The logic of sequences

Cian Dorr and Matt Mandelkern*

May 23, 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. We also show that when sequence semantics is generalized to arbitrary (transfinite) ordinal sequences, the result is the logic that adds only Flattening to C2. We also explore the logics of a few other interesting restrictions of ordinal sequence semantics, and explore 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 simple and compelling semantics for natural-language conditionals and a description of the corresponding logic C2. In 1970, Stalnaker and Thomason showed how to extend the theory to a language with quantifiers, and Stalnaker (1975) showed how to integrate his theory of the conditional, and the role of mood in interpreting conditionals, into the theory of communication he was developing. In parallel, Lewis (1973) developed a book-length defense of a theory of counterfactuals in a similar spirit, arguing for the philosophical significance of the approach and its applications. Kratzer (1981) and Kratzer (1986), in turn, developed

*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.

her own (logically quite different) variation on these approaches, bringing this style of analysis into the heart of linguistic theories of conditionals. Approaches in this spirit remain immensely popular, not only in the parts of philosophy, linguistics, and logic devoted to the study of conditionals, but throughout the many areas of philosophy where conditionals play a central role.

Our concern in this paper is with an intriguing semantic construction which grew out of a debate about the probabilities of conditionals. On top of the influential work cited above, 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]

While some clarification is called for, 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 the interpretation of the conditional can be coordinated with the interpretation of 'probability'—e.g., whether we are talking about credences before or after some update—so that Stalnaker's Thesis always holds *within a given 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, van Fraassen (1976) showed how to give 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

[1] See Stalnaker, 1970, p. 75. Like Popper (1959), Stalnaker 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 van Fraassen's published paper was in fact the version that Stalnaker's letter was responding to, despite the later publication date. Thanks to Bas van Fraassen for correspondence about this.

Thesis holds for conditionals whose antecedents do not themselves contain conditionals.[3] And, indeed, this may be enough to account for intuitions motivating Stalnaker's Thesis in the first place: intuitions about left-nested conditionals are generally not very clear, and, where clear, do not obviously favor Stalnaker's Thesis (as Kaufmann 2023 has recently argued).

This is not, however, a paper about the probabilities of conditionals (a topic we return to only briefly, in §13), 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 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 was true at w_0 according to the old valuation. We have the standard classical clauses for negation and conjunction; the interesting move comes in the treatment of conditionals. 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. Here, the *tails* of $\langle w_0, w_1, w_2, \ldots \rangle$ are the sequences $\langle w_0, w_1, w_2, \ldots \rangle, \langle w_1, w_2, w_3, \ldots \rangle, \langle w_2, w_3, w_4, \ldots \rangle, \ldots$.

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 axiomatize the logic of van Fraassen's ω-sequence models, as well as some interesting variants 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 interesting, 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, which, again, has been an increasingly popular approach in recent years.[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

[3]It also holds for some conditionals whose antecedents *do* contain conditionals: see section 13 for details.

[4]See e.g. Kaufmann, 2009; Kaufmann, 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.

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), 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 §8). 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 schemes.

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

We will argue, moreover, that Flattening is at least prima facie appealing for conditionals in natural language, while Sequentiality is, at best, too complex to assess, suggesting that ω-sequence semantics is not a strong contender for a logic of the natural language conditional. But, intriguingly, we will show that the logic of *ordinal* sequences is the more attractive logic comprising C2 together with just Flattening. Finally we explore the logics of a few other interesting restrictions of ordinal sequence semantics; and argue that, pace van Fraassen, sequence semantics cannot be motivated by considerations about the probabilities of conditionals.

2 The conditional logic C2

Before turning to sequence models and their logic, we will review 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 conditional language $\mathcal{L}$. Its syntax can be specified as follows, 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 $\rightarrow$, $\leftrightarrow$, and $\vee$ as abbreviations for the material conditional, material biconditional, and disjunction defined as usual. We sometimes use pq for $(p \wedge q)$ and $\overline{p}$ for $\neg p$. We sometimes omit parentheses: the order of operations is negation, then the conditional $>$, then $\wedge$ and $\vee$, and finally $\rightarrow$ and $\leftrightarrow$, so for instance $p > q \rightarrow \neg r > s \wedge t$ is to be read as $(p > q) \rightarrow (((\neg r) > s) \wedge t)$.

C2 is the closure of the following set of axiom-schemes:[7]

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

under the following two inference rules:

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

When p is a theorem of C2 we write $\vdash_{\mathsf{C2}} p$. As usual, when $\Gamma \subseteq \mathcal{L}$, we write $\Gamma \vdash_{\mathsf{C2}} p$ to mean that either $\vdash_{\mathsf{C2}} p$, or there is some non-empty finite subset $\Delta \subseteq \Gamma$ such that a material conditional whose antecedent is the conjunction of all the elements of Δ and whose consequent is p is derivable in C2.

Stalnaker's own axiomatization of C2 is somewhat different, and uses two further abbreviations: $\Box$, defined by $\Box p := \neg p > p$, and $\Diamond$, defined by $\Diamond p := \neg\Box\neg p$.[8] Stalnaker then defines C2 with the axioms PC, MP, and Reciprocity and the rule of Detachment as above, plus four further axioms and one further rule:

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

[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.

[7] 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.

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

definition of $\Box$ via $>$. For the other direction, the key principle to derive is:

Mod $$\Box p \to q > p$$

Stalnaker's axiomatization brings out the fact that C2 in fact contains a standard 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 scheme:

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, respectively.[9] Such connections would make the modal logics of various conditional logics especially interesting.

3 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, when modeling C2 with order models, we can use order models where each world is associate with a *well-ordering* of worlds, making the semantics particularly easy to state. Moreover, as we will see, van Fraassen's ω-sequence models can be very naturally viewed as a further special case of 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

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

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 $R(w)$.[10]

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 §8).

The order models characteristic of C2 are of a particularly constrained kind, where 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 can be obtained by relaxing this assumption. However, we will not have occasion to consider such models in this paper, so by 'order-model' we will always mean a well-order model. [11]

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

---

[10] In the general order-models considered by Lewis, 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.

[11] Order semantics for conditionals is equivalent to a semantics based on *selection functions*. A selection function is a function f which takes a world w and a set of worlds φ and returns a set $f(w, \varphi)$. 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 Identity, Reciprocity, MP, 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

We can move freely between selection functions obeying these constraints and order functions. Given a selection function f, we define an 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 a selection function f by saying that $f(\varphi, w)$ is the singleton of the first φ-world in $<_w$, if there is one, and otherwise the empty set. 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.

Definition 3.1. An *order frame* is a pair $\langle W, <\rangle$, where W is a nonempty 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 $w <_w x$.

We can read off the 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. In other words, $R(w) = \{w\} \cup \{v : w <_w v\}$.[12] 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 3.2. An *order model* is an order frame $\langle W, <\rangle$ together with a valuation function $V : At \to \mathcal{P}(W)$.

Definition 3.3. 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 $\Gamma \subseteq \mathcal{L}$, 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 3.4. 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, $p \nvdash_{\mathsf{C2}} \bot$, where $\bot := p_0 \wedge \neg p_0$) 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:

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

Theorem 3.5. 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 confirm our earlier assertion about the modal fragment of C2:

Theorem 3.6. 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 $\vdash_{\mathsf{KT}} p$.

□

4 A note on strong completeness

Theorem 3.5 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), it 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. Somewhat surprisingly, it turns out that when we define consistency in C2 for sets of sentences in the obvious way—Γ is C2-consistent iff there do not exist $p_1, \ldots, p_n$ in Γ such that $\vdash_{\mathsf{C2}} \neg(p_1 \wedge \cdots \wedge p_n)$—we do not have a strong completeness theorem analogous to Theorem 3.5.[13] The problem, of course, arises with certain infinite sets of sentences:

Theorem 4.1. There are C2-consistent sets $\Gamma \subseteq \mathcal{L}$ such that there is no pointed order model in which every member of Γ is true.

[13]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 any atom, i.e., $\phi = \bigcup_i V(p_i)$. We must have $\phi \cap R(w) \neq \varnothing$, 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 for every finite subset of Γ, we can find a pointed order model at which all elements of that subset are true, which with soundness shows that every finite subset of Γ, and hence Γ itself, is consistent. For given a finite $\Delta \subset \Gamma$, let p_k be the atom which appears in Δ whose index is the highest. Consider any set W with $k+1$ members, which we label $w_0, w_1, \dots w_k$. Δ is true at the pointed order model $\langle w_k, \langle W, <, V\rangle\rangle$, where $<$ is any order function such that $w_k <_{w_k} w_{k-1} <_{w_k} w_{k-2} <_{w_k} \cdots <_{w_k} w_0$, and V a 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 can be 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 domain, representing the allowable denotations for sentences—and then require that our orders are well-founded only relative to the elements of that parameter.

In more detail, let a generalized order frame be a triple $\langle W, \mathcal{B}, <\rangle$, where W is any non-empty set; $\mathcal{B}$ is a set of subsets of W, closed under the set-theoretic operations corresponding to $\wedge, \neg$, and $>$ in order semantics; and $<$ is a function which takes any $w \in W$ to a total linear order $<_w$ on a subset of W, such that whenever $\varphi \in \mathcal{B}$ and $\varphi \cap R(w) \neq \varnothing$, φ has a unique first element in $<_w$ (that is, $<_w$ is well-founded modulo $\mathcal{B}$). 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 an standard canonical model construction. To model a set of sentences like Γ which is not true in any order model, we can define 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.[14]

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.

5 ω-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-sequences is at least as strong as C2.

First, we will need 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, 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 5.1. Given a non-empty set P and an ordinal α, an *α-sequence* over P is a total 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, $\sigma(\beta)$. We write $\sigma^{[\beta:]}$ for the *βth tail* of σ, i.e., the $\alpha - \beta$ sequence such that $\sigma^{[\beta:]}(\gamma) = \sigma(\beta + \gamma)$ when $\beta + \gamma < \alpha$, and undefined otherwise.

When τ is a tail of σ, the *rank* of τ in σ is the least β such that τ is the βth tail of σ.

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

[14] The selection function models described in footnote 11 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 in 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 Stalnaker and Thomason place on such selection functions, such models are equivalent to generalized order models as we have defined them here.

natural way provided that it is closed under tailhood (that is, if it contains σ then it contains $\sigma^{[\beta:]}$ when the latter is defined):

Definition 5.2. The *tail order function* $<^S$ on a set of sequences S has $\tau <^S_\sigma \rho$ 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 5.3. 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; and $<$ is the tail order function $<^W$ as in Definition 5.2. 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 $\langle \sigma, \langle W, <^W, V \rangle \rangle$. 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 tails.

Van Fraassen's models can be characterised as a special case of ω-sequence models, namely those that are *full* and *categorical*:

Definition 5.4. An ω-sequence frame $\langle W, < \rangle$ is *full* when W is the set of *all and only* ω-sequences over some P.

Definition 5.5. An ω-sequence frame $\langle W, < \rangle$ is *categorical* iff $V(p_k)$ for each atom p_k is a *categorical* set of sequences: one that includes both or neither of any two sequences with the same first element

It should be clear how this semantics 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 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. Obviously, we are free to call them whatever we like. 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. The former questions are supposed settled by *how things are out there*, whereas the latter are in some sense mere expressions of the way we think, or artifacts of the way we talk. 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 a single element of P is enough to determine a truth value for any sentences with categorical denotations. But we will set aside these questions of metaphysical interpretation here, since they are irrelevant to our logical concerns: even if one regards the contrast between the factual and the non-factual as a chimera, one might see good reasons for accepting the logic of sequence models.

6 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 6.1. Pointed ω-sequence models $\langle \sigma, W, V \rangle$ and $\langle \tau, W', V' \rangle$ are modally equivalent 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 6.1, we have:

Fact 6.2. Any pointed ω-sequence model $\mathcal{M} = \langle \sigma, W, V \rangle$ is modally equivalent to the pointed ω-sequence model $\mathcal{M}_{\mathbb{N}} = \langle \mathrm{id}_\omega, V' \rangle$ where id_ω is the sequence $\langle 0, 1, 2, 3 \ldots \rangle$ of the natural numbers in their standard order, and $\mathrm{id}_\omega^{[j:]} \in V'(p_k)$ iff $\sigma^{[j:]} \in V(p_k)$ for all natural numbers j, k.

We can think of $\mathcal{M}_{\mathbb{N}}$ as a kind of minimal representation of $\mathcal{M}$.

From Fact 6.2 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 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 6.2 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" $\mathcal{M}_{\mathcal{N}}$, since in $\mathcal{M}_{\mathcal{N}}$, W is the set $\{\langle 0,1,2,\ldots\rangle, \langle 1,2,3,\ldots\rangle, \langle 2,3,4,\ldots\rangle, \ldots\}$, no two of whose members have the same first element. It follows that the logic of categorical ω-sequence models is the same as the logic of all ω-sequence models, and that the logic of full categorical models is the same as the logic of full models, and hence also the same as the logic of all ω-sequence models.

We can also use Fact 6.2 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 $\langle 3,1,2,1,2,\ldots\rangle$.[15] However, it would not matter if we ruled out eventually cyclic sequences, since none of the sequences in the minimal representation are eventually repeating.[16] Indeed, for exactly 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 Goldstein and Santorio (2021), who marry finite sequence semantics with the domain semantics from Yalcin (2007), resulting in an interesting extension of Yalcin's logic, which invalidates Strong Centering. 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.

7 Flattening

We turn now to 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 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

[15] An sequence σ is eventually cyclic iff for some n and m, for all j,k, $\sigma(n+jm+k) = \sigma(n+k)$.

[16] 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.

of σ. But 4 is not part of C2, whose modal logic, again, is KT, which does not include 4; equivalently, order frames in general need not have transitive accessibility.

Another modal schema that is not part of C2 (or of the result of adding 4 to C2) is

$$\text{H} \qquad (\Diamond p \wedge \Diamond q) \rightarrow (\Diamond(p \wedge \Diamond q) \vee \Diamond(q \wedge \Diamond p))$$

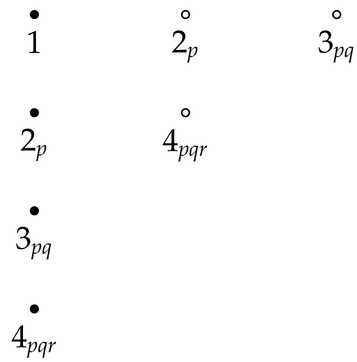
H is valid on a modal frame just in case its accessibility relation is *connected*: whenever wRv and wRu, either vRu or uRv. The accessibility relations of ω-sequence frames are connected: if τ and ρ are distinct tails of σ, whichever of them has greater rank is a tail of the other and hence accessible 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 in ways that cannot be captured in the modal fragment of $\mathcal{L}$. One such schema which we find especially interesting is the following:

$$\text{Flattening} \qquad p > (pq > r) \leftrightarrow pq > r$$

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, and its first p-tail has a pq-tail, and r is false at its first p-tail's first pq-tail. But any sequence with a pq-tail has a p-tail, and 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—while the logic of ω-sequences is strictly stronger.

On the other hand, there are pointed order models in which instances of Flattening are false, such as the following four-world model:



Here, 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 $1 \nVdash (p \wedge q) > r$ while $1 \Vdash p > ((p \wedge q) > r)$. We can obtain a counterexample to the opposite direction of Flattening by making r true at 3 but false at 4.

Flattening will play an important role in what follows, so let's give the name C2.F to the result of adding it to C2. 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 $\qquad p > (pq > r) \rightarrow pq > r$

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

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

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

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

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 pq is replaced by any q for which we have $\vdash q \to p$. For example, we can also characterise C2.F as the result of closing C2 under the following rule:

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

This implies Flattening since $\vdash pq \to p$, and follows from Flattening by the substitution of logical equivalents (since $\vdash p \to q$ means that p 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 $\neg p \to \neg\Box p$ to get

$$\neg\Diamond\neg p \to \neg\Box p > \neg\Diamond\neg p$$

which simplifies to $\Box p \to \neg\Box p > \Box p$, or more briefly $\Box p \to \Box\Box p$.

For H, we use Crashing Cautious Importation (in rule form, applied to the theorem $p \to (p \vee q)$) to get:

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

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

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

But by CEM and Normality, $((p \vee q) > p) \vee ((p \vee q) > q)$ is a theorem of C2. Since of course we also have $p > (p \vee q)$ and $q > (p \vee q)$, we can apply Reciprocity to derive

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

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

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

8 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: 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 (1-a) and (1-b):

(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.

(1-a) and (1-b) 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 (1-a) should be true while (1-b) 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 (1-a) 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 (1-b) 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 (1-a) and (1-b), 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 principle that $p > (p > q)$ is equivalent to $p > q$, since 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 the principle that $p > (p > q) \leftrightarrow p > q$ (a theorem of C2 as well as some weaker conditional logics), which 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 scheme:

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 the resulting logic of $>$ 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; and Materialism is not valid in ω-sequence models. 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 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 appear very 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.

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 simple axioms both of which are false but nomically possible, and 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.[17]

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.

9 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 turns out to be both sound and complete, the *flat* order frames.

Definition 9.1.

- order frame $\langle W, < \rangle$ is *collapsing* 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 $y \leq_x z$.
- $\langle W, < \rangle$ is *flat* iff it is collapsing and has a transitive accessibility relation.

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 to be flat is to be such that, for any $x, y, z, w \in W$, if $x <_w y$ and $y \leq_w z$, then $y \leq_x z$.

It may be helpful to think about this constraint as follows. An order function is equivalent to a function that associates with each world w a non-repeating sequence τ_w of worlds indexed by some positive ordinal, with w as its first member. In these terms, a flat order-model is one such that whenever x occurs in τ_w, τ_x can be obtained by first truncating τ_w at x, and

[17] 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.

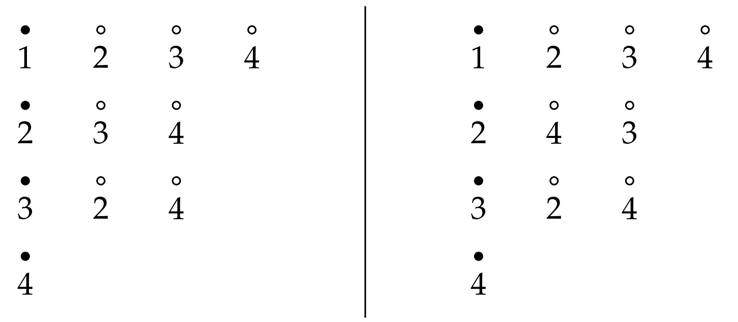


Figure 1: Illustrations of a flat order function (left) and a non-flat (because non-collapsing) order function (right) in frames with worlds {1,2,3,4}.

then optionally adding back worlds that come before x in τ_w. See Figure 1 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 9.2. Cautious Importation is valid on an order frame $\langle W, < \rangle$ iff it is collapsing.

Proof.

- $\Rightarrow$ Suppose that for some w, x, y, z, $x <_w y$ and either

  (a) $y <_w z$ and $y \not<_x z$; or
  (b) $z \in R(x) \setminus R(w)$ and $y \not<_x z$.

  In case (a), there are three possibilities:

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

  In case (b), let $V(p) = \{x, z, y\}, V(q) = \{z, y\}, V(r) = \{z\}$.

- $\Leftarrow$ Suppose that $\langle W, < \rangle$ satisfies is collapsing, and consider any V and $w \in W$ such that $w \Vdash p > (pq > r)$. If there is no pq-world in $R(w)$, then $w \Vdash pq > r$ and we are done; so we can assume that there is a pq-world in $R(w)$. Let u be the first pq-world in $<_w$, and let x be the first p-world in $<_w$. If $x \neq u$, then we have $x <_w u \leq_w u$, and hence $u \leq_x u$, i.e. $u \in R(x)$;

> so there is at least one pq-world in $R(x)$. Consider any pq-world z in $R(x)$. If $z \in R(w)$, we have $u \leq_w z$ and hence $u \leq_x z$ (since $x \leq_w$); if $z \in R(x) \setminus R(w)$, we also have $u \leq_x z$, since every world in $R(w) \cap R(x)$ precedes any world in $R(x) \setminus R(w)$. Thus u is also the first pq-world in $<_x$. Since $p > (pq > r)$ is true at w, $pq > r$ is true at x and so r is true at u; so $w \Vdash pq > r$. □

Since Flattening is the result of adding 4 to Cautious Importation, we can conclude that it is characterized by flatness:

Theorem 9.3. 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 collapsing by the lemma, and also 4 is, 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.[18]

With this characterization result in hand, we can turn to soundness and completeness results for C2.F. The right-to-left direction of Theorem 9.3 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 9.3 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 $\mathcal{F}_\mathsf{L}$ characterizes L but 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: intuitively, when the set $\mathcal{F}_\mathsf{L}$ is too small to find a countermodel to every non-theorem of L.

For a simple toy example to illustrate this sort of situation, consider the "logic" C2.A comprising the theorems of C2 together with a single atom, A. 'Logic' is in scare quotes since this logic is not closed under uniform substitution, and hence is not a logic in an ordinary sense; however, this provides a very simple illustration of the dialectical situation. If we view it as a logic, C2.A is characterized by the empty class of order frames, since A is not valid on any frame. However, C2.A is not complete with respect to the empty class of order frames, since it is not true that every non-theorem of this logic has a countermodel in that class.

[18] 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 $x <_w y$, $y \in R(x)$.

C2.A is, of course, not an interesting example of an incomplete modal logic, since, again, C2.A is not a normal modal logic in the usual sense because it is not closed under uniform substitution of sentence letters. However, there are in fact normal modal logics closed under uniform substitution which are not complete with respect to the class of modal frames that characterize them (see 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 many extensions of C2 (including C2.F) are not strongly complete for *any* class of order frames (as we discussed in Section 4.1) 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 (weak) completeness result for C2.F: in the appendix, we show that C2.F is *complete* for flat order frames. Indeed, we show that it is complete for *finite* such frames:

Theorem 9.4. C2.F is complete with respect to finite flat order frames: i.e., every sentence valid in all finite flat order frames is a theorem of C2.F.

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

One noteworthy corollary of the soundness direction of Theorem 9.3 is that the two modal schemas 4 and H we mentioned in section 7 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-schemes to KT is called S4.3, and it is well known that 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.) We can show that the modal logic in C2.F is exactly S4.3 by showing that any modal frame with such an accessibility relation can be endowed with a flat order function inducing the same accessibility relation.

Theorem 9.5. 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 section 7. By the completeness theorem Theorem 9.4, we can also give a simple model-theoretic proof: it is straightforward 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.

For the left-to-right direction, we need a recipe to transform a finite reflexive, transitive, and connected modal frame $\langle W, R\rangle$ into a flat order frame $\langle W, < \rangle$ whose accessibility relation is R. Fix a strict well-order $<$ of W, and let $x <_w y$ iff wRx and xRy and one of the following conditions obtains:

(i) not yRx

(ii) yRx and not xRw and $x < y$

(iii) yRx and xRw and $w < x < y$ or $x < y < w$ or $y < w < x$ or $w = x \neq y$.

It is easy to see that $w <_w x$ iff wRx and $x \neq w$, so that the accessibility relation defined in terms of $<$ coincides with R.

To show that each $<_w$ is connected on $R(w)$, suppose wRx and wRy and $x \neq y$ but not $x <_w y$. Then we must have yRx by condition (i). If not xRw, then not $x < y$ by condition (ii), so $y < x$ since $<$ is a total well-order, and also not yRw by transitivity, so $y <_w x$ by condition (ii). If xRw, then we cannot have $w < x < y$ or $x < y < w$ or $y < w < x$, so we must have $w < y < x$ or $y < x < w$ or $x < w < y$; moreover, yRw by transitivity, so $y <_w x$ by condition (iii). Since the model is finite, it follows from this that each $<_w$ is well-founded on $R(w)$.

To show that $<$ is collapsing, suppose $x <_w y$ and $y <_w z$. Then wRx, xRy, and yRz. If not zRy, then $y <_x z$ by case (i). If zRy but not yRx, then not yRw by transitivity, so we have $y < z$ since $y <_w z$, and hence also $y <_x z$ by case (ii). If zRy and yRx but not xRw, then again not yRw by transitivty, so we have $x < y$ and $y < z$ since $x <_w y$ and $y <_w z$, so $y <_x z$ by case (iii). Finally, if zRy, yRx, and xRw, also yRw by transitivity. If $w = x$, we immediately have $y <_x z$ by substitution, so assume $w \neq x$; also, $w \neq y$ since $x <_w y$. Then case (iii) obtains for both $x <_w y$ and $y <_w z$, so w, x, y are in the cyclic order $w < x < y < w$, and also w, y, z are in the cyclic order $w < y < z < w$. from which it follows that $x < y < z < x$ and so $y <_x z$ by case (iii). □

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 with a transitive and connected accessibility relation: the non-flat frame in Figure 1 is an example. Such frames do not validate all of C2.F, but do validate all its purely modal theorems.

Finally, we can use the soundness of C2.F for flat order frames to show that C2.F does not include all the sentences that are valid on all ω-sequence frames. Consider the flat order model in Figure 2.

It is easy to see that $\neg(pq) > \overline{p}$ and $\neg(p\overline{q}) > \overline{p}$ are both false at 1 and 2 and both true at 3. Thus all of the following are true at 1:

(a) $\neg p$
(b) $\Diamond pq$
(c) $(\neg(p\overline{q}) > q) > pq$
(d) $(p > q) > pq$

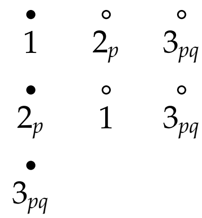


Figure 2: A flat order model which is not equivalent to any ω-sequence model.

But these four sentences cannot be true together in any ω-sequence model. For (a) and (b) to be true, the designated sequence σ must be a $\neg p$-sequence with a pq-tail. Let τ be σ's first pq-tail, and let ρ be the *last* tail of σ before τ. If ρ is an $p\overline{q}$-sequence, $\neg(p\overline{q}) > q$ is true at it, since its first pq-tail is τ. But then (c) is false at σ, since the antecedent is true at ρ, which comes earlier than any pq-tail. Likewise, if ρ is a $\neg p$-sequence, $p > q$ is true at it, since its first p-tail is τ. But then (d) is false at σ, for the same reason.

10 Sequentiality

We have now seen that the logic of ω-sequences is strictly stronger than C2.F. We will presently show that in fact the logic of ω-sequences is exactly C2.F together with the negation of the conjunction of (a)–(d); that is, all we need to do to get a logic that corresponds to ω-sequence frames is to rule out configurations like that in Figure 2.

To see why, let us try to understand what it takes for a flat order frame to be isomorphic to an ω-sequence frame. A distinctive feature of ω-sequence frames is that all the tails of a given sequence σ can be reached by some finite number of steps, where, starting with σ, at each step we lop a single protoworld off of the beginning of the sequence. This is an instance of a more general property that order frames in general can have: namely, that any world x accessible from a world w can be reached by a finite number of steps starting from w, where each step takes us from a world v to the first world after v in $<_v$. We call this property *being ancestral*.

Definition 10.1. Given an order frame $\langle W, < \rangle$:

- The *successor* of w, $\mathrm{succ}(w)$, is w if $R(w) = \{w\}$, and otherwise the first world after w in $<_w$.

- 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.

- x is *reachable* from w iff x is an element of α_w, i.e. w is related to x by the ancestral of the successor relation.

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

– Finally, a frame is *ancestral* iff every world accessible from any given world is also reachable from that world.

Note that in any flat frame, when v is the successor of w, $R(v)$ is either $R(w)$ or $R(w) \setminus \{w\}$, and $<_v$ agrees with $<_w$ on $R(w) \setminus \{w\}$. This means that the worlds reachable from w are always a subset of $R(w)$. Moreover, within the subset of reachable worlds, $x <_w y$ iff the rank of x from w is less than that of y. For if x precedes y according to w and $x \neq w$, then x also precedes y according to w's successor, and so on until we reach x; since x certainly does not precede y according to y, this means that as we take successor steps from w we must reach x before we reach y. If the frame is ancestral as well as flat, then, since reachability and accessibility coincide, this means that we can read off the order function from the successor function: that is, in a flat ancestral frame, α_w is exactly the sequence of worlds induced by $<_w$ (what we above called τ_w), followed by infinite repetitions of the last element of τ_w, if there is one.

By contrast, not every flat frame is ancestral. This is illustrated in the frame of the model in Figure 2, which is repeated as $<^1$ in Figure 3. Here accessibility and reachability do not coincide, even though the frame is flat. For instance, in this frame, 1 and 2 are each other's successors, so 3 is not reachable from 1, despite being accessible from 1: that is, $\alpha_1 = \langle 1, 2, 1, 2, 1, 2 \ldots \rangle$ and never contains 3.

As we have noted, every ω-sequence frame is ancestral as well as flat. Of course, not every flat ancestral frame is an ω-sequence frame; the worlds in a flat ancestral frame need not be sequences. But every flat ancestral frame $\langle W, < \rangle$ is *isomorphic* to an ω-sequence frame: the one we obtain by replacing each $w \in W$ with its successor-sequence α_w. For example, the frame $\langle \{1, 2, 3\}, <^2 \rangle$ illustrated in the center of Figure 3 is isomorphic to the ω-sequence frame with sequences $\langle 1, 2, 3, 3, 3 \ldots \rangle, \langle 2, 3, 3, 3, 3 \ldots \rangle, \langle 3, 3, 3 \ldots \rangle$; likewise, the frame $\langle \{1, 2, 3\}, <^3 \rangle$ in the right of Figure 3 is isomorphic to the ω-sequence frame with sequences $\langle 1, 2, 3, 1, 2, 3 \ldots \rangle$, $\langle 2, 3, 1, 2, 3, 1 \ldots \rangle$, $\langle 3, 1, 2, 3, 1, 2 \ldots \rangle$.

In general, $\alpha_{\text{succ}(w)}$ is obviously always the first tail of α_w, i.e. the result of deleting the first element from α_w. So, the set $\{\alpha_w \mid w \in W\}$ is closed under tails, meaning we can regard it as a sequence-frame with W as its set of protoworlds. Moreover, since the first tail of any sequence is also its successor, the function $w \mapsto \alpha_w$ preserves the successor-function on the original frame. Since the order function of any flat and ancestral frame can be read off the successor function, it follows that this function also preserves the order function, in the sense that $x <_w y$ iff $\alpha_x <_{\alpha_w} \alpha_y$. (We already noted that $x <_w y$ iff x first occurs before y in α_w; since α_x and α_y are respectively the first tails of α_w starting with x and y, this is true iff $\alpha_x <_{\alpha_w} \alpha_y$.) We thus have an isomorphism from $\langle W, < \rangle$ to the sequence-frame comprising the successor

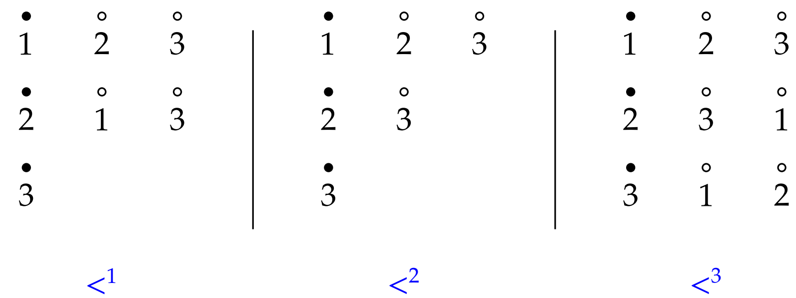


Figure 3: Illustrations of a flat but not ancestral frame $<^1$, and two ancestral frames $<^2, <^3$. In $<^1$ there is no way to get from 1 to 3 by the successor relation, since you end up stuck in a loop between 1 and 2.

sequences. So the task of axiomatizing the logic of ω-sequence frames is the same as the task of axiomatizing the logic of flat ancestral order frames.

To see how we can characterize the class of flat ancestral frames, let's start by considering *finite* flat frames where ancestrality fails, like $<^1$ in Figure 3. In fact, *every* flat finite frame where ancestrality fails will have to embed a configuration like that in $<^1$. The failure of ancestrality means that there is a world (like 1 in this case) which can access some world which it can't reach (in this case, 3). In a frame which is both flat and finite, this can happen only when, in taking successor steps from the given world, we eventually find ourselves trapped on a loop, where the unreachable world is accessible from but not reachable from any world in the loop. In $<^1$, the loop comprises just worlds 1 and 2. In general, it can have any finite size, and the world we started with need not be part of the loop: for example, in the frame in Figure 4, when we start from world 1, we end up on the loop comprising worlds 3, 4, and 5. But somewhere in the frame, we will have to find the configuration illustrated in Figure 5: take b to be any world on the loop, a to be b's successor, and c to be the first world after b in $<_a$; then c is also not reachable from b, since as we take successor steps from b we will inevitably find ourselves back at b before we come to c. If we pick a valuation where atom p is true only at b and c and atom q is true only at c, our sentences (a)–(d) (repeated here) will then be true at a for the same reason as before:

(a) $\neg p$
(b) $\Diamond pq$
(c) $(\neg(p\overline{q}) > pq) > pq$
(d) $(p > pq) > pq$

(c) and (d) are both true at a because whenever $w <_a c$, we have both $a <_w c$ and $b <_w c$; hence neither $(\neg p\overline{q} > pq)$ nor $p > pq$ is true at any such w, meaning that c is the first world in $<_a$ where either of these conditionals is true.

If the number of worlds is infinite, we could have failures of ancestral-

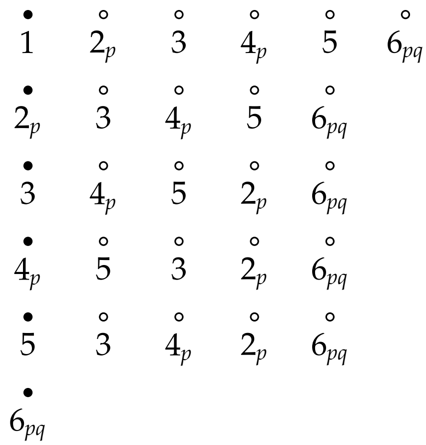


Figure 4: A more complicated flat non-ancestral frame, with a valuation that makes (a)–(d) true at 1

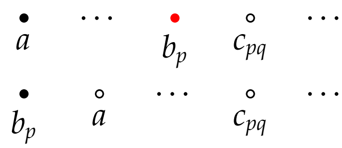


Figure 5: Every finite order frame which is flat but not ancestral contains a configuration with this structure: given the subscripted valuation, Sequentiality fails at a.

ity that don't involve loops (and hence don't involve the configuration in Figure 5)—for example, in the frame whose worlds are the natural numbers together with the first infinite ordinal ω, where each world sees exactly the worlds that are above it, with their standard ordering, ω is accessible but not reachable from any other world. But so long as there are some v and w such that v is accessible but not reachable from w, we can always find a valuation that makes (a)–(d) all true at w: let the valuation make atom q true only at v, and atom p true only at v together with those worlds reachable from w whose rank from w is odd. (Figure 4 shows what this looks like taking $w = 1$ and $v = 6$.) Then again (a)–(d) will all be true at w, since $\neg p\overline{q} > pq$ and $p > pq$ are true at v but not at any world earlier than v in $<_w$. By contrast, in a flat and ancestral model, the conjunction of (a–d) cannot be true at any world w: by (a) and (b), there must be a first pq-world in $<_w$, which must be reachable from w in some nonzero number of successor steps. If it was immediately preceded by a p-world, $\neg p\overline{q} > pq$ is true there, while if it was immediately preceded by a $\overline{p}$-world, $p > pq$ is true there. In the first case, (c) is false at w, and in the latter, (d) is false at w. So, the negation of the conjunction of (a)–(d) gives us what we are looking for: a formula valid on every flat ancestral frame, and not valid on any flat non-ancestral frame. We can simplify this negated conjunction a little by re-expressing it as (c) $\to \neg$(a) $\vee \neg$((b) $\wedge$ (d)), and noting that $\neg$((b) $\wedge$ (d)) is equivalent (in C2) to $(p > q) > \overline{p}$. This gives us the following characterizing axiom:

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

Let C2.FS be the logic derived from C2.F by adding Sequentiality as an additional axiom-scheme. Then we have shown:

Theorem 10.2. C2.FS is valid on an order frame iff it is flat and ancestral (and hence iff it is isomorphic to an ω-sequence frame).

Proof. We build on Theorem 9.3 by showing that a flat order frame validates every instance of Sequentiality iff it is also ancestral. We have already proved this in the running text, but we will briefly recapitulate the proof here.

- $\Rightarrow$ Suppose $\langle W, < \rangle$ is flat but not ancestral, where w can access but not reach v. Then value p true at v and at worlds whose rank from w is odd, and q true just at v. w verifies $((\neg p \vee q) > p) > q$, since the first accessible world in w verifying the antecedent is v. But w doesn't verify p (since its rank from w is 0), nor does it verify $(p > q)\neg p$, since v is the first $p > q$-world in $R(w)$.

- $\Leftarrow$ Suppose $\langle W, < \rangle$ is flat and ancestral. Consider $w \in W$. Suppose w verifies the premise of Sequentiality, $((\neg p \vee q) > p) > q$, but doesn't verify p. If $R(w)$ contains no pq-world, then either (i) $p > q$ is true nowhere in $R(w)$, in which case $(p > q) > \neg p$ is trivially true at w; or (ii) $p > q$ is first true at some $\neg p$-world in $R(w)$, so $(p > q) > \neg p$ is true at w. If $R(w)$ contains a pq-world, let z be the first such world in $<_w$, and let y be the world just before z in $<_w$ (the existence of such a y is guaranteed by ancestrality). We show that $y \Vdash \neg p$; by ancestrality, the successor of y is z, and hence $y \Vdash p > q$; hence the first $p > q$-world in $<_w$ is either y or some earlier $\neg p$-world, so $w \Vdash (p > q) > \neg p$. To see that $y \Vdash \neg p$, suppose otherwise, so $y \Vdash p \wedge \neg q$. But then since the successor of y is z, $y \Vdash (\neg p \vee q) > p$. Hence either y or some other $\neg q$-world is the first $(\neg p \vee q) > p$-world in $<_w$, meaning that $w \nVdash ((\neg p \vee q) > p) > q$ contrary to assumption.

□

10.1 The completeness of C2.FS

Theorem 10.2 tells us that C2.FS is sound for ω-sequence frames. As we emphasized in the discussion before Theorem 9.4, 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 the appendix, we show that C2.FS is complete for ω-sequence frames by showing the stronger result in Theorem 10.3:

Theorem 10.3. C2.FS is complete for finite ω-sequence frames.

This immediately implies:

Theorem 10.4. C2.FS is complete for finite flat ancestral order frames.

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

10.2 The modal logic of C2.FS

S4.3.1 is the result of adding every instance of Dum to S4.3:

$$\text{Dum: } \Box(\Box(p \to \Box p) \to p) \to (\Diamond\Box p \to p)$$

S4.3.1 is sound for frames in which the accessibility relation (quotiented on accessibility-equivalent classes) has no final segments with order type greater than ω, and (weakly) complete for $\{\langle \mathbb{N}, \leq \rangle\}$.[19]

Theorem 10.5. When p is a modal sentence, $\vdash_{\mathsf{C2.FS}} p$ iff $\vdash_{\mathsf{S4.3.1}} p$.

Proof.

- $\Leftarrow$ By the completeness of C2.FS with respect to flat ancestral order models.
- $\Rightarrow$ Suppose p is consistent in S4.3.1; then it is true at 0 in some model M based on $\langle \mathbb{N}, \leq \rangle$. This can be extended into the minimal ω-sequence frame, which has the same domain and has $\leq$ as its accessibility relation; p will also be true in that frame, so by the soundness of C2.FS for ω-sequence frames, p is consistent in C2.FS. □

10.3 Sequentiality in natural language

As with Flattening, we'd like to know whether Sequentiality is in fact valid. Unfortunately, this is not easy to assess; the inference pattern is too complex, and in particular involves too much left-nesting, to easily judge. Hence consider:

(11) If Mark is going if either Mark isn't going or Sue is going, then Sue is going. So, either Mark is going, or else, if Sue is going if Mark is going, then Mark isn't going.

(12) If the espresso would have been overextracted if either the espresso wasn't overextracted or he had gotten mad, then he would have gotten mad. So, either the expresso was overextracted, or if he would have gotten mad if it was overextracted, then it wasn't overextracted.

It seems hopeless to try to figure out whether these inferences are valid by consulting empirical judgments.

[19] Bull 1966; Segerberg 1970

10.4 A simpler axiomatization?

Given how difficult it is to grok English instances of Sequentiality, it is natural to ask is 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. Part of what makes Sequentiality especially hard to assess is the left-nested conditionals it contains: left-nesting is generally hard for English speakers to process (see Kaufmann 2023 for some recent discussion). But in the language restricted to disallow left-nesting, the logic of ω-sequence models is C2.F. That is:

Theorem 10.6. When $p \in \mathcal{L}_{BA}$, $\vdash_{\mathsf{C2.F}} p$ iff $\vdash_{\mathsf{C2.FS}} p$.

where

Definition 10.7. 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 $\mathcal{L}_B$ but only when the antecedent is conditional-free

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

This is proved in Appendix E. This makes us doubt we'll find an axiom scheme (or schemes) which distinguish C2.FS from C2.F about which we have clear intuitions.[20]

11 Ordinal sequence frames

Sequentiality doesn't look particularly well-motivated. This is somewhat surprising, since ω-sequence semantics looks very simple and natural, and contrasts with Flattening, which looks like a serious contender to be part of the correct conditional logic.

Of course, one reaction to the complexity of Sequentiality is this: we certainly won't find any compelling counterexamples to Sequentiality. So we should validate it, since we should in general validate the strongest logic compatible with our reflective intuitive judgments.

On reflection, however, this is a bad argument. There are infinitely many logical principles whose shortest statement is too complex for any human to process and hence such that we can be sure no compelling counterexample to them will ever be discovered. Should we add all of them to our logic? Surely not, since some of them will be inconsistent with each other. But then which ones should we choose? There is no principled way of answering this. So we think that C2.FS is not a compelling candidate conditional logic.

[20] A different possibility, which Wesley H. Holliday suggested, is to look for a rule or rules instead, which may be easier to assess. This is an interesting option which we leave open.

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. From the other direction, we have seen that C2.F is sound and complete with respect to flat order models, a somewhat finicky class of order models.

On reflection, however, the class of ω-sequence frames is restricted in a somewhat arbitrary way: namely, by the restriction to ω-sequences in particular. There is a natural generalization of this approach which bases the same kind of semantics on arbitrary ordinal sequences. As we will now see, the class of ordinal sequence frames is sound and complete for C2.F rather than C2.FS.

Definition 11.1. An *ordinal sequence frame* is an order frame $\langle W, < \rangle$ where W is a set of (possibly transfinite, possibly finite) sequences closed under tailhood, and $<$ is the tail order function on W.

An ordinal sequence model is an ordinal sequence frame with a valuation; a pointed ordinal sequence model adds a distinguished sequence. Note that as we already saw with ω-sequence frames, an ordinal sequence frame can be finite (i.e., have a finite W) even if some or all of the sequences that comprise W are transfinite. (We will always use 'finite ordinal sequence frame' to refer to an ordinal sequence frame with a finite set of sequences.)

Theorem 11.2. C2.F is sound for ordinal sequence models, and weakly complete for *finite* ordinal sequence models in which the domain of every sequence is an ordinal less than ω^ω.

For soundness, it suffices to show that all ordinal sequence models are flat, which is easy. The completeness result is proved in **??**. 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 **??**, the completeness of C2.F for flat order models.

Indeed, there is a natural sense in which there is no difference between the two kinds of model: every ordinal sequence frame is already a flat order frame, and every flat order frame is isomorphic to an ordinal sequence frame. Intuitively, one can turn a flat order frame $\langle W, < \rangle$ into an ordinal sequence frames by replacing each world w with an ordinal sequence defined by proceeding out along $<_w$, but inserting a block of ω many loops whenever we meet worlds v for which $<_v$ begins with worlds we have already encountered. The formal details of this procedure are somewhat involved, so we omit them here.

12 List frames and successor-ordinal frames

Just as we can expand the notion of an ω-sequence frame to an ordinal sequence frame, likewise it is interesting to explore various restrictions on ordinal sequence frames. In this section we will explore just two of these which are naturally related: ordinal sequence frames whose sequences all have a *finite* domain, and ordinal sequence frames whose sequences all have as their domain a *successor ordinal*, i.e., an ordinal with a final element. 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 class, as we will see, is related to the first as ordinal sequence frames are to ω-sequence frames.

Definition 12.1. Given a non-empty set P, a *list* over P is a sequence over P whose domain is a finite ordinal.

An ordinal sequence frame $\langle W, < \rangle$ is a list frame when W is a set of lists.

The logic of list frames is obviously at least as strong as C2.FS, because any k-length list $\langle w_0, \ldots w_{k-1} \rangle$ is isomorphic to the ω-sequence $\langle w_0, \ldots w_{k-1}, w_{k-1}, w_{k-1} \ldots \rangle$ which transforms the list into an ω-sequence by repeating the last world of the list infinitely. However, the logic of lists is strictly stronger than C2.FS: it is the logic which strengthens C2.FS with every instance of the McKinsey axiom scheme:[21]

$$\text{McKinsey: } \Box\Diamond p \to \Diamond\Box p$$

McKinsey fails in ω-sequence frames: for instance, in the ω-sequence model based on $\langle 1, 2, 1, 2, \ldots \rangle$, where p is true $\langle 1, 2, 1, 2, \ldots \rangle$ 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 on lists, since every list comes to an end: so if $\Box\Diamond p$ is true at a list, so that every tail of the list can access some p-tail, then the last (singleton) tail of the list must be a p-tail which can only access itself, and hence is an accessible $\Box p$-tail.

Theorem 12.2. Let C2.FSM be C2.FS plus every instance of the McKinsey axiom scheme. C2.FSM is sound and complete with respect to list models.

Soundness follows from the soundness of C2.F with respect to all ordinal sequence frames, of which list frames are a special case, together with the reasoning in the text just now. For completeness, see appendix.

Note that the reasoning that shows that McKinsey is sound for list models is essentially about the fact that a list has a *final* tail, and not about the fact that the list is finite. Hence C2.FSM is also sound and complete with respect the class of ω-frames whose sequences all comprise some list followed by a constant ω-sequence.

But, likewise, McKinsey is valid over the class of *ordinal* sequence frames which have a final tail, no matter how long they are. Let a *successor* sequence frame be an ordinal sequence frame $\langle W, < \rangle$ where W contains only sequences whose domain is a successor ordinal, i.e., an ordinal with a final element.

[21] An equivalent axiomatization can be given with the Grzegorczyk Axiom $\Box(\Box(\phi \to \Box\phi) \to \phi) \to \phi$.

Theorem 12.3. Let C2.FM be C2.F plus every instance of the McKinsey axiom scheme. C2.FM is sound and complete with respect to successor ordinal sequence frames.

Soundness is by the reasoning just as above; completeness is proved in the appendix. Indeed, once more, what is crucial to this result is not really being a successor ordinal but rather having a final tail in the sense that there is some β such that $\sigma^{[\beta:]}$ is defined and for any α, $\sigma^{[\beta:]} = (\sigma^{[\beta:]})^{[\alpha:]}$ whenever the right-hand side is defined (that is, σ's domain is the sum of some ordinal with a second ordinal which is additive principal, i.e. not the sum of two ordinals; and σ's value on the second ordinal is constant). So, equally we have that C2.FM is sound and complete for the class of sequence frames $\langle W, <\rangle$ such that every $\sigma \in W$ has a final tail.

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. But we have only brushed the surface: for every class of ordinal sequence models, we can ask whether it corresponds to an interesting logic, potentially revealing infinitely many new interesting conditional logics. Benjamin Przybocki (p.c.) has reported interesting results on the axiomatization of classes of ordinal frames in which the length limit is any ordinal strictly between ω and ω^ω.

13 Is there a probability-theoretic argument for 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 consistency of a restricted version of Stalnaker's Thesis, one might wonder whether there is some interesting argument from some version of that Thesis to Flatness and/or Sequentiality.

For Sequentiality, the answer is a straightforward *no*. What van Fraassen actually shows in his paper is the following:

Fact 13.1. If π is a probability measure on P and π^* is the induced product measure on P^ω, and V is a valuation on the sequence frame P^ω in which denotations of atoms 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 each of p and q is either 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).[22]

[22]Here, a probability measure on P is a (countably-additive) probability measure whose 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 $Y = \{f \mid f(i) \in g(i) \text{ for all } i \in X\}$. The domain of π^* is defined to be the smallest σ-algebra of

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

Fact 13.2. Where π, π^*, and V are as in the previous fact, $\pi^*([\![p > q]\!] \mid [\![r]\!]) = \pi^*([\![q]\!] \mid [\![p]\!])$ whenever $\pi^*([\![p]\!]) > 0$, $\pi^*([\![r]\!] \mid [\![p]\!]) = 1$ and each of p and r is a conjunction in which each conjunct is either Boolean or a zero-degree conditional.[23]

The proof of Fact 13.2 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 matters 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 such that 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, which are the only ones relevant for Fact 13.2.[24]

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 the distinguished w, since this will still let us represent worlds as ordinal sequences with an order-function that agrees with the sequence-model order function as regards the initial ω-subsequences. Still, it would be interesting if some weakening of flatness played a crucial 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 (or at least for the truth of its instances involving positive-probability propositions).[25]

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 Flatness, or anything beyond C2, that provides something pretty close to Fact 13.2.

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 $Y = \{f \mid f(i) \in g(i) \text{ for all } i \in X\}$, $\pi^*(Y) = \prod_{i \in X} \pi(g(i))$.

[23] Dorr and Hawthorne (2022) discuss several empirical reasons for being interested in the extra strength that comes from allowing $r \neq \top$.

[24] 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.

[25] Note that in any model with a probability function satisfying Fact 13.2 (or even the weakening below that requires r and p to be Boolean), the two sides of any instance of Flattening where p and q are Boolean and pq has positive probability will have to have the same 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 one-premise inference rules corresponding to the two directions of Flattening. However, there is no obvious route from 13.2 to the claim that the instances of Flattening have probability one, even for Boolean p and q; and the tree models introduced below show that if we weaken 13.2 to require p and r to be Boolean, we can have models where instances of Flattening have probability less than one.

The only way in which we weaken Fact 13.2 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. Formally, we can construct these as functions from lists of natural numbers to protoworlds:

Definition 13.3.

- For any set X, X^* is the set of lists over X.
- For a set P, the set of *trees* over P is $P^{\mathbb{N}^*}$: 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 $\alpha \in \mathbb{N}^*$.
- We can introduce the following order function on any $P^{\mathbb{N}^*}$: When $\rho, \sigma, \tau \in P^{\mathbb{N}^*}$, $\rho <_\tau \sigma$ iff either $\rho = \tau$ and σ is the nth branch of τ for some n, or 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 nonempty set, the *tree frame* over P is the order model with domain $P^{\mathbb{N}^*}$, with $<$ as above. A *tree model* is an order model based on a tree frame.

Theorem 13.4. The logic of tree models is C2.

Proof. Given a finite order model $\langle W, <, V\rangle$, we can associate each world w with a tree τ_w over W by 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 $<_w$. If we assign each atom the same truth value at τ_w that V assigns it at w, the same sentences will be true at τ_w as at w. The completeness of C2 for finite order frames then yields the completeness of C2 for tree frames. □

We can now prove our tree-frame analogue of Fact 13.2:

Fact 13.5. Where π is a probability measure on P, $\pi^\dagger$ is the probability measure on $P^{\mathbb{N}^*}$ derived from π by the product measure construction, and V is a valuation on the tree frame $P^{\mathbb{N}^*}$ 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 $\pi^*([\![r]\!] \mid [\![p]\!]) = 1$ and both p and r are Boolean.

Proof. First some definitions which apply to any function space P^I (including $P^{\mathbb{N}^*}$). 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 is a $f : I \to I$ and $Y \subseteq P^I$, $f^{-1}(Y)$ is $\{g \mid g \circ f \in Y\}$; f is *measurable* iff $f^{-1}(Y)$ is in the domain of π^* whenever Y is. We rely on two standard facts about the product measure π^* on P^I derived from some probability measure π on P:

(i) If Y and Z are orthogonal and in the domain of π^*, $\pi^*(Y \cap Z) = \pi^*(Y)\pi^*(Z)$.

(ii) If $f : I \to I$ is injective and measurable and Y is in the domain of π, $\pi^\dagger(f^{-1}(Y)) = \pi^\dagger(Y)$.

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 $f, g : \mathbb{N}^* \to \mathbb{N}^*$ by $f(\langle\rangle) = \langle 0\rangle$, $f(\langle n\rangle + \alpha) = \langle n+1\rangle + \alpha$; $g(\alpha) = \langle 0\rangle + \alpha$. Set:

$$\begin{aligned} Z &:= f^{-1}(Y) \\ U &:= g^{-1}([\![p \wedge q]\!]) \\ W &:= g^{-1}([\![\overline{p}]\!]) \end{aligned}$$

That is, Z is the set of all trees such that either none of their branches except perhaps for their zeroth branch is in $[\![p]\!]$, or their first positive-indexed branch in $[\![p]\!]$ is in $[\![q]\!]$; U is the set of all trees such that $p \wedge q$ is true at their zeroth branch, and W is the set of all trees such that p is false at their zeroth branch. Then $Y = U \cup (Z \cap W)$. 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. So

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

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

This gives us something 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 $\pi^\dagger([\![r]\!] \mid [\![p]\!]) = 1$. And Y is orthogonal to $[\![\overline{p} \wedge r]\!]$ (since the latter supervenes on $\{\langle\rangle\}$ while Y supervenes on the set of nonempty lists), and hence $\pi^\dagger(Y \mid [\![\overline{p} \wedge r]\!]) = \pi^\dagger(Y) = \pi^\dagger([\![q]\!] \mid [\![p]\!])$. So we have:

$$\begin{aligned} \pi^\dagger([\![p > q]\!] \mid [\![r]\!]) &= \pi^\dagger([\![q]\!] \mid [\![p]\!])\pi^\dagger([\![p]\!] \mid [\![r]\!]) + \pi^\dagger([\![q]\!] \mid [\![p]\!])(1 - \pi^\dagger([\![p]\!] \mid [\![r]\!])) \\ &= \pi^\dagger([\![q]\!] \mid [\![p]\!]) \end{aligned}$$

□

By contrast, tree models do *not* sustain a version of Stalnaker's Thesis that allows zero-degree conditionals in the antecedent. For example, suppose we build a tree model over $P = \{0, 1, 2, 3\}$, with π as the indifferent prior over these four protoworlds. Let $V(p) = \{\tau \mid \tau(\langle\rangle) \in \{0, 1\}\}$ and $V(q) = \{\tau \mid \tau(\langle\rangle) = 0\}$. Then $\pi^\dagger(p \mid p > q) = \pi^\dagger(pq)/\pi^\dagger(p > q) = \pi(pq)/\pi(p \mid q) = \pi^\dagger(p) = 1/2$. But with some tedious algebraic calculations we will spare the reader, we have shown that $\pi^\dagger((p > q) > p) = 7/12$ in this tree model.

So, the part of the proof of Fact 13.2 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, which is intimately connected to the validity of Flattening in the former. If one thought there were good reasons to want some version of Stalnaker's thesis that applies to conditionals whose antecedents are zero-degree conditionals (or conjunctions thereof), that might potentially yield an interesting argument for the validity of Flattening, or at least for the truth of Flattening-instances where pq has positive probability. But it is not clear that there are such reasons available to those who accept C2. For as Stalnaker showed in his 1974 letter to van Fraassen, proponents of C2 must reject Stalnaker's Thesis for certain conditionals whose antecedent is a *disjunction* one of whose disjuncts is a zero-degree conditional: in C2, $(p \vee (p > q)) > pq$ is equivalent to pq, so $\pi((p \vee (p > q)) > pq) = \pi(pq)$; but whenever $0 < \pi(pq) < \pi(p) < 1$, given that $\pi(p > q) = \pi(q \mid p)$, it follows that $\pi(pq) \neq \pi(pq \mid p \vee (p > q))$. Given this, those of us who (unlike Bacon 2015) are not willing to give up on C2 will need to have some strategy for explaining away the 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 seems likely that it will generalize in such a way that it can *also* explain away the appeal of Stalnaker's Thesis for conditionals whose antecedent is a single zero-degree conditional on its own, or a conjunction of such conditionals. Conversely, if such a story is unconvincing, then we will presumably want to follow Bacon 2015 in rejecting C2, in particular Reciprocity, so that we can vindicate a version of Stalnaker's Thesis for the full language. It seems unlikely that the extra strength of Fact 13.2 over Fact 13.5, lying as it does in an awkward terrain strictly between a natural limitation of Stalnaker's Thesis to Boolean antecedents and a vindication of Stalnaker's Thesis for all antecedents, could form the basis for a compelling argument for Flattening.

14 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 (while still not collapsing into the material conditional). While some of these logics seem too complex and arbitrary to be plausible candidates for the logic of natural-language conditionals, at least one of them, namely C2.F, enjoys considerable *prima facie* plausibility.

There are several further questions about this space of logics which one could explore.

- Exploration of the logics of classes of ordinal sequence models with given upper bound on the length of the sequences, either finite or strictly between ω and ω^ω.

- We characterized Sequentiality relative to the *flat* order frames. It is also possible to characterize Sequentiality directly, relative to all order frames, yielding a characterization result for the logic C2.S obtained by adding every instance of Sequentiality to C2. Given an order frame $\langle <, W\rangle$, and $P \subseteq W$, write P_w for $P \cap R(w)$.

  Theorem 14.1. $\langle <, W\rangle$ validates Sequentiality iff

  $$\forall P \subseteq W, w \in \overline{P}, x \in P_w : \exists y <_w x : (\forall z \in P_y : x \leq_y z) \vee (\forall z \in \overline{P}_y : x <_y z)$$

  The proof is in a footnote.[26] An interesting question is whether this condition is essentially second-order or can be restated in a first-order way.

[26] $\Rightarrow$: Suppose the RHS fails and let P, w, x witness the failure, and pick a valuation V with $V(p) = P$ and $V(q) = \{x\}$. Then $\forall y <_w x : (\exists z \in P_y : x \not\leq_y z) \wedge (\exists z \in \overline{P}_y : x \not<_y z)$. This ensures that the premise of Sequentiality is true and the conclusion false.

$\Leftarrow$: Suppose the RHS holds and consider any valuation V. Let us use P for $V(p)$ and Q for $V(q)$. Consider any $w \in W$:

- If $R(w) \cap P \cap Q = \varnothing$ then $w \Vdash (p > q) > \neg p$.
- If $R(w) \cap P \cap Q \neq \varnothing$ and $w \in P$ then $w \Vdash p \vee ((p > q) > \neg p)$.
- If $R(w) \cap P \cap Q \neq \varnothing$ and $w \notin P$, let x be the first PQ-world in $<_w$. Then by the condition, for some $y <_w x$, either:
  - the left disjunct holds, so $y \Vdash \neg p \wedge (p > q)$; and if $p > q$ is true at any world preceding y in $<_w$ it is true at a $\neg p$-world (since all the p-worlds before y in $<_w$ are $\neg q$-worlds), so $w \Vdash (p > q) > \neg p$; or
  - the right disjunct holds, so $y \Vdash ((\neg p \vee q) > p) \wedge \neg q$; again, if $(\neg p \vee q) > p$ is true at any world prior to y in $<_w$ it is a $\neg q$-world, since all p-worlds preceding y in $<_w$ are $\neg q$-worlds; so $w \nVdash ((\neg p \vee q) > p) > q$.

– From the other direction, we have given the logic of order frames that are ancestral *and flat*, namely, C2.FS. An interesting question is what logic ancestrality on its own corresponds to, that is: what is the logic of ancestral order frames? One coud ask similar questions about several other conditions that are equivalent to being ancestral in the presence of flatness, but not in general.

– We can explore the results of adding Sequentiality and/or Flatness to conditional logics weaker than C2.

– We can explore alternate axiomatizations of Sequentiality that may be easier to assess, including statements of Sequentiality as a rule rather than axiom.

– Fine (1971) establishes the striking fact that every modal logic that includes the modal logic S4.3 is decidable. Given the association between C2.F and S4.3, it is natural to conjecture that C2.F has the same property. However, Fine's proof is quite complex and does not immediately suggest a strategy for proving this.

A Preliminaries

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

We continue to use *sequence* for any function to an underlying non-empty set whose domain is an ordinal at least as great as ω, and *list* for any such from a finite ordinal.

- σ^α is the value of σ on α, where defined; so e.g. $\sigma^0, \sigma^1, \sigma^2$ are the first, second, and third element of σ.
- $\sigma^{-1}, \sigma^{-2}, \sigma^{-3}$ are the last, second last, and third last elements of σ, and so on, when they exist. (They may not exist, e.g. if it's an ordinal like ω that doesn't have a last element).
- $\sigma^{[\alpha:]}$ is the α-tail of σ, that is, $\langle \sigma^\alpha, \sigma^{\alpha+1} \dots \rangle$.
- $\sigma^{[\alpha:\beta]}$ is the result of deleting $\sigma^{[\beta:]}$ from the end of $\sigma^{[\alpha:]}$. We write $\sigma^{[:\beta]}$ for $\sigma^{[0:\beta]}$, that is, the result of deleting $\sigma^{[\beta:]}$ from σ.
- $(\sigma + \sigma')$ is the result of concatenating σ with σ': that is, where $l(\sigma)$ and $l(\sigma')$ are the respective domains of σ and σ':

$$(\sigma + \sigma')(\alpha) = \begin{cases} \sigma(\alpha) & \alpha \leq l(\sigma) \\ \sigma'(\alpha - l(\sigma)) & l(\sigma) < \alpha \leq l(\sigma) + l(\sigma') \\ \textit{undefined} & \textit{otherwise} \end{cases}$$

Recall some abbreviations for our conditional language:

- $p \gg q := \neg(p > \neg q)$, or, equivalently in $\mathsf{C2}$, $\Diamond p \wedge p > q$;
- $\Box p := \neg p > p$;
- $\Diamond p := \neg\Box\neg p$, or equivalently, $\neg(p > \neg p)$.

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 in this appendix 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 unique function from lists of sentences to sentences such that:

$$\underline{\langle\rangle} := \top$$

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

Example A.2. If $p, q \neq \bot$, 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 $p \wedge (\overline{p} > q) \wedge \Diamond\overline{p} \wedge \Box(p \vee q)$.

We call a list τ "consistent" iff $\underline{\tau}$ is consistent (in L). 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 $j < k$, $(\neg\bigvee\tau^{[:k]}) > \tau^k$ would be inconsistent with $\Diamond\tau^k$. Also, if $\bot$ occurs anywhere in a consistent list, every subsequent element of the list must also be $\bot$: if $p \nvdash \bot$, $\underline{\tau + \langle \bot, p \rangle}$ 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 $\neg(\bigvee\tau > \neg p)$. The consistent lists we will be concerned with all contain an inconsistency, if at all, only as the last element.

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

Lemma A.3. If τ^k entails $p \wedge q$ and every element of $\tau^{[:k]}$ entails $\overline{p}$, then $\underline{\tau}$ entails $p > q$.

Proof. We use the following facts about any logic including C2:

CMon	$(p > qr) \to (pq > r)$
Antecedent Substitution	If $\vdash p \leftrightarrow q$ then $\vdash (p > r) \leftrightarrow (q > r)$

$\underline{\tau}$ entails $\neg\bigvee\tau^{[:k]} > \tau^k$ and hence both $\neg\bigvee\tau^{[:k]} > p$ and $\neg\bigvee\tau^{[:k]} > q$. 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 antecedent substitution, $\underline{\tau}$ entails $p > q$. □

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 $\underline{\tau + \langle q_i \rangle}$ for some q_i, or p is consistent with $\underline{\tau + \langle \bot \rangle}$.

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

∨-Distribution	$(p > (q_1 \vee \cdots \vee q_n)) \to ((p > q_1) \vee \cdots \vee (p > q_n))$

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 $\underline{\tau + \langle q_i \rangle}$; otherwise, p is consistent with $\underline{\tau + \langle \bot \rangle}$. □

Lemma A.5. If $q \neq \bot$ and p is consistent with $(\neg\bigvee X \gg q)$, or $q = \bot$ and p is consistent with $(\neg\bigvee X > q)$, there is a consistent list τ of elements of X such that p is consistent with $\underline{\tau + \langle q \rangle}$.

Proof. We cover the case where $q \neq \bot$; the other case is similar. We first show that for any sequence θ consisting of members of X other than $\bot$, if $p \wedge (\neg \bigvee X \gg q) \wedge \underline{\theta}$ is consistent, then there is some p in $X \cup \{\bot\}$ but not in θ such that $p \wedge (\neg \bigvee \tau X \gg q) \wedge \underline{\theta + \langle p \rangle}$ is consistent. This follows from the previous lemma, since the disjunction of all elements of X with $\neg \bigvee X$ is a theorem, and $\underline{\theta + \langle \neg \bigvee X \rangle}$ is inconsistent with $\Diamond \neg \bigvee X$ which follows from $(\neg \bigvee X \gg q)$, while $\underline{\theta + \langle p \rangle}$ is inconsistent when p is in θ. But if p was not consistent with $\underline{\tau + \langle q \rangle}$ for any τ consisting entirely members of X, then the relevant p can never be q, so we have that every θ can be extended with some further element of X, which is impossible since X is finite. □

Thanks to these nice properties, we can use the list-to-sentence operation 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 finite set of atoms A, the sets $Y_n(A)$ (the "depth-n state descriptions over A") are defined as follows.

- $Y_0(A)$ is the set of all consistent conjunctions that include exactly one of p and $\overline{p}$ for each $p \in A$.

- $Y_{n+1}(A)$ is the set of all consistent sentences of the form $\underline{\tau + \langle \bot \rangle}$, where τ is any list of elements of $Y_n(A)$.

Example A.7. Suppose our logic L is C2 and A is $\{p, q\}$. Then $Y_0(A)$ is the 4-member set $\{pq, p\overline{q}, \overline{p}q, \overline{pq}\}$.

$Y_1(A)$ contains $\underline{\tau + \langle \bot \rangle}$ for each nonempty, nonrepeating sequence τ of elements of $Y_0(A)$. There are 64 such sequences (24 of length 4, 24 of length 3, 12 of length 2, and 4 of length 1). $\underline{\tau + \langle \bot \rangle}$ is consistent in C2 for each such sequence, although this is not true for all extensions of C2: for example, if we added the axiom $p \to \Box p$ to C2, $Y_1(A)$ 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 $\underline{\langle \overline{pq}, \bot \rangle}$.

$Y_2(A)$ is much bigger. However, it does not contain $\underline{\langle \tau, \bot \rangle}$ for every nonrepeating, nonempty list τ of elements of $Y_1(A)$. For example, it does not contain

$$\underline{\langle \underline{\langle pq, p\overline{q}, \bot \rangle}, \underline{\langle \overline{p}q, \bot \rangle}, \bot \rangle}$$

since this is inconsistent in C2: $\underline{\langle pq, p\overline{q}, \bot \rangle}$ entails $\neg(pq) > p\overline{q}$, and hence also $\neg\underline{\langle pq, p\overline{q}, \bot \rangle} > p\overline{q}$ by CMon; this is inconsistent with $\neg\underline{\langle pq, p\overline{q}, \bot \rangle} \gg \underline{\langle \overline{p}q, \bot \rangle}$.

For $\underline{\langle \tau_1, \ldots, \tau_n, \bot \rangle}$ to be consistent (where each τ_i is in Y_n), the list derived from this by first replacing each $\underline{\tau_i}$ with the first element of τ_i, and then deleting all but the first occurrence of every element in the result, must be τ_1. (In C2, this is in fact the only constraint.)

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

Proof. By induction on n. The base case is true since the elements of $Y_0(A)$ 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 $s \in Y_{n+1}(A)$ and p and q have modal depth $\leq n$, s entails one of $p > q$ and $\neg(p > q)$. Any such s will be of the form $\underline{\tau + \langle\bot\rangle}$ where each element of τ is in Y_{n+1}. Suppose that the first element of $\tau + \langle\bot\rangle$ 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, $\underline{\tau + \langle\bot\rangle}$ entails $p > q$. Otherwise, the first element of $\tau + \langle\bot\rangle$ that entails p does not entail q. Call this element t. By the induction hypothesis, t entails $\overline{q}$ and all of its predecessors entail $\overline{p}$, so $\underline{\tau + \langle\bot\rangle}$ entails $p > \overline{q}$ by Lemma A.3. Moreover, t is not $\bot$ since it does not entail q, so $\underline{\tau + \langle\bot\rangle}$ entails $\Diamond t$ and hence $\Diamond p$, which together with $p > \overline{q}$ entails $\neg(p > q)$. □

Lemma A.9. If p is consistent, it is consistent with some element of $Y_n(A)$ for every $n \geq 0$.

Proof. By induction on n. The base case holds since $\bigvee Y_0(A)$ is a tautology. For the induction step, we note that since the disjunction of $Y_n(A)$ is a theorem by the induction hypothesis, $p \wedge (\neg\bigvee Y_n(A) > \bot)$ is consistent whenever p is, so by Lemma A.5, there is a sequence τ of elements of $Y_n(A)$ such that p is consistent with $\underline{\tau + \langle\bot\rangle}$. $\underline{\tau + \langle\bot\rangle}$ is our desired element of $Y_{n+1}(A)$. □

The upshot of this is that if we want to show that every L-consistent sentence has a model of a certain sort, it suffices to show that every member of every $Y_n(A)$ has a model of that sort.

A.1 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 will provide an introduction to the proof techniques that we will wheel out for proving the completeness of the successively stronger logics C2.F and C2.FS.

Definition A.10. For a given finite set of atoms A and natural number n, we define an order-model $M_{A,n}$.

- The worlds of $M_{A,n}$ are all of the depth m state descriptions over A for $m \leq n$.
- The order relation of $M_{A,n}$ is defined as follows: where $s = \underline{\tau + \langle\bot\rangle}$, $t <_s u$ iff for some i, $u = \tau^i$, and either $t = s$ or $t = \tau^j$ for some $j < i$.
- The valuation of $M_{A,n}$ is the obvious one: p_i is true at s iff s entails p_i. (Atoms not in A are thus false everywhere.)

Lemma A.11. Each world in $M_{A,n}$ is true at itself.

Proof. We prove, by induction on n, that every sentence of (modal) depth $\leq m$ with atoms in A is true in $M_{A,n}$ at a state description s of depth $\geq m$ iff s entails p. The lemma follows immediately from this.

– *Atoms:* immediate from the definition of $M_{A,n}$.

– *Conjunction:* obvious.

– *Negation:* by Lemma A.8, s entails $\neg p$ when p is a depth $\leq m$ sentence that s doesn't entail p.

– *Conditional:* Suppose $s = \overline{\tau + \langle \bot \rangle}$ is a state-description of depth $\geq m+1$ and $p > q$ is a conditional of depth $\leq m+1$, meaning that p and q must have depth $\leq m$.

  (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 $\geq m$ state description, we have 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 Modus Ponens). 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 $\leq m$ sentences s entails. So by Lemma A.3, s entails $p \gg \neg q$, hence 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 $p > \neg q$ and thus does entail $p > q$ by Lemma A.8. 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 $\leq m$ sentences. But s entails $\Box(\tau^0 \vee \cdots \vee \tau^{-1})$, so s entails $\Box \neg p$, and hence also $p > q$. □

The completeness of C2 for finite order models is immediate from Lemma A.11. Suppose p is a depth n sentence that is consistent in C2. Then p is equivalent to a normalized depth n term over the set of atoms appearing in p. Hence there is at least one depth n state description that entails p in C2 (for instance, the first disjunct of the normalized depth n term); each such state-description is true at itself in $M_{A,n}$, and hence p is true at all of them in $M_{A,n}$.

B Completeness of 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, and hence for finite flat order frames. We build on our earlier definitions, now assuming that our underlying logic L includes C2.F.

The key new results we get by adding Flattening are as follows.

Lemma B.1. If τ is consistent, τ^{-1} entails $q > r$, and every element of $\tau^{[:-1]}$ entails $\overline{q}$, then every consistent element of $\tau^{[:-1]}$ is consistent with $q > r$.

Proof. Suppose for contradiction that τ is consistent, τ^{-1} entails $q > r$, every member of $\tau^{[:-1]}$ entails $\overline{q}$, and τ^k is consistent, but not consistent with $q > r$. $\underline{\tau}$ entails $\underline{\tau^{[:-1]}} \wedge ((\neg \bigvee \tau^{[:-1]}) > \tau^{-1})$, 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. $\underline{\tau^{[:k+1]}}$ 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 this is ruled out by the hypothesis that τ^k is inconsistent with $(q > r)$. ◻

Lemma B.2. If τ is consistent, then for each element τ^k other than τ^{-1}, there is a consistent list θ of elements of τ such that $\theta^0 = \tau^k$, $\theta^{-1} = \tau^{-1}$, and the elements of $\tau^{[k+1:]}$ all occur, in the same order, in θ.

Proof. We may suppose without loss of generality that the elements of τ are pairwise inconsistent, since if not we can just replace each one with its conjunction with the negations of all its predecessors. If τ is consistent and of length $j+1$ and $\tau^{-1} \neq \bot$, then for each $k < j$,

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

is consistent. Since $\neg \bigvee \tau^{[:-1]} \vdash \neg \bigvee \tau^{[:k]}$, the $\gg$-Flattening Rule says that this formula entails

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

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, there must be a list θ of elements of τ, ending with τ^{-1}, such the conjunction of this sentence with $\underline{\theta}$ is consistent. And clearly, for this to be consistent, π^0 must be τ^k and all of $\tau^{[k+1:-1]}$ must occur in π in the same order. ◻

The previous lemma suggests what will turn out to be a key contrast, between two kinds of consistent lists:

Definition B.3.

- τ is *direct* iff there is a consistent 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 consistent and not direct.

Note that given Lemma B.2, when τ is circuitous, we can associate each element t of τ other than the last with a permutation π_t of τ that begins with t and has the same last element as τ.

Example B.4. Recall the depth-1 state descriptions from **??**:

$$A := \underline{\langle pq, \overline{p}q, p\overline{q}, \bot\rangle}$$
$$B := \underline{\langle \overline{p}q, p\overline{q}, \bot\rangle}$$
$$C := \underline{\langle p\overline{q}, \bot\rangle}$$
$$D := \underline{\langle \overline{p}q, pq, p\overline{q}, \bot\rangle}$$

$\langle A, B, C\rangle$ is direct, since $\langle B, C\rangle$ is also consistent, has the same last element, but does not include A. By contrast, $\langle A, D, C\rangle$ is circuitous, since $\underline{\langle D, C\rangle}$ and $\underline{\langle A, C\rangle}$ are both inconsistent.

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

Definition B.5. We define $\uparrow\tau$ recursively, based on the length of τ. For the base cases, when the length of τ is 0 or 1, $\uparrow\tau := \langle\rangle$. When the length of τ is 2, $\uparrow\tau := \tau^{[:1]}$ (i.e. $\langle\tau^0\rangle$).

For the recursion step, when τ is of length $k > 2$, there are two cases, depending on whether τ is direct or circuitous.

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

  If $j = k-2$, let $\theta = \langle\tau^j\rangle$. If $j < k-2$, then we know from Lemma B.2 that there is a consistent sequence that contains every element of τ, begins with τ^{-2} and ends with τ^{-1}. Let θ^+ be the first such sequence, and let θ be its initial segment up to and including the occurrence of τ^j; note that this is also consistent and has length $< k$.

  Then define
  $$\uparrow\tau := \uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\pi$$

– Case 2: τ is circuitous. Then by Lemma B.2, for each element t of τ other than the last, there is a consistent list that begins with t, ends with $\tau^{[:-1]}$, and contains every element of τ. Define a function π such that for each t in τ other than its last element, $\pi(t)$ is such a list: τ if t is the first element of τ; otherwise, the alphabetically earliest such list. Let $\pi^-(t) := \pi(t)^{[:-1]}$, $g(t) := \pi(t)^{-1}$. Then define:

$$\uparrow\tau := \uparrow\pi^-(\tau^0) + \uparrow\pi^-(g(\tau^0)) + \uparrow\pi^-(g(g(\tau^0))) + \uparrow\pi^-(g(g(g(\tau^0)))) + \cdots$$

We also define $\uparrow^+\tau$ to be $\uparrow\tau + \langle\tau^{-1}\rangle$.

Lemma B.6. Whenever τ is consistent and of length at least 1, the elements of $\uparrow\tau$ are exactly those of $\tau^{[:-1]}$, and their first occurrences in $\uparrow\tau$ are in the same order as in $\tau^{[:-1]}$. (And hence the elements of $\uparrow^+\tau$ are exactly those of τ, and their first occurences are in the same order as in τ.)

Proof. By induction on the length of τ.

Base cases (1, 2): obvious.

Induction step: Suppose τ is consistent and of length k. If it isdirect, $\uparrow\tau$ is $\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\pi$, where θ and π are sequences of elements of τ of length $< k$. By the induction hypothesis, all elements of τ except the last two already occur in $\uparrow\tau^{[:-1]}$, in the same order in which they occur in τ (i.e. in $\tau^{[:-1]}$). Moreover, the penultimate element of τ occurs later in $\uparrow\tau$,either as the first element of $\uparrow\theta$ (if θ has length at least 2) or else as the first element of $\uparrow\pi$ (if θ has length 1). And furthermore, neither $\uparrow\theta$ nor $\uparrow\pi$ has any elements not in $\theta^{[:-1]}$ or $\pi^{[:-1]}$ 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.

Meanwhile, if τ approaches p circuitously, $\uparrow\tau$ is

$$\uparrow\pi^-(\tau^0) + \uparrow\pi^-(g(\tau^0)) + \uparrow\pi^-(g(g(\tau^0))) + \uparrow\pi^-(g(g(g(\tau^0)))) + \cdots$$

where π^- and g are as defined above. By the induction hypothesis, $\uparrow\pi^-(\tau^0)$, i.e. $\uparrow\tau^{[:-1]}$, 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 $\pi^-(g(\tau^0))$ and hence of $\uparrow\pi^-(g(\tau^0))$, 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. ◻

Lemma B.7 (Backtracking)**.** Suppose τ is a consistent sequence of Y_n state descriptions, $\sigma + \langle s\rangle$ is a subsequence of $\uparrow^+\tau$, and q, r are sentences of modal depth $< n$ such that s entails $q > r$ and every element of σ entails $\neg q$. Then every element of σ entails $q > r$.

Proof. By induction on j. Base cases for 0 and 1 are are trivial. Base case for 2: $\uparrow^+\tau = \tau$, so the only nontrivial case is where $s = \tau^1$ and $\sigma = \langle\tau^0\rangle$. Suppose τ^1 entails $q > r$ and τ^0 entails $\neg q$. Then $q > r$ is equivalent to $(\neg\tau^0 \wedge q) > r$, which is equivalent by Flattening to $\neg\tau^0 > ((\neg\tau^0 \wedge q) > r)$, which is equivalent to $\neg\tau^0 > (q > r)$, which is entailed by $\neg\tau^0 > \tau^1$. Since τ^0 is consistent with $\neg\tau^0 > \tau^1$, it is thus consistent with $q > r$. Since it's a depth n state description and q and r are depth $< n$, we can conclude that it entails $q > r$, i.e. that every element of σ entails $q > r$.

For the induction step, suppose the claim holds for sequences of length $\leq k$, and suppose τ is consistent and of length $k+1$.

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

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

with θ, t, and π sequences of length $\leq k$ as in Definition B.5. If σ is a subsequence of $\uparrow\tau^{[:-1]}$ or $\uparrow\theta$ or $\uparrow\pi$, $\sigma + \langle s\rangle$ is a subsequence of $\uparrow^+\tau^{[:-1]}$, $\uparrow^+\theta$, or $\uparrow^+\pi$, so the claim follows from the induction hypothesis. If σ is of the form $\sigma_1 + \sigma_2$ where σ_1 is a final subsequence of $\uparrow\theta$ and σ_2 is an initial subsequence of $\uparrow\pi$, then we first appeal to the induction hypothesis for π to show that every element of σ_2 entails $q > r$. In particular, σ_2^0 entails $q > r$, and $\sigma_1 + \langle\sigma_2^0\rangle$ is a subsequence of $\uparrow^+\theta$, so by the induction hypothesis for θ every element of σ_1 also entails $q > r$. Finally, if σ is of the form $\sigma_1 + \sigma_2$ where σ_1 is a final subsequence of $\uparrow\tau^{[:-1]}$ and σ_2 is an initial subsequence of $\uparrow\theta + \uparrow^+\pi$, then every element of σ_2 entails $q > r$ by what we just showed. But $\sigma_1 + \langle\sigma_2^0\rangle$ is a subsequence of $\uparrow^+\tau^{[:-1]}$, so by the induction hypothesis applied to $\tau^{[:-1]}$, every element of σ_1 also entails $q > r$.

Case 2: τ is circuitous. Let π^- and g be as in the definition of $\uparrow$. Then either (i) σ is a subsequence of $\uparrow\pi^-(t)$ for some element t of $\tau^{[:-1]}$, or (ii) σ is of the form $\sigma_1 + \sigma_2$, where for some t, σ_1 is a tail of $\uparrow\pi^-(t)$ and σ_2 is an initial subsequence of $\uparrow\pi^-(g(t))$, or (iii) σ is of the form

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

for some t, σ_1 which is a tail of $\uparrow\pi^-(t)$ and σ_2 which is an initial subsequence of $\uparrow\pi^-(g^{n+1}(t))$, or (iv) s is the last element of τ and σ is of the form

$$\sigma_1 + \uparrow\pi^-(g(t)) + \uparrow\pi^-(g(g(t))) + \cdots$$

where for some t, σ_1 is a tail of $\pi^-(t)$. In situation (i), we can appeal directly to the induction hypothesis for $\pi^-(t)$. In situation (ii), we first use the induction hypothesis for $\pi^-(g(t))$ to show that every element of σ_2 entails $q > r$, and then use the induction hypothesis for $\pi^-(t)$ and the fact that $\sigma_1 + \langle\sigma_2^0\rangle$ is a subsequence of $\uparrow^+\pi^(t)$ to show that every element of σ_1 also entails $q > r$. In situation (iii), we first use the same method to show that every element of $\uparrow\pi^-(g^n(t)) + \sigma_2$ entails $q > r$. Since every element of $\tau^{[:-1]}$ except $g^{n+1}(t)$ occurs in $\pi^-(g^n(t))$, 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 entails $q > r$. Finally, in situation (iv), we first that since every element of $\tau^{[:-1]}$ other than $g(g(t))$ occurs in $\uparrow\pi^-(g(t))$, and $g(g(t))$ is the first element of $\uparrow\pi^-(g(g(t)))$, the elements of σ are exactly the elements of $\tau^{[:-1]}$. Thus every element of $\tau^{[:-1]}$ entails $\neg q$, while τ^{-1} entails $q > r$, and so by Lemma B.1, every element of $\tau^{[:-1]}$, and hence every element of σ, entails $q > r$. ◻

Definition B.8. Given a sequence τ of depth-n state-descriptions such that $\underline{s\tau + \langle\bot\rangle}$ is consistent, $\mathfrak{M}_\tau$ is the ordinal sequence-model whose set of protoworlds is the set of depth n state descriptions over the atoms of s, and its sequences are the tails of $\uparrow\tau + \langle\bot\rangle$. The valuation is the obvious one: atom p_i is true at tail σ iff the first element of σ entails p_i (or equivalently, has it as a conjunct).

Lemma B.9. In $\mathfrak{M}_\tau$, every sentence p of depth $\leq n$ is true at a sequence σ iff it is entailed by σ^0.

Proof. By induction on complexity.

Atoms: given by the valuation.

Boolean operations: obvious.

Conditional. Suppose q and r are of depth $\leq k$. Suppose $q > r$ is not true at σ. Then for some β, $q \wedge \neg r$ is true at $\sigma^{[\beta:]}$ and 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 And-to-if, σ^β entails $q > \neg r$. So by Lemma B.7, σ^α entails $q > \neg r$ for all $\alpha < \beta$, and in particular σ^0 entails $q > \neg r$. 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 does not entail $q > \neg r$ by what we just proved ,and hence σ^0 entails $q > r$ (because it's a depth n state-description). 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 $\sigma + \langle\bot\rangle$ is a subsequence of $\uparrow^+\tau + \langle\bot\rangle$ and $\bot$ entails $q > r$, we can apply Lemma B.7 again to conclude that every element of σ, and thus in particular σ^0, entails $q > r$. ◻

Lemma B.10. If $\underline{\tau + \langle\bot\rangle}$ is a depth-$n + 1$ state description, it is true in $\mathfrak{M}_\tau$.

Proof. Given the definition of $\underline{(\tau + \langle\bot\rangle)}$, it is easy to see that it is true at a sequence in an ordinal sequence model iff the state-descriptions true at tails of that sequence are exactly those that occur in τ, and the order of their first occurrences is given by τ. Given the previous lemma, this means it is true in $\mathfrak{M}_\tau$ so long as the elements of $\uparrow(\tau + \langle\bot\rangle)$ 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 B.6. ◻

This establishes our first result:

Theorem B.11. If L extends C2.F, each consistent sentence of L is true in some ordinal sequence model.

Proof. Given a consistent sentence p of modal depth n, there must be a depth-n state description $\underline{\tau + \langle\bot\rangle}$ over the atoms used in p that entails p; then by the previous lemma, p will be true in $\mathfrak{M}_\tau$. □

As a bonus, our proof also yields a second completeness result:

Theorem B.12. If L extends C2.F, each consistent sentence of L is true in some finite flat order-model.

Proof. Ordinal sequence models are flat order-models, so it suffices to check that $\mathfrak{M}_\tau$ is always *finite*. (Finite in the sense that the domain contains finitely many worlds = sequences; the sequences themselves may be infinite sequences in the sense of having an infinite ordinal as their domain). We prove this by proving by induction on the length of τ that whenever τ is consistent, $\uparrow\tau$ has only finitely many tails.

Base cases (0, 1, 2): trivial.

Induction step: Suppose τ is of length $k+1$. If it is directly, then $\uparrow\tau p$ is of the form $\uparrow\tau^{[:-1]} + \uparrow\theta + \uparrow\pi$, where $\tau^{[:-1]}$, θ, and π are all of length $\leq k$. By the induction hypothesis, each of $\uparrow\tau^{[:-1]}$, $\uparrow\theta$, and $\uparrow\pi$ has only finitely many tails. But every tail of $\uparrow\tau$ is either (i) a tail of $\uparrow\pi$, or (ii) of the form $\sigma + \uparrow\pi$, where σ is a tail of $\uparrow\theta$, or (iii) of the form $\sigma + \uparrow\theta + \uparrow\pi$, where σ is a tail of $\uparrow\tau^{[:-1]}$. So there are only finitely many such tails.

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

$$\uparrow\pi^-(\tau^0) + \uparrow\pi^-(g(\tau^0)) + \uparrow\pi^-(g(g(\tau^0))) + \cdots$$

where g is a function that maps elements of $\tau^{[:-1]}$ to other elements of $\tau^{[:-1]}$, and π^- is a function that maps each element s of $\tau^{[:-1]}$ to a consistent list of length k. By the induction hypothesis, each of these has only finitely many tails. And there are only finitely many of them, so there are only finitely many ways of picking a tail of $\uparrow\tau$. □

C Completeness for C2.FS

Suppose now that our logic L includes C2.FS, i.e. the result of adding Sequentiality to C2.F:

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

Or, equivalently in C2, in contraposed form:

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

With this extra assumption, we can now show the following:

Lemma C.1. If τ is circuitous, τ^{-1} is $\bot$.

Proof. Suppose for contradiction that τ is circuitous and its last element is some $q \neq \bot$. Define:

$$\begin{aligned} s &:= \tau^0 \\ d &:= \bigvee \tau^{[1:-2]} \\ s^+ &:= s \wedge (\neg(s \vee d) \gg q) \\ d^+ &:= d \wedge (\neg(s \vee d) \gg q) \\ p &:= d^+ \vee q \end{aligned}$$

Note that s and d are both inconsistent with q. Further, d^+ is obviously inconsistent with $p > q$, since it entails $p \wedge \neg q$. We can also show that s^+ is inconsistent with $p > q$. Since τ is circuitous, $\langle s, q \rangle$ is inconsistent, meaning that $s \wedge (\neg s \gg q)$ is inconsistent. If so,

$$\text{(I)} \qquad s^+ \wedge ((d \vee q) \gg q)$$

must also be inconsistent: since $\neg s$ is equivalent to $\neg(s \vee d) \vee d \vee q$, $\neg(s \vee d) \gg q$ and $(d \vee q) \gg q$ jointly entail $\neg s \gg q$ by OR. By Flattening, s^+ entails $\neg s > (\neg(s \vee d) \gg q)$. It also entails $\neg s > (d \vee q)$ (by OR, since $\neg s$ is equivalent to $d \vee \neg(s \vee d)$); so by CMon, we have $(\neg s \wedge (d \vee q)) > (\neg(s \vee d) \gg q)$, or more simply, $(d \vee q) > (\neg(s \vee d) \gg q)$. Looking at the definitions of s^+ and d^+, we see that this is equivalent to $(d \vee q) > (d^+ \vee q)$. So by CTrans,

$$\text{(II)} \qquad s^+ \wedge ((d^+ \vee q) \gg q)$$

entails I and is thus also inconsistent. So $s^+ \vee d^+$ is inconsistent with $p > q$.

Since $\neg(s^+ \vee d^+)$ is weaker than $\neg(s \vee d)$, by Flattening s^+ entails $\neg(s^+ \vee d^+) > (\neg(s \vee d) \gg q)$, and hence $\neg(s^+ \vee d^+) \gg \neg(s \vee d)$. So by CTrans, s^+ also entails $\neg(s^+ \vee d^+) \gg q$, and hence $(p > q) \gg q$. and hence $(p > q) \gg p$.

Now finally we can appeal to the contraposed form of Sequentiality, to get that s^+ also entails $(p \wedge (\neg p\overline{q} > q)) \gg \neg q$, or equivalently, $(q \vee (d^+ \wedge (\neg d^+ > q))) \gg \neg q$. Since d^+ entails $\neg d^+ > \neg d$ by Flattening, by the same reasoning as above, this is also equivalent to $(q \vee (d^+ \wedge (\neg d > q))) \gg \neg q$. But the stipulation that τ is circuitous means that $d^+ \wedge (\neg d \gg q)$ is *not* consistent: if this were consistent, there would have to be some consistent subsequence of τ ending in q and excluding s. Thus, $d^+ \wedge (\neg d > q)$ is equivalent to $d^+ \wedge \Box d$, meaning that s^+ entails $(q \vee (d^+ \wedge \Box d)) \gg \neg q$, and hence $(q \vee \neg\Diamond q) \gg \neg q$, i.e. $\neg((q \vee \neg\Diamond q) > q)$. But by Flattening, $\Diamond q$ entails $(q \vee x > \Diamond q)$ for all x, and in particular $(q \vee \neg\Diamond q) > q$. We can conclude that s^+ entails $\neg\Diamond q$. But this contradicts the stipulation that τ is consistent. □

Using this, we can show

Lemma C.2. Whenever τ is consistent and does not end with $\bot$, $\uparrow\tau$ is finite.

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

Lemma C.3. Whenever τ is consistent, $\upharpoonleft\tau$ is at most of length ω.

Proof. Given the previous lemma, it suffices to prove the result in the last element of τ is $\bot$. We do so by induction on the length of τ. The base cases (0,1,2) are trivial. For the first part of the induction step, suppose τ is direct, so that $\upharpoonleft\tau$ is of the form $\upharpoonleft\tau^{[:-1]} + \upharpoonleft\theta + \upharpoonleft\pi$. 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 the $\upharpoonleft\tau$ is at most of length ω. For the second part of the induction step, suppose τ is circuitous. Then $\upharpoonleft\tau$ is

$$\upharpoonleft[\pi^-(\tau^0), +]\upharpoonleft\pi^-(g(\tau^0)) + \upharpoonleft\pi^-(g(g(\tau^0))) + \cdots$$

where each $\pi^-(t)$ is a sequence not ending in $\bot$. By the previous lemma, all these sequences are finite; so the sum of all of them has order type ω. □

Hence:

Theorem C.4. C2.2 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 C.5. C2.2 is complete for finite ordinal sequence models of order-type *exactly* ω.

Proof. For any 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 (nonempty) tail of σ iff the ω-padding of τ is a tail of the ω-padding of σ, and the omega-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 omega-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:

Theorem C.6. C2.FS is complete for finite flat ancestral order models.

D The McKinsey axiom

In this section we consider logics L that include, along with C2.F, the McKinsey axiom

M $$\Diamond\Box p \vee \Diamond\Box\neg p$$

We'll use the following consequence of the axiom in the context of S4:

M′ $$\Box(p_1 \vee \cdots \vee p_n) \to (\Diamond\Box p_1 \vee \cdots \vee \Diamond\Box p_n)$$

Proof. By induction on n. Base case trivial. Induction step: suppose $\Box(p_1 \vee \cdots \vee p_n)$. By M we have $\Diamond\Box(p_1 \vee \cdots \vee p_{n-1}) \vee \Diamond\Box p_n$. 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, $\Diamond\Box p_1 \vee \cdots \vee \Diamond\Box p_{n-1} \vee \Diamond\Box p_n$. □

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

Lemma D.1. If τ is circuitous, it does not end with $\bot$.

Proof. Suppose that τ is consistent and ends with $\bot$; then in particular $\neg(\bigvee \tau^{[0:-1]}) > \bot$, i.e. $\Box(\tau^0 \vee \cdots \vee \tau^{-2})$, is consistent. So by the previous lemma, $\Diamond\Box\tau^0 \vee \cdots \vee \Diamond\Box\tau^{-2}$ 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 consistent, so τ is not circuitous. □

Using this, we can show

Lemma D.2. Whenever τ is consistent and ends with $\bot$, the domain of $\uparrow\tau$ is a successor ordinal.

Proof. An obvious induction on the length of τ. □

Hence

Theorem D.3. C2.FM is complete for finite ordinal sequence models in which the domains of all sequences are successor ordinals.

And putting together this lemma with the one from the previous section, we have

Theorem D.4. C2.FSM is complete for finite list-models, i.e. models whose domain consists of finitely many ordinal-sequences, each of which has a finite ordinal as its domain.

E 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.

Suppose we start with a valuation V on the set of protoworlds P, and the ordinal-sequence model $\mathcal{M}_{V,\alpha}$ whose domain is the set of all ordinal-sequences over P whose domain is less than α (for some given ordinal α). We define a function h that takes a sequence σ in this model's domain and a $\mathcal{L}_{BA}$-sentence p to a set $h(\sigma, p)$ of ordinals in the domain of σ—intuitively, the ones 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.

Any set X of ordinals is well-ordered by $\leq$, and hence there is an order-preserving mapping f_X between X and some ordinal. Thus, for any ordinal sequence σ and a set X of ordinals, we can construct a new ordinal sequence $\sigma\restriction X$, defined by $(\sigma\restriction X)(\alpha) = \sigma(f_X^{-1}(\alpha))$ (if α is in the range of f, else undefined). Note that when $\alpha \in X$, we have $\sigma^{[\alpha:]}\{\beta : \alpha + \beta \in X\}\restriction = \sigma\restriction X[f_X(\alpha) :]$. 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 E.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. Then $\alpha \in h(\sigma, p > q) \subseteq X$. Then $h(\sigma, p > q) = \{0\} \cup \{\alpha + \beta : \beta \in h(\sigma^{[\alpha:]}, q)\}$, so $h(\sigma^{[\alpha:]}, q) \subseteq \{\beta : \alpha + \beta \in X\}$. So we can see that the truth value of $p > q$ at σ is the same as the truth value of q at $\sigma^{[\alpha:]}$, which (by the induction hypothesis) is the same as its truth value at $\sigma^{[\alpha:]}\restriction\{\beta : \alpha + \beta \in X\}$, i.e. $\sigma\restriction X[f_X^{-1}(\alpha) :]$. Since no p-protoworlds occur in σ before position α, none occur in $\sigma\restriction X$ before position $f_X^{-1}(\alpha)$, and so this is the same as the truth value of $p > q$ at $\sigma\restriction X$. □

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

Corollary E.2. For each p, the truth value of p at σ in $\mathcal{M}_{V,\alpha}$ is its truth value at the finite sequence $\sigma h(\sigma, p)\restriction$.

Finally, since throwing all sequences other than $\sigma h(\sigma, p)\restriction$ and its (finitely many) tails out of the domain will not affect the truth value of any sentence, we can conclude that:

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

References

Bacon, Andrew (2015). "Stalnaker's Thesis in Context". In: *Review of Symbolic Logic* 8.1, pp. 131–163.

— (2020). "Logical Combinatorialism". In: *The Philosophical Review* 129.4, pp. 537–589.

Bacon, Andrew and Cian Dorr (2024). "Classicism". In: *Higher-Order Metaphysics*. Ed. by Peter Fritz and Nicholas K. Jones. Oxford: Oxford University Press, pp. 109–190.

Bull, R. A. (1966). “That all Normal Extensions of S4.3 have the Finite Model Property”. In: *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 12, pp. 341–344.

Dale, A. J. (1974). “A Defence of Material Implication”. In: *Analysis* 34.3, pp. 91–95.

Dale, Tony (1979). “A natural deduction system for ‘If Then’”. In: *Logique & Analyse* 22.87, pp. 339–345.

Dorr, Cian, Jeremy Goodman, and John Hawthorne (2014). “Knowing Against the Odds”. In: *Philosophical Studies* 170.2, pp. 277–287.

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

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

Douven, Igor and Sara Verbrugge (2013). “The Probabilities of Conditionals Revisited”. In: *Cognitive Science* 37.4, pp. 711–730.

Edgington, Dorothy (1995). “On Conditionals”. In: *Mind* 104.414, pp. 235–329.

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

Fine, Kit (1971). “The Logics Containing S4.3”. In: *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 17, pp. 371–376.

— (1974). “An incomplete logic containing S4”. In: *Theoria* 40.1, pp. 23–29.

van Fraassen, Bas (1976). “Probabilities of Conditionals”. In: *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Ed. by Harper and Hooker. Vol. I. Dordrecht-Holland: D. Reidel Publishing Company, pp. 261–308.

Gibbard, Allan (1981). “Two Recent Theories of Conditionals”. In: *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Ed. by William L. Harper, Robert Stalnaker, and Glenn Pearce. Dordrecht: Reidel, pp. 211–247.

Goldstein, Simon and Paolo Santorio (2021). “Probability for Epistemic Modalities”. In: *Philosophers’ Imprint* 33.

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

Holliday, Wesley H. and Tadeusz Litak (2019). “Complete Additivity and Modal Incompleteness”. In: *The Review of Symbolic Logic* 12.3, pp. 487–535.

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

— (2015). “Conditionals, conditional probabilities, and conditionalization”. In: *Bayesian Natural Language Semantics and Pragmatics*. Ed. by H. Zeevat and H.-C. Schmitz. Springer International Publishing Switzerland, pp. 71–94.

— (2017). “The Limit Assumption”. In: *Semantics & Pragmatics* 10.

Kaufmann, Stefan (2023). "Bernoulli semantics and ordinal semantics for conditionals". In: *Journal of Philosophical Logic* 52.1, pp. 199–220.
Khoo, Justin (2022). *The Meaning of* If. Oxford University Press.
Khoo, Justin and Paolo Santorio (2018). "Lecture Notes: Probability of Conditionals in Modal Semantics". Unpublished manuscript. July 2018.
Kratzer, Angelika (1981). "The Notional Category of Modality". In: *Words, Worlds, and Contexts: New Approaches in Word Semantics*. Ed. by H. Eikmeyer and H. Rieser. de Gruyter, pp. 38–74.
— (1986). "Conditionals". In: *Chicago Linguistics Society* 22.2, pp. 1–15.
Lewis, David (1973). *Counterfactuals*. Oxford: Blackwell.
— (1976). "Probabilities of Conditionals and Conditional Probabilities". In: *The Philosophical Review* 85.3, pp. 297–315. ISSN: 00318108.
— (1994). "Humean Supervenience Debugged". In: *Mind* 103, pp. 473–90.
Mandelkern, Matthew (2024). *Bounded meaning: The dynamics of interpretation*. In press. Oxford University Press.
McGee, Vann (1985). "A Counterexample to Modus Ponens". In: *The Journal of Philosophy* 82.9, pp. 462–471.
Popper, Karl (1959). *The Logic of Scientific Discovery*. Hutchinson.
Salmon, Nathan U. (2005). *Reference and Essence*. 2nd ed. Amherst, NY: Prometheus Books. Repr. of *Reference and Essence*. 2nd ed. Amherst, New York: Prometheus Books, 2006.
Santorio, Paolo (2021). "Path Semantics for Indicative Conditionals". In: *Mind*.
Schultheis, Ginger (2022). "Counterfactual Probability". In: *Journal of Philosophy*.
Segerberg, Krister (1970). "Modal logics with linear alternative relations". In: *Theoria* 36.3, pp. 301–322.
— (1989). "Notes on Conditional Logic". In: *Studia Logica: An International Journal for Symbolic Logic* 48.2, pp. 157–168.
Stalnaker, Robert (1968). "A Theory of Conditionals". In: *Studies in Logical Theory*. Ed. by Nicholas Rescher. Oxford: Blackwell, pp. 98–112. URL: https://philpapers.org/rec/STAATO-5.
— (1974). *Letter to Bas Van Fraassen*. Jan. 1974.
— (1975). "Indicative Conditionals". In: *Philosophia* 5.3, pp. 269–86. URL: https://philpapers.org/rec/STAIC.
Stalnaker, Robert C. (1970). "Probability and Conditionals". In: *Philosophy of Science* 37.1, pp. 64–80.
Stalnaker, Robert C. and Richmond H. Thomason (1970). "A semantic analysis of conditional logic". In: *Theoria* 36.1, pp. 23–42.
Thomason, S. K. (1974). "An incompleteness theorem in modal logic". In: *Theoria* 40.1, pp. 30–34.
Van Benthem, J. F. A. K. (1978). "Two Simple Incomplete Modal Logics". In: *Theoria* 44.1, pp. 25–37.
Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.

Yalcin, Seth (2007). “Epistemic Modals”. In: *Mind* 116.464, pp. 983–1026.