Counterfactual Worlds

Benjamin Brast-McKie

March 9, 2024

Abstract

This paper extends Kit Fine's (2017a,b,c) truthmaker framework to provide a novel semantics for counterfactual conditionals where possible worlds will be constructed out of *states* and *state transitions*. Rather than taking possible worlds to be points in a model, possible worlds will include both temporal and mereological structure. Accordingly, possible worlds may be compared at a time independent of that time's past and future where the comparison will be carried out in mereological terms instead of invoking a primitive similarity ordering. These elements will make for an elegant account of counterfactual conditionals that avoids the problems that similarity theories face. After reviewing the motivations for this approach, I will provide a compositional semantics for a language with counterfactual and extensional operators, presenting its logic. By extending this language to include temporal operators, I will analyse forwards, backwards, and backtracking counterfactual conditional claims.

Keywords: Counterfactual Conditionals, State Semantics, Hyperintensionality

1 Introduction

Drawing on Kripke's (1963) semantics for the metaphysical modal operators, intensional theories of counterfactual conditionals model *possible worlds* by structureless points. Although treating possible worlds as points is a harmless idealisation in the case of the metaphysical modal operators, the same cannot be said for counterfactual claims of the form ⌜If it were the case that A, then it would be the case that B⌝, or in symbols ⌜$A \boxright B$⌝. In order to motivate a hyperintensional semantics for counterfactuals which takes the possible worlds specified by a model to have both temporal and mereological structure, this section considers two familiar problems facing the *similarity theories* that Stalnaker (1968) and Lewis (1973) developed. Roughly, similarity theories take $A \boxright B$ to be true at a world w just in case B is true at the most similar accessible world(s) to w in which A is true. I will take issue with the following features of views of this kind:

Totality: Similarity orders worlds considered in their entirety.

Restriction: The consequent is only required to be true at the world(s) in which the antecedent is true that are the most similar to the world of evaluation.

Whereas *Totality* leads to counterexamples of the kind raised by Fine (1975a), *Restriction* invalidates the principle of *Simplification of Disjunctive Antecedents* (SDA). Rather than considering attempts to address these issues within an intensional framework, I will draw lessons from the problems that these features raise in order to motivate a hyperintensional alternative. As I will show, the resulting framework can be used to define a variant of the *outcome relation* that Fine (2012) introduces, and so in this right may be considered an extension of Fine's truthmaker semantics for counterfactual conditionals.

The paper will be organised around two central aims: (1) to provide a logic for counterfactual reasoning **CL**; and (2) to provide an intuitively compelling semantic theory. In order to identify a counterfactual logic, I will construct a compositional semantics for a propositional language that validates SDA without also validating *Strengthening the Antecedent* (STA). In addition to offering a means by which to identify valid forms of counterfactual reasoning, the semantics provides an intuitive method for evaluating counterfactual judgements in particular cases. This takes two forms. First, I will require the primitives included in the definition of a model to be easier to understand than the counterfactual claims which these primitives will be used to interpret. Additionally, I will show that the semantic clause for the counterfactual conditional both aligns with and clarifies the pre-theoretic means by which we interpret counterfactual claims. I will take these desiderata to provide broad criteria by which to evaluate the success of a semantic theory for counterfactuals. Rather than attempting to survey the full range of semantic theories that have been developed for counterfactuals, the following subsections will consider three different proposals, drawing lessons from the shortcomings of each in order to help motivate and illuminate what is distinctive about the present approach.

1.1 Totality

In addition to a set of worlds W and an accessibility relation R, similarity theories equip models with a primitive similarity relation (or selection function) so that one may identify the most similar accessible worlds to the world of evaluation which make the antecedent true.[1] Insofar as the worlds in W do not have any internal structure, similarity theories can only identify the most similar worlds to a given world by comparing worlds in their entirety. This does not stop similarity theorists from doing otherwise while working over an intended model which includes the possible worlds themselves along with all of their spatio-temporal structure and various features. However, the features that one might take possible worlds to have in an intended model are not included in the definition of a model. For the purposes of logic, the models of a given object language only have the structure that they are given, and standard intensional models do not assign any internal structure to the possible worlds in W. Rather, W is taken to be any nonempty set whatsoever where the similarity relation stipulates an ordering over the worlds in W by treating each world as a structureless point.

Instead of taking W to be any nonempty set, the present treatment will take worlds to be structured into a history of moments, where each moment is rich in mereological structure. By including these elements in the definition of a world, we may evaluate counterfactuals by considering worlds that differ from the evaluation world at a given time while disregarding their differences at other times. In order to motivate this approach, consider the counterexample that Fine's (1975a, p. 425) raised against Lewis' account:

(`N`) If Nixon were to press the button, there would be a nuclear holocaust.[2]

Here we are to imagine Nixon hovering over the detonator, where one advisor asserts `N` to another. As Fine points out, Lewis' theory predicts that `N` is false since for any world in which Nixon presses the button and there is a nuclear holocaust, there will be

[1] Although I will focus on Lewis' account, similar considerations will apply to Stalnaker's original view. For simplicity, I will often leave the restriction to *accessible* worlds implicit in what follows.

[2] I have modified the original example to maintain present tense. See §4.2 for an analysis of the original.

a much more similar world which includes, "a change that prevents the holocaust but that does not require such a great divergence from reality." For instance, perhaps the button is disconnected. Even so, `N` is true since the differences that result from Nixon pressing the button are irrelevant to the moment we are considering in which he did not press the button but could have. Rather than measuring the similarity of the actual world in its entirety against all of what would result were Nixon to press the button, it is natural to restrict consideration to the moments that result from making minimal changes to the actual moment in which Nixon did not press the button to moments in which he did. Given such a range of minimally changed moments, we may take `N` to make a claim about the disaster that would have resulted even if those possibilities end up diverging widely from the actual course of events. So long as the moments that result from changing the actual moment in which Nixon did not press the button to one in which he did do not include any other changes (e.g., a disconnected button, etc.), `N` may be expected to be true if a nuclear holocaust is guaranteed to follow.

However natural it may be to restrict consideration to the moment in which the antecedent is to be evaluated while ignoring other moments, standard intensional models do not include the resources to do so. Since possible worlds are modeled by points that do not have any temporal structure, a similarity relation can only order possible worlds considered in their entirety. It is this limitation which makes it easy to imagine possible worlds in which Nixon pushed the button that are much more similar to the actual world than to possible worlds in which he pushed the button and a nuclear holocaust took place. Lewis (1979) responded to Fine's challenge by dispensing with our pre-theoretic similarity judgements, opting instead for weighted rules which constrain how the similarity of worlds is to be measured. Instead of precise rules which are easy to apply, Lewis' weighting system consists of open-ended recommendations for how to adjust our untutored intuitions about similarity. Whether one adopts Lewis' rules or not, it is natural to worry that the metalanguage Lewis used to articulate a semantics for counterfactuals is considerably harder to understand than the object language that we intend to interpret.[3] Rather, we should hope to develop the semantics for counterfactuals in a metalanguage for which we have a better understanding than the object language under study. Instead of following Lewis' attempt to maintain an intensional similarity semantics, I will extend Fine's (2017a,b) state semantics to include a primitive *task relation* which encodes the possible transitions between states. By constructing possible worlds from states and tasks, I will present a semantics that appeals to the resulting temporal and mereological structure of possible worlds to accommodate `N` and to provide a compelling theory of logical consequence for counterfactuals.

In order to motivate the mereological account given in §2, the following subsection considers a similarity view that captures the truth of `N` by giving up *Totality*. Instead of taking possible worlds to be structureless, I will take possible worlds to be histories of moments. Assuming that moments are structureless points for the time being, we may order moments by their similarity to a given moment. Relating moments in this way has a number of advantages. To begin with, we may capture the means by which we intuitively evaluate counterfactuals by restricting consideration to the moment in which a counterfactual is evaluated and considering the most similar possible worlds at just that moment. As brought out above, the moments following the moment in which Nixon did not press the button do not enter into our consideration of which possibilities to entertain in evaluating `N`. Additionally, it is by taking possible worlds to be histories of

[3] Something similar may be said for Stalnaker's (1968) view which posits a primitive selection function.

moments that we may hope to interpret tense operators. Since counterfactuals often combine with tense operators, providing a unified semantics for tense and counterfactual operators is extremely natural. Setting these merits aside, a more pressing reason to encode temporal structure into possible worlds is that it is by slicing worlds into moments that we may take each moment to have mereological structure. Given any moment m in which a counterfactual is evaluated, we may appeal to the *parts* of m in order to identify the moments that result from minimally changing m to make the antecedent true without positing a primitive similarity relation. In order to further motivate the mereological structure that I will take each moment to include, the following section will review an important problem for similarity theories stemming from *Restriction*.

1.2 Restriction

Suppose that a similarity theorist were to attempt to accommodate the truth of `N` by giving up *Totality*. Rather than taking possible worlds to be points in a model, possible worlds may be constructed from *possible moments*.[4] Whereas possible worlds are conceived of as complete histories, possible moments are identified with the time slices of those histories, amounting to *maximal possible ways for things to be at an instant*. A moment m is *accessible* to n just in case it is possible for m to transition to n. Letting time have the structure of the integers for simplicity, we may take a *possible world* to be any function from times to moments where every moment is accessible from its predecessor. By taking similarity to hold between possible moments instead of worlds, one may take $A \mathbin{\Box\!\!\rightarrow} B$ to be true in a world w at a time x just in case B is true in w' at x whenever $w'(x)$ is one of the most similar moments to $w(x)$ and A is true in w' at x.[5] Given the actual moment in which Nixon did not press the button, the most similar moments in which Nixon did push the button do not include any broken wires or blown fuses, and so are sure to lead to a nuclear holocaust. Thus `N` is predicted to come out true. Even so, this semantics invalidates *Simplification of Disjunctive Antecedents*:

$$(\textsc{sda})\quad A \vee B \mathbin{\Box\!\!\rightarrow} C \;\vdash (A \mathbin{\Box\!\!\rightarrow} C) \wedge (B \mathbin{\Box\!\!\rightarrow} C).^{6}$$

The same reason that standard similarity theories invalidate SDA continues to apply here: that the consequent is true in only the most similar worlds at the time of evaluation in which the antecedent is true does not require the consequent to be true in all of the most similar moments in which a disjunct of the antecedent is true.[7] For instance, suppose that $A \vee B \mathbin{\Box\!\!\rightarrow} C$ is true in w at x. Since the most similar $A \vee B$-moments to $w(x)$ may all be $A \wedge C$-moments without also all being B-moments, there is no guarantee that $B \mathbin{\Box\!\!\rightarrow} C$ will hold, and so the conclusion of SDA may be false when the premise is true. The trouble is caused by the restriction to the most similar $A \vee B$-moments, since the assumption that the most similar $A \vee B$-moments are C-moments does not imply anything about the most similar B-moments (similarly, A-moments).

For the sake of generality, let *points* be generic elements at which sentences are to be evaluated. We may then observe that the invalidity of SDA does not turn on anything peculiar to possible worlds, moments, or similarity. Rather, any theory of counterfactuals

[4] Alternatively, a similarity theorist may replace worlds with moments so as to maintain an intensional framework. This comes at the cost of forgoing the integration of tense operators that I provide in §4.

[5] Since it does not matter for present purposes whether we employ a selection function or similarity ordering over moments, I will leave 'most similar moments' undefined to accommodate both views.

[6] I will assume that the extensional operators bind more tightly than the counterfactual conditional.

[7] Fine (1975a, 2012) makes this observation for standard intensional theories of counterfactuals.

which imposes restrictions on the set of antecedent points which are required to make the consequent true is liable to invalidate SDA for the same reason given above.

If SDA is to be maintained, no restriction should be imposed on the range of antecedent points in which the consequent is required to be true. *Strict theories* of counterfactuals take this lesson to heart by requiring the consequent to hold in every accessible world in which the antecedent holds. We may extend this approach to encode temporal structure as above by taking $A \boxright B$ to be true in w at x just in case B is true in w' at x whenever $w'(x)$ is accessible from $w(x)$ and A is true in w' at x. It follows that SDA is valid: if C is true in all accessible moments in which $A \vee B$ is true, then C is true in all accessible moments in which A (similarly B) is true. Even so, strict theories validate the principle:

(STA) $A \boxright C \ \vdash A \wedge B \boxright C$.

Given any world w and time x in which $A \boxright C$ is true, it follows by the strict theory that C is true in every accessible moment in which A is true, and so C is also true in every accessible moment in which $A \wedge B$, thereby validating STA. However, there are natural counterexamples to STA. Just because Judy would be angry if Matan were to attend the party, it does not follow that Judy would be angry if Matan and Stav were to both attend the party, since Stav would prevent Matan from arguing with Judy. In response, defenders of STA provide pragmatic explanations for why the counterexamples are only apparent by appealing to contextual shifts in R between the evaluation of the premise and conclusion. Nevertheless, strict theories admit that in any particular context, counterfactual conditionals have the same logic as the strict conditional.[8]

Rather that accepting an error theory of this kind, I will assume for the purposes of this paper that the counterexamples to STA show that the logic for the counterfactual conditional cannot be identified with the logic of the strict conditional. However, if the judgments above are to be preserved, there would seem to be little space between imposing a restriction on the antecedent points in which the consequent is to be evaluated and thereby invalidating SDA, and imposing no restriction at all on the antecedent points in which the consequent is to be evaluated and thereby validating STA. The following subsection considers an alternative that avoids both problems.

1.3 Outcomes

Rather than restricting the set of antecedent points in which the consequent of a counterfactual is required to be true, Fine (2012, p. 237) introduces a primitive *outcome* relation $\rightarrowtriangle$ where $w \rightarrowtriangle_t u$ indicates that, "u is a possible outcome of imposing the change t on the world w." Whereas w and u are possible worlds, t is taken to be a *state* which Fine conceives of as a "fragment" or *part* of a possible world. Instead of settling the truth-value of any sentence at a time the way possible worlds do, states only settle the truth-values of those sentences that they *exactly verify* or *exactly falsify* where states, "tell us what it is *in* the world that makes the statement true if it is true or what it is *in* the world that makes it false if it is false" (p. 235). Building on Fine's (2017a,b) recent work, I will have more to say about states in the following section. For the time being, it will be enough to consider the broad role that states play in his semantics.

Given the addition of states together with the outcome relation, Fine takes $A \boxright C$ to be true in a world w just in case for any exact verifier state t for A, the consequent C is true in any world u which is the outcome of imposing the change t on w. Fine is clear that this *truthmaker semantics* for counterfactuals is based on the following ideas:

[8] Warmbrōd (1981) and Gillies (2007, 2009) provide strict conditional views in intensional frameworks.

(URA) A counterfactual will only be taken to be true when it is true for *any way* in which its antecedent A might be true.

(UVC) A counterfactual will be taken to be true, given some way in which its antecedent might be true, only when its consequent is made true under *any outcome* of the way in which its antecedent is true.

Whereas URA forbids any restriction from being put in place on the set of antecedent states under consideration, UVC forbids any limitation on the range of outcomes that result from imposing a given antecedent state on the world of evaluation. Nevertheless, the antecedent states differ from the outcomes that they induce. Outcome theories of this kind are to be contrasted with strict theories of counterfactuals which require the consequent to be true at all accessible antecedent moments (or worlds) *themselves*. Even so, URA and UVC preserve the unrestricted spirit of strict theories of counterfactuals, indicating a means by which to validate SDA without also validating STA. Since the exact verifiers for $A \vee B$ include the exact verifiers for A as a subset, it follows that if C is true in every outcome that results from imposing an exact verifier for $A \vee B$ on w, then C is also true in every outcome that results from imposing an exact verifier for A (similarly B) on w, and so SDA is valid. By contrast, the outcomes that result from imposing an exact verifier for A on a world w may differ considerably from the outcomes that result from imposing an exact verifier for $A \wedge B$ on w, and so even if C is true in all of the former, C may fail to be true in all of the latter, thereby invalidating STA.[9]

Although Fine's truthmaker semantics for counterfactual conditionals validates SDA without also validating STA, a number of challenging questions remain. Postponing discussion of the precise relationship between worlds and states to the following section, I will conclude the present subsection by considering an objection that Fine raises:

> It has been argued, in the second place, that the present semantics is relatively problematic in its conceptual commitments. For the transition relation must itself be understood in terms of counterfactuals. Thus to say that u is a possible outcome of t in w is just to say that, in w, u might obtain if t were to obtain (and also that u is maximal in this respect). (p. 241)

In short, the objection is that Fine's truthmaker semantics is homophonic, employing an analogue of the counterfactual conditional in the metalanguage. Although homophonic semantic clauses may be tolerated for relatively simple operators such as conjunction, negation, and the quantifiers when interpreted in clear formal contexts and there is no other means by which to provide their semantics, the same cannot so easily be said for the modal operators, much less counterfactual conditionals. At least it is desirable to provide a semantic theory for counterfactuals that does not rely on counterfactual primitives.[10] In response, Fine points out that Lewis (1979) may be accused of something similar given that counterfactual assumptions will be difficult to avoid in providing a "suitably doctored" notion of similarity. Instead of requiring a semantics to, "provide an analysis of the locutions with which it deals," Fine settles for the position that his truthmaker semantics provides, "a perspicuous account of how the truth-conditions of the sentences containing the [counterfactual] locutions are to be determined."

Although analysis is by no means the only standard for productive semantic theorising, it is important to establish a solid understanding of the primitive terms used to articulate

[9] I will use the semantics that I defend to provide a countermodel to STA in §3.3.

[10] A similar complaint may be raised for theories that employ causal notions in the metalanguage.

a semantics in order to provide an unambiguous specification of the truth-conditions for the locutions in question. In the case of Fine's outcome relation, it is not clear at what point in a world's history the states are to be imposed. Continuing with the example from before, suppose that the detonator button was only engaged after a security meeting was held between Nixon and his advisors and they settled on a range of defence strategies. After finishing the meeting, Nixon is hovering over the button when one of his advisors asserts `N` to another. Given that the button is engaged at the time of the assertion, it is natural to take `N` to be true. However, given the world w in which `N` is evaluated along with any exact verifier b for 'Nixon presses the button', it is unclear when to impose b on w. In particular, were b to be imposed before the button is engaged, no nuclear holocaust will result. If the range of outcomes that result from imposing b on w are to include such premature button pressings, then `N` will come out false contrary to expectation.

1.4 Plan

In order to avoid imposing states at the wrong time so as to preserve the truth of `N`, we may amend Fine's semantics by taking the outcome relation to specify the time of imposition. Accordingly, $A \boxright B$ is true in w at x just in case for any state t that exactly verifies A and any outcome u that results from imposing t on w at x (i.e., $w \rightarrowtriangle_t^x u$), B is true in u at x. Although the addition of a temporal parameter to the outcome relation avoids the ambiguity brought out above, it remains to clarify how the amended outcome relation is to be understood. At the very least, we should hope to specify principles that constrain the interpretation of '$\rightarrowtriangle$' so as to clarify its intended meaning.

Instead of axiomatising the outcome predicate, the following section will be devoted to providing its definition in terms that we better understand.[11] I will begin by extending Fine's (2017a,b,c) recent version of the truthmaker framework to include a *task relation* $\rightarrow$ that encodes the possible transitions between states. Intuitively, there is a task from the state s to the state t— i.e., $s \rightarrow t$— just in case it is possible for the state s to transition to the state t. Accordingly, the present development will rely on a modal interpretation of the task relation in order to provide a semantics for counterfactuals rather than positing a counterfactual primitive in the metalanguage. Given a discrete theory of time for simplicity, I will show how to construct possible worlds out of states and tasks. Not only will worlds include temporal structure by assigning moments to times, each moment will be taken to have mereological structure and will correspond to what Fine refers to as *world states*. After presenting the construction of possible worlds in the following section, I will draw on these resources in order to provide a semantics and logic for counterfactual conditionals in §3. Since counterfactuals often occur together with tense operators, §4 extends the semantics to include clauses for a number of tense operators in order to regiment forwards, backwards, and backtracking counterfactuals. The paper will conclude in §5 by considering objections and extensions.

2 Constructing Possible Worlds

Although states may be interpreted in a number of different ways, I will take states to be *some things being a specific way* for the purposes of the present application. Whereas possible worlds are complete histories of everything, states are static and typically partial,

[11] I developed this framework to model causal claims and only later applied it to counterfactuals. In attempting to present this semantics, I realised that a modified form of Fine's (2012) outcome relation could be defined, and that providing this definition helps to bring out the contrast with Fine's account.

concerning some limited way for certain things to be at an instant. For instance, given any possible world in which Sanna is sitting at time x, restricting to the part of that world which occurs at time x and makes it true that Sanna is sitting is a state of Sanna sitting. But there is more than just one state of Sanna sitting. Just as the world-time pairs in which Sanna is sitting are intended to cover the many different ways for it to be true that Sanna is sitting, the states of Sanna sitting are the *parts* of those world moments which make it true that Sanna is sitting in all of those different ways. Despite being partial, states are *specific*: just as there is exactly one way for a possible world to be *actual*, there is exactly one way for a state to *obtain*. Put otherwise, neither worlds nor states are multiply realisable but rather model specific realisations.

In addition to being parts of possible worlds at times, states may bear parthood relations to each other, where '$s \sqsubseteq t$' reads 's is a *part* of t'. For instance, the state of Sanna sitting s includes a state of Sanna's legs being bent t as a proper part, where t is a *proper part* of s— i.e., $t \sqsubset s$— just in case $t \sqsubseteq s$ and $s \not\sqsubseteq t$. Following Fine (2017a,b,c), I will take a *state space* $\mathcal{S} = \langle S, \sqsubseteq \rangle$ to be any complete lattice where S is the set of all *states* and $\sqsubseteq$ is the *parthood relation*.[12] Since the least upper bound of any set of states $X \subseteq S$ is unique, we may refer to the least upper bound of X as the *fusion* $\bigsqcup X$ of the states in X. In particular, $\bigsqcup \varnothing \coloneqq \square$ is the *null state* and $\bigsqcup S \coloneqq \blacksquare$ is the *full state*. When $X = \{s, t, \ldots\}$ is finite, it will be convenient to represent the fusion of the states in X as $s.t.\cdots \coloneqq \bigsqcup\{s, t, \ldots\}$. For instance, the fusion of a state s of Sanna sitting and a state k of Kim smiling is the state $s.k$ of Sanna sitting and Kim smiling. In general, a fusion will obtain just in case all of its parts obtain.

Fine goes on to take a *modalized state space* to be any ordered triple $\langle S, P, \sqsubseteq \rangle$ where $\langle S, \sqsubseteq \rangle$ is a state space, $P \subseteq S$ is a non-empty set of *possible states*, and s is possible for every part $s \sqsubseteq t$ of a possible state t. Whereas every state in P can obtain, the states in S/P are *impossible* and so cannot obtain.[13] Fine takes the states s and t to be *compatible* just in case their fusion is possible, i.e., $s \circ t \coloneqq s.t \in P$. For instance, although any state of Sanna sitting is compatible with any state of Kim smiling, the same cannot be said for the states of Sanna standing which are incompatible with the states of her sitting. Having introduced the notion of compatibility, Fine takes s to be a *world state* just in case s is possible and includes all states compatible with s as parts. A *world space* is any modalized state space where every possible state is part of a world state.

Although the world spaces that Fine introduces include both mereological and modal structure, they do not include any temporal structure. For instance, assuming a discrete theory of time for simplicity, suppose that it is possible for Sanna to transition from her sitting state s at a time x to a standing state t at $x + 1$. Moreover, suppose that Sanna may transition from the standing state t at time $x + 1$ to a jumping state j at $x + 2$, but cannot transition directly from s at x to j at $x + 1$. Or to take a more dramatic example, Sanna cannot transition in any amount of time from her sitting state s to a state k of Kim smiling. So far, nothing included in the definition of a world space permits us to distinguish between possible and impossible transitions between states. In particular, we should like to identify which transitions between world states are possible. Even though Fine provides the resources to define world states, there is no way to distinguish sequences of world states which count as possible worlds from sequences of world states

[12] If $\langle S, \sqsubseteq \rangle$ is a partial order, $s \in S$ is an *upper bound* of $X \subseteq S$ just in case $x \sqsubseteq s$ for all $x \in S$, and a *least upper bound* of X just in case s is an upper bound of X and $s \sqsubseteq y$ for every upper bound y of X. A *complete lattice* is any partial order $\langle S, \sqsubseteq \rangle$ where every subset $X \subseteq S$ has a least upper bound.

[13] Without admitting impossible states, the exact verifiers for a sentence could not be closed under fusion, and so $A \equiv A \wedge A$ may fail to hold. I provide further discussion of idempotence in Brast-McKie (2021).

that do not. Accordingly, the following subsection will introduce a primitive task relation $\rightarrow$ in order to distinguish possible from impossible state transitions.

2.1 Task Space

Given any states s and t, the ordered pair $\langle s, t\rangle$ represents the *state transition* from s to t where t obtains as a result of s obtaining. However, not all state transitions are possible. Rather, we may say that there is a *task* $s \rightarrow t$ just in case it is possible for t to obtain as a result of s obtaining. For any state s, there is a *trivial transition* $\langle s, s\rangle$ which leaves s unchanged. So long as it is possible for s to obtain, s may obtain as a result of s obtaining by making no change to s, and so every possible state s has a *trivial task* $s \rightarrow s$. Conversely, if s cannot obtain, then it is impossible for s to obtain as a result of s obtaining even by making no change, and so $s \nrightarrow s$. Thus it is possible for s obtain just in case there is a trivial task $s \rightarrow s$, and so we may define the set of *possible states* $P \coloneqq \{s \in S : s \rightarrow s\}$ to include all and only the states which have trivial tasks.

It remains to constrain the interpretation of $\rightarrow$ in order to support the intended reading of P as the set of possible states. Since $s \rightarrow t$ indicates that it is possible for the state t to obtain as a result of the state s obtaining, I will assume that it is possible for both s and t to obtain if $s \rightarrow t$. Accordingly, I will impose the following constraint:

Quasi-Reflexivity: If $s \rightarrow t$, then $s \rightarrow s$ and $t \rightarrow t$.

Continuing with the example above, we may imagine that Sanna stands up, transitioning from sitting s to standing t. Insofar as this transition is possible— i.e., $s \rightarrow t$— we may conclude that it is also possible for Sanna to have continued sitting $s \rightarrow s$, and that once she is standing she may go on doing so $t \rightarrow t$.[14] Despite being quasi-reflexive, we may observe that the task relation is not reflexive since impossible states do not have trivial tasks, and so are inaccessible even to themselves. For instance, consider the fusion state $s.t$ of Sanna sitting and standing: since $s.t$ cannot obtain, no possible transition will result in the state $s.t$ obtaining. Accordingly, $s.t$ is inaccessible from all states, including from itself. Something similar may be said for state transitions from $s.t$. Since $s.t$ cannot obtain, no state ends up obtaining as the result of $s.t$ obtaining.

Recall that a state fusion is said to obtain just in case all of its parts obtain. Since a possible state can obtain, we ought to expect that each of its parts can obtain as a result. As brought out before, Fine requires modalized state spaces to satisfy the constraint:

Possibility: If $s \in P$ and $t \sqsubseteq s$, then $t \in P$.

Given that P is defined rather than primitive, I will derive Possibility by constraining the interaction between the task and parthood relations. In particular, tasks between fusions must be decomposable into subtasks between their respective parts, where $s \rightarrow t$ is a *subtask* of $s' \rightarrow t'$ just in case $s \sqsubseteq s'$ and $t \sqsubseteq t'$. More specifically, consider:

Parthood: If $d \sqsubseteq s$ and $s \rightarrow t$, then $d \rightarrow r$ for some $r \sqsubseteq t$.[15]
If $r \sqsubseteq t$ and $s \rightarrow t$, then $d \rightarrow r$ for some $d \sqsubseteq s$.[16]

[14] A different approach admits possible states that are *transient* on account of failing to have a trivial task. Accordingly, we would have to take both P and $\rightarrow$ to be primitive elements of the model.

[15] It is also natural to impose the following containment constraints: (F) if $s \in P$ where $d \sqsubseteq s$ and $d \rightarrow r$, then $s \rightarrow t$ for some t where $r \sqsubseteq t$; and (B) if $t \in P$ where $r \sqsubseteq t$ and $d \rightarrow r$, then $s \rightarrow t$ for some s where $d \sqsubseteq s$. However plausible these constraints may be, they will not be required for what follows.

[16] These principles are interderivable if $\rightarrow$ is symmetric. Although it is natural to take the task relation to be symmetric and transitive for certain applications, these constraints will not be required.

The constraints above ensure that every part is accounted for in any task between fusions. For example, given the wall of brake lights ahead (state l), we may take there to be a task from Priya's driving state $d.l$ to a state $b.l'$ in which she breaks (state b) and there is a similar arrangement of break lights l'. Since $d \sqsubseteq d.l$, it follows from Parthood that there is some suitable $t \sqsubseteq b.l'$ where $d \rightarrow t$. Moreover, since $b \sqsubseteq b.l'$, there is some $s \sqsubseteq d.l$ where $s \rightarrow b$. Although Parthood only specifies the existence of such states, we may expect that $s = d$ and $t = b$, and so $d \rightarrow b$ where, similarly, $l \rightarrow l'$.

Given that the null state $\square$ is a part of every state, Parthood is trivialized if $s \rightarrow \square$ and $\square \rightarrow s$ for any state s. In order to avoid instances of this scenario, it suffices to require the null state $\square$ to be necessary. I will define the set of necessary states by first letting s and t be *connected* $s \sim t$ just in case either $s \rightarrow t$ or $t \rightarrow s$. A state s is *contingent* just in case s is connected to a distinct state, i.e., $s \sim t$ for some $t \neq s$. It follows from Quasi-Reflexivity that contingent states are possible and so may obtain, and yet may also cease to obtain on account of transitioning to or from distinct states. By contrast, there are no tasks between impossible states, and so no way for impossible states to change. Rather, impossible states are non-contingent insofar as they cannot transition to or from any state at all, much less to or from a distinct state. Despite being possible, necessary states are also non-contingent since they must obtain. Thus we may define a state s to be *necessary* just in case it is only connected to itself where $N \coloneqq \{s \in S : \forall t \in S(s = t \Leftrightarrow s \sim t)\}$ is the set of necessary states. Given this definition, we may avoid trivializing instances of Parthood with the following constraint:

Nullity: $\square \in N$.[17]

Requiring the null state $\square$ to be necessary fits with its definition as the fusion of the empty set. Since a fusion obtains just in case all of its parts obtain, nothing is required for $\square$ to obtain, and so $\square$ obtains trivially. Moreover, we may derive the following:

Nonempty: $\square \in P$.

In addition to avoiding trivializing instances of Parthood, it follows from the Nullity constraint that the null state $\square$ is possible. As a result, the set of possible states is nonempty in accordance with Fine's definition of a modalized state space.

It remains to show that every possible state is a part of a world state in accordance with Fine's definition of a world space. As before, s and t are *compatible* $s \circ t$ just in case their fusion $s.t$ is possible. Given any state s and possible state t, we may consider the parts of s that are compatible with t. Rather than permitting there to be bigger and bigger parts $r_1 \sqsubset r_2 \sqsubset \ldots$ of s without end which are each compatible with t, I will impose the following constraint on the compatible parts of possible states:

Maximality: If $s \in S$ and $t \in P$, there is some *maximal part* $r \sqsubseteq s$ that is compatible with t, i.e., some $r \sqsubseteq s$ where $r \circ t$ and for any $q \sqsubseteq s$, if $q \circ t$ and $r \sqsubseteq q$, then $r = q$.

Consider the total state s of the strategy room at the moment Nixon nearly pressed the button. We may imagine s to include parts that fix all the features of the objects which make up the room and its occupants. Although any state of Nixon pressing the button

[17] Since a necessary state must obtain, all of its parts must also obtain, and so it is natural to impose the necessity constraint: if $s \in N$ and $t \sqsubseteq s$, then $t \in N$. Additionally, since any necessary state t must obtain, any possible state s that can obtain can only do so alongside t, and so it is natural to impose the Compatibility constraint: if $s \in N$ and $t \in P$, then $s \circ t$ since $s \circ t \coloneqq s.t \in P$. However natural these constraints may be, neither will be required for what follows.

is incompatible with s, this is not true for the parts of s. For instance, the state a of Nixon's advisors sitting just as they are is a part of s and perfectly compatible with any state of Nixon pressing the button. Given a state b of Nixon pressing the button, Maximality requires there to be a maximal part of s which is compatible with b, ruling out unending chains of bigger and bigger parts of s that are compatible with b.[18]

As above, a *world state* is any possible state that includes all compatible states as parts where $W := \{w \in P : s \sqsubseteq w$ whenever $s \circ w\}$ is the set of all world states. Whereas Fine restricts attention to world spaces in which every possible state is part of a world state, **P1** in the *Appendix* derives this constraint from Maximality:

Worldhood: If $s \in P$, then $s \sqsubseteq w$ for some $w \in W$.[19]

Following Fine (2017a), §3 observes that Worldhood ensures that every proposition has exactly one truth-value at every world state. Since I will identify possible worlds at a time with world states, it follows that the tautologies of classical logic are true in every model at every world at every time. This result provides a powerful reason to maintain Worldhood and given that we may derive Worldhood from Maximality, a corresponding reason to adopt Maximality. Moreover, Maximality is in keeping with an intuitive understanding of states, ruling out exotic cases in which there is no biggest part of a state s that is compatible with a state possible t.

A *task space* may then be defined as any ordered triple $\mathcal{T} = \langle S, \sqsubseteq, \rightarrow \rangle$ where $\langle S, \sqsubseteq \rangle$ is a state space and $\rightarrow$ satisfies Quasi-Reflexivity, Parthood, Nullity, and Maximality. Since Possibility, Nonempty, and Worldhood may be derived, every task space determines a world space. Whereas the world spaces that Fine defines are rich with mereological and modal structure, task spaces include the primitive resources needed to encode both the modal and temporal structure of possible worlds.

2.2 Possible Worlds

Assuming a discrete theory of time for simplicity, I will model times by the integers in $\mathbb{Z}$ where every time x will be interpreted as differing in time from $x + 1$ by a unit value specified for a given application.[20] We may then define an *evolution* $\tau : \mathbb{Z} \rightarrow S$ to be any assignment of states to times where $\tau(x) \rightarrow \tau(x + 1)$ for all times $x \in \mathbb{Z}$. Put otherwise, every state in the evolution is required to be accessible from its predecessor. Given Quasi-Reflexivity, all states that belong to an evolution are possible.

A *world history* is any evolution α where $\alpha(x) \in W$ for every time $x \in \mathbb{Z}$. Accordingly, we may take $H_{\mathbb{Z}}$ to be the set of all world histories parametrised by $\mathbb{Z}$. Given a history α, we may refer to the range of values that α occupies at different times as the *moments* of that history. Officially, the moments of a history α are its members which take the form $\langle x, \alpha(x)\rangle$, though it is convenient to write them as $\alpha(x)$. For instance, suppose that α is a history which occupies the same world state $\alpha(x) = \alpha(y)$ at different times $x \neq y$. Even though the world state is the same at both times, the moments differ since they occur at different times in α's history. Letting $\alpha \sim \beta$ just in case there is some $z \in \mathbb{Z}$ where

[18] Recall that states are static. Since pushing the button is not static, one might deny that any state makes it true that Nixon pushed the button. I will return to this objection in the conclusion. For now, think of the state of Nixon pushing the button as the point of no return in which the circuit is closed.

[19] We may also consider the Determinism constraint: if $s \rightarrow t$ and $s \rightarrow r$, then $t = r$. Additionally, there is the World Determinism constraint: if $s \rightarrow t$ and $s \rightarrow r$ where $s, t, r \in W$, then $t = r$. These constraints are less natural given the present application, but may be of interest elsewhere.

[20] I will consider an objection to the arbitrariness of the unit in §5.

$\alpha(x) = \beta(x + z)$ for all $x \in \mathbb{Z}$, we may identify the *possible worlds* parametrised by $\mathbb{Z}$ with the set of equivalence classes $W_{\mathbb{Z}} = \{[\alpha]_{\mathbb{Z}} : \alpha \in H_{\mathbb{Z}}\}$ where $[\alpha]_{\mathbb{Z}} = \{\beta \in H_{\mathbb{Z}} : \alpha \sim \beta\}$. By contrast with histories, possible worlds represents genuinely distinct orderings of world states. Although the set of possible worlds may claim to hold a metaphysical standing that the set of histories cannot, it is the histories that will play an important role in the semantics for counterfactuals rather than possible worlds. Nevertheless, I will refer to histories as worlds, though officially they are to be distinguished.[21]

Given a world α and time x, we may consider the result of making a minimal change to $\alpha(x)$ so as to include a state s without crossing over into impossibility. Letting an *s-part* of $\alpha(x)$ be any state that is part of $\alpha(x)$ and compatible with s, we may take a state t to be a *maximal s-part of* $\alpha(x)$ just in case t is not a proper part of any s-parts of $\alpha(x)$. Assuming s is possible, Maximality requires there to be at least one maximal s-part of $\alpha(x)$ where in general there may be many. We may then define:

> Outcomes: A world β is an *outcome* of imposing a state s on a world α at time x— i.e., $\alpha \to_s^x \beta$— just in case $s.t \sqsubseteq \beta(x)$ for some t that is a maximal s-part of $\alpha(x)$.

Since there may be a range of maximal s-parts of $\alpha(x)$ as well as a range of histories that include s together with any given maximal s-part of $\alpha(x)$, there will typically be more than one outcome of imposing s on α at x. It is by quantifying over all outcomes that the following section will provide a semantics for counterfactual conditionals.

3 Task Semantics

Let $\mathcal{L} = \langle \mathbb{L}, \neg, \wedge, \vee, \boxright \rangle$ be a propositional language where $\mathbb{L} := \{p_i : i \in \mathbb{N}\}$ is the set of *sentence letters*. We may then define the *extensional sentences* of $\mathcal{L}$ as follows:

$$\varphi \ ::= \ p_i \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi.$$

Letting $\texttt{ext}(\mathcal{L})$ be the set of extensional sentences of $\mathcal{L}$, we may define the *well-formed sentences* of $\mathcal{L}$ by restricting the antecedent of a counterfactual to $\varphi \in \texttt{ext}(\mathcal{L})$:

$$A \ ::= \ p_i \mid \neg A \mid A \wedge A \mid A \vee A \mid \varphi \boxright A.$$

Although counterfactual operators may occur in the consequent of a counterfactual, only extensional operators may occur in the antecedent of a counterfactual. Letting $\texttt{wfs}(\mathcal{L})$ be the set of all well-formed sentences of $\mathcal{L}$, I will assume standard metalinguistic conventions for the other extensional operators. In order to interpret $\mathcal{L}$, the following subsection will present the version of Fine's bilateral theory of propositions that I defend in Brast-McKie (2021). As I will argue, this hyperintensional theory is well motivated for the present application. I will then provide a theory of logical consequence along with a proof system in §3.2, discussing a number of countermodels in §3.3.

3.1 Propositions

Consider the manner in which states were imposed on worlds at a time. Intuitively, we are to imagine making a minimal change to the moment at which a counterfactual is evaluated in order to make the antecedent true, checking to see if the consequent is made

[21] World histories may be compared to the coordinate plane which stipulates an origin while acknowledging that nothing of significance turns on where the origin is chosen to be.

true at the result. If the antecedent were to be interpreted by a set of moments or set of possible worlds, it is unclear what the result of a minimal change would be. After all, any moment in which the antecedent is true may differ from the evaluation moment in many completely unrelated ways, where something similar may be said for possible worlds. As observed in §1.2, merely replacing the evaluation moment with an antecedent moment and requiring the consequent to be true in all antecedent moments validates STA contrary to expectation. Rather, it is by taking moments to have mereological structure that we may make minimal changes to the evaluation moment, checking to see if the consequent is made true as a result. The reason the exact verifier states for the antecedent are of greater use in this application than moments or possible worlds is that they amount to specific ways for the antecedent to be true without including anything irrelevant. Put otherwise, the exact verifier states for the antecedent belong to the *subject-matter* of the antecedent. By contrast, both possible worlds and moments contain elements reaching far beyond the subject-matter of any given sentence. This motivates a departure from intensional theories of propositions to a hyperintensional theory of propositions which is sensitive to differences in subject-matter in addition to modal profile.[22]

In Brast-McKie (2021), I draw on Fine's (2017a,b) recent work in order to identify and defend a bilateral theory of propositions which distinguishes propositions that differ in either subject-matter or modal profile without positing any unnecessary distinctions between propositions. Given that the present application calls for a theory of propositions which is sensitive to both modal profile and subject-matter, I will employ that bilateral theory here without reproducing its defence. The theory is said to be *bilateral* on account of taking propositions to consist of both a positive and negative content. For instance, 'John is sitting down' expresses a proposition that consists of a set of exact verifier states which make that sentence true along with a set of exact falsifier states which make that sentence false. Nevertheless, not all ordered pairs of sets of states may be said to count as a proposition. In order to define the space of propositions, let a set of states X be *closed under fusion* just in case the fusion of any non-empty subset $Y \subseteq X$ belongs to X. Given any sets of states V and F, the ordered pair $\langle V, F\rangle$ is *exclusive* just in case the states in V are incompatible with the states in F, and *exhaustive* just in case every possible state is compatible with some state in either V or F. An ordered pair $\langle V, F\rangle$ is a *bilateral proposition* just in case $\langle V, F\rangle$ is exclusive and exhaustive where V and F are both closed under fusion.[23] Given any proposition $\langle V, F\rangle$, every world state will include a part which belongs to either V or F but not both, thereby validating classical logic.[24] Letting $\mathbb{P}$ be the set of bilateral propositions, we may now turn to interpret $\mathcal{L}$.

3.2 Counterfactual Logic

In order to provide a theory of logical consequence for the well-formed sentences of $\mathcal{L}$, we may take a *model* $\mathcal{M} = \langle S, \sqsubseteq, \rightarrow, |\cdot|\rangle$ of $\mathcal{L}$ to be any ordered tuple where $\langle S, \sqsubseteq, \rightarrow\rangle$ is a task space and $|p_i| \in \mathbb{P}$ for all $p_i \in \mathbb{L}$. Given the ambition to preserve differences in both subject-matter and modal profile, I defend Fine's *exact inclusive semantics* for the extensional operators in Brast-McKie (2021). In order to reproduce that semantics

[22] Further motivation may be drawn from SDA which, as Fine (2012, p. 231) observes, entails STA given the assumption that 'A' and '$A \vee (B \wedge A)$' express the same proposition as intensional theories assume.

[23] For any proposition $\langle V, F\rangle$, if $s \in V$ is necessary, then no state in F is possible. In the other direction, if no state in F is possible, then every world state w includes a part in V. Since V is not required to include a necessary state, it is consistent to claim that $\square$ is the only necessary state.

[24] See Fine (2017a, p. 630) for this observation and relevant definitions.

here, we may let $X \otimes Y \coloneqq \{s.t : s \in X,\ t \in Y\}$ be set of all pairwise fusions of the states in X with states in Y, and $X \oplus Y \coloneqq X \cup Y \cup (X \otimes Y)$ be the set of all states in either X, Y, or $X \otimes Y$. We may then define the propositional analogues for conjunction $\langle V, F\rangle \wedge \langle V', F'\rangle \coloneqq \langle V \otimes V', F \oplus F'\rangle$, disjunction $\langle V, F\rangle \vee \langle V', F'\rangle \coloneqq \langle V \oplus V', F \otimes F'\rangle$, and negation $\neg\langle V, F\rangle \coloneqq \langle F, V\rangle$.[25] With these definitions in place, we may extend the interpretation provided by any model $\mathcal{M}$ to include all extensional sentences in $\mathcal{L}$:

[$\neg$] $|\neg\varphi| = \neg|\varphi|$.

[$\wedge$] $|\varphi \wedge \psi| = |\varphi| \wedge |\psi|$.

[$\vee$] $|\varphi \vee \psi| = |\varphi| \vee |\psi|$.

It is easy to show that $\neg$, $\wedge$, and $\vee$ are propositional operators insofar as the conjunction, disjunction, or negation of any propositions is guaranteed to be a proposition.[26] Having provided a means by which to assign the extensional sentences to propositions, we may now turn to evaluate the well-formed sentences of $\mathcal{L}$ at world-time pairs.

Given any extensional sentence φ together with the outcome relation defined in §2.2, we may take $[\![\varphi]\!]^{\alpha}_{x} = \{\beta \in H_{\mathbb{Z}} : \alpha \to^{x}_{s} \beta \text{ for some } s \in |\varphi|^{+}\}$ to be the set of all outcomes that result from imposing any exact verifier for φ on the world α at time x. We may then present the *task semantics* below, where $\mathcal{M}$ is any model of $\mathcal{L}$:

$\mathcal{M}, \alpha, x \vDash p_i$ *iff* there is some $s \sqsubseteq \alpha(x)$ where $s \in |p_i|^{+}$.

$\mathcal{M}, \alpha, x \vDash \neg A$ *iff* $\mathcal{M}, \alpha, x \nvDash A$.

$\mathcal{M}, \alpha, x \vDash A \wedge B$ *iff* $\mathcal{M}, \alpha, x \vDash A$ and $\mathcal{M}, \alpha, x \vDash B$.

$\mathcal{M}, \alpha, x \vDash A \vee B$ *iff* $\mathcal{M}, \alpha, x \vDash A$ or $\mathcal{M}, \alpha, x \vDash B$.

$\mathcal{M}, \alpha, x \vDash \varphi \mathrel{\Box\!\!\rightarrow} C$ *iff* $\mathcal{M}, \beta, x \vDash C$ whenever $\beta \in [\![\varphi]\!]^{\alpha}_{x}$.

Since the semantic clauses given above for the extensional operators differ from the exact semantics, it is worth investigating whether there is a corresponding difference in the meaning assigned to the extensional operators. In order to assuage these doubts, let a world state *inexactly verify* an extensional sentence just in case it includes a part that exactly verifies that sentence and, similarly, *inexactly falsify* an extensional sentence just in case it includes a part that exactly falsifies that sentence. We may then show that an extensional sentence φ of any complexity will be true in a world α at time x just in case $\alpha(x)$ inexactly verifies φ. Letting an extensional sentence φ be *false* in a world α at time x just in case $\alpha(x)$ inexactly falsifies φ, we may likewise show that φ is false in α at x just in case φ is not true in α at x. Whereas the exact semantics took a bilateral form in order for negation to be a propositional function while maintaining relevance, relevance no longer holds in evaluating sentences at world-time pairs. Moreover, nothing new is gained by defining falsity at a world and time in addition to truth.

The semantics for the counterfactual conditional accommodates both URA and UVC brought out in §1.3: not only is every way for the antecedent to be true accounted for, the consequent is required to be true at every resulting outcome. It follows that the

[25] Since there is no confusing 'φ' with '$|\varphi|$', the abuse of notion for negation, conjunction, and disjunction may be forgiven for the sake of the syntactic simplicity.

[26] In contrast to $\mathcal{P}(W)$ which forms a complete lattice, $\mathbb{P}$ may be shown to form a *non-interlaced bilattice*. See Ginsberg (1988) and Fitting (1990) for definitions and Brast-McKie (2021) for further discussion.

task semantics validates SDA without also validating STA. In order to make these claims precise, I will define logical consequence for $\mathcal{L}$ as follows where $\Gamma \subseteq \texttt{wfs}(\mathbb{L})$:

Logical Consequence: $\Gamma \vDash C$ *iff* for any any model $\mathcal{M}$ of $\mathcal{L}$, world α, and time x, if $\mathcal{M}, \alpha, x \vDash A$ for all $A \in \Gamma$, then $\mathcal{M}, \alpha, x \vDash C$.

An inference rule is *valid* just in case its conclusion is a logical consequence of its premises. Letting **PL** be the set of tautologies of classical logic, we may show that the following rules are valid over the space of all models, where **M1** and **M2** preserve validity:

M1 $\vdash A$ *where* $A \in$ **PL**. **M2** *If* $A \vdash B$, *then* $\vdash A \boxright B$.

R1 $A \to B, A \vdash B$. **R2** $A \boxright B \vdash A \to B$.

R3 $A \vee B \boxright C \vdash A \boxright C$. **R4** $A \vee B \boxright C \vdash B \boxright C$.

R5 $A \vee B \boxright C \vdash A \wedge B \boxright C$. **R6** $A \boxright C, B \boxright C, A \wedge B \boxright C \vdash A \vee B \boxright C$.

R7 $A \boxright B \wedge C \vdash A \boxright B$. **R8** $A \boxright B \wedge C \vdash A \boxright C$.

R9 $A \boxright B, A \boxright C \vdash A \boxright B \wedge C$. **R10** $A \boxright B, A \wedge B \boxright C \vdash A \boxright C$.

I will take $\vdash_{\text{CL}}$ to be the smallest relation to include all instances of **R1** – **R10** which is closed under **M1**, **M2**, and the standard structural rules. It is easy to see that SDA follows from **R3** and **R4**.[27] Rather than discussing the validity of each rule, I will examine a number of invalid inferences in the following subsection by providing countermodels.

3.3 Countermodels

Although **CL** provides a sense of the deductive power which the task semantics for $\mathcal{L}$ supports, it will help to consider a number of invalid rules of inference in order to get a better understanding of how the semantics works. In particular, the task semantics invalidates the following rules of inference where STA is included at **#9**:

#1 $A, B \vdash A \boxright B$. **#2** $A \boxright B, B \boxright C \vdash A \boxright C$.

#3 $(A \boxright B) \vee (A \boxright \neg B)$. **#4** $A \boxright B \vee C \vdash (A \boxright B) \vee (A \boxright C)$.

#5 $A \boxright C, B \boxright C \vdash A \wedge B \boxright C$. **#6** $A \boxright B, \neg B \vdash \neg B \boxright \neg A$.

#7 $A \wedge B \boxright C \vdash A \boxright (B \boxright C)$. **#8** $A \boxright (B \boxright C) \vdash A \wedge B \boxright C$.

#9 $A \boxright C \vdash A \wedge B \boxright C$. **#10** *If* $\Gamma, A \vdash B$, *then* $\Gamma \vdash A \boxright B$.

For brevity, I will focus attention on **#1**, **#6**, **#9**, and **#10** where the discussion of a countermodel for each will help to shed light on both the semantics for counterfactual conditionals as well as the nature of counterfactual reasoning.

[27] I prove **R2** and **R3** are valid in the *Appendix* where **R4** is similar.

#1 $A, B \vdash A \boxright B$

The ball is red and Mary likes it. Even so, it would be wrong to claim that if the ball were red Mary would like it since there are certain shades of red Mary does not like. In accordance with these assumptions, we may draw on the task semantics in order to provide a countermodel for the instance $p_1, p_2 \vdash p_1 \boxright p_2$:

$$|p_1|^+ = \{a, c, a.c\} \qquad \alpha(x) = a.b \qquad W = \{a.b, c.d\}$$

$$|p_2| = \langle\{b\}, \{d\}\rangle \qquad \beta(x) = c.d$$

The world state $a.b$ includes an exact verifier for both p_1 and p_2, making both sentences true in α and x. However, c is also an exact verifier for p_1 and is incompatible with any exact verifier for p_2. Thus the maximal c-part of $\alpha(x)$ is $\square$, and so β is an outcome of imposing c on α at x which does not include a part that exactly verifies p_2. Adding color, we may take a to be a state of the ball being a shade of red that Mary likes, b to be the state of Mary liking the ball, c to be a state of the ball being a shade of red that Mary does not like, and d to be the state of Mary disliking the ball. It follows by the semantics for counterfactuals that $p_1 \boxright p_2$ is not true in α at x. The rest of the details needed to complete the model do not matter and so will be omitted.

#6 $A \boxright B, \neg B \vdash \neg B \boxright \neg A$

If Boris had gone to the party, Olga would have gone too. Neither Olga nor Boris went in the end. Even so, it would be wrong to claim that if Olga were to not go to the party then Boris would not go to the party. For even though Olga likes to go to the parties Boris attends, Boris prefers to socialise without Olga. Drawing on the task semantics, we may present a countermodel to the instance $p_1 \boxright p_2, \neg p_2 \vdash \neg p_2 \boxright \neg p_1$ as follows:

$$|p_1| = \langle\{a\}, \{c\}\rangle \qquad \alpha(x) = c.d.f \qquad \gamma(x) = a.e.g$$

$$|p_2| = \langle\{b\}, \{d, e, d.e\}\rangle \qquad \beta(x) = a.b.f \qquad W = \{a.b.f, c.d.f, e.a.g\}$$

Since $\alpha(x)$ contains the exact falsifier d for p_2, we know that $\neg p_2$ is true in α at x. Moreover, the only outcome of imposing an exact verifier for p_1 on $\alpha(x)$ is $\beta(x)$ which includes an exact verifier for p_2, and so $p_1 \boxright p_2$ is true in α at x. However, e is an exact verifier for $\neg p_2$ which, if imposed on $\alpha(x)$, will result in the outcome $\gamma(x)$ which does not include an exact verifier for $\neg p_1$, and so $\neg p_2 \boxright \neg p_1$ is false in α at x.

Adding substance, we may take d and e to be states of Olga staying home where d is only compatible with the state f of Boris not knowing where Olga is and e is only compatible with the state g of Boris believing her to be home. Since the state of Boris being at the party a is compatible with him not knowing where Olga is f and the only world state to include both a and f as parts also includes the state of Olga going to the party b, imposing a on $\alpha(x)$ results in the world state $\beta(x)$ in which Olga goes to the party. Thus it is true at α at x that if Boris had gone to the party, then Olga would have gone. However, the state of Olga staying at home e is only compatible with Boris believing her to be home g, where the only world state $\gamma(x)$ to include both e and g as parts also includes Boris going to the party a. Thus it is false to claim that if Olga were to not go to the party, then Boris would not go either. All that is required is for Olga to not go to the party in a way that is compatible with Boris believing her to be home. For instance, perhaps the lights are on in her house which Boris can see.

The countermodel to **#6** also invalidates the principle of *counterfactual contraposition* (CC) $A \boxright B \vdash \neg B \boxright \neg A$. Nevertheless, I show in the *Appendix* that **R2** is valid where this entails *counterfactual modus tollens* (CFMT) $\neg B, A \boxright B \vdash \neg A$. It follows that **#10** does not preserve validity for otherwise we may derive CC from CFMT. However, **#10** is sometimes defended, at least for certain restricted subject-matters.[28] Were one to defend **#10**, then CC must also be taken to be valid unless **R2** is given up.

#9 $\boldsymbol{A \boxright C \vdash A \wedge B \boxright C}$

Judy and Joey are at the party. If Matan were to go to the party, he would argue with Judy, making her angry. Even so, if both Matan and Stav were to go to the party, Stav would keep Matan occupied, making Joey jealous. Consider the following countermodel:

$$|p_1| = \langle\{a\},\{b\}\rangle \qquad |p_2| = \langle\{c\},\{d\}\rangle \qquad |p_3| = \langle\{e\},\{f\}\rangle$$
$$\alpha(x) = b.d.f.g \qquad \beta(x) = a.d.e.g \qquad \gamma(x) = a.c.f.h$$
$$W = \{b.d.f.g, a.d.e.g, a.c.f.h\}$$

Letting $p_1 \boxright p_3 \vdash p_1 \wedge p_2 \boxright p_3$ be the instance of **#9**, we may take a to be a state of Matan being at the party, b to be a state of Matan being at home, c to be a state of Stav being at the party, d to be a state of Stav being at home, e to be a state of Judy being angry, f to be a state of Judy being happy, g to be a state of Joey being content, and h to be a state of Joey being jealous. Since $\beta(x)$ is the only world state to result from imposing an exact verifier for p_1 on $\alpha(x)$ and includes an exact verifier for p_3, the counterfactual $p_1 \boxright p_3$ is true in α at x. Nevertheless, the world state $\gamma(x)$ which results from imposing an exact verifier for $p_1 \wedge p_2$ on $\alpha(x)$ does not include an exact verifier for p_3, and so $p_1 \wedge p_c \boxright p_3$ is false in α at x.[29] Thus **#9** (STA) is invalid.[30]

It is worth pausing to consider the role that states as opposed to moments play in invalidating STA. Although all $A \wedge B$-moments are A-moments, the outcomes of imposing an $A \wedge B$-state on an evaluation moment $\alpha(x)$ may fail to be an outcome of imposing an A-state of $\alpha(x)$. The sensitivity of the outcomes induced by imposing a state on a moment of evaluation is made possible not only by the mereological structure of each moment, but by the requirement that the exact verifier states for the antecedent belong to the subject-matter of the antecedent, where this requirement constitutes the *exactness* of the state semantics. It is for this reason that a hyperintensional theory of propositions which is sensitive to differences in subject-matter is appropriate in this application.

[28] See Yli-Vakkuri and Hawthorne (2020) for such a view and Elgin (2021) for further discussion.

[29] This model can be extended to make $p_1 \wedge p_2 \wedge p_4 \boxright p_3$ true while still making $p_1 \boxright p_3$ true and $p_1 \wedge p_2 \boxright p_3$ false. To do so, let $|p_4| = \langle\{i\},\{j\}\rangle$ where i is the state of Daniel attending the party, j is the state of Daniel being at home, and $\alpha(x) = b.d.f.g.j$, $\beta(x) = a.d.e.g.j$, $\gamma(x) = a.c.f.h.j$, and $\delta(x) = a.c.e.h.i$ are world states. Repeating this strategy generates a Sobel sequence.

[30] A similar countermodel cannot so plausibly be constructed for Goodman's case: the match would light if struck, but not if it were struck and wet. Whereas the moment of evaluation $\alpha(x)$ above includes Joey being happy which is incompatible with Stav being at the party c, it is not clear what part of the moment of evaluation besides lighting is incompatible with the match being wet. For Goodman's case to hold, more would have to change for the match to be wet when struck than for it to light when struck.

4 Counterfactuals and Tense

Even though many of the examples discussed above were stated in simple present tense, this was by contrivance rather than a reflection of standard usage. For instance, consider Fine's (1975b) original example stated in the past tense:

(`N′`) If Nixon had pressed the button there would have been a nuclear holocaust.

Suppose that one of Nixon's advisors were to assert `N′` a year after the incident. Since `N′` is concerned with a past time, there is no easy way to regiment `N′` with the expressive resources included in $\mathcal{L}$. Instead of accepting this limitation, the following subsection will extend $\mathcal{L}$ to include tense operators, drawing on the task semantics to provide appropriate semantic clauses. Given these resources, I will analyse forwards counterfactuals in §4.2 as well as both backwards and backtracking counterfactuals in §4.3.

4.1 Temporal Operators

When interpreting claims like `N′`, it is important to know the time at which the antecedent is to be evaluated. Although there would have been a nuclear holocaust if Nixon pressed the button after the security meeting had finished and the button was engaged, there would not have been a Nuclear holocaust had Nixon pressed the button earlier that day. In order to account for these differences, I will extend $\mathcal{L}$ to include a unary operator $\downarrow^i$ for all $i \in \mathbb{N}$ which shifts the time of evaluation to a value specified by the point of evaluation. Additionally, I will include Priorian tense operators $\Diamond\mkern-11mu{\scriptscriptstyle \mathrm{P}}\,$ and $\Diamond\mkern-11mu{\scriptscriptstyle \mathrm{F}}\,$ for the past and future, respectively. We may then define the *well-formed sentences* of $\mathcal{L}^T$ where the antecedent of a conditional is restricted to extensional sentences $\varphi \in \mathtt{ext}(\mathcal{L})$ as before:

$$A ::= p_i \mid \neg A \mid A \wedge A \mid A \vee A \mid \varphi \mathrel{\Box\!\!\rightarrow} A \mid \downarrow^i A \mid \Diamond\mkern-11mu{\scriptscriptstyle \mathrm{P}}\, A \mid \Diamond\mkern-11mu{\scriptscriptstyle \mathrm{F}}\, A.$$

In order to interpret sentences of the form $\ulcorner \downarrow^i A \urcorner$, I will include a vector $\vec{v} = \langle v_1, v_2, \ldots \rangle$ of stored times in the point of evaluation. We may then provide the following clauses:

($\downarrow$) $\mathcal{M}, \alpha, x, \vec{v} \vDash \downarrow^i A$ *iff* $\mathcal{M}, \alpha, v_i, \vec{v} \vDash A$.

($\Diamond\mkern-11mu{\scriptscriptstyle \mathrm{P}}\,$) $\mathcal{M}, \alpha, x, \vec{v} \vDash \Diamond\mkern-11mu{\scriptscriptstyle \mathrm{P}}\, A$ *iff* $\mathcal{M}, \alpha, y, \vec{v} \vDash A$ for some $y < x$

($\Diamond\mkern-11mu{\scriptscriptstyle \mathrm{F}}\,$) $\mathcal{M}, \alpha, x, \vec{v} \vDash \Diamond\mkern-11mu{\scriptscriptstyle \mathrm{F}}\, A$ *iff* $\mathcal{M}, \alpha, y, \vec{v} \vDash A$ for some $y > x$.

Whereas $\downarrow^i$ shifts the time of evaluation to the i^{th} value stored in $\vec{v}$, the operators $\Diamond\mkern-11mu{\scriptscriptstyle \mathrm{P}}\,$ and $\Diamond\mkern-11mu{\scriptscriptstyle \mathrm{F}}\,$ require their arguments to be true at some time in the past or future, respectively. Although $\vec{v}$ must be added to the points of evaluation, the semantic clauses given for $\mathcal{L}$ may otherwise be maintained. We may now proceed to apply this semantics in order interpret a range of tensed counterfactual conditional claims.

4.2 Tensed Counterfactuals

Let v_1 be the time at which Nixon almost pressed the button. Although this time may be forever imprinted on the minds of Nixon's advisors, the time at which the nuclear holocaust would have taken place may not be known. It might have taken days for the disaster to unfold, or perhaps only hours. Rather than assuming there is a specific time, we may take `N′` to assert that there is a future time at which the nuclear holocaust would have taken place. We may capture this reading with the following regimentation:

(N″) $\downarrow^1(B \boxright \Diamond H)$.

The semantics for $\mathcal{L}^T$ takes N″ to be true at a world α, time x, and vector $\vec{v}$ just in case for every world β which results from imposing an exact verifier for B on α at v_1, there is some future time $y > v_1$ at which H is true in β. Alternatively, if there is a stored time v_2 at which the consequent is to be evaluated— perhaps Nixon's past advisor intends to indicate the time at which N′ is asserted a year later— then '$\Diamond$' may be replaced with '$\downarrow^2$'. Which regimentation is right may depend on what the speaker intends.

Although typical, first going back in time and then going forward is not the only way to interpret tensed counterfactual claims, and sometimes not the most natural. The following subsection will consider both backwards and backtracking counterfactuals, providing their analysis with the resources of the present framework.

4.3 Backtracking

Following tradition, I will focus on an example presented by Jackson (1977) in which Smith is standing on the edge of a building threatening to jump. Standing behind him in safety, Beth and Bill watch in fear for their friend. Moments later, Smith steps back from the edge. We may then consider the following claims:

(J) If Smith had jumped, he would have died.

(U) If Smith had jumped, there would have been a net under him.

(L) If Smith had jumped, he would have lived.

Kicking things off, suppose that Bill asserts J. Beth responds by insisting on U. After all, they both know that Smith wants to live so the only way Smith would jump is if jumping would not kill him. Having made her case for U, Beth goes on to assert L. Unconvinced, Bill points out that nobody would install a net on the building, reasserting J. This prompts Beth to enumerate further reasons why Smith would never jump without a net there to catch him, insisting on U and L once more. This may be imagined to continue round after round. The challenge is to identify what they are disagreeing about and to provide a theory of counterfactuals that adequately models their disagreement.

Whereas J is characteristic of forward counterfactuals and can be regimented in a manner similar to N″ above, U and L express *backwards* and *backtracking* counterfactuals, respectively. Rather than changing the world of evaluation to make the antecedent true at one time and evaluating the consequent at a later time, backwards counterfactuals such as U evaluate the consequent at an earlier time than the time at which the antecedent is taken to be true. Consider the following regimentations:

(D′) $\downarrow^1(J \boxright \downarrow^2 D)$.

(U′) $\downarrow^1(J \boxright \downarrow^3 U)$.

(L′) $\downarrow^1(J \boxright \downarrow^2 L)$.

Holding the world α, time x, and vector of stored times $\vec{v}$ fixed for the purposes of evaluating Bill and Beth's conversation, I will assume that $v_3 < v_1 < v_2$, where Bill and Beth agree about this much. Nevertheless, Bill takes D′ to be true since any world to result from imposing a state in which Smith jumps in α at v_1 will be one in which Smith dies at v_2. Beth disagrees, claiming that L′ is true since any world to result from imposing a state in which Smith jumps in α at v_1 will be one in which Smith lives at v_2. Since the states of Smith being alive are incompatible with the states of Smith being dead, the worlds that Beth predicts will result from imposing a state of Smith jumping at v_1 make it false that Smith dies at v_2, and so Beth and Bill disagree.

In defence of her view, Beth asserts $\mathtt{U}'$ which requires the worlds that result from imposing a state of Smith jumping in α at v_1 to include the placement of a net at v_3 prior to the time v_1 of Smith's jump. Bill disagrees. Beth's claim $\mathtt{U}'$ helps to bring the source of disagreement to light: whereas Beth assumes that making minimal changes to the actual moment $\alpha(v_1)$ to include a state of Smith jumping will also include a net having been placed there at an earlier time v_3, Bill does not. What they disagree about is which possible worlds are the outcome of imposing a state of Smith jumping on the actual world at the time v_1 that Beth and Bill are considering. Perhaps Beth's optimism reflects the emphasis she places on Smith's love for live, assuming that more would have to change about the actual world at v_1 for Smith to willingly die than merely including a net there to catch him. Whereas jumping without a net is incompatible with Smith's love for life, jumping with a net is perfectly compatible. By contrast, Bill's pessimism may be attributed to his willingness to see Smith change his outlook on life, taking this to be a smaller change than including a net there beneath him.

Although the task semantics can help model our counterfactual claims, the semantics cannot help us make up our minds about which world states are the result of making minimal changes to an actual moment. In the disagreement between Bill and Beth, each believes different worlds are the outcomes of imposing a state of Smith jumping in the actual world α at v_1 when he was poised to do so. Nevertheless, the task semantics helps to shed light on the nature of backtracking counterfactuals. Even though backtracking counterfactuals like $\mathtt{L}'$ are forward counterfactuals just as much as $\mathtt{D}'$, the proposed changes at the time of evaluation propagate backwards on account of requiring an altered past. Assuming that a net is present at v_1 requires that it was installed on the building at an earlier time v_3. By contrast, agreeing with Bill that Smith would have died only requires that Smith give up his love for life in the moment of jumping without requiring any substantial changes to the past. Given that the task semantics only selects outcome worlds by requiring them to occupy moments that differ minimally from the moment of evaluation, there is no requirement whatsoever that the worlds which result from imposing some state on the world of evaluation at a given time agree with the world of evaluation at earlier times. Indeed, small changes at the time of evaluation might well require massive changes to the past leading up to the point of evaluation.

In order to emphasise this final point, I will conclude the present section with another backwards counterfactual along similar lines. To begin with, suppose that there are no advanced alien civilisations anywhere in the universe. Consider the following case:

> `Icosahedron:` Deep inside an Egyptian tomb, Harry discovered a perfectly formed icosahedron that appeared to be made out of pure titanium. In fact, the icosahedron was only a crude iron alloy. Relieved, Harry couldn't help but say out loud, "If the icosahedron had been titanium, advanced aliens put it there."

Although the icosahedron was not made of titanium, supposing otherwise requires massive changes to the past. Even if aliens did not put it there, supposing that humans had developed the metallurgy needed to refine pure titanium at the time of the pyramids still requires a radical shift from the past of the actual world. Nevertheless, changing the moment of Harry's discovery to include a titanium icosahedron instead of the crude alloy that he found does not require much change at all. It is worth comparing such a case to

Fine's Nixon example given above. Whereas `N′` makes a small change to the moment of supposition that results in a drastically altered future, `Icosahedron` makes a small change to the moment of supposition which requires a drastically altered past. Despite these differences, the two cases are handled by the task semantics in the same way: any world which differs minimally from the actual world at the time of evaluation is to be considered, no matter how far that world may diverge in its past or future.

5 Conclusion

The task semantics posits three primitives: states, parthood, and tasks. Whereas Lewis' revised account of similarity dispenses with common intuition, the primitive elements that I have included encode judgements that we often rely on to navigate the world. Not only may I attend to my present sitting state, I may consider a range of states to which I may transition, as well as the evolutions by which I might hope to arrive at those states. Although the primitive elements included in the task semantics are much more intuitive than Lewisian similarity, this is by no means a high bar, nor the only metric of success. Rather, I hope to have made the advantages of the present approach plain to see, where it is by taking possible worlds to have temporal and mereological structure that we may consider the possible worlds that result from making minimal changes to the moment of evaluation. Instead of relying on the structure included in an intended model, temporal and mereological structure has been encoded into the definition of a model itself. By contrast, it is difficult to see how a theory which takes possible worlds to be structureless points could succeed in identifying minimal changes to a moment.

Although speaking in terms of the elements that we may imagine to be included in an intended task space can help to keep track of the working parts of the theory, nothing about the semantics turns on this reading or its supposed metaphysics. Nevertheless, a number of difficulties remain, inspiring further extensions of the present framework. I will conclude by briefly mentioning two loose ends in the following subsections.

5.1 Events

It is easy to consider states which exactly verify sentences like 'The icosahedron is pure titanium' or 'The coin is heads up'. Exact verifier states for such sentences are specific, perfectly static ways for some things to be. For instance, we may consider the precise arrangement of titanium atoms which make up the icosahedron, or the exact orientation of the coin relative to the table. By contrast, consider the sentences B and J from before:

(B) Nixon presses the button.

(J) Smith jumps (off the building).

It is natural to deny that there are any individual states of Nixon pressing the button or Smith jumping off the building. For instance, suppose Nixon presses the button. His doing so consists of some sequence of states in which he moves his finger ever closer to the button, slowly depresses the button, and then pulls his finger back. Accordingly, one may doubt that there is any single state which exactly verifies B. In place of a single state, we may model his action by a *process* $\bar{e} \coloneqq \langle e_1, \ldots, e_n \rangle$ consisting of a sequence of states where $e_i \to e_{i+1}$ for all $1 \leqslant i < n$. We may then model the *event* of Nixon pressing the button by the set $\overline{V}$ of all processes which make the sentence B true together with the set $\overline{F}$ of all processes which make B false. In interpreting counterfactuals whose antecedent is an event, we cannot identify outcome worlds by merely imposing some single state on the moment of evaluation. Rather we ought to adapt the world of evaluation to incorporate a process which exactly verifies the antecedent.

Before attempting to explain how this is to be done, it is important to evaluate what the merits are of doing so. For instance, if the result is a satisfying metaphysics which nicely distinguishes events from static propositions but which does not otherwise effect the semantics for counterfactuals, no harm will come in taking at least some states to represent processes as a simplifying idealisation. Nevertheless, it is natural to worry that conflating states with processes is incompatible with the intuitive vision guiding the construction of the semantics as well as the resulting principles. After all, it is not clear than any part of a possible process will be possible, where the analogous principle for states played an important role in the semantics above. Moreover, sets of processes do not have determinate fusions since these processes may be fused in different ways. At the very least, considerably more would have to be said in order to spell out a coherent process-theoretic analogue of the task spaces presented above.

Instead of facing this challenge, one might take events to be a convenient proxy for closely related static propositions. For instance, we might take any process of Nixon pressing the button $\bar{e}$ to include a critical state e_i in which the electrodes in the button first complete a circuit. More generally, such an account will have to provide a general means by which to identify the critical state in any given process, where a *critical state* amounts to the point of no return in that process. Given my purposes here, I will leave the development of a theory of events and critical states for another time.

5.2 Continuous Time

Taking time to be discrete introduces a degree of arbitrariness, for in defining a world history we must also fix the unit. Whereas some histories may string together moments separated by mere seconds, others may be much lower resolution, separating moments by minutes or even hours. If time is discrete as some contemporary theories of physics claim, then no change to the present theory is needed. Moreover, there may even be a correct answer to the question how far apart moments are in time.[31] Even so, histories with a high resolution are unlikely to be necessary for the purposes of semantics. Instead of taking the choice of unit to be arbitrary, we may take this decision to be practical in nature: we may choose a temporal resolution which is as low as possible while nevertheless preserving enough temporal features for the application at hand.

By contrast, taking time to be continuous requires amending the present theory. For instance, suppose that time has the structure of the real numbers $\mathbb{R}$. Given any state at a time, there is no next time which we may assign to an accessible state. As a result, the definition of an evolution cannot be maintained in its present form. Letting $\xrightarrow{*}$ be the transitive closure of $\rightarrow$, we may define a *continuous evolution* to be any function $\tau : \mathbb{R} \rightarrow S$ where $\tau(x) \xrightarrow{*} \tau(y)$ whenever $x \leqslant y$. Although amending the present framework to accommodate continuous time may make for an interesting project in metaphysics, I suspect there is little to be gained for the purposes of semantics. For this reason, I have not developed a continuous analogue of the theory here though I hope to do so elsewhere.

[31] For instance, see Garay (1995) for discussion of discrete theories of time in recent physics.

Appendix

P1 *If $s \in P$, then $s \sqsubseteq w$ for some $w \in W$.*

Proof. Let $s \in P$. Since $\bigsqcup S \in S$, MAXIMALITY requires there to be some $w \sqsubseteq \bigsqcup S$ where $w \circ s$ and for any $q \sqsubseteq \bigsqcup S$, if $q \circ s$ and $w \sqsubseteq q$, then $w = q$. Given that $q \sqsubseteq \bigsqcup S$ just in case $q \in S$, it follows that for any $q \in S$, if $q \circ s$ and $w \sqsubseteq q$, then $w = q$.

Let $q = w.s$. Since $w \circ s$, we know that $q \circ s$. Additionally $w \sqsubseteq q$, and so $w = q$ by the universal claim above. Thus $w = w.s$, and so $s \sqsubseteq w$.

Let $u \in S$ where $u \circ w$. So $u.w \circ s$ where $w \sqsubseteq u.w$. It follows that $w = u.w$, and so $u \sqsubseteq w$. Generalising on u, we may conclude that $w \in W$ where $s \sqsubseteq w$. ☐

L1 *Letting (m/s) be the set of maximal s-parts of m, if $s \circ m$, then $(m/s) = \{m\}$.*

Proof. Let $m \circ s$. Thus m is an s-part of m. Observe that m cannot be a proper part of any s-part of m since $r \sqsubseteq m$ for every r that is an s-part of m. Thus m is a maximal s-part of m, and so $m \in (m/s)$.

Let $n \in (m/s)$. So $n \sqsubseteq m$ where $n \circ s$. Moreover, $n \not\sqsubset k$ for any k which is an s-part of m. Thus $n \not\sqsubset m$ since m is an s-part of m, and so $n \not\sqsubseteq m$ or $m \sqsubseteq n$. Given that $n \sqsubseteq m$, we may conclude that $m \sqsubseteq n$, and so $m = n$. ☐

L2 *If $s \sqsubseteq \alpha(x)$ for $\alpha \in H_{\mathbb{Z}}$ and $x \in \mathbb{Z}$, then $\alpha \rightarrow_s^x \alpha$.*

Proof. Let $s \sqsubseteq \alpha(x)$ where $\alpha \in H_{\mathbb{Z}}$ and $x \in \mathbb{Z}$. Thus $s.\alpha(x) \sqsubseteq \alpha(x)$ where it follows by **L1** that $(\alpha(x)/s) = \{\alpha(x)\}$. Since there is some $t \in (\alpha(x)/s)$ where $s.t \sqsubseteq \alpha(x)$, namely where $t = \alpha(x)$, we may conclude that $\alpha \rightarrow_s^x \alpha$. ☐

P2 $A \mathrel{\Box\!\!\rightarrow} B \vDash A \rightarrow B$.

Proof. Let $\mathcal{M}, \alpha, x \vDash A \mathrel{\Box\!\!\rightarrow} B$ and assume $\mathcal{M}, \alpha, x \nvDash A \rightarrow B$ where we may take $A \rightarrow B \coloneqq \neg A \vee B$ to fix conventions. Thus $\mathcal{M}, \alpha, x \vDash A$ and $\mathcal{M}, \alpha, x \nvDash B$. So there is some $s \sqsubseteq \alpha(x)$ where $s \in |A|^+$. By **L2**, $\alpha \rightarrow_s^x \alpha$, and so $\alpha \in [\![A]\!]_x^\alpha$. By assumption, $\mathcal{M}, \alpha, x \vDash B$, contradicting the above. Thus $\mathcal{M}, \alpha, x \vDash A \rightarrow B$ as desired. ☐

P3 $A \vee B \mathrel{\Box\!\!\rightarrow} C \vDash A \mathrel{\Box\!\!\rightarrow} C$.

Proof. Let $\mathcal{M}, \alpha, x \vDash A \vee B \mathrel{\Box\!\!\rightarrow} C$ and $\beta \in [\![A]\!]_x^\alpha$. So $\alpha \rightarrow_s^x \beta$ for some $s \in |A|^+$. Thus $s \in |A \vee B|^+$, and so $\alpha \in [\![A \vee B]\!]_x^\alpha$. By assumption, $\mathcal{M}, \alpha, x \vDash C$. Since $\beta \in [\![A]\!]_x^\alpha$ was arbitrary, we may conclude that $\mathcal{M}, \alpha, x \vDash A \mathrel{\Box\!\!\rightarrow} C$. ☐

References

Brast-McKie, Benjamin. 2021. “Identity and Aboutness.” *Journal of Philosophical Logic* 50:1471–1503. ISSN 1573-0433. doi:10.1007/s10992-021-09612-w.

Elgin, Samuel Z. 2021. “Counterfactual Logic and the Necessity of Mathematics.” *Journal of Philosophical Logic* 50:97–115. ISSN 1573-0433. doi:10.1007/s10992-020-09563-8.

Fine, Kit. 1975a. “Critical Notice.” *Mind* 84:451–458. ISSN 0026-4423.

—. 1975b. “Vagueness, Truth and Logic.” *Synthese* 30:265–300. ISSN 0039-7857, 1573-0964. doi: 10.1007/BF00485047.

—. 2012. “Counterfactuals Without Possible Worlds.” *Journal of Philosophy* 109:221–246. ISSN 0022-362X. doi:10.5840/jphil201210938.

—. 2017a. “A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation.” *Journal of Philosophical Logic* 46:625–674. ISSN 0022-3611, 1573-0433. doi:10.1007/s10992-016-9413-y.

—. 2017b. “A Theory of Truthmaker Content II: Subject-Matter, Common Content, Remainder and Ground.” *Journal of Philosophical Logic* 46:675–702. ISSN 0022-3611, 1573-0433. doi:10.1007/s10992-016-9419-5.

—. 2017c. “Truthmaker Semantics.” In *A Companion to the Philosophy of Language*, 556–577. Malden, MA: Wiley-Blackwell, 2nd edition edition. ISBN 978-1-118-97209-0. doi:10.1002/9781118972090.ch22.

Fitting, Melvin. 1990. “Bilattices in Logic Programming.” In *, Proceedings of the Twentieth International Symposium on Multiple-Valued Logic, 1990*, 238–246. doi:10.1109/ISMVL.1990.122627.

Garay, Luis J. 1995. “Quantum Gravity and Minimum Length.” *International Journal of Modern Physics A* 10:145–165. ISSN 0217-751X, 1793-656X. doi:10.1142/S0217751X95000085.

Gillies, Anthony S. 2007. “Counterfactual Scorekeeping.” *Linguistics and Philosophy* 30:329–360. ISSN 1573-0549. doi:10.1007/s10988-007-9018-6.

—. 2009. “On Truth-Conditions for If (but Not Quite Only If).” *The Philosophical Review* 118:325–349. ISSN 0031-8108. doi:10.1215/00318108-2009-002.

Ginsberg, Matthew L. 1988. “Multivalued Logics: A Uniform Approach to Inference in Artificial Intelligence.” *Computational Intelegence* 4:265–316.

Jackson, Frank. 1977. “A Causal Theory of Counterfactuals.” *Australasian Journal of Philosophy* 55:3–21. ISSN 0004-8402. doi:10.1080/00048407712341001.

Kripke, Saul A. 1963. “Semantical Analysis of Modal Logic I Normal Modal Propositional Calculi.” *Mathematical Logic Quarterly* 9:67–96. ISSN 1521-3870. doi:10.1002/malq.19630090502.

Lewis, David. 1973. *Counterfactuals*. Harvard: Harvard University Press.

—. 1979. “Counterfactual Dependence and Time’s Arrow.” *Noûs* 13:455–476. ISSN 0029-4624. doi: 10.2307/2215339.

Stalnaker, Robert C. 1968. “A Theory of Conditionals.” In Nicholas Rescher (ed.), *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, 98–112. Oxford: Blackwell.

Warmbrōd, Ken. 1981. “An Indexical Theory of Conditionals.” *Dialogue: Canadian Philosophical Review / Revue canadienne de philosophie* 20:644–664. ISSN 1759-0949, 0012-2173. doi:10.1017/S0012217300021399.

Yli-Vakkuri, Juhani and Hawthorne, John. 2020. “The Necessity of Mathematics.” *Noûs* 54:549–577. ISSN 1468-0068. doi:10.1111/nous.12268.