ω-sequence model $\mathcal{M}$ may violate this requirement, the categoricity requirement is automatically satisfied in that model's minimal representation $\mathcal{M}_\omega$, since the sequences in $\mathcal{M}_\omega$ never share an initial element. It follows that the logics of categorical ω-sequence models, full categorical models, full models, and ω-sequence models are identical.

We can also use Fact 3.7 to identify some further conditions which we could, if we wished, impose on ω-sequence models without making a difference to the logic. In our models, we allow the domain to include *eventually cyclic* sequences some of whose tails are identical: for instance, $\langle 1, 2, 1, 2, \ldots \rangle$ is the first, third, fifth, ... tail of $\langle 3, 1, 2, 1, 2, \ldots \rangle$.[16] However, it would not matter if we ruled out eventually cyclic sequences, since none of the sequences in the minimal representation are eventually repeating.[17] For the same reason, it would make no difference if we ruled out *all* repetition of protoworlds within a sequence, so that (e.g.) we cannot have a sequence beginning $\langle 1, 2, 1, \ldots \rangle$.

With all this in hand, it is worth noting from the opposite direction that some approaches which bear a close resemblance to ω-sequence semantics have logics that are very different from the logics we consider, and indeed are orthogonal to C2, rather than strengthening C2. Two noteworthy recent examples are the approach of Bacon (2015), who develops a version of sequence semantics which gives up Reciprocity; and that of Goldstein and Santorio (2021), which marries finite sequence semantics with the domain semantics from Yalcin (2007). We will set aside these approaches, as well as other variants whose logic is orthogonal to C2, focusing instead on semantics corresponding to logics that extend C2.

4 Flattening

We are now in a position to begin answering our main question: what is the logic of ω-sequence models?

Since we were able to present ω-sequence models as a special case of order models, it is immediate from the soundness of C2 with respect to order models that the logic of ω-sequence models includes C2. But it includes

[16] A sequence σ is eventually cyclic iff for some n and m, for all $j : \sigma^{[n+jm:]} = \sigma^{[n:]}$.

[17] It turns out that we also get the same logic if we *require* the sequences to be eventually cyclic. This follows from our completeness theorem for C2.FS, which works by generating models all of whose sequences are eventually cyclic.

more as well. For an especially obvious example of how it goes beyond C2, consider the following modal schema:

4 $$\Box p \to \Box\Box p$$

As is well known, 4 is valid on a modal frame just in case its accessibility relation is transitive. It is consequently valid on ω-sequence frames: τ is accessible from σ just in case τ is a tail of σ, and any tail of a tail of σ is a tail of σ. But 4 is not part of C2, whose frames need not be transitive; again, the modal logic of C2 is KT, which does not include 4.

Another modal schema that is not part of C2 is

H $$(\Diamond p \wedge \Diamond q) \to (\Diamond(p \wedge \Diamond q) \vee \Diamond(q \wedge \Diamond p))$$

H is valid on a modal frame just in case it's *connected*: whenever wRv and wRu, either vRu or uRv. The accessibility relations of ω-sequence frames are connected: if τ and ρ are distinct tails of σ, whichever of them has greater rank is a tail of the other and hence accessible from it. But H is not part of C2 or of even of C2 extended with the 4 axiom, since order frames with transitive accessibility relations need not have connected accessibility relations.

In addition to 4 and H, the logic of ω-sequence frames also includes schemas which essentially involve the conditional. We will discuss two such schemas, which together yield an axiomatization of the the logic of ω-sequences models. In this section, we begin with the following:

Flattening $$p > (pq > r) \leftrightarrow pq > r$$

Let C2.F be the result of adding Flattening to C2 (i.e., the smallest extension of C2 including every instance of Flattening and closed under Detachment and Normality).

It is easy to see that Flattening is valid on ω-sequence frames. The right-hand side is false at a sequence just in case it has a pq-tail and r is false at its first pq-tail. The left-hand side is false just in case it has a p-tail with a pq-tail, and r is false at the first such pq-tail. But any sequence with a pq-tail thus has a p-tail, and in any sequence with a pq-tail, its first pq-tail is identical to the first pq-tail of its first p-tail. So the two sides of Flattening have the same truth-value in any pointed ω-sequence model.

Indeed, to foreshadow a bit, note that this reasoning depends just on the structure of the tailhood relation, not on the ordinal structure of ω-sequences. That means that Flattening is valid on sequence semantics *whatever the domain of the underlying sequence*. Indeed, we will see that C2.F is the logic of ordinal sequence semantics—a variant of ω-sequence semantics where the underlying sequences can take any ordinal as their domain—whereas the logic of ω-sequences is strictly stronger than C2.F.

On the other hand, there are pointed order models in which instances of Flattening are false, such as in the model in Figure 1.

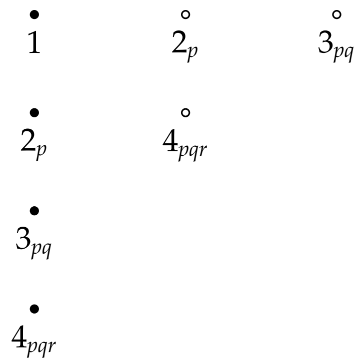

Figure 1: A countermodel to Flattening. Each horizontal line represents the order induced at the left-most (shaded) world, with a world appearing to the left of another just in case the first precedes the second in the relevant ordering, so, e.g., $<_3$ is the empty order, while $<_2$ is the order $\{\langle 2, 4 \rangle\}$. Subscripts indicate atomic valuations. Thus $1 \nVdash pq > r$ while $1 \Vdash p > (pq > r)$. For a counterexample to the opposite direction of the Flattening biconditional, make r true at 3 but false at 4.

It is worth noting a few alternative axiomatizations of this logic. First, Flattening is equivalent to the corresponding schema using the strong conditional connective $\gg$, defined by $p \gg q := \neg(p > \neg q))$ (or equivalently, $p \gg q := \Diamond p \wedge (p > q)$):

$\gg$-Flattening $\qquad p \gg (pq \gg r) \leftrightarrow pq \gg r$

$\gg$-Flattening can be obtained from Flattening by replacing r in Flattening with $\neg r$ and negating both sides. Conversely, $\gg$-Flattening entails that $pq > r$ and $p > (pq > r)$ are equivalent modulo $\Diamond(pq)$; but they are also obviously equivalent when $\Diamond pq$ is false, in which case they are both trivially true.

Second, we can obviously break up Flattening into its two directions

Cautious Importation $\qquad p > (pq > r) \to pq > r$

Cautious Exportation $\qquad (pq > r) \to p > (pq > r)$

We can also give names to the special cases of these principles where r is a contradiction $\bot$. These can be written using $\Diamond$ as

Crashing Cautious Importation $\qquad p > \neg\Diamond pq \to \neg\Diamond pq$

Crashing Cautious Exportation $\qquad \neg\Diamond pq \to p > \neg\Diamond pq$

It turns out that given either one of Cautious Importation and Cautious Exportation, we only need the 'Crashing' restriction of the other to get back the full strength of Flattening. For example, to derive Cautious Exportation from Cautious Importation plus Crashing Cautious Exportation, suppose $pq > r$. If $\neg\Diamond pq$, then $p > \neg\Diamond pq$ and hence $p > (pq > r)$. Otherwise, $\neg(pq > \neg r)$, so by Cautious Importation $\neg(p > (pq > \neg r))$, so by CEM $p > \neg(pq > \neg r)$, so by CEM and Normality, $p > (pq > r)$. The other derivation is analogous.

Third, it is often convenient (especially in working with natural language examples) to use "rule" forms of these axioms, where the conjunction pq is replaced by any q for which we have $\vdash q \rightarrow p$. For example, we can also characterise C2.F as the result of closing C2 under the following rule:

Flattening Rule   If $\vdash q \rightarrow p$ then $\vdash p > (q > r) \leftrightarrow q > r$

This implies Flattening since $\vdash pq \rightarrow p$, and follows from Flattening by the substitution of logical equivalents (since $\vdash q \rightarrow p$ entails that that q is logically equivalent to pq).

C2.F turns out to include many—though not all—of the distinctive principles that hold in ω-sequence models but not all order models. For example, it includes both the modal axioms 4 and H. 4 is actually equivalent to Crashing Cautious Exportation, as we can see by applying the rule form of that axiom to the C2-theorem $\neg p \rightarrow \neg\Box p$ to get

$$\neg\Diamond\neg p \rightarrow \neg\Box p > \neg\Diamond\neg p$$

which simplifies to $\Box p \rightarrow \Box\Box p$. For H, we use Crashing Cautious Importation (in rule form, applied to the tautology $p \rightarrow (p \vee q)$) to get:

$$(p \vee q) > \neg\Diamond p \rightarrow \neg\Diamond p$$

Contraposing and applying CEM, this implies $\Diamond p \rightarrow (p \vee q) > \Diamond p$. By parallel reasoning we also have $\Diamond q \rightarrow (p \vee q) > \Diamond q$, and hence by Normality

$$\Diamond p \wedge \Diamond q \rightarrow (p \vee q) > (\Diamond p \wedge \Diamond q)$$

But CEM, Normality, and Reciprocity together yield the entailment from $(p \vee q) > r$ to $p > r \vee q > r$; applying this to the above, we have:

$$\Diamond p \wedge \Diamond q \rightarrow (p > (\Diamond p \wedge \Diamond q)) \vee (q > (\Diamond p \wedge \Diamond q))$$

Since $\Diamond p$ and $p > q$ entail $\Diamond(p \wedge q)$ in C2, this implies the H axiom:

$$\Diamond p \wedge \Diamond q \rightarrow \Diamond(p \wedge \Diamond q) \vee \Diamond(q \wedge \Diamond p)$$

4.1 Evaluating Flattening

As we mentioned above, one way to interpret order models is as representing relative similarity between worlds. From the point of view of that interpretation, it is no accident that Flattening fails in those models: Flattening is in clear tension with that interpretation of order models.

Schematically, similarity-based theories of the conditional predict Flattening can fail because the most similar pq-world(s) to actuality need not be the most similar pq-world(s) to the p-world(s) most similar to actuality. For a simple concrete example of this, consider a variation on a toy example of Lewis's involving a line L. Lewis's similarity-based intuition was that, given a world w where L has length n, if x and y are otherwise exactly alike, except that in x the length of L is closer to n than it is in y, then x is more similar to w than y is. Now suppose that L is in fact 10 inches long, and compare (1a) and (1b):

(1) a. If L hadn't been strictly between 8–11 inches, then if it hadn't been strictly between 8–13 inches, it would have been 13 inches.
b. If L hadn't been strictly between 8–13 inches, then it would have been 13 inches.

(1a) and (1b) instantiate the two conditionals in Flattening (in the rule-based formulation from the last section) since not being strictly between 8–13 inches entails not being strictly between 8–11 inches. But, if we interpret conditionals via similarity, in particular with Lewis's simple assumption above, then (1a) should be true while (1b) is false. The world x most similar to actuality where the line isn't strictly between 8 and 11 inches is one where it's 11 inches. The world y most similar to x where the line isn't strictly between 8–13 inches is one where it's 13 inches. So (1a) is true. By contrast, the world most like actuality where the line isn't between 8 and 13 inches is one where it's 8 inches, so (1b) is false. (For a counterexample in the opposite direction, change 'it would have been 13 inches' to 'it would have been 8 inches'.)

Is this counterexample convincing? We find it difficult to hear a clear divergence between (1a) and (1b), except by doggedly holding in mind the Lewisian interpretation of 'if p. . . ' as a proxy for 'in the world most similar to actuality where p is true. . . '. Of course, a defender of a similarity-based view could claim that we simply fail to clearly see a contrast which does exist here. But they would need a story about why we make an error here, whereas we have clear intuitions about many other subtle judgments recorded in the literature on conditionals. Barring such a theory, the apparent validity of Flattening might provide a new argument in the battery of well-known arguments against similarity theories of conditionals.

Setting aside the baggage of similarity, we can try to evaluate Flattening on its own terms, by considering pairs of sentences that would be logically equivalent according to Flattening and seeing whether they in fact seem equivalent. It seems to us that the results of this exercise speak in favor of Flattening. For instance, compare these pairs:

(2) a. If Mark and Sue are at the party, there will be a conflagration.
b. If Mark is at the party, then if Mark and Sue are at the party, there will be a conflagration.

(3) a. If he had gotten an espresso and it had been overextracted, he would have had a fit.

b. If he had gotten an espresso, then if he had gotten an espresso and it had been overextracted, he would have had a fit.

These feel pairwise equivalent. At best, the (b)-variants feel *redundant*; the first antecedent feels like it's doing nothing. This is not explained by order semantics, according to which the two variants have logically orthogonal meanings. But this intuition is explained (given standard theories of redundancy) if Flattening is valid, since then the (b)-variants are equivalent to their consequents.

By relying on the rule form of Flattening, we can formulate test pairs that feel somewhat less clunky:

(4) a. If he had gotten an espresso and it had been overextracted, he would have had a fit.
b. If he had gotten an espresso, then he would have had a fit if he'd gotten an overextracted one.

(5) a. If he had been in the south of France, he'd have had a great time.
b. If he had been in France, he'd have had a great time if he had been in the south of France.

Again, these feel pairwise equivalent. We have checked many instances of Flattening, in both the indicative and subjunctive mood, and have not found clear counterexamples.

To be sure, there are superficial counterexamples to Flattening involving tense and anaphora:

(6) a. If John wins, then if John and Sue win, John will have won twice.
b. If John and Sue win, John will have won twice.

(7) a. If a man came in, then if a man came in and a man came in, then three men came in.
b. If a man came in and a man came in, then three men came in.

But it seems implausible that these are counterexamples to Flattening, and more plausible that the felt inequivalence in these pairs arises from different indexing of tense/anaphora in the two pairs. This is a somewhat delicate issue, involving questions about the representation of context-sensitivity that are beyond our scope. But it is worth noting that if we accept these as counterexamples to Flattening, then we also have to accept that there are counterexamples to the very widely accepted 'Contraction' principle $p > (p > q) \leftrightarrow p > q$. For the following also feel pairwise inequivalent:

(8) a. If a man came in, then if a man came in, then two men came in.
b. If a man came in, then two men came in.

Thus Flattening seems, from the point of view of natural language, in at least as good *prima facie* standing as Contraction (a theorem of C2 as well as many weaker logics of conditionals, though one that must notoriously be rejected by non-classical logicians attempting to maintain a naïve theory of truth). This is a strong position to be in.

However, there are reasons for caution about taking these appearances at face value. Flattening is a "cautious" cousin of the well-known Import-Export axiom schema:

Import-Export (IE) $$p > (q > r) \leftrightarrow pq > r$$

The only difference between Flattening and Import-Export is that, in Flattening, p recurs in the antecedent of the conditional consequent on the left-hand side, so we have $p > (pq > r)$, rather than $p > (q > r)$ as in IE. From a logical point of view, this small difference is crucial, for, as Dale (1974), Dale (1979), and Gibbard (1981), showed, adding IE to C2 (or, indeed, to many weaker conditional logics) collapses $>$ to the material conditional, that is, results in a logic that validates:

Materialism $$p > q \leftrightarrow (p \rightarrow q)$$

Materialism is, however, widely rejected in the literature on conditionals, as we noted above; see Edgington 1995 for many arguments against it. For a brief argument, consider the claim that no tree is deciduous if it keeps its leaves through the winter. According to materialism this entails that every tree keeps its leaves through the winter, since the negation of $p \rightarrow q$ entails p. But this is obviously wrong.

However, Flattening does not have the same suspect logical status: we have already seen that it is valid in ω-sequence models; but Materialism is not. For instance, in the ω-sequence model generated from $\sigma = \langle 1, 2, 1, 2, \ldots \rangle$, where p is false at σ and true at $\langle 2, 1, 2, 1, \ldots \rangle$, while q is false at both sequences, the material implication $p \rightarrow q$ is true at σ while the conditional $p > q$ is false. Nor does Flattening lead to any other troubling form of triviality, as ω-sequence models show. Moreover, at least one error theory of the apparent validity of IE precisely relies, in part, on the validity of Flattening (Mandelkern 2024). So validating Flattening may turn out to be a key stepping stone towards explaining the apparent validity of IE.

A final relevant observation is that there are compelling counterexamples to IE in the case of subjunctive conditionals: e.g. the sentences in (9) can intuitively diverge in meaning (Etlin 2008).

(9) a. If the match had lit and it had been soaked in water, then it would have lit.
b. If the match had lit, then it would have lit if it had been soaked in water.

A standard desideratum in the theory of conditionals is to give a unified theory of indicative and subjunctive conditionals: there is one word 'if' which can express both conditionals, depending on the mood of the rest of the sentence. So we have positive reason *not* to validate IE as a matter of the logic of 'if', and instead to explain its apparent validity for indicatives, and lack thereof for subjunctives, as arising from the interaction of the meaning of 'if' with mood. But matters are different for Flattening, which appears valid for both indicatives and subjunctives. The pairs which look like counterexamples to IE for subjunctive conditionals still look equivalent when we change them to instantiate Flattening:

(10) a. If the match had lit and it had been soaked in water, then it would have lit.
b. If the match had lit, then it would have lit if it had been soaked in water and it had lit.

In sum, despite their superficial similarity, Flattening and IE have very different statuses vis-à-vis the theory of conditionals, and reasonable arguments against validating IE, and instead giving some kind of error theory to explain its apparent validity, do not extend to Flattening.

Stepping back, we think the relationship of Flattening to Import-Export is reminiscent of the relationship of the following pairs of principles:

Cautious Transitivity	$(p > q) \wedge (pq > r) \rightarrow (p > r)$
Transitivity	$(p > q) \wedge (q > r) \rightarrow (p > r)$

Cautious Monotonicity	$(p > qr) \rightarrow (pr > q)$
Monotonicity	$(p > q) \rightarrow (pr > q)$

The "incautious" principles Transitivity and Monotonicity are widely rejected (Stalnaker 1968); and, just as for Import-Export, adding either of them to C2 results in Materialism. By contrast, the "cautious" principles are widely accepted, and are theorems of C2 and of many other popular conditional logics. In this light, we might naturally regard Flattening as a plausible *cautious* cousin of the implausible, incautious Import-Export principle.

Nevertheless, we do not want to suggest that the case for the validity of Flattening is anything like watertight. The strongest reason we see to worry about involves the fact that, as already noted, it implies the 4 and H principles for the $\Box$ defined in terms of $>$. This is a potential warning sign, since there are well known arguments against the 4 principle for many seemingly relevant interpretations of $\Box$, and many of these arguments also extend to the H principle. Williamson (2000) and Dorr, Goodman, and Hawthorne (2014) argue against 4 on an interpretation where $\Box$ means '*a* is in a position to know that. . . '. This suggests that 4 might also fail for the epistemic 'must' if its meaning is related to that of 'know'; and the arguments can also be adapted to directly use epistemic 'must'. Insofar as the $\Box$ defined in terms of $>$ interpreted as an indicative conditional is equivalent to, or otherwise intimately connected to, the epistemic modal, these considerations may also threaten the 4 axiom for that $\Box$. Meanwhile, when we turn to the notion of *nomic* necessity—which might be thought to be identical to the defined $\Box$ on some counterfactual interpretations of $>$—we find influential forms of Humeanism which motivate rejection of at least H and perhaps also 4. On the 'best system' theory of laws (Lewis 1994), being a nomic necessity is being entailed by whatever collection of true axioms achieves the best balance of simplicity and strength. On this picture, there could be a complex world where there are two jointly inconsistent simple propositions both of which are false but nomically possible, and (because of their simplicity) such that necessarily, if they are true, they are nomically necessary. This is inconsistent with the connectedness of nomic accessibility, and thus with H. (It is also arguable that Humeans should reject 4 for nomic necessity, though we will not go into that here.) While it is not so plausible that nomically necessary truths are counterfactually necessary on *every* interpretation of counterfactuals, one might think that there are *some* salient of interpretations of counterfactuals that involve "holding the laws fixed" in such a way that these Humean worries would carry over to 4 and H for the defined $\Box$. There are also potential reasons for doubting both 4 and H for *metaphysical* modality, which many take to be equivalent to $\Box$ defined in terms of counterfactuals: Salmon (2005) rejects 4 for metaphysical necessity in order to solve certain puzzles of Tolerance (though see Dorr, Hawthorne, and Yli-Vakkuri 2021 for an alternative approach to those puzzles which preserves 4); meanwhile Bacon (2020) and Bacon and Dorr (2024) explore versions of "combinatorialism" on which metaphysical modality would not obey H.[18]

We will not here undertake to evaluate these arguments against 4 and H for various familiar interpretations of necessity, or adjudicate the question to what extent they carry over to the $\Box$ defined in terms of 'if'. If we accept these arguments (for that defined $\Box$), we will have to reject Flattening, and develop an error theory of its apparent validity. But the appearances favoring Flattening are quite strong, so the difficulty of this task should not be underestimated. In any case, we think it's clear that the logic C2.F has strong prima facie appeal as (at least part of) the logic of the natural language conditional.

4.2 The logic C2.F

As we have already asserted, the logic of ω-sequence frames is not exhausted by C2.F. To see why this is the case, and get some intuition of what is missing, it will be useful to introduce a different class of order frames with respect to which C2.F is complete as well as sound, the *flat* order frames.

Definition 4.1. Order frame $\langle W, < \rangle$ is

[18] On these views, we can have propositions p and q—e.g., the results of predicating two different fundamental properties of some fundamental individual—such that it is metaphysically possible both that $p = q$ and that $p = \neg q$. But $p = q$ entails $\neg\Diamond(p = \neg q)$ and $p = \neg q$ entails $\neg\Diamond(p = q)$, so this violates H.

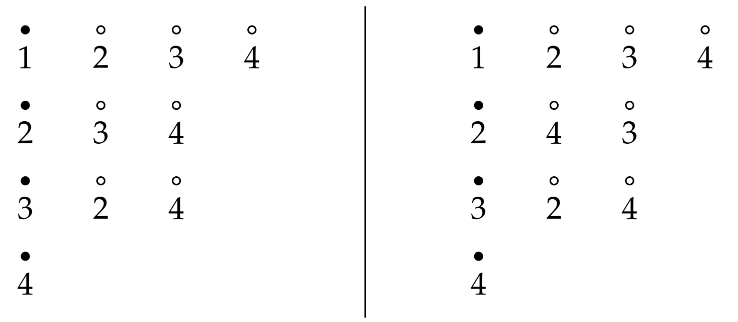

Figure 2: Illustrations of a flat order function (left) and a non-flat (because not semi-flat) order function (right) in frames with worlds {1,2,3,4}.

- *semi-flat* iff for any $x, y, z, w \in W$, if $x <_w y$ and either $y \leq_w z$ or $z \in R(x) \setminus R(w)$, then $y \leq_x z$.
- and *flat* iff it is semi-flat and transitive

Note that in a transitive frame, the case where $x <_w y$ and $z \in R(x) \setminus R(w)$ cannot arise, since $x \in R(w)$ guarantees $R(x) \subseteq R(w)$. So an order frame is flat iff it is transitive and for any $x, y, z, w \in W : x <_w y \leq_w z$ implies $y \leq_x z$. In other words, a flat order frame is a transitive frame where, whenever x is accessible from w, $<_x$ orders all the worlds that come *after* x in $<_w$ in just the same way they were ordered in $<_w$. By contrast worlds that came *before* x in $<_w$ may occur at any position in $<_x$, or not be accessible from x at all. See Figure 2 for an illustration of a flat and non-flat order frame.

We can show that Flattening is valid on all flat order frames. In fact, flatness *characterizes* Flattening, in the sense that the order frames on which Flattening is valid are exactly the flat ones. Since we already know that Flattening is equivalent to the combination of Cautious Importation with Crashing Cautious Exportation, and that the latter is equivalent to 4 (which is characterized by transitivity), it suffices to prove the following lemma:

Lemma 4.2. Cautious Importation is valid on an order frame $\langle W, < \rangle$ iff it is semi-flat.

Proof.

⇒ Suppose $x <_w y$ and for some z, either

(a) $y \leq_w z$ and $y \not\leq_x z$; then

- if $y \notin R(x)$, let $V(p) = \{x, y\}$, $V(q) = \{y\}$, $V(r) = \emptyset$;
- if $y \in R(x)$ and $z \notin R(x)$, let $V(p) = \{x, z\}$, $V(q) = \{z\}$, $V(r) = \emptyset$;
- if $z <_x y$, then let $V(p) = \{x, y, z\}$, $V(q) = \{y, z\}$, $V(r) = \{z\}$.

(b) or $z \in R(x) \setminus R(w)$ and $y \not\leq_x z$. Then let $V(p) = \{x, z, y\}$, $V(q) = \{z, y\}$, $V(r) = \{z\}$.

$\Leftarrow$ Suppose that $\langle W, < \rangle$ is semi-flat, and consider any V and $w \in W$ such that $w \Vdash p > (pq > r)$. If there is no pq-world in $R(w)$, or the first p-world in $<_w$ is a pq-world, then $w \Vdash pq > r$and we are done. Otherwise, let x be the first p-world in $<_w$ and let y be the first pq-world in $<_w$. Since $x <_w y$, $y \in R(x)$, so $R(x)$ contains a pq-world; let u be the first of them according to $<_x$. We know $u \Vdash r$ so it suffices to show $y = u$. We cannot have $u <_w y$ since then y would not be the first pq-world in $<_w$. And we also cannot have $y <_w u$ or $u \notin R(w)$, since in either case semi-flatness would yield $y <_x u$, so u would not be the first pq-world in $<_x$. The only remaining possibility is that $y = u$.

□

Since Flattening is equivalent (modulo C2) to 4 together with Cautious Importation, Flattening is characterized by flatness:

Theorem 4.3. Flattening is valid on an order frame $\langle W, < \rangle$ iff $\langle W, < \rangle$ is flat.

Proof. If $\langle W, < \rangle$ is flat, then Cautious Importation is valid on it by the lemma, and by transitivity 4 and hence Crashing Cautious Exportation are also valid on it, hence Flattening is valid on it. Conversely, if Flattening is valid on the frame, then Cautious Importation is, so it is semi-flat by the lemma; and also 4 is, since 4 is equivalent to Crashing Cautious Exportation, so it is transitive. □

We can also formulate frame conditions that characterize Cautious Exportation and Crashing Cautious Importation; the combination of these conditions is also equivalent to flatness, since the conjunction of the axioms is equivalent to Flattening.[19]

With this characterization result in hand, we can turn to soundness and completeness results for C2.F. The right-to-left direction of Theorem 4.3 says that C2.F is sound for flat order frames: that is, all the theorems of C2.F are valid on every flat order frame. However, completeness is another matter: a characterization result like Theorem 4.3 does not entail a completeness result. Abstractly, a characterization result for a logic L against a background class of frames $\mathcal{F}$ specifies the subset $\mathcal{F}_\mathsf{L}$ of $\mathcal{F}$ such that for all $F \in \mathcal{F}$, $F \in \mathcal{F}_\mathsf{L}$ iff L is valid on F. However, it is possible that L is not complete with respect to $\mathcal{F}_\mathsf{L}$, when there is some sentence p which is valid on every $F \in \mathcal{F}_\mathsf{L}$ but is not a theorem of L: i.e., when the set $\mathcal{F}_\mathsf{L}$ is "too small" to find a countermodel to every non-theorem of L.

[19] Cautious Exportation is valid on $\langle W, < \rangle$ iff it is transitive and for all w, x, y, z, if $x <_w y <_w z$ and $z \in R(x)$, then $y <_x z$. Crashing Cautious Importation is valid on $\langle W, < \rangle$ iff whenever $x <_w y$, $y \in R(x)$.

And indeed, it is well known that there are normal modal logics which are not complete with respect to the class of modal frames that characterize them: in other words, certain formulae that are not theorems of the logic are nevertheless valid on every Kripke frame (Fine 1974; Thomason 1974; van Benthem 1978; see Holliday and Litak 2019 for a helpful recent overview and discussion). So we cannot simply assume that a characterization result yields a corresponding completeness result. Moreover, the fact that C2.F, like many extensions of C2, is not strongly complete for *any* class of order frames (as we showed in Theorem 2.7) means that the standard canonical model method for proving completeness will not work for these logics.

With all that said, we do in fact have a completeness result for C2.F: in the appendix, we show that C2.F is (weakly) complete for flat order frames. Indeed, we show that it is complete for *finite* such frames:

Theorem 4.4. C2.F is weakly complete with respect to finite flat order frames.

Because of its restriction to finite frames (and the fact that flatness is a decidable property of finite frames), this result has the corollary that C2.F is decidable.

One noteworthy corollary of the soundness direction of Theorem 4.3 is that the two modal schemas 4 and H we mentioned in §4 in fact *exhaust* the purely modal content of C2.F. Here are the two axioms again:

4	$\Box p \to \Box\Box p$
H	$(\Diamond p \wedge \Diamond q) \to (\Diamond(p \wedge \Diamond q) \vee \Diamond(q \wedge \Diamond p))$

The modal logic that adds these two axiom schemas to KT is called S4.3, and it is sound and complete (indeed strongly complete) for modal frames with a reflexive, transitive, and connected accessibility relation. It is also weakly complete for *finite* such frames (Bull 1966), which we use to show:

Theorem 4.5. When p belongs to the modal fragment of $\mathcal{L}$, $\vdash_{\mathsf{C2.F}} p$ iff $\vdash_{\mathsf{S4.3}} p$.

Proof. We already established the right-to-left direction in section §4.[20] For the left-to-right direction, suppose p is false in a finite reflexive, transitive, and connected model $v, \langle W, R, V\rangle$. Then (by a standard induction, relying on the transitivity of R) p is also false in the generated submodel $v, \langle R(v), R_v, V_v\rangle$, where R_v and V_v are respectively R and V restricted to $R(v)$. Next we give a recipe to transform $v, \langle R(v), R_v, V_v\rangle$ into an equivalent flat order model $v, \langle R(v), <, V_v\rangle$, yielding a flat order model where p is also false.

To construct $<$, choose a strict linear order $\prec$ on $R(v)$ such that $x \prec y$ implies xRy. Then set $x <_w y$ iff $x \neq y$ and either

[20]By the completeness theorem 4.4, we can also give a simple model-theoretic proof: it is easy to see that every flat order model has an accessibility relation that is reflexive, transitive, and connected; thus if p is true in every reflexive transitive connected modal model, it is also true in every flat order model, and thus a theorem of C2.F by completeness.

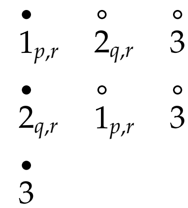

Figure 3: A flat order model which is not equivalent to any ω-sequence model, verifying (a)–(c) but not (d).

- $x = w$ and wRy; or
- wRx, $x \prec y$, and $y \neq w$.

The result obviously agrees with R_v (since, when wRx and $x \prec y$, xRy by our construction of $\prec$ and hence wRy by transitivity). It is obviously irreflexive. It is well-founded by the finitude of $R(v)$. Transitivity of $<_w$ follows from the transitivity of $\prec$. And it is connected: when $w \in R(v)$: for any distinct $x, y \in R(w)$, if $x = w$ or $y = w$ then $x <_w y$ or $y <_w x$ respectively; otherwise by the totality of $\prec$ either $x \prec y$ or $y \prec x$ and hence $x <_w y$ or $y <_w x$, respectively.

This order function is also semi-flat: if $x <_w y$ and either $y \leq_w z$ or $z \in R(x) \setminus R(w)$, then $y \leq_x z$. Since R is transitive, we may ignore the second case, so we must show that if $x <_w y \leq_w z$ then $y \leq_x z$. If $x = w$ this holds trivially. If $x \neq w$ and $y = z$, we just need to show that $y \in R(x)$; but since we have $x <_w y$ and $x \neq w$, we have $x \prec y$ and hence xRy. If $x \neq w$ and $y \neq z$, we must show $y <_x z$. The truth of $x <_w y$ and $y <_w z$ implies that $y \neq w$ and $y \neq x$ and $z \neq w$ and $z \neq y$, and so $x \prec y$ and $y \prec z$, hence $x \prec z$ and so $z \neq x$. And xRy, since $x <_w y$.

□

A corollary of this result is that C2.F cannot be axiomatized by adding any purely modal principles (such as 4 or H) to C2. For it is easy to see that there are non-flat order frames that are reflexive, transitive and connected: the non-flat frame in Figure 2 is an example. Such frames do not validate all of C2.F, but do validate all its purely modal theorems.

5 Sequentiality

C2.F does not include all the sentences that are valid on all ω-sequence frames. To see this, we can appeal to the fact that C2.F is sound for flat order frames (the right-to-left direction of Theorem 4.3). Consider the flat order model in Figure 3. It is easy to see that (a)–(c) are true at 1 while (d) is false at 1:

$$\begin{array}{ll} \text{(a)} & \Box(p \to (\neg p > r)) \\ \text{(b)} & \Box(q \to (\neg q > r)) \\ \text{(c)} & p \vee q \\ \text{(d)} & \neg(p \vee q) > r \end{array}$$

But in any ω-sequence model where (a)–(c) are true, (d) must be true as well. Suppose for contradiction that (a)–(c) are true at σ but (d) is false there. Then σ has a first $\neg(p \vee q)$-tail, at which r is false; let n be the rank of that tail. Since (c) is true at σ, n cannot be 0, so $\sigma^{[n-1:]}$ exists. Either p or q is true at $\sigma^{[n-1:]}$. If p is, then $\sigma^{[n:]}$ is the first $\neg p$-tail of $\sigma^{[n-1:]}$, so $\neg p > r$ is false at $\sigma^{[n-1:]}$, contradicting (a). Similarly, if q is true at $\sigma^{[n-1:]}$, $\sigma^{[n:]}$ is the first $\neg q$ tail of $\sigma^{[n-1:]}$, so $\neg q > r$ is false at $\sigma^{[n-1:]}$, contradicting (b).

Thus the following axiom schema, which says that (d) follows from (a)–(c), is valid on ω-sequence models:

Sequentiality $\Box(p \to (\neg p > r)) \wedge \Box(q \to (\neg q > r)) \to ((p \vee q) \to (\neg(p \vee q) > r))$

Let C2.FS be the result of adding Sequentiality to C2.F (i.e., the smallest extension of C2.F including every instance of Sequentiality and closed under Detachment and Normality). We have just shown that C2.FS is sound for ω-sequence frames; we will presently see that it is also complete for them.

Note that that given the validity of the 4 axiom in C2.F, the necessitation of Sequentiality is equivalent to the following variant version which turns the final material conditional into a strict conditional:

$$\Box(p \to (\neg p > r)) \wedge \Box(q \to (\neg q > r)) \to \Box((p \vee q) \to (\neg(p \vee q) > r))$$

This principle (which hence gives an equivalent axiomatization of C2.F) shows that Sequentiality can be seen as saying that a certain property of propositions—the property of being a proposition p such that $\Box(p \to (\neg p > r))$—is closed under finite disjunction.

Since Sequentiality will play an important role in what follows, it is instructive to consider a special case of it, namely where r is simply the disjunction $p \vee q$. Given our definition of $\Box$ and Normality, this is equivalent to the following:

Restricted Sequentiality

$$\Box(p \to (\neg p > q)) \wedge \Box(q \to (\neg q > p)) \to ((p \vee q) \to \Box(p \vee q))$$

Surprisingly, Restricted Sequentiality turns out to be equivalent to the full strength Sequentiality schema against the background of C2.F: see Appendix G for a proof. Restricted Sequentiality can be false in flat models for the same reason as Sequentiality: in fact since r is necessarily equivalent to $p \vee q$ in the model from Figure 3, this model also demonstrates the invalidity of Restricted Sequentiality in C2.F. Meanwhile, seeing why Restricted Sequentiality holds in ω-sequence models is perhaps even easier than Sequentiality itself. For suppose the antecedent and $p \vee q$ are both true at σ; without loss of generality, we may suppose that p is true at σ. Then as we pop worlds off σ one at a time, if we ever reach a tail where p is false, q will have to be true there; as we continue popping, if we ever reach a tail where q is false, p will have to be true *there*, and so on. Since all the tails of σ will eventually be reached in this process, either p or q (or both) is true at all of them, meaning that $\Box(p \vee q)$ is true at σ.

5.1 Ancestral order models

To better understand what property of ω-sequence models is responsible for the validity of Sequentiality, we can characterize the class of order frames on which C2.FS is valid. In the reasoning we just gave explaining why Sequentiality is valid over ω-sequence frames, the only fact about these frames (apart from their flatness) that we relied on is that every sequence accessible from a given sequence can be reached from it in finitely many steps, where at each step we go from a sequence to its first tail (or, equivalently, that every tail of a sequence has an immediate predecessor). In other words, the key property of ω-sequence frames, which (together with flatness) explains why they validate Sequentiality, is their *ancestrality*, where this is defined as follows:

Definition 5.1. Given an order frame $\langle W, < \rangle$:

- The *successor* of w, $\text{succ}(w)$, is w if $R(w) = \{w\}$, and otherwise the first world after w in $<_w$.
- x is *reachable* from w iff w is related to x by the ancestral of the successor relation.
- Finally, the frame is *ancestral* iff every world accessible from any given world is also reachable from that world.

It is easy to see that every ω-sequence frame is ancestral as well as flat. And while of course not every flat ancestral order frame is an ω-sequence frame, we can show every flat ancestral frame is isomorphic to an ω-sequence frame, so the logic of ω-sequence frames is the logic of flat ancestral frames.

To construct an ω-sequence frame isomorphic to a given flat ancestral frame $\langle W, < \rangle$, we will replace each world in W with its *successor-sequence*, defined as follows:

Definition 5.2. In an order frame $\langle W, < \rangle$, the *successor-sequence* α_w of w is the ω-sequence starting with w where each element after the first is the successor of the previous element.

The set $\{\alpha_w \mid w \in W\}$ is obviously closed under tailhood, and thus is an ω-sequence frame. A few examples should give an intuitive sense for why, when $\langle W, < \rangle$ is flat and ancestral, the function mapping each world to its successor-sequence is an isomorphism from the starting frame to the corresponding ω-sequence frame. Consider the frame $\langle \{1,2,3\}, <^2 \rangle$ illustrated in the center of Figure 4: replacing each member with its successor-sequence yields an ω-sequence frame comprising the three sequences $\langle 1,2,3,3,3\ldots \rangle$, $\langle 2,3,3,3,3\ldots \rangle$,$\langle 3,3,3\ldots \rangle$, which is indeed isomorphic to the one we started with. Likewise, if we start with the frame $\langle \{1,2,3\}, <^3 \rangle$ on the right of Figure 4, we get the isomorphic ω-sequence frame with sequences $\langle 1,2,3,1,2,3\ldots \rangle$, $\langle 2,3,1,2,3,1\ldots \rangle$, $\langle 3,1,2,3,1,2\ldots \rangle$.

By contrast, this procedure *won't* work if we start with a non-ancestral order frame. To get a feel for why not, consider the flat but non-ancestral order frame $<^1$ in Figure 4. The successor-sequence of 1 in this order frame is $\langle 1,2,1,2\ldots \rangle$, since 1 and 2 are each other's successors. The successor sequence of 3 is $\langle 3,3,3,3\ldots \rangle$, since 3 is its own successor. α_3 is hence not accessible from (i.e., not a tail of) α_1, even though 3 *is* accessible from 1. Hence the sequence frame constructed this way is not isomorphic to the starting frame under the successor sequence function. More generally, whenever v is accessible but not reachable from w, α_v will not be accessible from α_w.

To show that flat ancestral frames are always isomorphic to the corresponding successor-sequence frame, consider any flat ancestral frame $F = \langle W, < \rangle$ and $w \in W$. First we show that $u \in R(w)$ iff $\alpha_u \in R(\alpha_w)$. The fact that F is ancestral means that every world accessible from w is reachable from it. And the fact that F is flat, and hence transitive, means every world reachable from w is accessible from it. So $u \in R(w)$ just in case u is reachable from w. But now note that, likewise, α_u is a tail of α_w iff u is reachable from w: if u is reachable from w, u will appear in α_w, and the truncation of α_w at u is obviously α_u; and if u is not reachable from w, u will never appear in α_w, but α_u always begins with u.

Now we need to show that $x <_w y$ iff $\alpha_x \prec_{\alpha_w} \alpha_y$, i.e., iff the first occurrence of x in α_w precedes the first occurrence of y in α_w. To show this, the following definition will be helpful:

- When x is in α_w, the *rank* of x from w is the least n such that x is the nth element of α_w.

Hence we must show that $x <_w y$ iff the rank of x from w is smaller than the rank of y from w:

⇒ Suppose $x <_w y$. If $x = w$, then the claim holds trivially since the rank of x from w is 0. If $x \neq w$, then (by flatness) x also precedes y according to w's successor, and so on until we reach x; since x cannot precede y according to $<_y$, that means that as we take successor steps from w, we must reach x before we reach y.

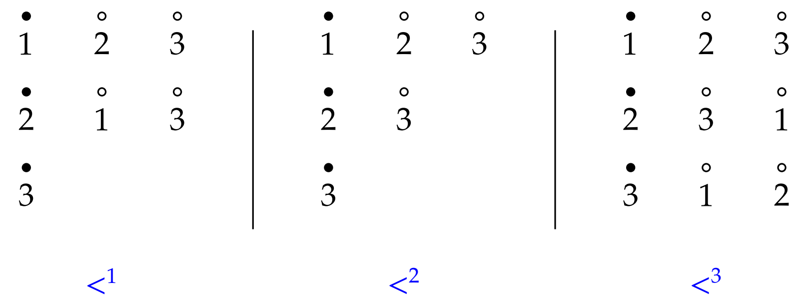

Figure 4: Illustrations of a flat but not ancestral frame $<^1$, and two ancestral frames $<^2, <^3$. In $<^1$ there is no way to get from 1 to 3 by taking successor steps, since you end up stuck in a loop between 1 and 2.

- $\Leftarrow$ Suppose for contradiction that $x \not<_w y$ although the rank of x from w is less than the rank of y from w. Then by connectedness we must have $y <_w x$. But by flatness, for any $u \neq y$, if $y <_u x$, we have $y <_{\text{succ}(u)} x$. So reasoning inductively, we have $y <_u x$ for every u that appears in α_w before y, and in particular $y <_x x$, which is ruled out by the definition of order model.

This isomorphism result both shows the interest of, and puts us in a position to prove, the following characterization result:

Theorem 5.3. C2.FS is valid on an order frame iff it is flat and ancestral.

Proof. Given Theorem 4.3, it suffices to show that a flat order frame validates every instance of Sequentiality iff it is also ancestral.

- $\Rightarrow$ Suppose $\langle W, < \rangle$ is flat but not ancestral, so for some w, v, w can access but not reach v. Set p true at exactly the worlds whose rank from w is even, q true at exactly the worlds whose rank from w is odd, and r true at exactly the worlds reachable from w. Then:
  - w verifies $\Box(p \to \neg p > r)$. Consider any p-world y accessible from w; by our choice of valuation, y is reachable from w. There must be a world reachable from y which has odd rank and hence verifies $\neg p$. There are two options.
    - y is not the highest-ranked world reachable from w, so y's successor is also the next world in α_w and hence has odd rank.
    - y is the highest-ranked world reachable from w. Since v is accessible but unreachable from w, v is accessible but unreachable from y, so y has a successor u other than itself. u's rank from w is strictly less than y's, and hence must be able to reach a world t with odd rank from w, which must hence be reachable from y.

Hence y can reach a $\neg p$-world; but obviously any world reachable from y is reachable from w, and hence (by our choice of valuation) verifies r; and, by flatness, all worlds reachable from y precede in $<_y$ any worlds unreachable from w, so that y verifies $\neg p > r$.

- w verifies $\Box(q \to \neg q > r)$, by parallel reasoning.
- w verifies $p \vee q$, since w verifies p
- w falsifies $\neg(p \vee q) > r$, since the first accessible $\neg(p \vee q)$ worlds in $<_w$ are not reachable from w and hence, by our choice of valuation, falsify r.

$\Leftarrow$ Suppose $\langle W, < \rangle$ is flat and ancestral. Then it is isomorphic to the corresponding successor sequence frame, and we have already shown that every sequence frame validates Sequentiality. □

Since we showed above that an order frame is isomorphic to an ω-sequence frame if it is flat and ancestral, Theorem 5.3 implies that any order frame which is not flat and ancestral is not isomorphic to an ω-sequence frame. So a corollary is that an order frame validates C2.FS iff it is isomorphic to an ω-sequence frame.[21]

Turning to soundness and completeness: the right-to-left direction of Theorem 5.3 is equivalent to the claim that C2.FS is sound for ω-sequence frames. As we emphasized in the discussion before Theorem 4.4, characterization results do not always yield corresponding completeness results. However, once more, our characterization result does indeed point the way towards a completeness result. In Appendix D, we prove a stronger claim which implies that C2.FS is complete for ω-sequence frames:

Theorem 5.4. C2.FS is complete for finite ω-sequence frames.

This immediately implies:

Theorem 5.5. C2.FS is complete for finite flat ancestral order frames.

As before, the restriction to finite frames (and the fact that the question whether a finite order frame is flat and ancestral is decidable) gives us the decidability of C2.FS as a corollary.

One noteworthy application of these results is that they let us exactly pin down the purely modal fragment of C2.FS. The modal logic S4.3.1 is the result of adding every instance of Dum to S4.3:

(Dum) $$\Box(\Box(p \to \Box p) \to p) \to (\Diamond\Box p \to p)$$

[21] Another corollary of this result is that C2.FS is sound over frames $\langle W, \prec \rangle$ where W is a set of ω-sequences *not necessarily closed under tailhood*; such frames are obviously still flat and ancestral, and hence by this result validate C2.FS. This shows that the assumption that ω-sequence frames are closed under tailhood, while convenient in our discussion and faithful to van Fraassen's exposition, is logically irrelevant.

S4.3.1 is sound and (weakly) complete for the singleton of the Kripke frame $\langle \mathbb{N}, \leq \rangle$ (Segerberg 1970). We saw in §3.1 that every pointed ω-sequence model is equivalent to a pointed model based on the frame whose worlds are $\mathbb{N}$ with the order function $j <_i k$ whenever $i \leq j < k$. But this order function corresponds to the accessibility relation $\leq$. Hence a modal sentence is true in some ω-sequence model iff it is true in some model based on $\langle \mathbb{N}, \leq \rangle$, which with our soundness and completeness results yields:

Theorem 5.6. When p is a modal sentence, $\vdash_{\mathsf{C2.FS}} p$ iff $\vdash_{\mathsf{S4.3.1}} p$.

5.2 Evaluating Sequentiality

As with Flattening, we'd like to know whether Sequentiality is in fact valid for the natural language conditional connective. Here we will tentatively argue it is not, highlighting some methodological difficulties as we go.

We'll focus on Restricted Sequentiality, repeated below, which, recall, is equivalent, modulo C2.F, to Sequentiality:

Restricted Sequentiality

$$\Box(p \to (\neg p > q)) \wedge \Box(q \to (\neg q > p)) \to ((p \vee q) \to \Box(p \vee q))$$

To assess an axiom schema like this, the standard methodology would be to find sentences of natural language that instantiate the schema under a plausible translation, and assess whether they strike reflective speakers as valid. But this is difficult to do in the present case. There are two options. We could unpack the $\Box$'s in Restricted Sequentiality into the left-nested conditionals that they abbreviate, and then consider corresponding translations into natural language. However, humans generally struggle to evaluate complex left-nested conditionals, and indeed, in this particular case, the resulting sentences of English are little more than word salad.

The other option is to try to find natural language expressions that express the target $\Box$ defined out of $>$. But it is controversial whether there are words of natural language that express this $\Box$. Even if there are, there is plausibly context-sensitivity *both* in the interpretation of 'if' *and* in the interpretation of natural language necessity modals. To assess Restricted Sequentiality, we must find a natural language modal that is interpreted *throughout the particular context of our counterexample* in the target way.

There are various options to try. Here is an attempt using the adverbial phrase 'in every case' as our necessity operator, which at least makes for a readable example. In particular, we gloss $\Box(p \to q)$ as 'in every case where p, q'.

(11) *The Sequential Assassin's Guild (SAG) comprises two seasoned assassins, p and q. They have a strict back-up policy in place: whenever p kills someone, they make sure that q would otherwise have carried out the assassination, and vice versa. The SAG is deciding whether to kill Nero, and will flip a fair coin to determine whether they in fact do so. Consider:*

a. In every case where p kills Nero, if p hadn't done it, q would have. $\Box(p \rightarrow (\neg p > q))$
b. In every case where q kills Nero, if q hadn't done it, p would have. $\Box(q \rightarrow (\neg q > p))$
c. p or q will kill Nero. $p \vee q$
d. In every case, p or q kills Nero. $\Box(p \vee q)$

Given SAG's iron-clad back-up policies, we can be certain of (11a) and (11b). (11c) has .5 chance: it will be true just in case the fair coin lands heads. But (11d) is certainly false, since killing Nero is a contingent matter: the coin flip could come up either way.

But observe that (11a)–(11c) together entail (11d) via Restricted Sequentiality, so if Restricted Sequentiality were valid, we could not rationally be sure of (11a) and (11b), assign credence .5 to (11c), and assign credence less than .5 to (11d). Hence if 'in every case' and 'if' are coordinated in the right way (so that 'in every case, p' is true iff 'if not p, then p' is true), and have the same interpretations throughout (11), then Restricted Sequentiality is not valid.

Similar cases are easy to multiply (we encourage readers to experiment with other modal phrases to capture the target notion of necessity).[22]

However, it is hard to be sure that the interpretation of 'in every case' is *univocally* tied to 'if' in the right way throughout (11). An alternative hypothesis is that, in (11a) and (11b), we interpret 'in every case' so that Nero being killed *is* a necessity, given that he in fact is killed (since, had the person who killed him not done it, the other one would have), while in (11d) we have a more expansive notion of necessity in mind. It is difficult to rule out this possibility.

So, while (Restricted) Sequentiality has apparent counterexamples in natural language, there is plenty of room to resist them. However, it is hard to see how a clear argument could be formulated *for* the validity of Sequentiality in natural language. And in general, considerations of parsimony suggest that there is no reason to adopt a logic that validates principles for which we lack a positive case. There are countless logical principles that are too complex for humans to assess; no one would advocate simply adopting them all on the grounds that there is no clear case against them.

Given the difficulties of asessing Sequentiality directly, it is natural to wonder whether there might be an axiomatization of C2.FS which is easier to assess. While we cannot rule this out, we can rule out one possibility. As we have seen, part of what makes Sequentiality especially hard to assess is the left-nested conditionals (or, equivalently, $\Box$'s) that it contains. Could we axiomatize C2.FS without left-nesting? The answer is no: for in the fragment

[22] An interesting question is whether conditional mood matters in these cases. One obstacle to assessing Sequentiality for indicatives is that material conditionals with the form $p \rightarrow \neg p > q$—or the corresponding disjunctions $\neg p \vee \neg p > q$—are generally infelicitous when $>$ is indicative, making the premises difficult to assess.

of our language without left-nesting, the logic of ω-sequence models is in fact C2.F. More carefully:

Definition 5.7. The *Boolean language* $\mathcal{L}_B$ is the standard language of propositional logic:

$$q ::= p_k \in At \mid \neg q \mid (q \wedge q)$$

The *Boolean-antecedent* language $\mathcal{L}_{BA}$ is the language which adds conditionals to $\mathcal{L}_B$ exactly when the antecedent is conditional-free

$$p ::= p_k \in At \mid \neg p \mid (p \wedge p) \mid (q > p) : q \in \mathcal{L}_B$$

Theorem 5.8. When $p \in \mathcal{L}_{BA}$, $\vdash_{\mathsf{C2.F}} p$ iff $\vdash_{\mathsf{C2.FS}} p$.

Theorem 5.8 is proved in Appendix F, and shows that the difficulties for evaluating Sequentiality which we have just encountered are ineliminable: the characterization of C2.FS essentially involves left-nested conditionals (or their abbreviation in terms of the defined $\Box$).

A different, less linguistic argument against Sequentiality turns on the observation that, given the characterization result above, any world in any model for C2.FS can access at most countably many worlds under the accessibility relation relevant to the $\Box$ defined out of $>$. If that interpretation of $\Box$ is taken seriously as a model of metaphysical or epistemic modality, or indeed any ordinary interpretation of natural language necessity modals, then this limitation to countable infinity seems implausible. Clearly, a dart can be thrown in such a way that it could hit any point on a dartboard, in any relevant sense of 'could' (epistemic, metaphysical, circumstantial); but there are uncountably many points on a dartboard. This limitation to countable accessible worlds isn't an artifact of the model theory: in a suitably rich object language (for example, higher order logic), one could simply prove that the only points that could have been hit are the one that would have been hit if the actual one hadn't been, the one that would have been hit if that one hadn't been either, and so on, up to at most countably many.

Together with the natural language argument given above, and the methodological case against adopting complex logical principles that lack a positive motivation, this constitutes at least a prima facie case that Sequentiality is an undesirable artifact of ω-sequence semantics—unlike Flattening, which is at least prima facie attractive.

6 Ordinal sequence frames

It is prima facie surprising that a semantics as apparently natural as ω-sequence semantics should give rise to such a peculiar logic as C2.FS, especially when the first step towards our axiomatization, namely the addition of Flattening to C2, is on the face of it much more compelling.

On reflection, however, the class of ω-sequence frames is restricted in a somewhat arbitrary way: namely, to sequences of length ω. There is a natural generalization of this approach which bases the same kind of semantics on *arbitrary ordinal sequences*, which turns out to be sound and complete for C2.F rather than C2.FS.

Definition 6.1. An *ordinal sequence frame* is an order frame $\langle W, \prec^W \rangle$ where W is a set of (possibly transfinite, possibly finite) sequences, and $\prec^W$ is the tail order function on W.

Theorem 6.2. C2.F is sound and weakly complete for ordinal sequence models, and in fact, for ordinal sequence models which are finite, closed under non-empty tailhood, and in which every sequence has length less than ω^ω.

For soundness, it suffices to show that all ordinal sequence models are flat, which is just as for the case of ω-sequences. The completeness result is proved in Appendix C. In fact, it is by showing the completeness of C2.F for ordinal sequence models (and appealing to the fact that these models are flat) that we prove Theorem 4.4 (the completeness of C2.F for flat order models).

Above we showed that flat ancestral order frames are in fact isomorphic to ω-sequence frames. This means that not only do these give rise to the same set of validities, but also to the same binary consequence relation (i.e., the same set of pairs Γ, p such that p is true in every pointed model in the relevant class where Γ is). This is a substantive fact, since in general, two classes of model may agree on a set of validities while disagreeing on a binary consequence relation, in particular for pairs Γ, p where Γ is infinite. And indeed, it turns out that flat order frames and ordinal sequence frames do *not* give rise to the same binary consequence relations.

For an illustration of this, consider the frame in Figure 5, with the domain is $\mathbb{N}$, with $j \prec_i k$ iff (i) $i \leq j$, and $j < k$ or $k < i$; or (ii) $k < j < i$. This frame is not isomorphic to any ordinal sequence frame. And indeed, equipping it with a valuation that sets each atom p_i true at exactly the world i results in a set of sentences which is not true in any ordinal sequence model.

To get a sense for the obstacle that arises if we try to transform this frame into an ordinal sequence frame, think about how you would go about constructing a sequence corresponding to world 0. The natural idea would be to start with the natural numbers in order, followed by a copy of the natural numbers in reverse order. The problem, however, is that such a construction is not a sequence, since there is no ordinal with this order structure.

To make this thought precise, consider the following sets of sentences, whose union is true at 0 in this frame, given the valuation above:

$$\Gamma = \{\neg((p_j \vee p_k) > p_k) \mid j < k\}$$

$$\Delta = \{p_i > \neg((p_j \vee p_k) > p_j) \mid j < i < k\}$$

$$K = \{p_i > \neg((p_j \vee p_k) > p_j) \mid j < k < i\}$$

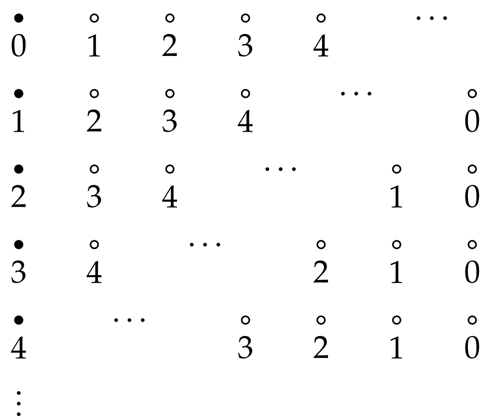

Figure 5: A flat order frame not isomorphic to any ordinal sequence frame.

$$\Lambda = \{\neg\Diamond(p_i \wedge p_j) \mid i \neq j\}$$

Now suppose for contradiction that $\Gamma \cup \Delta \cup K \cup \Lambda$ is true at some ordinal sequence σ. Λ ensures no two atoms are ever true at the same tail of σ. Γ ensures that some first p_i tail precedes some first p_j tail, whenever $i < j$. Δ ensures that when $i < k$, between the first p_i tail and the first p_k tail after it, there is no p_j tail when $j < i$. So, together, $\Gamma \cup \Delta \cup \Lambda$ ensure that σ starts with a stretch of tails verifying just p_0, followed by a stretch verifying just p_1, followed by a stretch verifying just p_2, and so on (possibly interspersed with some stretches verifying no atoms at all).

Finally, K ensures that, after the first p_i tail of σ, for any $j < k < i$, there is a first p_k tail before any p_j tail; in other words, we have a sequence that *descends* towards i of tails that first verify p_{i-1}, then p_{i-2}, then p_{i-3}, and so on, down to p_1 (again, possibly also with stretches verifying no atoms). But since $\Gamma \cup \Delta \cup \Lambda$ ensure that σ starts with an *ascending* sequence, so that the first p_j tail always precedes the first p_k tail when $j < k$, these descending sequences must come after every atom has appeared at least once. Since the ordinals are well-founded, there is a *least* ordinal α such that some atom is true at $\sigma^{[\alpha:]}$ and every atom is true at some earlier tail in σ. Let p_k be the atom true at $\sigma^{[\alpha:]}$. Then there is no p_{k+1}-tail between the first p_{k+2}-tail and $\sigma^{[\alpha:]}$, so the conditional $p_{k+2} > \neg((p_k \vee p_{k+1}) > p_k)$ is false. But this is in K, and so our assumption that σ verifies all these sentences leads to contradiction.

In combination with our incompleteness result in §2.2 (which we framed for C2, but which extends immediately to C2.F), this shows that we have identified at least four, nested notions of consistency for sets of sentences, corresponding to four kinds of frame we have considered, all of which correspond to C2.F and agree as regards finite sets: namely, sets which hold in some finite ordinal sequence or flat order frame;[23] sets which hold in some ordinal sequence frame; sets which hold in some flat order frame; and sets

[23]Where a set like Γ above is inconsistent.

which hold in some *generalized* flat order frame, or, equivalently, from which no contradiction can be derived in C2.F.

6.1 List frames and successor-ordinal frames

In this section we consider two interesting restrictions of ordinal sequence semantics: first, to ordinal sequence frames in which all sequences have finite length; second, to ordinal sequence frames whose sequences all have a *final tail* in the domain (in a sense to be specified). The first class is of obvious interest for reasons of simplicity, and has been discussed in the recent literature (Khoo and Santorio 2018; Khoo 2022); the second, as we will see, is related to the first as ordinal sequence frames are to ω-sequence frames.

Definition 6.3. Given a non-empty set P, a *list* over P is a sequence over P whose domain is a finite ordinal.

A *list frame* is any ordinal sequence frame $\langle W, <^W \rangle$ in which W is a set of lists closed under non-empty tailhood.

The logic of list frames is at least as strong as C2.FS, because any such frame is isomorphic to an ω-sequence frame generated from the ω sequences we obtain by replacing every k-length list with an ω-sequence which agrees with that list up to $k-1$ and then repeats the last world of the list infinitely. In fact, the logic of lists is strictly stronger than C2.FS: it is the logic which strengthens C2.FS with every instance of McKinsey:[24]

McKinsey $\qquad\qquad \Box\Diamond p \to \Diamond\Box p$

McKinsey is not valid in ω-sequence frames: for instance, in the ω-sequence model based on $\langle 1, 2, 1, 2, \ldots \rangle$, where p is true at $\langle 1, 2, 1, 2, \ldots \rangle$ but false at $\langle 2, 1, 2, 1 \ldots \rangle$, $\Box\Diamond p$ is true at $\langle 1, 2, 1, 2, \ldots \rangle$ while $\Diamond\Box p$ is false there. But McKinsey is valid in list frames, since for every list τ and set X of lists containing τ, τ has a tail in X which has no tails in X other than itself. If $\Box\Diamond p$ is true at τ, every tail of τ in X must have some p-tail in X, so the last tail of τ which is in X must be a p-tail which can only access itself.

Theorem 6.4. Let C2.FSM be C2.FS plus every instance of the McKinsey axiom schema. C2.FSM is sound and complete with respect to list models.

For completeness, see Appendix E.

The reasoning that shows that McKinsey is sound for list models will generalize to any ordinal sequence frame where all the sequences have a *final tail* in the domain, no matter how long they are. Hence we can validate McKinsey without also validating Sequentiality. Let a *final* sequence frame be an ordinal sequence frame $\langle W, <^W \rangle$ such that whenever $\sigma \in W$, there

[24] Equivalently, we could add the Grzegorczyk Axiom, $\Box(\Box(p \to \Box p) \to p) \to p$.

exists a β such that $\sigma^{[\beta:]} \in W$ and whenever $\alpha \geq \beta$ and $\sigma^{[\alpha:]}$ is defined and in W, $\sigma^{[\alpha:]} = \sigma^{[\beta:]}$. Natural examples are ordinal sequence frames closed under non-empty tailhood where every sequence's domain is a successor ordinal (i.e., an ordinal with a final element), as well as ordinal sequence frames whose domain includes the empty sequence (which will be a trivial final tail of every sequence).

Theorem 6.5. Let C2.FM be C2.F plus every instance of the McKinsey axiom schema. C2.FM is sound and complete with respect to final ordinal sequence frames.

Completeness is proved in the appendix.

We have now seen four logics which are sound and complete for different classes of ordinal sequence models: namely, C2.F, C2.FS, C2.FM, and C2.FSM. And we have only brushed the surface: for every class of ordinal sequence models, we can ask whether it corresponds to a logic, potentially revealing infinitely many new interesting conditional logics. Indeed, Benjamin Przybocki (p.c.) has reported interesting results along these lines on the axiomatization of classes of ordinal frames in which the length limit is some ordinal strictly between ω and ω^ω.

7 Does Stalnaker's Thesis motivate C2.FS?

Since van Fraassen came up with a class of models that validates both Flatness and Sequentiality as a byproduct of trying to show the nontriviality of a restricted version of Stalnaker's Thesis, one might wonder whether there is some interesting argument from some version of that Thesis to Flattening and/or Sequentiality.[25]

Here is the consistency result van Fraassen proves:

Fact 7.1. Suppose P is a non-empty set, π is a probability measure on P, π^* is the induced product measure on P^ω, and V is a valuation on the ω-sequence frame P^ω in which the denotations of atoms are in the domain of π^* and never distinguish between sequences with the same first element. Then $\pi^*(\llbracket p > q \rrbracket) = \pi^*(\llbracket q \rrbracket \mid \llbracket p \rrbracket)$ whenever $\pi^*(\llbracket p \rrbracket) > 0$ and either both p and q are Boolean (a truth-functional combination of atoms), or one is Boolean and the other is a zero-degree conditional (a conditional with a Boolean antecedent and consequent).[26]

[25] A more direct kind of probabilistic argument for Flattening would follow from a probabilistic argument for the validity of IE (from which Flattening follows). However, while McGee (1989) develops a model for the probabilities of conditionals that validate IE, most have not followed McGee in taking a probabilistic argument for IE seriously, since the resulting construction requires a non-classical interpretation of conditional probability, as well as an unusual conditional logic. We are interested here in whether there is a more conservative probabilistic case for Flattening and/or Sequentiality.

[26] Here, a probability measure on P is a (countably-additive) probability measure whose

Using a straightforward generalization of van Fraassen's proof, we can prove a stronger result along the same lines:

Fact 7.2. Where P, π, π^*, and V are as in Fact 7.1, $\pi^*(\llbracket p > q\rrbracket \mid \llbracket r\rrbracket) = \pi^*(\llbracket q\rrbracket \mid \llbracket p\rrbracket)$ whenever $\pi^*(\llbracket p\rrbracket) > 0$, $\pi^*(\llbracket r\rrbracket \mid \llbracket p\rrbracket) = 1$, and each of p and r is a conjunction of Booleans and zero-degree conditionals.[27]

The proof of Fact 7.2 goes through without modification if, instead of the set of all ω-sequences over P, we consider the set of all α-sequences over P for any other transfinite ordinal α. From the point of view of the probabilities of conditionals, all that will matter about the sequences is their first ω elements. So long as the antecedent p has positive probability and is a conjunction of Booleans and zero-degree conditionals, the set of all sequences for which p is true at their nth tail for some finite n has probability 1, so the new distinctions introduced by allowing transfinite sequences make no difference when it comes to conditionals with positive-probability antecedents.[28]

This also shows that getting the above facts will not require models that strictly validate Flattening—we can easily allow failures of flatness, so long as they only show up among worlds accessible but not reachable from any given world w, since this will still let us represent worlds as ordinal sequences with an order-function that agrees with the tail-order function as regards the sequences *reachable* from any given sequence. This will let us model van Fraassen's result consistent with failures of Flattening having positive probability. Those failures, however, will have to involve conditionals with very strange antecedents (amounting to something equivalent to the negation of Sequentiality) in order to get us out past the initial ω-sequence of worlds in the relevant sequences. So it would be interesting to see if instances of Flattening with Boolean substitution instances play a role in securing the relevant restriction of Stalnaker's Thesis, since this might provide the basis for a simplicity argument for the validity of Flattening.[29]

domain is some σ-algebra of subsets of P. Given any probability measure on P and any index set I, there is a natural product measure π^* on P^I. Informally, π^* treats each $i \in I$ like a fresh draw of a member of P from an urn, with the probabilities on each draw given by π. More carefully, we say that $Y \subseteq P^I$ is a *cylinder set* iff there is a finite set $X \subseteq I$, and a function $g : X \to \mathcal{P}(P)$, such that $g(i)$ is in the domain of π for all $i \in X$, and $Y = \{f \in P^I \mid f(i) \in g(i) \text{ for all } i \in X\}$. The domain of π^* is defined to be the smallest σ-algebra of subsets of P^I that contains all the cylinder sets. π^* is the unique probability function on this σ-algebra such that for any cylinder set $Y = \{f \mid f(i) \in g(i) \text{ for all } i \in X\}$, $\pi^*(Y) = \prod_{i \in X} \pi(g(i))$.

[27] Dorr and Hawthorne (2022) discuss a range of empirical phenomena that can be explained by the extra strength that comes from allowing $r \neq \top$.

[28] The move to transfinite sequences may however introduce new possibilities if we move to a theory of primitive conditional probability like Popper's, or allow for infinitesimal probabilities.

[29] Note that in any model with a probability function satisfying Fact 7.2 (or even the weakening below that requires r and p to be Boolean), the two sides of any instance of Flattening where p, q are Boolean and pq has positive probability will have to have the same

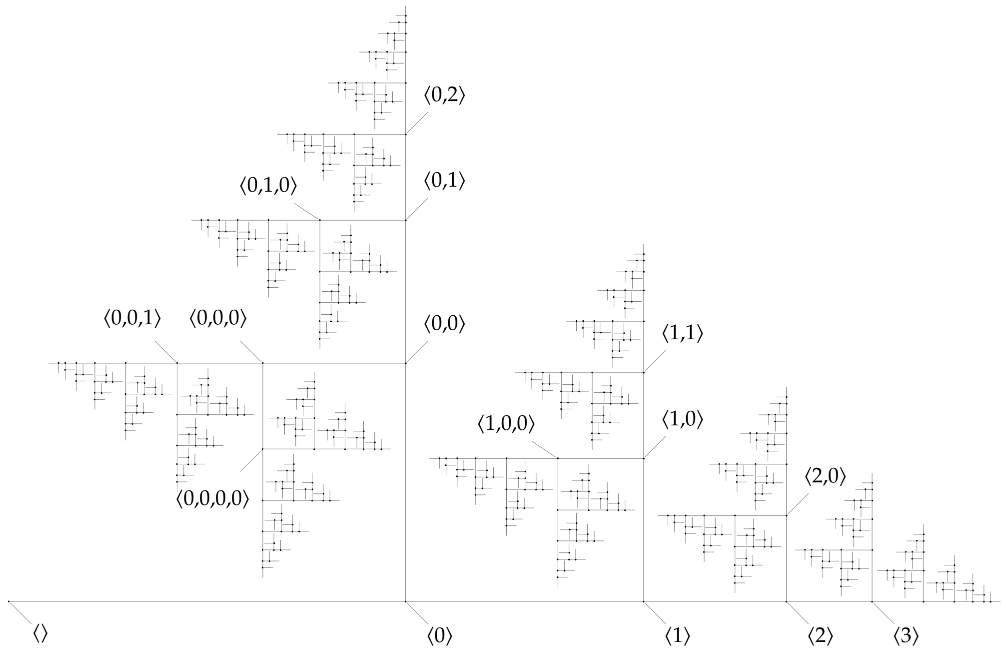

Figure 6: The set $\mathbb{N}^{<\omega}$ of lists of natural numbers, depicted as the domain of a tree.

As it turns out, however, there is variant of van Fraassen's consistency result that uses a different kind of model that does *not* validate Flattening, or anything beyond C2, that provides something pretty close to Fact 7.2. The only way in which we weaken Fact 7.2 is that the antecedent p and the background condition r now both have to be Boolean: we no longer allow zero-degree conditionals or conjunctions thereof. The idea of the variant result is to build an order model whose worlds are not *sequences* of protoworlds, but infinitely-branching *trees* of protoworlds—structures that consist of a "root" protoworld together with a countable infinity of "branches", each of which is itself a tree. Figure 6 depicts the structure of such a tree: a particular tree will attach a protoworld to the root and to every branching point. Formally, we can identify such trees with functions from *lists* of natural numbers to protoworlds:

Definition 7.3. *Trees:*

– For any set X, $X^{<\omega}$ is the set of lists over X.

probability. For $\pi^*([\![p > (pq > r)]\!]) = \pi^*([\![pq > r]\!] \mid [\![p]\!]) = \pi^*([\![r]\!] \mid [\![pq]\!]) = \pi^*([\![pq > r]\!])$. One might see this securing a limited kind of "probabilistic validity" for the two single-premise inference rules corresponding to the two directions of Flattening. However, there is no obvious route from Fact 7.2 to the claim that the instances of Flattening have probability one, even for Boolean p and q. The tree models introduced below show that if we weaken Fact 7.2 to require p and r to be Boolean, we can have models where instances of Flattening with Boolean p and q have probability less than one.

– For a set P, the set of *trees* over P is $P^{\mathbb{N}^{<\omega}}$: i.e., the set of functions from lists of naturals to members of P.

– When τ is a tree over P, the *root* of τ is $\tau(\langle\rangle)$, and the *nth branch* of τ (for any $n \in \mathbb{N}$) is the tree τ' such that $\tau'(\alpha) = \tau(\langle n\rangle + \alpha)$ for any $\alpha \in \mathbb{N}^{<\omega}$, where + is sequence concatenation.

– We can define an order function $\prec^T$ on any set T of trees analogous to the tail order function on sequences: $\rho \prec^T_\tau \sigma$ iff either $\rho = \tau$ and σ is the nth branch of τ for some n, or $\rho \neq \tau$ and $\sigma \neq \tau$ and for some n and m, ρ is the nth branch of τ, and σ is the mth branch of τ, and there is no $k \leq n$ such that σ is the kth branch of τ.

– When P is any non-empty set, a *tree frame* over P is an order model whose domain W is some set of trees over P, with the tree order function $\prec^W$ as above. A *tree model* is an order model based on a tree frame.

– Where π is a probability measure on P, we denote the product measure on $P^{\mathbb{N}^{<\omega}}$ by $\pi^\dagger$.

Theorem 7.4. The logic of tree models is C2.

Proof sketch. Given a finite order frame $\langle W, <\rangle$, we generate an isomorphic tree frame by recursively associating each world w with a tree τ_w over W, setting $\tau_w(\langle\rangle) = w$ and $\tau_w(\langle n\rangle + \alpha) = \tau_v(\alpha)$, where v is the world $n + 1$ steps out from w in $<_w$ if there is one, otherwise the last world in $<_w$. Then we can show that $u <_w v$ iff $\tau_u \prec_{\tau_w} \tau_v$. □

We conjecture moreover that every C2-consistent sentence has nonzero probability (according to some product measure) in some tree frame. Certainly we will have tree models where failures of Flattening (even with Boolean p, q, r) have non-zero probability. For instance, consider three atoms p, q, r and a tree where p is false at the root and q is false at the first p-branch; suppose n is the number of that branch. For the p, q, r-instance of Flattening to be true at such a tree, we would need that the first branch of the n^{th} branch where pq is true and the first branch of the starting tree where pq is true have the same valuation for r. But these valuations are fully independent, and so there is non-zero probability that they differ.

We can now prove our tree-frame analogue of Fact 7.2—exactly the same except that p and r have to be Boolean (whereas before they could be conjunctions of Booleans and zero-degree conditionals).

Fact 7.5. Where π is a probability measure on P, $\pi^\dagger$ is the corresponding product measure on $P^{\mathbb{N}^{<\omega}}$, and V is a valuation on the tree frame $P^{\mathbb{N}^{<\omega}}$ in which denotations of atoms never distinguish between trees with the same root, then $\pi^\dagger([\![p > q]\!] \mid [\![r]\!]) = \pi^\dagger([\![q]\!] \mid [\![p]\!])$ whenever $\pi^\dagger([\![r]\!] \mid [\![p]\!]) = 1$ and both p and r are Boolean.

Proof. We begin with some definitions which apply to any function space P^I (including $P^{\mathbb{N}^{<\omega}}$) and product measure $\pi^\dagger$ on P^I derived from a probability measure π on P. When $X \subseteq I$ and $Y \subseteq P^I$, say that Y *supervenes on* X iff for any $f, g \in P^I$, if $f(i) = g(i)$ for all $i \in X$, then $f \in Y$ iff $g \in Y$. When $Y, Z \subseteq P^I$, say that Y is *orthogonal* to Z iff for some $X \subseteq I$, Y supervenes on X and Z supervenes on $I \setminus X$. When $f : I \to I$ and $Y \subseteq P^I$, $f^*(Y)$ is $\{g \in P^I \mid g \circ f \in Y\}$. Finally, f is *measurable* iff $f^*(Y)$ is in the domain of $\pi^\dagger$ whenever Y is. We rely on two standard facts about product measures:

(i) If $Y, Z \subseteq P^I$ are orthogonal and in the domain of $\pi^\dagger$, $\pi^\dagger(Y \cap Z) = \pi^\dagger(Y)\pi^\dagger(Z)$.

(ii) If $f : I \to I$ is injective and measurable, and $Y \subseteq P^I$ is in the domain of π, $\pi^\dagger(f^*(Y)) = \pi^\dagger(Y)$.

Fix P, π, V, p, q, r such that p and r are Boolean and $\pi^\dagger([\![r]\!] \mid [\![p]\!]) = 1$. If $\pi^\dagger([\![p]\!] \mid [\![r]\!]) = 1$, we are done, since $\pi^\dagger([\![p > q]\!] \mid [\![r]\!]) = \pi^\dagger([\![p \wedge q]\!] \mid [\![r]\!]) = \pi^\dagger([\![q]\!] \mid [\![p]\!])$; so we may assume $\pi^\dagger([\![\overline{p}]\!] \mid [\![r]\!])$ is positive.

Let Y be the set of all trees such that either they have no branch in $[\![p]\!]$, or their first branch in $[\![p]\!]$ is in $[\![q]\!]$. Define two injective functions $f, g : \mathbb{N}^{<\omega} \to \mathbb{N}^{<\omega}$ by $f(\langle\rangle) = \langle 0\rangle$, $f(\langle n\rangle + \alpha) = \langle n + 1\rangle + \alpha$; $g(\alpha) = \langle 0\rangle + \alpha$. Lset:

$Z := f^*(Y)$	i.e., the set of all trees such that either none of their positive-indexed branches is in $[\![p]\!]$, or their first positive-indexed branch in $[\![p]\!]$ is in $[\![q]\!]$
$U := g^*([\![p \wedge q]\!])$	i.e., the set of all trees such that p and q are both true at their zeroth branch
$W := g^*([\![\overline{p}]\!])$	i.e., the set of all trees such that p is false at their zeroth branch

Thus $Y = U \cup (Z \cap W)$. U and W are disjoint, and Z is orthogonal to W, since Z supervenes on the set of lists beginning with a positive integer, while W supervenes on the set of lists beginning with 0. Thus, since g^* and f^* are measurable, we have by the two facts above:

$$\pi^\dagger(Y) = \pi^\dagger(U) + \pi^\dagger(Z)\pi^\dagger(W) = \pi^\dagger([\![p \wedge q]\!]) + \pi^\dagger(Y)\pi^\dagger([\![\overline{p}]\!])$$

and hence

$$\pi^\dagger(Y) = \pi^\dagger([\![p \wedge q]\!])/(1 - \pi^\dagger([\![\overline{p}]\!])) = \pi^\dagger([\![q]\!] \mid [\![p]\!])$$

This gets us close to what we need. Looking at the definition of Y, we see that $p > q$ is true at a tree τ iff either $p \wedge q$ is true at τ, or p is false at τ and $\tau \in Y$. That is, $[\![p > q]\!] = [\![p \wedge q]\!] \cup (Y \cap [\![\overline{p}]\!])$. So,

$$\begin{aligned}\pi^\dagger([\![p > q]\!] \mid [\![r]\!]) &= \pi^\dagger([\![p \wedge q]\!] \mid [\![r]\!]) + \pi^\dagger(Y \cap [\![\overline{p}]\!] \mid [\![r]\!]) \\ &= \pi^\dagger([\![q]\!] \mid [\![p \wedge r]\!])\pi^\dagger([\![p]\!] \mid [\![r]\!]) + \pi^\dagger(Y \mid [\![\overline{p} \wedge r]\!])\pi^\dagger([\![\overline{p}]\!] \mid [\![r]\!])\end{aligned}$$

But $\pi^\dagger([\![q]\!] \mid [\![p \wedge r]\!]) = \pi^\dagger([\![q]\!] \mid [\![p]\!])$, since by hypothesis $\pi^\dagger([\![r]\!] \mid [\![p]\!]) = 1$. And Y is orthogonal to $[\![\overline{p} \wedge r]\!]$ (since the latter supervenes on $\{\langle\rangle\}$, while Y supervenes on the set of non-empty lists), and thus $\pi^\dagger(Y \mid [\![\overline{p} \wedge r]\!]) = \pi^\dagger(Y)$, which is identical to $\pi^\dagger([\![q]\!] \mid [\![p]\!])$ by what we showed above. So we have:

$$\begin{aligned}\pi^\dagger([\![p > q]\!] \mid [\![r]\!]) &= \pi^\dagger([\![q]\!] \mid [\![p]\!])\pi^\dagger([\![p]\!] \mid [\![r]\!]) + \pi^\dagger([\![q]\!] \mid [\![p]\!])\pi^\dagger([\![\overline{p}]\!] \mid [\![r]\!]) \\ &= \pi^\dagger([\![q]\!] \mid [\![p]\!])(\pi^\dagger([\![p]\!] \mid [\![r]\!]) + \pi^\dagger([\![\overline{p}]\!] \mid [\![r]\!])) \\ &= \pi^\dagger([\![q]\!] \mid [\![p]\!]) \qquad \square\end{aligned}$$

By contrast, tree models (with product measures) do *not* sustain a version of Stalnaker's Thesis that allows zero-degree conditionals in the antecedent. For example, $\pi^\dagger([\![p]\!] \mid [\![p > q]\!]) = \pi^\dagger([\![pq]\!])/\pi^\dagger([\![p > q]\!]) = \pi([\![pq]\!])/\pi([\![q]\!] \mid [\![p]\!]) = \pi^\dagger([\![p]\!])$. But $\pi^\dagger([\![(p > q) > p]\!])$ can be strictly less than $\pi^\dagger([\![p]\!])$. We omit the tedious algebraic proof, but to get a sense for why this is, it's helpful to compare the situation with the product measure on ω-sequence models. In that construction, a simple calculation shows that this equality does hold.[30] The pq indices and the $(p > q) \wedge \neg p$ indices get the same probability mass in the two constructions, so we need only compare the $\neg(p > q)$ indices. The key observation is that the information that the first $(p > q)$-tail of a sequence is also a p-tail tells us nothing at all about the tails strictly preceding that one, beyond what we already knew (that they are all $\neg(p > q)$-tails). By contrast, finding out that the first $(p > q)$ branch of a tree is also a p-branch (and thus a pq branch) tells us that some $p\overline{q}$ branch must have preceded this one (since the root verifies $\neg(p > q)$), whereas if the the first $(p > q)$-branch is a $\neg p$-branch, then the first $p\overline{q}$ branch of the tree can come either before or after that branch. So there are fewer $\neg(p > q)$-trees where the first $(p > q)$-branch is a p-branch than there are such sequences.

So, the part of the proof of Fact 7.2 that involves antecedents that are zero-degree conditionals (or conjunctions thereof) turns essentially on the extra strength of sequence models as opposed to tree models. If one thought there were good reasons to want a version of Stalnaker's thesis with the extra strength of Fact 7.2 over Fact 7.5, that might potentially yield an interesting argument for the validity of Flattening. But it is not clear that there are good reasons for wanting this. Note that it is out of the question to have a version of Fact 7.2 that allows r to be a *disjunction* of a Boolean and a zero-degree conditional. Then whenever p and q are Boolean and $\pi([\![pq]\!]) > 0$, we would have $\pi([\![p > q]\!]) = \pi([\![q]\!] \mid [\![p]\!]) = \pi([\![p > q]\!] \mid [\![(p \vee (p > q))]\!])$, and hence $\pi([\![p \vee p > q]\!]) = 1$ and $\pi([\![p > q]\!] \mid \overline{p}) = 1$. This implies that $\pi([\![q]\!] \mid [\![p]\!]) \geq 1 - \pi([\![p]\!])$, which cannot hold for all Boolean p and q unless π is trivial.[31] And

[30] $\pi^*([\![(p > q) > p]\!]) = \pi([\![pq]\!]) + \sum_{n \geq 0} \pi([\![\neg p \vee \neg q]\!])^n \pi([\![p\overline{q}]\!])\pi([\![pq]\!]) = \pi([\![pq]\!]) + \pi([\![p\overline{q}]\!])\pi([\![pq]\!])/(1 - \pi([\![\neg p \vee \neg q]\!])) = \pi([\![p]\!])$.

[31] To specific: it would have to be the case that whenever $0 < \pi([\![p]\!]) < 1/2$, every other q is such that either $\pi([\![q]\!] \mid [\![p]\!]) = 1$ or $\pi([\![q]\!] \mid [\![p]\!]) = 0$. Note that this reasoning did not depend on any assumptions about the logic of the conditional.

even if we forget about r (i.e. restrict the claim to the case where $r = \top$), as Stalnaker showed in his 1974 letter to van Fraassen, we still cannot have the version of Fact 7.2 in which the antecedent p can be a disjunction of a Boolean and a zero-degree conditional unless we are willing to give up C2. For in C2 (and indeed in many popular systems weaker than C2), $(p \vee (p > q)) > pq$ is equivalent to pq, so $\pi([\![(p \vee (p > q)) > pq]\!]) = \pi([\![pq]\!])$ for any π. But so long as $\pi([\![pq]\!]) > 0$, the only way we could have $\pi([\![pq]\!]) = \pi([\![pq]\!] \mid [\![p \vee (p > q)]\!])$ would be for $\pi([\![p \vee (p > q)]\!])$ to be 1, which as we saw above leads to triviality.

Given these limitative results, those of us who (unlike Bacon 2015) are not willing to give up on C2 will need to have some strategy for explaining away any prima facie appeal of strong versions of Stalnaker's Thesis that apply even to antecedents with the forbidden disjunctive form. Perhaps, for example, they will appeal to some special factors that influence the resolution of context-sensitivity in such a way that conditionals embedded in the antecedents of other conditionals tend to be interpreted in some special way, maybe differently from the conditional in which they are embedded (cf. Kaufmann 2023). Whatever we end up saying in response to this challenge, it is a reasonable guess that it will generalize in such a way that it could *also* explain away any remaining appeal of Stalnaker's Thesis for conditionals whose antecedent is a single zero-degree conditional on its own, or a conjunction of such conditionals and Booleans. It seems unlikely that the extra strength of Fact 7.2 over Fact 7.5—lying, as it does, in an awkward terrain strictly between the full strength of unrestricted Stalnaker's Thesis and the natural restriction to Boolean antecedents—could form the basis for a compelling argument for Flattening.

Indeed, as Kaufmann (2023) argues, it is not at all clear that sequence models make the right predictions even about conditionals with left-nesting limited to conjunctions of zero-degree conditionals. Suppose you think John is very likely to go to the party, but dislikes Liam so much that he is very unlikely to go if Liam is going. Then it seems that you could reasonably think that it is *also* quite unlikely that John will go to the party if Liam will go if John goes. But this runs contrary to the application of Stalnaker's Thesis in this case, which, as we saw above, requires $(p > q) > p$ always to have the same probability as p. So one might even see the validation of Stalnaker's Thesis in this case as a *drawback* of sequence models vis-à-vis tree models.

Evaluating this argument raises tricky questions about context-sensitivity which we will not try to settle here. Our main conclusion is that it seems unlikely that apparently unrelated considerations about Stalnaker's Thesis could motivate either Flattening or Sequentiality.

8 Conclusion

Setting aside the logics of material and strict conditionals, the study of classical conditional logic has focused almost exclusively on logics of which C2 is an extension. But we have seen in this paper that van Fraassen's models point the way to a rich array of conditional logics which properly extend C2, without collapsing into the material conditional. Although the original motivation for these models—namely, sustaining a version of Stalnaker's Thesis—does little to support the features of the models responsible for their additional logical strength, the logics they generate are nevertheless quite interesting. Moreover, at least one of these logics, namely C2.F, enjoys considerable *prima facie* plausibility for conditionals in natural language.

Appendix A Preliminaries

We start with some definitions that we will use throughout this appendix.

We continue to use *sequence* for any function whose domain (or 'length') is an ordinal, and *list* for a sequence of finite length.

- $\ell(\sigma)$ is the length of σ.
- For $\alpha < \ell(\sigma)$, σ^α is the value of σ on α; so e.g. $\sigma^0, \sigma^1, \sigma^2$ are the first, second, and third element of σ.
- For $\alpha \le \ell(\sigma)$, $\sigma^{[\alpha:]}$ is the α-tail of σ: that is, $\sigma^{[\alpha:]\beta} = \sigma^{\alpha+\beta}$ for $\beta < \ell(\sigma) - \alpha$.
- For $\alpha \le \ell(\sigma)$, $\sigma^{[:\alpha]}$ is the initial segment of σ of length at most α, i.e. the restriction of σ to α.
- $\sigma^{[\alpha:\beta]}$ is $\sigma^{[:\beta][\alpha:]}$: the length $\beta - \alpha$ segment of σ that starts at α.
- $\sigma^{-1}, \sigma^{-2}, \ldots$ are the last, second last, ... elements of σ: i.e., $\sigma^{-n} = \sigma^\alpha$ where $\ell(\sigma) = \alpha + n$, if such an n exists. (It may not exist, e.g. if the length of σ is an ordinal like ω that doesn't have a last element). Similarly, when $\ell(\sigma) = \alpha + n$, $\sigma^{[:-n]}$ is $\sigma^{[:\alpha]}$.
- $\sigma + \rho$ is the result of concatenating σ with ρ: that is:

$$(\sigma + \rho)(\alpha) = \begin{cases} \sigma^\alpha & \text{when } \alpha < \ell(\sigma) \\ \rho^{\alpha - \ell(\sigma)} & \text{when } \ell(\sigma) \le \alpha < \ell(\sigma) + \ell(\rho) \\ \text{undefined} & \text{otherwise} \end{cases}$$

- $\sigma :: x$ is $\sigma + \langle x \rangle$.
- ρ is a *segment* of σ iff $\rho = \sigma^{[\alpha:\beta]}$ for some α and β; a *tail* of σ iff $\rho = \sigma^{[\alpha:]}$ for some α; and an *initial segment* of σ iff $\rho = \sigma^{[:\alpha]}$ for some α.

Recall some abbreviations for our conditional language:

- $p \gg q := \neg(p > \neg q)$ (which is equivalent in $\mathsf{C2}$ to $\Diamond p \wedge p > q$);
- $\Box p := \neg p > p$;
- $\Diamond p := \neg\Box\neg p$ (which is equivalent in $\mathsf{C2}$ to $\neg(p > \neg p)$).
- $\bot := p_0 \wedge \neg p_0$

Fix a logic L containing $\mathsf{C2}$; talk of consistency, entailment, equivalence, and so on throughout the appendix are relative to L (we will make successively stronger assumptions about L as we go).

For a list of sentences τ, set $\bigwedge \tau = \tau^0 \wedge \cdots \wedge \tau^{-1}$ and $\bigvee \tau = \tau^0 \vee \cdots \vee \tau^{-1}$, with $\bigwedge\langle\rangle = \top$ and $\bigvee\langle\rangle = \bot$. Fix a standard ordering on the sentences, so we can extend the use of $\bigwedge$ and $\bigvee$ from lists to finite sets.

Our key tool throughout the appendices will be a function that takes a list of sentences τ and makes a single sentence $\underline{\tau}$. We define this recursively:

Definition A.1. The function $\underline{\cdot}$ is the function from lists of sentences to sentences such that:

$$\underline{\langle\rangle} := \top$$
$$\underline{\tau :: p} := \begin{cases} \underline{\tau} \wedge (\neg\bigvee\tau \gg p) & \text{if } p \neq \bot \\ \underline{\tau} \wedge (\neg\bigvee\tau > p) & \text{if } p = \bot \end{cases}$$

Example A.2. Suppose $p, q \neq \bot$. Then:

1. $\underline{\langle p\rangle}$ is $\top \wedge (\neg\bot \gg p)$, which is C2-equivalent to p.
2. $\underline{\langle p, q\rangle}$ is $\underline{\langle p\rangle} \wedge (\neg\bigvee\langle p\rangle \gg q)$, equivalent to $p \wedge (\overline{p} > q) \wedge \Diamond\overline{p}$.
3. $\underline{\langle p, q, \bot\rangle}$ is $\underline{\langle p, q\rangle} \wedge (\neg\bigvee\langle p, q\rangle > \bot)$, equivalent to $p \wedge (\overline{p} > q) \wedge \Diamond\overline{p} \wedge \Box(p \vee q)$.

Note that for $\underline{\tau}$ to be consistent, no later element other than $\bot$ can entail any earlier element, since if τ^j entailed τ^k for $j < k$, $(\neg\bigvee\tau^{[:k]}) \gg \tau^k$ would be inconsistent. Also, for $\underline{\tau}$ to be consistent, τ cannot contain any inconsistent sentences other than $\bot$, and if it does include $\bot$, every subsequent element of the list must also be $\bot$: if $p \neq \bot$, $\underline{\tau + \langle\bot, p\rangle}$ is equivalent to $\underline{\tau} \wedge ((\neg\bigvee\tau) > \bot) \wedge ((\neg(\bigvee\tau \vee \bot) \gg p)$, which is inconsistent since the second conjunct entails $\neg\bigvee\tau > \neg p$ while the last conjunct is equivalent to $\neg(\neg\bigvee\tau > \neg p)$.

The interest of this list-to-sentence operation turns on the following basic facts.

Lemma A.3. If τ^k entails pq and every element of $\tau^{[:k]}$ entails $\overline{p}$, then $\underline{\tau}$ entails $p > q$; if moreover $\tau^k \neq \bot$, $\underline{\tau}$ entails $p \gg q$.

Proof. We use the following 'Catious Monotonicity' and 'Left Logical Equivalence' properties of any logic including C2:

CMon	$\vdash (p > qr) \rightarrow (pq > r)$
CMon$_\gg$	$\vdash (p \gg qr) \rightarrow (pq \gg r)$
LLE	If $\vdash p \leftrightarrow q$ then $\vdash (p > r) \leftrightarrow (q > r)$
LLE$_\gg$	If $\vdash p \leftrightarrow q$ then $\vdash (p \gg r) \leftrightarrow (q \gg r)$

For the first part, note that if τ^k entails pq, then since $\underline{\tau}$ entails $\neg\bigvee\tau^{[:k]} > \tau^k$, it also entails $\neg\bigvee\tau^{[:k]} > pq$. So by CMon, it entails $((\neg\bigvee\tau^{[:k]}) \wedge p) > q$. But since every member of $\tau^{[:k]}$ entails $\overline{p}$, $(\neg\bigvee\tau^{[:k]}) \wedge p$ is equivalent to p. So by LLE, $\underline{\tau}$ entails $p > q$. The second part is similar using CMon$_\gg$ and LLE$_\gg$. ☐

Lemma A.4. If p is consistent with $\underline{\tau}$, τ does not end with $\bot$, and $\vdash \bigvee\tau \vee q_1 \vee \cdots \vee q_n$, then either p is consistent with $\underline{\tau :: q_i}$ for some q_i, or p is consistent with $\underline{\tau :: \bot}$.

Proof. By induction on the length of τ. The claim holds trivially when τ is $\langle\rangle$, since $\underline{\langle q_i\rangle}$ is equivalent to q_i. For the induction step, we use the following theorem of C2:

$\vee$-Distribution $\quad \vdash (p > (q_1 \vee \cdots \vee q_n)) \rightarrow ((p > q_1) \vee \cdots \vee (p > q_n))$

If $\vdash \bigvee\tau \vee q_1 \vee \cdots \vee q_n$, then $\neg\bigvee\tau > (q_1 \vee \cdots \vee q_n)$ is a theorem, so by Distribution, $p \wedge \underline{\tau}$ must be consistent with $\neg\bigvee\tau > q_i$ for some q_i. If it is moreover consistent with $\neg\bigvee\tau \gg q_i$, that means that p is consistent with $\underline{\tau :: q_i}$; otherwise, p is consistent with $\underline{\tau :: \bot}$. □

Lemma A.5. Suppose X is a finite set of consistent sentences and p and q are sentences such that either $q \neq \bot$ and p is consistent with $\neg\bigvee X \gg q$, or $q = \bot$ and p is consistent with $\neg\bigvee X > q$. Then there is a list τ of elements of X such that p is consistent with $\underline{\tau :: q}$.

Proof. We cover the case where $q \neq \bot$; the other case is similar. We first show that for any list θ of members of X, if $p \wedge (\neg\bigvee X \gg q) \wedge \underline{\theta}$ is consistent, then there is some r in $X \cup \{q\}$ but not in θ such that $p \wedge (\neg\bigvee X \gg q) \wedge \underline{\theta :: r}$ is consistent. This follows from Lemma A.4, since $\bigvee(X \cup \{q, \neg q \wedge \neg\bigvee X\})$ is a theorem, and $\underline{\theta :: r}$ is inconsistent when r is in θ, while $\underline{\theta :: \neg q \wedge \neg\bigvee X}$ and $\underline{\theta :: \bot}$ are both inconsistent with $\neg\bigvee X \gg q$.

But if p were not consistent with $\underline{\tau :: q}$ for any τ consisting entirely of members of X, the relevant r could never be q, so it would have to be true that any θ of elements of X for which $p \wedge (\neg\bigvee X \gg q) \wedge \underline{\theta}$ is consistent can be extended to a longer such list by adding some r in X but not in θ. This is obviously impossible, since X is finite. □

Thanks to these nice properties, we can use the sentence-forming operation $\underline{\cdot}$ to define a hierarchy of 'state descriptions' over a given set of atoms, where the state descriptions of a given depth n consistently settle the truth value of all sentences of modal depth no greater than n that can be built out of those atoms.

Definition A.6. For a given logic L containing $\mathsf{C2}$ and finite non-empty set of atoms A, the sets $Y_{\mathsf{L}}(A, n)$ (the "depth-n L-state descriptions over A") are defined as follows.

- $Y_{\mathsf{L}}(A, 0)$ is the set of all consistent conjunctions that include exactly one of p and $\overline{p}$ for each atom $p \in A$.
- $Y_{\mathsf{L}}(A, n+1)$ is the set of all consistent sentences of the form $\underline{\tau :: \bot}$, where τ is a list of elements of $Y_{\mathsf{L}}(A, n)$.

Example A.7. Let p and q be atoms. Then $Y_{\mathsf{C2}}(\{p\}, 0)$ is $\{p, \overline{p}\}$, and $Y_{\mathsf{C2}}(\{p, q\}, 0)$ is $\{pq, p\overline{q}, \overline{p}q, \overline{pq}\}$.

$Y_{\mathsf{C2}}(\{p\}, 1)$ has four members, $\underline{\langle p, \bot\rangle}$, $\underline{\langle \overline{p}, \bot\rangle}$, $\underline{\langle p, \overline{p}, \bot\rangle}$, and $\underline{\langle \overline{p}, p, \bot\rangle}$, equivalent respectively to $\Box p$, $\Box\overline{p}$, $p \wedge \Diamond\overline{p}$, and $\overline{p} \wedge \Diamond p$. $Y_{\mathsf{C2}}(\{p, q\}, 1)$ contains $\underline{\tau :: \bot}$ for each of the 64 non-empty, non-repeating sequence τ of elements of $Y_{\mathsf{C2}}(\{p, q\}, 0)$.[32] Note that for certain logics L extending $\mathsf{C2}$, some of these

[32] $24 = \frac{4!}{(4-4)!}$ of length 4, $24 = \frac{4!}{(4-3)!}$ of length 3, $12 = \frac{4!}{(4-2)!}$ of length 2, and $4 = \frac{4!}{(4-1)!}$ of length 1.

elements would be inconsistent and hence absent from $Y_{\mathsf{L}}(\{p,q\},1)$: for example, if L includes the axiom schema $p \to \Box p$, $Y_{\mathsf{C2}}(\{p,q\},1)$ would just contain the four sentences $\underline{\langle pq,\bot\rangle}$, $\underline{\langle p\overline{q},\bot\rangle}$, $\underline{\langle \overline{p}q,\bot\rangle}$, and $\underline{\langle \overline{pq},\bot\rangle}$.

Each member of $Y_{\mathsf{C2}}(\{p\},2)$ is of the form $\underline{\tau :: \bot}$ for some list τ of elements of $Y_{\mathsf{C2}}(\{p\},1)$; but not every non-repeating, non-empty τ can appear in this role. If τ begins with $\underline{\langle p,\bot\rangle}$, it cannot contain any elements entailing $\overline{p}$, so the only other element that could appear is $\underline{\langle p,\overline{p},\bot\rangle}$; similarly, if τ begins with $\underline{\langle \overline{p},\bot\rangle}$, the only other element that can appear is $\underline{\langle \overline{p},p,\bot\rangle}$. Meanwhile, if τ begins with $\underline{\langle p,\overline{p},\bot\rangle}$, at least one element entailing $\overline{p}$—either $\underline{\langle \overline{p},\bot\rangle}$or $\underline{\langle \overline{p},p,\bot\rangle}$—must appear later in τ; similarly, if it begins with $\underline{\langle \overline{p},p,\bot\rangle}$, an element entailing p must appear later. 32 lists meet these constraints.[33]

More generally: where $\rho = \langle \tau_0,\dots,\tau_n\rangle$ is a non-repeating list of elements of $Y_{\mathsf{C2}}(A,n)$, $\underline{\rho :: \bot}$ is consistent in $\mathsf{C2}$ (and hence a member of $Y_{\mathsf{C2}}(A,n+1)$) only if τ_0 is the result of deleting all but the first occurrence of each element in $\langle \tau^0_0,\dots,\tau^0_n\rangle$.[34]

Lemma A.8. If $s \in Y_{\mathsf{L}}(A,n)$ and p is a sentence of modal depth $\leq n$ with atoms from A, then either s entails p in $\mathsf{C2}$ or s entails $\overline{p}$ in $\mathsf{C2}$.

Proof. By induction on n. The base case is true since the elements of $Y_{\mathsf{L}}(A,0)$ settle the truth value of every atom in A, hence every Boolean combination of atoms in A. For the induction step, it suffices to show that when $s \in Y_{\mathsf{L}}(A,n+1)$ and p and q have modal depth $\leq n$, s entails one of $p > q$ and $\neg(p > q)$ in $\mathsf{C2}$. Any such s will be of the form $\underline{\tau :: \bot}$ where each element of τ is in $Y_{\mathsf{L}}(A,n)$. Suppose that the first element of $\tau :: \bot$ that entails p also entails q. Then since no previous element entails p, all of them entail $\overline{p}$ by the induction hypothesis; so by Lemma A.3, $\underline{\tau :: \bot}$ entails $p > q$. Otherwise, the first element of $\tau :: \bot$ that entails p does not entail q. This element must be τ^j for some j, since $\bot$ does entail q. By the induction hypothesis, τ^j entails $\overline{q}$ and all elements of $\tau^{[:j]}$ entail $\overline{p}$, so by Lemma A.3, $\underline{\tau :: \bot}$ entails $p \gg \overline{q}$ and hence $\neg(p > q)$. □

Lemma A.9. If p is consistent, it is consistent with some element of $Y_{\mathsf{L}}(A,n)$ for every A and n.

Proof. By induction on n. The base case holds since $\bigvee Y_{\mathsf{L}}(A,0)$ is a tautology. For the induction step, we note that since $\bigvee Y_{\mathsf{L}}(A,n)$ is a theorem by the induction hypothesis, $p \wedge (\neg \bigvee Y_{\mathsf{L}}(A,n) > \bot)$ is consistent whenever p is, so

[33] Two beginning with $\underline{\langle p,\bot\rangle}$, two beginning with $\underline{\langle p,\bot\rangle}$, 14 ($= \frac{3!}{(3-3)!} + \frac{3!}{(3-2)!} + 2$) beginning with $\underline{\langle p,\overline{p},\bot\rangle}$, and 14 beginning with $\underline{\langle \overline{p},p,\bot\rangle}$.

[34] In fact this is the only constraint. Given disjoint pointed order-models $\mathcal{M}_1 \dots \mathcal{M}_\backslash$ for $\underline{\tau_1},\dots,\underline{\tau_n}$, we can construct a new pointed order model by taking the union $\mathcal{M}_1 \dots \mathcal{M}_\backslash$ together with one new world w—the distinguished world—where $<_w$ comprises w followed by the distinguished worlds of $\mathcal{M}_1 \dots \mathcal{M}_\backslash$, and atom p_i is true at w iff entailed by $\underline{\tau_0}$. Then so long as ρ obeys the given constraint, $\underline{\tau_0}$ and hence also $\underline{\rho}$ are true at w.

by Lemma A.5, there is a sequence τ of elements of $Y_{\mathsf{L}}(A, n)$ such that p is consistent with $\underline{\tau :: \bot}$. $\underline{\tau :: \bot}$ is our desired element of $Y_{\mathsf{L}}(A, n+1)$. □

Lemmas A.8 and A.9 together imply that if want to show that every L-consistent sentence has an order model of a certain sort, it suffices to show that every member of every $Y_{\mathsf{L}}(A, n)$ has a model of that sort. For when p has modal depth n and atoms from A, it will be L-consistent with some $s \in Y_{\mathsf{L}}(A, n)$ by Lemma A.9, hence entailed in C2 by this s by Lemma A.8, hence true in any model where s is true (by the soundness of C2 for order models), and hence true in some model since s is.

Appendix B C2 is weakly complete for finite order models

We now turn to our first result: C2 is weakly complete for finite order models. While this result is not new (or at least, is part of the conditionals folklore), proving it now provides an opportunity to showcase the use of some of the definitions from the previous section, which will also be needed for the later completeness theorems for logics which strengthen C2.

Definition B.1. For a given finite set of atoms A and natural number n, we define an order model $\mathcal{M}_{A,n} = \langle W, <, V \rangle$:

- W is $\bigcup_{m \leq n} Y_{\mathsf{C2}}(A, m)$: all of the C2-state descriptions over A of depth no greater than n.
- When s is a depth-0 state description, $t <_s u$ is never true (so $R(s) = \{s\}$). When $s = \underline{\tau :: \bot}$ is a state description of positive depth, $t <_s u$ iff for some i, $u = \tau^i$, and either $t = s$ or t is in $\tau^{[1:i]}$.
- V is the obvious valuation which has s in $V(p_i)$ iff s C2-entails p_i. (Atoms not in A are thus false everywhere.)

Lemma B.2. Each world in $\mathcal{M}_{A,n}$ is true at itself.

Proof. We prove, by induction on complexity, that for every sentence p with atoms in A and every state description s whose depth is not less than the modal depth of p, p is true at s in $M_{A,n}$ iff s entails p. The claim follows as the special case where $p = s$.

- *Atoms:* immediate from the definition of $\mathcal{M}_{A,n}$.
- *Conjunction:* obvious.
- *Negation:* If $\neg p$ is true at s, then p is not true at s, so by the induction hypothesis s does not entail p; since p is of modal depth $\leq m$, it follows by Lemma A.8 that s entails $\neg p$. Conversely, if s entails $\neg p$, then since s is consistent, s does not entail p, so p is false at s by the induction hypothesis, so $\neg p$ is true at s.

– *Conditional:* Suppose $p > q$ is a conditional of modal depth $\leq m$, meaning that p and q must have modal depth $< m$, and $s = \underline{\tau :: \bot}$ is a state-description of depth m.

(i) First suppose $p > q$ is false at s. Then there is some $u \in R(s)$ such that p and $\neg q$ are true at u, while $\neg p$ is true at t whenever $t <_s u$. Since every member of $R(s)$ is a depth m or depth $m - 1$ state description, the induction hypothesis implies that u entails $p \wedge \neg q$ while every t such that $t <_s u$ entails $\neg p$. If $u = s$, s does not entail $p > q$ (since if it did it would be inconsistent, by MP). Otherwise, $u = \tau^i$ for some $i \geq 1$, and we have that τ^j entails $\neg p$ for all $j < i$—including $j = 0$, since τ^0 entails all the depth $< m$ sentences s entails. So by Lemma A.3, s entails $p \gg \neg q$. Since s is consistent, it does not entail $p > q$.

(ii) Next, suppose $p > q$ is true at s. Then there are two cases: either $p > \neg q$ is false at s, or $\Box\neg p$ is true at s. In the former case, by part (i), s does not entail $p > \neg q$. Since $p > \neg q$ has modal depth $\leq m$, it follows by Lemma A.8 that s entails $\neg(p > \neg q)$ and hence also $p > q$ (by CEM). In the latter case, $\neg p$ is true at every world in $R(s)$, so by the induction hypothesis, all of these worlds entail $\neg p$. Hence every member of τ entails $\neg p$. (τ^0 does too because it agrees with s on depth $< m$ sentences.) But s entails $\neg \bigvee \tau > \bot$, so by CMon, s entails $p > \bot$ and hence also $p > q$. □

Given Lemmas A.8 and A.9, this result immediately establishes:

Theorem B.3 (=Theorem 2.5)**.** C2 is complete for finite order models.

Appendix C Completeness for C2.F

Now we turn towards our completeness result for C2.F. We will prove that C2.F is complete for finite ordinal sequence frames. We build on our earlier definitions, now assuming that our underlying logic L includes C2.F. The following definitions will turn out to be helpful, where τ is any list:

Definition C.1.

– τ is *orderly* iff $\underline{\tau}$ is consistent, τ is non-empty, any two elements of τ are jointly inconsistent, and $\bot$ never occurs in τ except possibly as the final element.

– τ is *direct* iff τ is orderly and there is a orderly list of elements of τ that has the same last element as τ and includes at least two elements of τ, but does not include every element of τ.

– τ is *circuitous* iff τ is orderly and not direct and has length at least 3.

Example C.2. Let the logic L be $\mathsf{C2.F}$, let $a := pq, b := \overline{p}q, c := p\overline{q}$ (where p and q are atoms), and consider the following length-6 list (:

$$\tau := \langle \underbrace{\langle a,b,c,\bot\rangle}_{\tau^0}, \underbrace{\langle b,a,c,\bot\rangle}_{\tau^1}, \underbrace{\langle c,a,b,\bot\rangle}_{\tau^2}, \underbrace{\langle a,c,\bot\rangle}_{\tau^3}, \underbrace{\langle c,\bot\rangle}_{\tau^4}, \underbrace{\bot}_{\tau^5}\rangle$$

τ is orderly, as can be seen by considering the ordinal sequence model whose domain comprises the following length $\omega + 5$ sequence over $\{a, b, c\}$ and its non-empty tails—six sequences in all—with the valuation where an atom is true at a sequence iff it is entailed by its first element:

$$\langle a, b, a, b, \ldots, c, a, b, a, c\rangle$$

(In fact, $\underline{\tau}$ is an element of $Y_{\mathsf{C2.F}}(\{p,q\}, 2)$—a depth-2 state-description over those atoms.) Moreover, τ is direct, since $\langle \tau^4, \bot\rangle$ is consistent. ($\underline{\langle \tau^4, \bot\rangle}$ is true at the final tail $\langle c\rangle$ in this model.) The initial segments of τ are also orderly, since obviously any initial segment of a orderly list is orderly. $\tau^{[:5]}$ is direct, since $\langle \tau^3, \tau^4\rangle$ is orderly (as witnessed by the tail $\langle a, c\rangle$). $\tau^{[:4]}$ is also direct, since $\langle \tau^1, \tau^3\rangle$ is orderly (as witnessed by the tail $\langle b, a, c\rangle$). $\tau^{[:3]}$, by contrast, is circuitous, since neither $\langle \tau^0, \tau^2\rangle$ nor $\langle \tau^1, \tau^2\rangle$ is orderly. (τ^0 entails $\overline{a} \gg b$ and hence $\overline{\tau^0} \gg b$ (by CMon), which is inconsistent with $\overline{\tau^0} \gg \tau^2$, and likewise τ^1 entails $\overline{b} \gg a$ and hence $\overline{\tau^1} \gg a$ which is inconsistent with $\overline{\tau^1} \gg \tau^2$.) Finally, $\tau^{[:2]}$ and $\tau^{[:1]}$ are orderly but neither direct nor circuitous, since their lengths are less than 3.

The key new facts secured by adding Flattening are as follows.

Lemma C.3. Suppose the logic L includes $\mathsf{C2.F}$, τ is orderly, and every element of $\tau^{[:-1]}$ entails $\overline{q}$. Then if τ^{-1} entails $q > r$, every element of $\tau^{[:-1]}$ is consistent with $q > r$, and if τ^{-1} is consistent and entails $\neg(q > r)$, every element of $\tau^{[:-1]}$ is consistent with $\neg(q > r)$.

Proof. Suppose $k < \ell(\tau)$, every member of $\tau^{[:-1]}$ entails $\overline{q}$, and τ^{-1} entails $q > r$. $\underline{\tau}$ entails $\underline{\tau^{[:k+1]}}$ and $\neg\bigvee\tau^{[:-1]} > \tau^{-1}$, and hence $\underline{\tau^{[:k+1]}} \wedge (\neg\bigvee\tau^{[:-1]} > (q > r))$. Since q entails $\neg\bigvee\tau^{[:-1]}$ and $\neg\bigvee\tau^{[:k]}$, $\neg\bigvee\tau^{[:-1]} > (q > r)$ and $\neg\bigvee\tau^{[:k]} > (q > r)$ are both equivalent to $q > r$ by the Flattening Rule, hence equivalent to each other. Hence, $\underline{\tau^{[:k+1]}} \wedge (\neg\bigvee\tau^{[:k]} > (q > r))$ is consistent. Since $\tau^k \neq \bot$, $\underline{\tau^{[:k+1]}}$ is $\underline{\tau^{[:k]}} \wedge (\neg\bigvee\tau^{[:k]} \gg \tau^k)$. So we can conclude that $\neg\bigvee\tau^{[:k]} \gg (\tau^k \wedge (q > r))$ is consistent. But for this to be the case, $\tau^k \wedge (q > r)$ must also be consistent.

The case where τ^{-1} is consistent and entails $\neg(q > r)$ is similar. Then, $\underline{\tau}$ entails $\neg\bigvee\tau^{[:-1]} \gg \tau^{-1}$ and hence $\underline{\tau^{[:k+1]}} \wedge (\neg\bigvee\tau^{[:-1]} \gg \neg(q > r))$, i.e. $\underline{\tau^{[:k+1]}} \wedge \neg(\neg\bigvee\tau^{[:-1]} > (q > r))$. By the same reasoning as above, this is equivalent to $\underline{\tau^{[:k+1]}} \wedge \neg(\neg\bigvee\tau^{[:k]} > (q > r))$, i.e. $\underline{\tau^{[:k]}} \wedge (\neg\bigvee\tau^{[:k]} \gg \tau^k) \wedge (\neg\bigvee\tau^{[:k]} \gg \neg(q > r))$. This is consistent only if $\tau^k \wedge \neg(q > r)$ is. □

Lemma C.4. If the logic L includes $\mathsf{C2.F}$ and τ is orderly, then for each element $\tau^k \neq \bot$, there is a orderly list ρ of elements of τ such that $\rho^0 = \tau^k$, $\rho^{-1} = \tau^{-1}$, and the elements of $\tau^{[k+1:]}$ all occur, in the same order, in ρ.

Proof. First consider the case where $\tau^{-1} \neq \bot$. Then for each $k < \ell(\tau)$,

$$(\neg\bigvee\tau^{[:k]} \gg \tau^k) \wedge \cdots \wedge (\neg\bigvee\tau^{[:-1]} \gg \tau^{-1})$$

is consistent (since all its conjuncts are conjuncts of $\underline{\tau}$). Since $\neg\bigvee\tau^{[:j]} \vdash \neg\bigvee\tau^{[:k]}$ for all $j > k$, the $\gg$-Flattening Rule says that this is equivalent to

$$\neg\bigvee\tau^{[:k]} \gg (\tau^k \wedge (\neg\bigvee\tau^{[:k+1]} \gg \tau^{k+1}) \wedge \cdots \wedge (\neg\bigvee\tau^{[:-1]} \gg \tau^{-1}))$$

which is thus also consistent. Since $p \gg q$ is consistent only when q is, we can conclude that

(*) $$\tau^k \wedge (\neg\bigvee\tau^{[:k+1]} \gg \tau^{k+1}) \wedge \cdots \wedge (\neg\bigvee\tau^{[:-1]} \gg \tau^{-1})$$

is consistent too. But then, by Lemma A.5 (setting p in that lemma to be $\tau^k \wedge (\neg\bigvee\tau^{[:k+1]} \gg \tau^{k+1}) \wedge \cdots \wedge (\neg\bigvee\tau^{[:-2]} \gg \tau^{-2})$, q to be τ^{-1}, and X to be the set of elements of $\tau^{[:-1]}$), there must be a list ρ of elements of τ, ending with τ^{-1}, such that the conjunction of (*) with $\underline{\theta}$ is consistent. But clearly, given that the elements of τ are pairwise inconsistent, this conjunction can only be consistent if ρ^0 is τ^k and all of $\tau^{[k:]}$ occur in ρ in the same order as in $\tau^{[k:]}$.

The case where $\tau^{-1} = \bot$ is parallel. □

Lemma C.5. When τ is circuitous and t is in $\tau^{[:-1]}$, there is a orderly list that begins with t, ends with τ^{-1}, and contains all and only the elements of τ.

Proof. By Lemma C.4 there is a orderly list ρ beginning with t, ending with τ^{-1}, and containing only elements of τ. But since τ is circuitous, any such list must contain every element of τ. □

We will now describe a function that takes any orderly list τ and returns a (possibly-repeating, possibly-transfinite) sequence $\upharpoonleft\tau$, such that the elements of $\upharpoonleft\tau$ are exactly the elements of $\tau^{[:-1]}$, and the order of their first occurrences in $\upharpoonleft\tau$ is the same as their order in τ.

Definition C.6. We define $\upharpoonleft\tau$ recursively, based on the length of the orderly list τ.

For the base cases, when the length of τ is 1 or 2, $\upharpoonleft\tau := \tau^{[:-1]}$: that is, $\langle\rangle$ if $\ell(\tau) = 1$ and $\langle\tau^0\rangle$ if $\ell(\tau) = 2$.

For the recursion step, when $\ell(\tau) > 2$, there are two cases, depending on whether τ is direct or circuitous.

- Case 1: τ is direct, so there is a orderly list of elements of τ with the same last element as τ and length strictly between 1 and $\ell(\tau)$. Let j be the greatest number such that τ^j is the first element of such a sequence, and let ρ be such a sequence beginning with τ^j. (If there are multiple appropriate sequences beginning with τ^j, choose ρ to be the first one according to some fixed order on sequences.)

If $j = \ell(\tau) - 2$, note that since $\tau^{[:-1]}$ and ρ are both shorter than τ we may assume that $\uparrow$ is defined on them, and define:

$$\uparrow\tau := \uparrow\tau^{[:-1]} + \uparrow\rho$$

If $j < \ell(\tau) - 2$, then we know from Lemma C.4 that there is a orderly list that begins with τ^{-2}, ends with τ^{-1}, and contains exactly the elements of τ. Let θ^{+} be the first such list, and let θ be its initial segment up to and including the occurrence of τ^{j}; note that θ is also orderly and shorter than τ. Then define:

$$\uparrow\tau := \uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\rho$$

It will later be convenient to subsume the previous case (where $j = \ell(\tau) - 2$) to this one by defining θ to be $\langle \tau^{-2} \rangle$ (hence $\uparrow\theta = \langle\rangle$) when $j = k - 2$.

– Case 2: τ is circuitous. Then by Lemma C.5, for each element t of $\tau^{[:-1]}$, there is a orderly list that begins with t, ends with τ^{-1}, and contains exactly the elements of τ. Define a function π such that for each t in $\tau^{[:-1]}$, $\pi(t)$ is such a list: τ itself if t is τ^{0}; otherwise, the earliest such list according to our fixed order. Let $\rho(t) := \pi(t)^{[:-1]}$ (so each $\rho(t)$ is shorter than τ), and $g(t) := \pi(t)^{-2}$. Then define:

$$\uparrow\tau := \uparrow\rho(\tau^{0}) + \uparrow\rho(g(\tau^{0})) + \uparrow\rho(g(g(\tau^{0}))) + \uparrow\rho(g(g(g(\tau^{0})))) + \cdots$$

Example C.7. Fixing L as $\mathsf{C2.F}$, let us compute $\uparrow\tau$ where τ is the example from Example C.2:

$$\tau := \langle \underbrace{\langle a, b, c, \bot \rangle}_{\tau^0}, \underbrace{\langle b, a, c, \bot \rangle}_{\tau^1}, \underbrace{\langle c, a, b, \bot \rangle}_{\tau^2}, \underbrace{\langle a, c, \bot \rangle}_{\tau^3}, \underbrace{\langle c, \bot \rangle}_{\tau^4}, \underbrace{\bot}_{\tau^5} \rangle$$

As we noted in Example C.2, τ is direct (when the logic is $\mathsf{C2.F}$), so we are in case 1. Since $\langle \tau^4, \bot \rangle$ is orderly, $j = 4$ and $\rho = \langle \tau^4, \bot \rangle$ so

$$\uparrow\tau = \uparrow\tau^{[:5]} + \uparrow\langle \tau^4, \bot \rangle = \uparrow\tau^{[:5]} :: \tau^4$$

using the base case for length-2 lists for the second identity. Proceeding to calculate $\uparrow\tau^{[:5]}$, we note that since $\langle \tau^3, \tau^4 \rangle$ is orderly, $\tau^{[:5]}$ is also direct, $j = 3$, $\rho = \langle \tau^3, \tau^4 \rangle$, so

$$\uparrow\tau^{[:5]} = \uparrow\tau^{[:4]} + \uparrow\langle \tau^3, \tau^4 \rangle = \uparrow\tau^{[:4]} :: \tau^3$$

Turning to $\tau^{[:4]}$, we find that this is also direct, since $\langle \tau^1, \tau^3 \rangle$ is orderly. There are no orderly lists of elements that begin with τ^2 and end with τ^3 and have length less than 4, since the only $i < 4$ for which $\langle \tau^2, \tau^i \rangle$ is orderly is 0, and the only $i < 4$ for which $\langle \tau^2, \tau^0 \rangle, \tau^i$ is orderly is 1. So in this case, $j = 1$ and $\rho = \langle \tau^1, \tau^3 \rangle$. θ^{+} is the only orderly list of elements of $\tau^{[:4]}$ beginning with τ^2

and ending with τ^3, namely $\langle\tau^2, \tau^0, \tau^1, \tau^3\rangle$, and θ is thus its initial segment $\langle\tau^2, \tau^0, \tau^1\rangle$. So,

$$\begin{aligned}\uparrow\tau^{[:4]} &= \uparrow\tau^{[:3]} + \uparrow\langle\tau^2, \tau^0, \tau^1\rangle + \uparrow\langle\tau^1, \tau^3\rangle \\ &= \uparrow\tau^{[:3]} + \uparrow\langle\tau^2, \tau^0\rangle + \uparrow\langle\tau^0, \tau^1\rangle + \uparrow\langle\tau^1, \tau^3\rangle \\ &= \uparrow\tau^{[:3]} + \langle\tau^2, \tau^0, \tau^1\rangle\end{aligned}$$

Turning finally to computing $\uparrow\tau^{[:3]}$, we already noted that $\tau^{[:3]}$ is circuitous, so we will be in case 2. The only function meeting the requirements on π is as follows:

$$\begin{aligned}\pi(\tau^0) &= \langle\tau^0, \tau^1, \tau^2\rangle \\ \pi(\tau^1) &= \langle\tau^1, \tau^0, \tau^2\rangle\end{aligned}$$

So, $\rho(\tau^0) = \langle\tau^0, \tau^1\rangle$, $\rho(\tau^1) = \langle\tau^1, \tau^0\rangle$, $g(\tau^0) = \tau^1$, $g(\tau^1) = \tau^0$, and

$$\begin{aligned}\uparrow\tau^{[:3]} &= \uparrow\langle\tau^0, \tau^1\rangle + \uparrow\langle\tau^1, \tau^0\rangle + \uparrow\langle\tau^0, \tau^1\rangle + \uparrow\langle\tau^1, \tau^0\rangle + \cdots \\ &= \langle\tau^0, \tau^1, \tau^0, \tau^1, \ldots\rangle\end{aligned}$$

Combining all of the above, we have

$$\uparrow\tau = \langle\tau^0, \tau^1, \tau^0, \tau^1, \ldots, \tau^2, \tau^0, \tau^1, \tau^3, \tau^4\rangle$$

Note that if we replace each element of this $\omega + 5$-sequence with the depth-0 state-description it entails, we get the sequence

$$\langle a, b, a, b, \ldots, c, a, b, a, c\rangle$$

which we used in Example C.2 as the basis for the ordinal sequence model verifying all the relevant consistency claims. In particular, $\underline{\tau}$ is true in the ordinal sequence model based on either of these sequences, with the obvious valuation. This will be our general strategy: given a depth-n state description $s = \underline{\tau :: \bot}$, we will turn $\uparrow\tau$ into an ordinal sequence model in the obvious way, and we will be able to show that s is true in this model.

Definition C.8. Where L includes $\mathsf{C2.F}$, $n > 0$ and $s = \underline{\tau :: \bot} \in Y_{\mathsf{L}}(A, n)$ (the set of depth-n state descriptions with atoms A), $\mathcal{M}_s$ is the ordinal sequence-model whose sequences are the non-empty tails of $\uparrow(\tau :: \bot)$, with the natural valuation: atom p_i is true at tail σ iff σ^0 entails p_i.

The key thing we need to show is that s is true in $\mathcal{M}_s$. To get there, we will need a few more lemmas. First, we check that $\uparrow$ behaves as advertised:

Lemma C.9. When τ is orderly, the elements of $\uparrow\tau$ are exactly those of $\tau^{[:-1]}$, and their first occurrences in $\uparrow\tau$ come in the same order as in $\tau^{[:-1]}$.

Proof. By induction on the length of τ.

Base cases (1, 2): obvious.

Induction step: If τ is direct, $\uparrow\tau$ is $\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\pi$, where θ and ρ are lists of elements of τ of length $< \ell(\tau)$. By the induction hypothesis, all elements of $\tau^{[:-1]}$ except τ^{-2}already occur in $\uparrow\tau^{[:-1]}$, in the same order in which they occur in $\tau^{[:-1]}$. Moreover, τ^{-2} occurs later in $\uparrow\tau$, either as the first element of $\uparrow\theta$ (if $\uparrow\theta$ is non-empty) or else as the first element of $\uparrow\rho$. And furthermore, neither $\uparrow\theta$ nor $\uparrow\rho$ has any elements not in $\theta^{[:-1]}$ or $\rho^{[:-1]}$ respectively, hence neither has any elements not in $\tau^{[:-1]}$. So all the elements of $\tau^{[:-1]}$ occur in $\uparrow\tau$, in the right order.

If τ is circuitous, $\uparrow\tau$ is

$$\uparrow\rho(\tau^0) + \uparrow\rho(g(\tau^0)) + \uparrow\rho(g(g(\tau^0))) + \uparrow\rho(g(g(g(\tau^0)))) + \cdots$$

where ρ and g are as in Definition C.6, and $\uparrow\rho(\tau^0) = \uparrow\tau^{[:-1]}$. By the induction hypothesis, $\uparrow\tau^{[:-1]}$ comprises exactly the elements of $\tau^{[:-2]}$, with the same order of first occurrence. Meanwhile, τ^{-2} is $g(\tau^0)$, which is the first element of $\rho(g(\tau^0))$ and hence of $\uparrow\rho(g(\tau^0))$, and thus also occurs in $\uparrow\tau$, after all elements of $\tau^{[:-2]}$. And since each subsequent term in the infinite sum is derived by applying $\uparrow$ to a sequence of elements of $\tau^{[:-1]}$, nothing not in $\tau^{[:-1]}$ occurs in any of them. □

We can also observe some tight limits on the 'complexity' of the sequences output by $\uparrow$:

Lemma C.10. For any orderly τ of length $k \geq 2$, $\uparrow\tau$ has length at most ω^{k-2}, and has at most $\frac{3}{2}(k-1)!$ non-empty tails.

Proof. By induction on k.

Base case: when τ has length 2, $\uparrow\tau$ has length 1 with one non-empty tail.

Induction step: Suppose τ is of length $k+1$. If it is direct, then $\uparrow\tau$ is $\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\rho$, where $\tau^{[:-1]}$, θ, and ρ are all of length $\leq k$. By the induction hypothesis, each of $\uparrow\tau^{[:-1]}$, $\uparrow\theta$, and $\uparrow\pi$ is of length at most ω^{k-2} with at most $\frac{3}{2}(k-1)!$ non-empty tails. Thus the length of $\uparrow\tau$ is at most $\omega^{k-2}\cdot 3 \leq \omega^{k-1}$. Also, every non-empty tail of $\uparrow\tau$ is either (i) a non-empty tail of $\uparrow\pi$, or (ii) of the form $\sigma + \uparrow\pi$, where σ is a non-empty tail of $\uparrow\theta$, or (iii) of the form $\sigma + \uparrow\theta + \uparrow\pi$, where σ is a non-empty tail of $\uparrow\tau^{[:-1]}$. So the number of such tails is at most $\frac{9}{2}(k-1)! \leq \frac{3}{2}k!$ (since $k > 2$).

Meanwhile, if τ is circuitous, $\uparrow\tau$ is of the form

$$\uparrow\rho(\tau^0) + \uparrow\rho(g(\tau^0)) + \uparrow\rho(g(g(\tau^0))) + \cdots$$

where g is a function that maps elements of $\tau^{[:-1]}$ to other elements of $\tau^{[:-1]}$, and ρ is a function that maps each element t of $\tau^{[:-1]}$ to a orderly list of length k. By the induction hypothesis, each $\uparrow\rho(t)$ is of length at most ω^{k-2}, so $\uparrow\tau$ has length at most ω^{k-1}. Also, every non-empty tail of $\uparrow\tau$ is of the form

$$\theta + \uparrow\rho(g(t)) + \uparrow\rho(g(g(t))) + \cdots$$

where θ is a tail of $\uparrow\rho(t)$ for some t in $\tau^{[:-1]}$. There are only k such elements t, and by the induction hypothesis, each $\uparrow\rho(t)$ has at most $\frac{3}{2}(k-1)!$ tails. So $\uparrow\tau$ has at most $\frac{3}{2}k!$ tails. □

Definition C.11. For any orderly τ, we define $\uparrow^{+}\tau$ to be $\uparrow\tau :: \tau^{-1}$.

Lemma C.12. Suppose τ is a orderly list of elements of $Y_{\mathsf{L}}(A, n)$, $\sigma :: s$ is a segment of $\uparrow^{+}\tau$, and q, r are sentences of modal depth $< n$ such that every element of σ entails $\neg q$. Then if s entails $q > r$, every element of σ entails $q > r$; and if s entails $\neg(q > r)$ and is consistent, every element of σ entails $\neg(q > r)$.

Proof. By induction on the length of τ. The base cases for 0 and 1 are trivial. Base case for 2: $\uparrow^{+}\tau = \tau$, so the only nontrivial case is where $s = \tau^1$ and $\sigma = \langle\tau^0\rangle$. If τ^1 entails $q > r$ and τ^0 entails $\neg q$, Lemma C.3 says that τ^0 is consistent with $q > r$. Since τ^0 is a depth n state description and q and r are depth $< n$, it follows by Lemma A.8 that it *entails* $q > r$. Similarly, if τ^1 is consistent and entails $\neg(q > r)$, Lemma C.3 implies that $\neg(q > r)$ is consistent with, and hence entailed by, τ^0.

For the induction step, suppose the claim holds for sequences of length $\leq k$, and suppose τ is orderly and of length $k+1$, and $\sigma :: s$ is a segment of $\uparrow^{+}\tau$ such that either s entails $q > r$ or s entails $\neg(q > r)$ and is consistent.

Case 1: τ is direct, so $\uparrow^{+}\tau$ has the form

$$\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow^{+}\rho$$

with θ, t, and ρ sequences of length $\leq k$, as in Case 1 of Definition C.6. If σ is a segment of $\uparrow\tau^{[:-1]}$ or $\uparrow\theta$ or $\uparrow\rho$, $\sigma :: s$ is a segment of $\uparrow^{+}\tau^{[:-1]}$, $\uparrow^{+}\theta$, or $\uparrow^{+}\rho$, so the claim follows from the induction hypothesis. If σ is of the form $\sigma_1 + \sigma_2$ where σ_1 is a tail of $\uparrow\theta$ and σ_2 is an initial segment of $\uparrow\rho$, then we first appeal to the induction hypothesis for ρ to show that every element of σ_2 agrees with s on $q > r$. In particular, σ_2^0 agrees with s on $q > r$, and $\sigma_1 :: \sigma_2^0$ is a segment of $\uparrow^{+}\theta$, so by the induction hypothesis for θ every element of σ_1 also agrees with s on $q > r$. Finally, if σ is of the form $\sigma_1 + \sigma_2$ where σ_1 is a tail of $\uparrow\tau^{[:-1]}$ and σ_2 is an initial segment of $\uparrow\theta + \uparrow^{+}\rho$, then every element of σ_2 agrees with s on $q > r$ by what we just showed. But $\sigma_1 :: \sigma_2^0$ is a segment of $\uparrow^{+}\tau^{[:-1]}$, so by the induction hypothesis applied to $\tau^{[:-1]}$, every element of σ_1 also agrees with s on $q > r$.

Case 2: τ is circuitous. Let the functions ρ and g be as in the definition of $\uparrow$. Then there are four possible subcases:

(i) σ is a segment of $\uparrow\rho(t)$ for some element t of $\tau^{[:-1]}$.

(ii) σ is of the form $\sigma_1 + \sigma_2$, where for some t in $\tau^{[:-1]}$, σ_1 is a tail of $\uparrow\rho(t)$ and σ_2 is an initial segment of $\uparrow\rho(g(t))$.

(iii) σ is of the form

$$\sigma_1 + \uparrow\rho(g(t)) + \cdots + \uparrow\rho(g^n(t)) + \sigma_2$$

where for some t in $\tau^{[:-1]}$, σ_1 is a tail of $\uparrow\rho(t)$ and σ_2 is an initial segment of $\uparrow\rho(g^{n+1}(t))$.

(iv) s is τ^{-1} and σ is of the form

$$\sigma_1 + \uparrow\rho(g(t)) + \uparrow\rho(g(g(t))) + \cdots$$

where for some t in $\tau^{[:-1]}$, σ_1 is a tail of $\rho(t)$.

In subcase (i), we can appeal directly to the induction hypothesis for $\rho(t)$. In subcase (ii), we first use the induction hypothesis for $\rho(g(t))$ to show that every element of σ_2 agrees with s on $q > r$, and then use the induction hypothesis for $\rho(t)$ and the fact that $\sigma_1 :: \sigma_2^0$ is a segment of $\uparrow^+\pi(t)$ to show that every element of σ_1 also agrees with s on $q > r$. In subcase (iii), we first use the same method to show that every element of $\uparrow\rho(g^n(t)) + \sigma_2$ agrees with s on $q > r$. Since every element of $\tau^{[:-1]}$ except $g^{n+1}(t)$ occurs in $\rho(g^n(t))$, and $g^{n+1}(t)$ is the first element of σ_2, and every element of σ is in $\tau^{[:-1]}$, this is already enough to show that every element of σ agrees with s on $q > r$. Finally, in subcase (iv), we first reason that since every element of $\tau^{[:-1]}$ other than $g(g(t))$ occurs in $\uparrow\rho(g(t))$, and $g(g(t))$ is the first element of $\uparrow\rho(g(g(t)))$, the elements of σ are exactly the elements of $\tau^{[:-1]}$. Thus every element of $\tau^{[:-1]}$ entails $\neg q$, and so by Lemma C.3, every element of $\tau^{[:-1]}$, and hence every element of σ, agrees with τ^{-1} ($= s$) on $q > r$. ◻

Lemma C.13. When $s \in Y_L(A, n+1)$, every sentence p of depth $\leq n$ is true at a tail σ in the model $\mathcal{M}_s$ iff p is entailed by σ^0.

Proof. By induction on the complexity of p.

- *Atoms:* given by the valuation.
- *Conjunction:* obvious.
- *Negation:* by Lemma A.8 and the fact that σ^0 is a depth-n state description.
- *Conditional:* Suppose $p = q > r$; then since p is of depth $\leq n$, q and r are of depth $< n$.

  Suppose first $q > r$ is *not* true at σ. Then for some β, $q \wedge \neg r$ is true at $\sigma^{[\beta:]}$, while q is not true at $\sigma^{[\alpha:]}$ for any $\alpha < \beta$. By the induction hypothesis, σ^β entails $q \wedge \neg r$, and σ^α entails $\neg q$ for all $\alpha < \beta$. By MP, σ^β also entails $\neg(q > r)$. Since σ^β is also consistent, by Lemma C.12 σ^α entails $\neg(q > r)$ for all $\alpha < \beta$. In particular σ^0 entails $\neg(q > r)$; since it is consistent, it does not also entail $q > r$.

Meanwhile, if $q > r$ is true at σ, there are two cases. In the first case, $q > \neg r$ is not true at σ, in which case σ^0 entails $\neg(q > \neg r)$ by what we just proved; by CEM, it also entails $q > r$. In the second case, $q > \bot$ is true at σ, meaning that q is false at every tail of σ. By the induction hypothesis, every element of σ entails $\neg q$. Since $\sigma :: \bot$ is a segment of $\uparrow^+(\tau :: \bot)$ and $\bot$ entails $q > r$, we can apply the first part of Lemma C.12 again, to conclude that every element of σ, in particular σ^0, entails $q > r$.

□

Lemma C.14. s is true in $\mathcal{M}_s$.

Proof. When $s \in Y_{\mathsf{L}}(A, n+1)$, $s = \underline{\tau :: \bot}$ for some orderly list τ of elements of $Y_{\mathsf{L}}(A, n)$. Looking back at the definition of $\underline{\cdot}$, it is easy to see that it is true at a sequence in an ordinal sequence model iff the depth-n state-descriptions true at tails of that sequence are exactly those that occur in τ, and the order of their first occurrences is their order in τ. Given the previous lemma, this means it is true in $\mathcal{M}_s$ so long as the elements of $\uparrow(\tau :: \bot)$ are exactly those of τ and their first occurrences are in the same order as in τ. But we already proved that this is the case, as Lemma C.9. □

This establishes our desired completeness result:

Theorem C.15 (= the completeness half of Theorem 6.2)**.** Every consistent sentence of C2.F is true in some finite ordinal sequence model, closed under non-empty tailhood, in which every sequence has length less than ω^ω.

Proof. Given a consistent sentence p of modal depth n with atoms from A, by Lemmas A.8 and A.9, there must be a depth-n state description s over A that entails p. By Lemma C.14, s will be true in $\mathcal{M}_s$, which by Lemma C.10 is finite and obeys the required length limit. By the soundness of C2.FS for ordinal sequence models, p is true in $\mathcal{M}_s$. □

Since ordinal sequence models are also flat order models, it follows that:

Theorem C.16 (=the completeness half of Theorem 4.4)**.** Every consistent sentence of C2.F is true in some finite flat order model.

Appendix D Completeness for C2.FS

We turn next to the stronger logic C2.FS, defined as the result of adding all instances of the following schema to C2.F:

Sequentiality $\quad \Box(p \to \overline{p} > r) \wedge \Box(q \to \overline{q} > r) \to ((p \vee q) \to \neg(p \vee q) > r)$.

Here is the key new fact about this logic:

Lemma D.1. In any logic L including $\mathsf{C2.FS}$, every circuitous list ends with $\bot$.

Proof. Suppose for contradiction that τ is circuitous and $\tau^{-1} \neq \bot$. Then

$$\tau^0 \wedge (\neg \bigvee \tau^{[:-1]} \gg \tau^{-1}) \tag{1}$$

is consistent, since both conjuncts are equivalent to conjuncts of $\underline{\tau}$. But

$$\tau^0 \wedge (\neg \tau^0 \gg \tau^{-1}) \tag{2}$$

is inconsistent, since it is equivalent to $\underline{\langle \tau^0, \tau^{-1} \rangle}$, and the circuitousness of τ means that $\langle \tau^0, \tau^{-1} \rangle$—being a sequence of some but not all elements of τ ending with τ^{-1}—must be inconsistent. Likewise,

$$\bigvee \tau^{[1:-1]} \wedge (\neg \bigvee \tau^{[1:-1]} \gg \tau^{-1}) \tag{3}$$

must also be inconsistent. For if it were consistent, then by Lemma A.5, there would have to be some orderly list of elements of $\tau^{[1:]}$ ending in τ^{-1}, which is also disallowed by the circuitousness of τ.

The inconsistency of (2) means that $\vdash \tau^0 \to (\neg \tau^0 > \neg \tau^{-1})$, and the inconsistency of (3) means that $\vdash \bigvee \tau^{[1:-1]} \to (\neg \bigvee \tau^{[1:-1]} > \neg \tau^{-1})$. Since the logic is closed under necessitation, it follows that $\vdash \Box(\tau^0 \to (\neg \tau^0 > \neg \tau^{-1}))$ and $\vdash \Box(\bigvee \tau^{[1:-1]} \to (\neg \bigvee \tau^{[1:-1]} > \neg \tau^{-1}))$. Noting that $\tau^0 \vee \bigvee \tau^{[1:-1]}$ is just $\bigvee \tau^{[:-1]}$, we can apply Sequentiality to conclude that $\vdash \bigvee \tau^{[:-1]} \to (\neg \bigvee \tau^{[:-1]} > \neg \tau^{-1})$. This implies $\vdash \tau^0 \to (\neg \bigvee \tau^{[:-1]} > \neg \tau^{-1})$, contradicting the consistency of (1). □

Using this, we can show

Lemma D.2. In any logic including $\mathsf{C2.FS}$, whenever τ is orderly and does not end with $\bot$, $\uparrow\tau$ is finite.

Proof. A straightforward induction on the length of τ. □

Lemma D.3. In any logic including $\mathsf{C2.FS}$, whenever τ is orderly, $\uparrow\tau$ is at most of length ω.

Proof. Given the previous lemma, it suffices to prove the result when the last element of τ is $\bot$. We do so by induction on the length of τ. The base cases (1 and 2) are trivial. For the first part of the induction step, suppose τ is direct, so that $\uparrow\tau$ is of the form $\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\pi$. Since neither $\tau^{[:-1]}$ nor θ ends with $\bot$, the first two summands are finite by the previous lemma, and the third summand is at most of length ω by the induction hypothesis, so $\uparrow\tau$ is at most of length ω. For the second part of the induction step, suppose τ is circuitous. Then $\uparrow\tau$ is

$$\uparrow\rho(\tau^0) + \uparrow\rho(g(\tau^0)) + \uparrow\rho(g(g(\tau^0))) + \cdots$$

where each $\rho(t)$ is a non-empty sequence not ending in $\bot$. By the previous lemma, all these sequences are finite; so their join has order type ω. □

Combining this with Lemma C.14, we have:

Theorem D.4. C2.FS is complete for finite ordinal sequence models in which all sequences have order-type at most ω.

In fact we can slightly strengthen this result:

Theorem D.5 (=Theorem 5.4)**.** C2.FS is complete for finite ordinal sequence models of order-type *exactly* ω.

Proof. For any non-empty list σ, let the ω-padding of σ be the ω-sequence that results from repeating the last element of σ ω times. Note that τ is a (non-empty) tail of σ iff the ω-padding of τ is a tail of the ω-padding of σ, and the ω-padding of any list has the same first element of that list. So in any ordinal sequence model that contains some finite sequences, we can replace every such sequence with its ω-padding without disrupting the order relation or the valuation. □

Note too that every ordinal sequence model whose sequences have length at most ω is *ancestral*: every tail of every sequence can be reached by successively deleting the initial element. So we can also draw the following corollary from Theorem D.4:

Theorem D.6 (=Theorem 5.5)**.** C2.FS is complete for finite flat ancestral order models.

Appendix E Adding the McKinsey axiom

In this section we consider logics L that extend C2.FM, the result of adding the McKinsey axiom to C2.F:

McKinsey $\quad\quad \Diamond\Box p \vee \Diamond\Box\neg p$

We'll use the following consequence of McKinsey in the context of S4 (which C2.FS includes):

M* $\quad\quad \vdash \Box(p_1 \vee \cdots \vee p_n) \to (\Diamond\Box p_1 \vee \cdots \vee \Diamond\Box p_n)$

Lemma E.1. Every instance of the schema M* is a theorem of the modal logic S4.M (the result of adding every instance of McKinsey to S4).

Proof. By induction on n. Base case trivial. Induction step: suppose $\Box(p_1 \vee \cdots \vee p_n)$. By McKinsey we have $\Diamond\Box\neg p_n \vee \Diamond\Box p_n$, hence $\Diamond\Box(p_1 \vee \cdots \vee p_{n-1}) \vee \Diamond\Box p_n$. By the induction hypothesis, $\Diamond\Diamond\Box p_1 \vee \cdots \vee \Diamond\Diamond\Box p_{n-1} \vee \Diamond\Box p_n$. So by 4, $\Diamond\Box p_1 \vee \cdots \vee \Diamond\Box p_{n-1} \vee \Diamond\Box p_n$. □

McKinsey gives us the following opposite number for the main lemma with Sequentiality:

Lemma E.2. In any logic L including C2.FM, no circuitous list ends with $\bot$.

Proof. Suppose that τ is orderly and ends with $\bot$; then in particular $\neg(\bigvee \tau^{[:-1]}) > \bot$, i.e. $\Box(\tau^0 \vee \cdots \vee \tau^{-2})$, is consistent. So by Lemma E.1, $\Diamond\Box\tau^0 \vee \cdots \vee \Diamond\Box\tau^{-2}$ is consistent, so there must be some p in $\tau^{[:-1]}$ such that $\Diamond\Box p$ and hence also $\Box p$ is consistent. In that case $\langle p, \bot\rangle$ is orderly, so τ is not circuitous. □

Using this, we can show:

Lemma E.3. In any logic including C2.FM, whenever τ is orderly and ends with $\bot$, the length of $\uparrow\tau$ is a successor ordinal.

Proof. Induction on the length of τ. The base cases (1 and 2) are trivial. For the induction step, suppose τ is orderly and ends with $\bot$: then it is direct by Lemma E.2, so $\uparrow\tau$ is of the form $\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\rho$, where ρ is shorter than τ and also ends with $\bot$; by the induction hypothesis, the length of $\uparrow\rho$ is a successor, so the length of $\uparrow\tau$ is a successor too. □

Combining this with Lemma C.14, we have:

Theorem E.4 (the completeness half of Theorem 6.5)**.** C2.FM is complete for finite ordinal sequence models closed under non-empty tailhood in which the domains of all sequences are successor ordinals.

And putting together this theorem with Theorem D.4, we have

Theorem E.5 (the completeness half of Theorem 6.4)**.** C2.FSM is complete for finite list-models (i.e., ordinal sequence models closed under non-empty tailhood whose domain consists of finitely many ordinal-sequences, each of finite length).

Appendix F Languages without left-nesting

In this section, we show that all theorems of C2.FSM in the language $\mathcal{L}_{BA}$ (in which conditionals are required to have Boolean antecedents) are already theorems of C2.F.

Given a valuation V on a set P, let $\mathcal{M}_{V,\alpha}$ be the ordinal sequence models whose domain is the set of all non-empty sequences over P with length $\leq \alpha$, with the valuation given by applying V to the first element of each sequence. We define a function h that takes a sequence σ in this model's domain and a $\mathcal{L}_{BA}$-sentence p, and returns a set $h(\sigma, p)$ of ordinals in the domain of σ—intuitively, the elements of σ that are "relevant" to the truth value of p at σ. Here is the definition:

$$h(\sigma, p_i) := \{0\} \text{ for } p_i \text{ an atom}$$
$$h(\sigma, \neg p) := h(\sigma, p)$$
$$h(\sigma, p \wedge q) := h(\sigma, p) \cup h(\sigma, q)$$
$$h(\sigma, p > q) := \begin{cases} \{0\} \cup \{\alpha + \beta \mid \beta \in h(\sigma^{[\alpha:]}, q)\} & \text{if } \sigma \text{ has a first } p\text{-tail, } \sigma^{[\alpha:]} \\ \{0\} & \text{otherwise} \end{cases}$$

Obviously $h(\sigma, p)$ is always a finite set of ordinals containing 0.

Any set X of ordinals is well-ordered by $\leq$, and hence there is an order-preserving bijection f_X from X to some ordinal. Thus, for any sequence σ and set X of ordinals in its domain, we can construct a new sequence $\sigma \restriction X$, defined by $(\sigma \restriction X)^\alpha = \sigma^{f_X^{-1}(\alpha)}$. Note that when $\alpha \in X$, we have $\sigma^{[\alpha:]} \restriction \{\beta : \alpha + \beta \in X\} = (\sigma \restriction X)^{[f_X(\alpha):]}$. Since $\sigma \restriction X$ cannot be longer than σ, it is guaranteed to be in the domain of $\mathcal{M}_{V,\alpha}$ if σ is.

Lemma F.1. Suppose X includes every member of $h(\sigma, p)$. Then, in $\mathcal{M}_{V,\alpha}$, the truth value of p at σ is the same as the truth value of p at $\sigma \restriction X$.

Proof. By induction on the complexity of p. For atoms, this follows from the fact that restricting any sequence by a set of ordinals that includes 0 yields a sequence with the same first element. For negation and conjunction it is obvious.

For a conditional $p > q$ (where p is Boolean), note first that if no protoworld where p is true occurs in σ, this will also be true of $\sigma \restriction X$. So suppose that a p-protoworld occurs for the first time at position α in σ. Since no p-protoworlds occur in σ before position α, none occur in $\sigma \restriction X$ before position $f_X(\alpha)$. And of course $(\sigma \restriction X)^{f_X(\alpha)} = \sigma^\alpha$, which is a p-protoworld; and so the truth value of $p > q$ at $\sigma \restriction X$ is the same as the truth value of q at $(\sigma \restriction X)^{[f_X(\alpha):]}$, i.e. at $\sigma^{[\alpha:]} \restriction \{\beta : \alpha + \beta \in X\}$. But $h(\sigma, p > q) = \{0\} \cup \{\alpha + \beta : \beta \in h(\sigma^{[\alpha:]}, q)\} \subseteq X$, so $h(\sigma^{[\alpha:]}, q) \subseteq \{\beta : \alpha + \beta \in X\}$. So by the induction hypothesis, the truth value of q at $\sigma^{[\alpha:]} \restriction \{\beta : \alpha + \beta \in X\}$ is identical to the truth value of q at $\sigma^{[\alpha:]}$, which is in turn identical to the truth value of $p > q$ at σ. □

Taking $X = h(\sigma, p)$, we have:

Corollary F.2. The truth value of any $\mathcal{L}_{BA}$-sentence p at σ in $\mathcal{M}_{V,\alpha}$ is the same as its truth value at the list $\sigma \restriction h(\sigma, p)$.

Finally, since throwing all sequences other than the list $\sigma \restriction h(\sigma, p)$ and its non-empty tails out of the domain of $\mathcal{M}_{V,\alpha}$ will not affect the truth value of any sentence, we can derive:

Theorem F.3 (=Theorem 5.8)**.** Every $\mathcal{L}_{BA}$ sentence p that is consistent in C2.F is true in some finite list-model, and hence also consistent in C2.FSM.

Appendix G Equivalence of different axiomatizations of C2.FS

In this section, we show that the following axiom schemas are equivalent over C2.F. The third, which was not mentioned in the main text, is useful for proving the equivalence of the other two. In earlier versions of this paper (before we hit on Sequentiality) we used a version of it in the official axiomatization of C2.FS.

Sequentiality $\Box(p \to (\overline{p} > r)) \wedge \Box(q \to (\overline{q} > r)) \to ((p \vee q) \to (\neg(p \vee q) > r))$

Restricted Sequentiality

$$\Box(p \to (\overline{p} > q)) \wedge \Box(q \to (\overline{q} > p)) \to ((p \vee q) \to \Box(p \vee q))$$

Conditional Sequentiality

$$((\overline{p} > \overline{q}) > \overline{p}) \wedge ((\overline{q} > \overline{p}) > \overline{q}) \to ((p \vee q) \to \Box(p \vee q))$$

Theorem G.1. For any logic L containing C2.F, the results of adding Sequentiality, Restricted Sequentiality, and Conditional Sequentiality are all identical.

Proof. For the purposes of this proof, let $s := (\overline{p} > \overline{q}) > \overline{p}$ and $t := (\overline{q} > \overline{p}) > \overline{q}$, so Conditional Sequentiality is $st \to ((p \vee q) \to \Box(p \vee q))$.

(a) To derive Conditional Sequentiality from Restricted Sequentiality, it suffices to derive

(1) $$\vdash pst \to \neg pst > qst$$

(2) $$\vdash qst \to \neg qst > pst$$

For given these, we can apply necessitation and then Restricted Sequentiality to derive

(3) $$\vdash (pst \vee qst) \to \Box(pst \vee qst)$$

which implies

(4) $$\vdash st \to ((p \vee q) \to \Box(p \vee q))$$

So, let's see how to establish (1); the proof of (2) will be parallel. For contradiction, assume p, s, t, and $\neg(\neg pst > qst)$. By CEM we have $\neg pst > \neg qst$ and hence (a) $\neg pst > (st \to \overline{p}\overline{q})$. The converse (b) $(st \to \overline{p}\overline{q}) > \neg pst$ is trivially true. Noting that the antecedents $\overline{p} > \overline{q}$ and $\overline{q} > \overline{p}$ of s and t each entail $st \to \overline{p}\overline{q}$, we can apply the Flattening (or Cautious Exportation) Rule to our assumptions s and t to derive $(st \to \overline{p}\overline{q}) > s$ and $(st \to \overline{p}\overline{q}) > t$, and hence (c) $(st \to \overline{p}\overline{q}) > \overline{p}\overline{q}$. By Reciprocity, (a)–(c) jointly imply $\neg pst > \overline{p}\overline{q}$, and hence $\overline{p} > \overline{p}\overline{q}$ by CMon. This implies $\overline{p} > \overline{q}$, so by s and MP we have $\overline{p}$, contradicting our assumption of p.

(b) To derive Sequentiality from Conditional Sequentiality, assume $\Box(p \to (\overline{p} > r))$, $\Box(q \to (\overline{q} > r))$, and $p \vee q$; we want to show $\overline{p}\overline{q} > r$. If $\Box(p \vee q)$ this is vacuously true, so we can further suppose $\Diamond(\overline{p}\,\overline{q})$. We can then apply Conditional Sequentiality to conclude that at least one of s and t is false. Suppose without loss of generality that it's s. Then by CEM, $(\overline{p} > \overline{q}) > p$. So by MOD and the first assumption, $(\overline{p} > \overline{q}) > (\overline{p} > r)$, hence $(\overline{p} > \overline{q}) > (\overline{p} > \overline{q}r)$. By CMon, this implies $(\overline{p} > \overline{q}) > (\overline{p}\,\overline{q} > r)$. Since the antecedent of the right-nested conditional entails the first antecedent $(\overline{p} > \overline{q})$, we can apply the Flattening (or Cautious Importation) Rule to infer that $\overline{p}\,\overline{q} > r$.

(c) To derive Restricted Sequentiality from Sequentiality, just let $r := p \vee q$ in Sequentiality. ◻

References

Bacon, Andrew (2015), "Stalnaker's Thesis in Context", *Review of Symbolic Logic*, 8/1: 131–63.

—— (2020), "Logical Combinatorialism", *The Philosophical Review*, 129/4: 537–89.

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

Bull, R. A. (1966), "That All Normal Extensions of S4.3 have the Finite Model Property", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 12: 341–4.

Dale, A. J. (1974), "A Defence of Material Implication", *Analysis*, 34/3: 91–5.

Dale, Tony (1979), "A Natural Deduction System for 'If Then'", *Logique & Analyse*, 22/87: 339–45.

Dorr, Cian, Goodman, Jeremy, and Hawthorne, John (2014), "Knowing Against the Odds", *Philosophical Studies*, 170/2: 277–87.

Dorr, Cian, and Hawthorne, John (2022), "If...: A theory of conditionals", Manuscript, NYU and USC.

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

Douven, Igor, and Verbrugge, Sara (2013), "The Probabilities of Conditionals Revisited", *Cognitive Science*, 37/4: 711–30.

Edgington, Dorothy (1995), "On Conditionals", *Mind*, 104/414: 235–329.

Etlin, David (2008), "Modus Ponens Revisited", Manuscript, MIT, Feb. 2008.

Fine, Kit (1974), "An Incomplete Logic Containing S4", *Theoria*, 40/1: 23–9.

van Fraassen, Bas (1976), "Probabilities of Conditionals", in Harper and Hooker (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, I (Dordrecht-Holland: D. Reidel Publishing Company), 261–308.

Gibbard, Allan (1981), "Two Recent Theories of Conditionals", in William L. Harper, Robert C. Stalnaker, and Glenn Pearce (eds.), *Ifs: Conditionals, Belief, Decision, Chance, and Time* (Dordrecht: Reidel), 211–47.

Goldstein, Simon, and Santorio, Paolo (2021), "Probability for Epistemic Modalities", *Philosophers' Imprint*, 33.

Holliday, Wesley H., and Icard, Thomas F. (2018), "Axiomatization in the Meaning Sciences", in *The Science of Meaning: Essays on the Metatheory of Natural Language Semantics*, ed. Derek Ball and Brian Rabern (Oxford: Oxford University Press), 73–97.

Holliday, Wesley H., and Litak, Tadeusz (2019), "Complete Additivity and Modal Incompleteness", *The Review of Symbolic Logic*, 12/3: 487–535.

Kaufmann, Stefan (2009), "Conditionals Right and Left: Probabilities for the Whole Family", *Journal of Philosophical Logic*, 38: 1–53.

—— (2015), "Conditionals, Conditional Probabilities, and Conditionalization", in H. Zeevat and H.-C. Schmitz (eds.), *Bayesian Natural Language Semantics and Pragmatics* (Springer International Publishing Switzerland), 71–94.

—— (2017), "The Limit Assumption", *Semantics & Pragmatics*, 10.

—— (2023), "Bernoulli Semantics and Ordinal Semantics for Conditionals", *Journal of Philosophical Logic*, 52/1: 199–220.

Khoo, Justin (2022), *The Meaning of* If (Oxford University Press).

Khoo, Justin, and Santorio, Paolo (2018), "Lecture Notes: Probability of Conditionals in Modal Semantics", Unpublished manuscript, July 2018.

Lewis, David (1973), *Counterfactuals* (Oxford: Blackwell).

—— (1976), "Probabilities of Conditionals and Conditional Probabilities", *The Philosophical Review*, 85/3: 297–315.

—— (1994), "Humean Supervenience Debugged", *Mind*, 103: 473–90.

Mandelkern, Matthew (2024), *Bounded Meaning: The Dynamics of Interpretation* (Oxford University Press).

McGee, Vann (1985), "A Counterexample to Modus Ponens", *The Journal of Philosophy*, 82/9: 462–71.

—— (1989), "Conditional Probabilities and Compounds of Conditionals", *The Philosophical Review*, 98/4: 485–541.

Popper, Karl (1959), *The Logic of Scientific Discovery* (Hutchinson).

Salmon, Nathan U. (2005), *Reference and Essence* (2nd edn., Amherst, NY: Prometheus Books); repr. of *Reference and Essence* (Oxford: Blackwell, 1981; 2nd edn., Amherst, New York: Prometheus Books, 2006).

Santorio, Paolo (2021), "Path Semantics for Indicative Conditionals", *Mind*.

Schultheis, Ginger (2022), "Counterfactual Probability", *Journal of Philosophy*.

Segerberg, Krister (1970), "Modal Logics with Linear Alternative Relations", *Theoria*, 36/3: 301–22.

—— (1989), "Notes on Conditional Logic", *Studia Logica: An International Journal for Symbolic Logic*, 48/2: 157–68.

Stalnaker, Robert C. (1968), "A Theory of Conditionals", in Nicholas Rescher (ed.), *Studies in Logical Theory* (Oxford: Blackwell), 98–112.

—— (1970), "Probability and Conditionals", *Philosophy of Science*, 37/1: 64–80.

Stalnaker, Robert C. (1974), *Letter to Bas van Fraassen*, Jan. 1974.

Stalnaker, Robert C., and Thomason, Richmond H. (1970), "A Semantic Analysis of Conditional Logic", *Theoria*, 36/1: 23–42.

Thomason, S. K. (1974), "An Incompleteness Theorem in Modal Logic", *Theoria*, 40/1: 30–4.

Van Benthem, J. F. A. K. (1978), "Two Simple Incomplete Modal Logics", *Theoria*, 44/1: 25–37.

Williamson, Timothy (2000), *Knowledge and its Limits* (Oxford University Press).

Yalcin, Seth (2007), "Epistemic Modals", *Mind*, 116/464: 983–1026.