Set-Theoretic Completeness for Epistemic and Conditional Logic*

Joseph Y. Halpern
Dept. Computer Science
Cornell University
Ithaca, NY 14853
halpern@cs.cornell.edu
http://www.cs.cornell.edu/home/halpern

September 7, 1998

Abstract

The standard approach to logic in the literature in philosophy and mathematics, which has also been adopted in computer science, is to define a language (the *syntax*), an appropriate class of models together with an interpretation of formulas in the language (the *semantics*), a collection of axioms and rules of inference characterizing reasoning (the *proof theory*), and then relate the proof theory to the semantics via soundness and completeness results. Here we consider an approach that is more common in the economics literature, which works purely at the semantic, set-theoretic level. We provide set-theoretic completeness results for a number of epistemic and conditional logics, and contrast the expressive power of the syntactic and set-theoretic approaches.

*The work was supported in part by NSF under grant IRI-96-25901 and by the Air Force Office of Scientific Research under grant F49620-96-1-0323.

1 Introduction

The standard approach to logic in the literature in philosophy and mathematics, which has also been adopted in computer science, is to define a language (the *syntax*), an appropriate class of models together with an interpretation of formulas in the language (the *semantics*), a collection of axioms and rules of inference characterizing reasoning (the *proof theory*), and then relate the proof theory to the semantics via soundness and completeness results.

The economics literature has also been interested in various logics, particularly logics of knowledge and belief, and more recently conditional logic for counterfactual reasoning. By and large, they dispense with syntax altogether, working purely at a set-theoretic, semantic level. To understand how this is done, it is perhaps best to take as an example epistemic logic.

Both approaches would start with what is called a *frame* in the logic literature, that is, a set W of possible worlds and a binary relation $\mathcal{K}$ on W. Intuitively, $(w, w') \in \mathcal{K}$ if in world w, the agent considers w' possible.[1] In the economics literature, the $\mathcal{K}$ relation is used to define an operator K on events (subsets of W). Taking $\mathcal{K}(w) = \{w' : (w, w') \in \mathcal{K}\}$, the operator $\mathsf{K} : 2^W \mapsto 2^W$ is defined as follows:

$$\mathsf{K}(E) = \{w : \mathcal{K}(w) \subseteq E\}. \tag{1}$$

We read $\mathsf{K}(E)$ as "the agent knows E".

The mathematical/philosophical approach adds an extra level of indirection to this more set-theoretic approach. A language for reasoning about knowledge is defined, starting in the usual way with a set Φ of primitive propositions, and closing off under conjunction, negation, and applications of the modal operator K. As is well known, to give semantics to formulas in this language, we use a Kripke structure $M = (W, \mathcal{K}, \pi)$, where $(W, \mathcal{K})$ is a frame, and π is an *interpretation* that associates with each primitive proposition and each world a truth value; that is, $\pi : \Phi \times W \mapsto \{\mathbf{true}, \mathbf{false}\}$. We then define what it means for a formula φ to be true at a world $w \in W$, written $(M, w) \models \varphi$, using the usual inductive definition.[2] We can think of $\models$ as just associating with each formula φ an event $[\![\varphi]\!]_M = \{w : (M, w) \models \varphi\}$ in structure M ($[\![\varphi]\!]_M$ is called the *intension* of φ in the literature [Fagin, Halpern, Moses, and Vardi 1995]). Not surprisingly, $[\![K\varphi]\!]_M = \mathsf{K}([\![\varphi]\!]_M)$; that is, we obtain the event associated with the formula $K\varphi$ by applying the operator K to the event associated with φ.

If all that is done with a formula is to translate it to an event, why bother with the overhead of formulas and $\models$? Would it not just be simpler to dispense with formulas and interpretations, and work directly with events? Syntax often plays an important role—for example, it allows us to express concepts in a model-independent way. On the

[1] Actually, in the economics literature, it is more standard to consider a partition of W; this is equivalent to the case where $\mathcal{K}$ is an equivalence relation.

[2] Readers unfamiliar with the definition can find details in Section 2.

other hand, if we have in mind an intended model all along, then perhaps it makes sense to just work directly with events. For example, the *model-checking* approach, which has been widely used in proving correctness of programs [Clarke, Emerson, and Sistla 1986], typically works with one fixed model, the one generated by the program whose correctness we are trying to prove. Model checking has been advocated for epistemic reasoning as well [Halpern and Vardi 1991]. Perhaps when using the model-checking approach, it might make sense to work at the set-theoretic level.

Probability provides another example. Probabilists start by defining a particular model—the probability space of interest—and then investigating its properties. As the many texts on probability demonstrate, they have been able to prove a great many results about probability by working purely at a set-theoretic level. While some logics for reasoning about probability have been proposed, both propositional [Fagin, Halpern, and Megiddo 1990] and first-order [Halpern 1990], they certainly do not begin to capture all the subtleties of the reasoning we find in probability texts. For example, typical logics of probability cannot express notions such as expectation and variance.

One of the apparent advantages of working with syntax is that we can define a proof theory, that allows us to manipulate formulas in order to prove properties of interest. We do not have to give up proof theory if we work at the set-theoretic level. For example, a standard property of knowledge is *introspection*: if an agent knows a fact, then he knows that he knows it, and if he does not know it, then he knows that he does not know it. Syntactically, these properties are expressed as

- $K\varphi \Rightarrow KK\varphi$, and
- $\neg K\varphi \Rightarrow K\neg K\varphi$.

These properties have immediate set-theoretic analogues:

- $\mathsf{K}(E) \subseteq \mathsf{K}(\mathsf{K}(E))$, and
- $\neg\mathsf{K}(E) \subseteq \mathsf{K}(\neg\mathsf{K}(E))$, where $\neg$ here denotes the set-theoretic complement

As this example suggests, we can translate a syntactic axiom to a set-theoretic axiom by

1. replacing formulas by events,
2. replacing the modal operator K by the set operator K,
3. replacing the Boolean operations $\neg, \wedge, \vee$ by their set-theoretic analogues $\neg, \cap, \cup$.

In this paper, I explore set-theoretic completeness proofs in the context of epistemic logics and conditional logics. Both of these logics were introduced in the philosophical literature [Hintikka 1962; Stalnaker 1968], but have been widely used in computer science and AI. Epistemic logic has been used as a tool for analyzing multi-agent systems [Fagin, Halpern, Moses, and Vardi 1995]; conditional logic has been used as a framework for analyzing nonmonotonic reasoning [Boutilier 1994] and counterfactual reasoning [Lewis 1973]. It also has an important role to play in the analysis of causality [Lewis 1973], which is becoming an increasingly important issue in AI as well [Pearl 1995]. Set-theoretic completeness proofs for logics of knowledge and common knowledge are standard in the economics literature (see, for example, [Aumann 1989; Milgrom 1981]). I compare them to the more familiar syntactic completeness proofs in the philosophical literature, and then do the same for conditional logic. For the logics considered here, every syntactic operator has a semantic counterpart; thus, every property expressible syntactically has a semantic counterpart. However, as we shall see, the converse is not always true. For the logics considered here, the set-theoretic approach gives us extra expressive power, allowing us to express more properties.

In part, this comes from the use of arbitrary (rather than just finite) unions and intersections over events. We can already see the use of countable intersections and unions in the context of probability theory. Although probability is typically taken to be countably additive, this fact is not expressible in propositional logics of probability, although it can be expressed once we allow countable operations.[3] To be precise, suppose that we have an operator of the form $pr^p(\varphi)$ in the language, which is to be interpreted as "the probability of φ is at least p", and a corresponding operator on events $\mu^p(E)$. The property of countable additivity cannot be expressed in a propositional logic, because we do not have countable disjunctions. Indeed, in [Fagin, Halpern, and Megiddo 1990], a complete axiomatization is given for a logic of probability where, semantically, probability is taken to be countably additive, but axiomatically, we require only finite additivity. (This means that there will be nonstandard models of the theory where the probability is finitely additive but not countably additive.) By way of contrast, countable additivity is immediately expressible using μ^p though, using countable unions: If $E_i, i \in I$ is a countable collection of pairwise disjoint sets, then countable additivity just says

$$\cap_{i\in I}\mu^{p_i}(E_i) \subseteq \mu^p(\cup_{i\in I}E_i), \text{where } p = \textstyle\sum_{i\in I} p_i.$$

In the case of knowledge, there is also a property that involves infinite intersection. With only finite intersection, we have:

$$\mathsf{K}(E) \cap \mathsf{K}(E') = \mathsf{K}(E \cap E'). \tag{2}$$

Once we allow infinite intersections, we have

$$\cap_{j\in J}\mathsf{K}_i(E_j) = \mathsf{K}_i(\cap_{j\in J}E_j) \text{ for any index set } J \text{ and events } E_j,\ j \in J. \tag{3}$$

Clearly (3) implies (2). Somewhat surprisingly, it can be shown that if $\mathcal{K}$ is an equivalence relation (that is, knowledge satisfies the properties of **S5**), then they are equivalent.

[3]For those readers unfamiliar with probability, finite additivity says that if E and F and disjoint sets, then $\Pr(E \cup F) = \Pr(E) + \Pr(F)$. Countable additivity says that if $E_1, E_2, \ldots$ is a sequence of pairwise disjoint sets, then $\Pr(\cup_i E_i) = \sum_i \Pr(E_i)$.

(This follows from Aumann's set-theoretic completeness proof [1989], although I provide a direct proof.) However, once we weaken the **S5** properties, for example, if we consider either **S4** or **K45**, then the equivalence no longer holds, and we need the full strength of (3) to get set-theoretic completeness proofs. At the syntactic level, this distinction is lost, because infinite conjunctions cannot be expressed.

Of course, it can be argued that if we extended propositional logic to allow infinite conjunctions, then these differences could be expressed perfectly well syntactically. However, issues of expressiveness do not arise only for infinite conjunctions and disjunctions. In the context of conditional logic, I show that there are properties that involve only finite intersections and unions that have no analogue at the syntactic level. Moreover, even properties that can be captured syntactically are more naturally expressed at the set-theoretic level. Finally, as we shall see, completeness proofs seem to be more straightforward and transparent at the set-theoretic level, at least in the case of the logics considered here.

The rest of this paper is organized as follows. In Section 2, I consider epistemic logic, while in Section 3, I consider conditional logic. I conclude in Section 4.

2 Epistemic Logic

I start by reviewing the syntactic approach to epistemic logic, and then I examine how the set-theoretic approach works.

2.1 The Syntactic Approach: A Review

For simplicity here, I consider single-agent epistemic logic; all the points I want to make already arise in the single-agent case. I briefly review the syntax and semantics here for those not familiar with it. We start with a nonempty set Φ of primitive propositions, and close off under negation, conjunction, and applications of the modal operator K. Let $\mathcal{L}^K$ be the language consisting of all formulas that can be built up this way. Thus, a typical formula in $\mathcal{L}^K$ is $K(\neg K(p \wedge q))$. We define implication and disjunction as usual.

A *Kripke structure* is a tuple $M = (W, \mathcal{K}, \pi)$, as discussed in the introduction. We define $(M, w) \models \varphi$ by induction as follows:

$(M, w) \models p$ (for a primitive proposition $p \in \Phi$) iff $\pi(w, p) = \mathbf{true}$

$(M, w) \models \varphi \wedge \varphi'$ iff $(M, w) \models \varphi$ and $(M, w) \models \varphi'$

$(M, w) \models \neg\varphi$ iff $(M, w) \not\models \varphi$

$(M, w) \models K\varphi$ iff $(M, w') \models \varphi$ for all w' such that $(w, w') \in \mathcal{K}$.

There are well-known soundness and completeness results for epistemic logic, that show the close connection between the assumptions we make about $\mathcal{K}$ and axiomatic properties. Consider the following axioms:

Prop. All substitution instances of tautologies of propositional calculus

K1. $(K\varphi \wedge K(\varphi \Rightarrow \psi)) \Rightarrow K\psi$, (Distribution Axiom)

K2. $K\varphi \Rightarrow \varphi$, (Knowledge Axiom)

K3. $K\varphi \Rightarrow KK\varphi$, (Positive Introspection Axiom)

K4. $\neg K\varphi \Rightarrow K\neg K\varphi$, (Negative Introspection Axiom)

MP. From φ and $\varphi \Rightarrow \psi$ infer ψ (Modus ponens)

Gen. From φ infer $K\varphi$ (Knowledge Generalization)

The system with axioms and rules Prop, K1, MP, and Gen has been called **K**. If we add K2 to **K**, we get **T**; if we add K3 to **T**, we get **S4**; if we add K4 to **S4**, we get **S5**; finally, if we add K3 and K4 to **K**, we get **K45**. (Other systems can also be formed; these are the ones I focus on here.)

Let $\mathcal{M}$ be the class of all Kripke structures. We are also interested in subclasses of $\mathcal{M}$ where the $\mathcal{K}$ relation has various properties of interest. Let $\mathcal{M}^r$ (resp., $\mathcal{M}^{rt}$; $\mathcal{M}^{et}$; $\mathcal{M}^{rst}$) consist of the Kripke structures where the $\mathcal{K}$ relation is reflexive (resp., reflexive and transitive; Euclidean[4] and transitive; reflexive, symmetric, and transitive, i.e., an equivalence relation).

Theorem 2.1: *For formulas in the language $\mathcal{L}^K$:*

(a) **K** *is a sound and complete axiomatization with respect to $\mathcal{M}$,*

(c) **T** *is a sound and complete axiomatization with respect to $\mathcal{M}^r$,*

(c) **S4** *is a sound and complete axiomatization with respect to $\mathcal{M}^{rt}$,*

(d) **K45** *is a sound and complete axiomatization with respect to $\mathcal{M}^{et}$.*

(e) **S5** *is a sound and complete axiomatization with respect to $\mathcal{M}^{rst}$.*

[4] A relation R is Euclidean if $(s,t),(s,u) \in R$ implies that $(t,u) \in R$.

2.2 The Set-Theoretic Approach

In the set-theoretic approach, we just start with a frame $(W, \mathcal{K})$. We can then define an operator K as in Equation 1 in the Introduction. Consider the following properties of the K operator:

A1. $\mathsf{K}(E) \cap \mathsf{K}(E') = \mathsf{K}(E \cap E')$

A2. $\mathsf{K}(E) \subseteq E$

A3. $\mathsf{K}(E) \subseteq \mathsf{K}(\mathsf{K}(E))$

A4. $\neg\mathsf{K}(E) \subseteq \mathsf{K}(\neg\mathsf{K}(E))$

A5. $\cap_{j \in J}\mathsf{K}(E_j) = \mathsf{K}(\cap_{j \in J} E_j)$ for any index set J and events $E_j, j \in J$[5]

A2, A3, and A4 are the obvious analogues of K2, K3, and K4, respectively. A1 can be viewed as an analogue of K1. In fact, it is not hard to show that

K1$'$. $K(\varphi \wedge \psi) \Leftrightarrow (K\varphi \wedge K\psi)$

is equivalent to K1 in the presence of MP and Prop. We could have replaced K1 by K1$'$ in all the axiom systems and still have obtained all the completeness proofs of Theorem 2.1. Instead of A1, Aumann uses the monotonicity property

A1$'$. If $E \subseteq F$, the $\mathsf{K}(E) \subseteq \mathsf{K}(F)$.

It is easy to see that A1$'$ follows from A1 (since if $E \subseteq F$, then $E \cap F = E$, so $\mathsf{K}(E) \subseteq \mathsf{K}(E) \cap \mathsf{K}(F) \subseteq \mathsf{K}(F)$). A1 does not follow from A1$'$, but it follows from Aumann's set-theoretic completeness theorem that A1 does follow from A1$'$, A2, A3, and A4; Proposition 2.2 provides a self-contained proof of this fact.

Note there is no analogue to Prop, MP, or Gen above; they turn out to be unnecessary at the set-theoretic level. (In particular, once we work at the level of sets, we do not need the Boolean equivalences encoded in Prop as axioms.) On the other hand, there is no analogue of A5 at the syntactic level; we cannot express infinite conjunctions in propositional logic, so it is unnecessary. It turns out that A5 is also unnecessary if we assume that $\mathcal{K}$ is an equivalence relation. This follows from the following result.

Proposition 2.2: *Any operator on events in W that satisfies A1', A2, and A4 must also satisfy A5 (and hence A1).*

[5]If $J = \emptyset$, we take the intersection over the empty set to be W, as usual. Thus, as a special case of this axiom, we get $W = \mathsf{K}(W)$.

Proof: Suppose K satisfies A1', A2, and A4. Consider the fixed points of K, that is, those sets E such that $\mathsf{K}(E) = E$. I first show that the set of fixed points of K is closed under negations and arbitrary unions.

Suppose $\mathsf{K}(E) = E$. Then $\neg E = \neg\mathsf{K}(E)$, so $\mathsf{K}(\neg E) = \mathsf{K}(\neg\mathsf{K}(E)) = \neg\mathsf{K}(E) = \neg E$, where the second equality follows from A2 and A4. Thus, $\neg E$ is a fixed point of K.

Next, suppose $\mathsf{K}(E_j) = E_j$ for all j in some index set J. Since $E_j \subseteq \cup_j E_j$, by A1$'$, we have $\mathsf{K}(E_j) \subseteq K(\cup_j E_j)$. Thus,

$$\cup_j \mathsf{K}(E_j) \subseteq \mathsf{K}(\cup_j E_j). \tag{4}$$

It now follows that

$$\begin{array}{lll} & \mathsf{K}(\cup_{j\in J} E_j) & \\ = & \mathsf{K}(\cup_{j\in J}\mathsf{K}(E_j)) & \text{since } E_j = \mathsf{K}(E_j) \\ \subseteq & \mathsf{K}(\mathsf{K}(\cup_{j\in J} E_j)) & \text{by A1}' \text{ and (4)} \\ \subseteq & \mathsf{K}(\cup_{j\in J} E_j) & \text{by A2.} \end{array}$$

Thus, $\mathsf{K}(\cup_{j\in J} E_j) = \mathsf{K}\mathsf{K}(\cup_{j\in J} E_j)$, and the set of fixed points of K is closed under arbitrary unions.

We are now ready to prove A5. Since $\cap_{j\in J} E_j \subseteq E_j$, it follows from A1$'$ that $\mathsf{K}(\cap_{j\in J} E_j) \subseteq \mathsf{K}(E_j)$. Thus, $\mathsf{K}(\cap_{j\in J} E_j) \subseteq \cap_{j\in J}\mathsf{K}(E_j)$. For the opposite inclusion, observe that by A2, we have $\cap_{j\in J}\mathsf{K}(E_j) \subseteq \cap_{j\in J} E_j$. Thus, by A1$'$, $\mathsf{K}(\cap_{j\in J}\mathsf{K}(E_j)) \subseteq \mathsf{K}(\cap_{j\in J} E_j)$. By A2 and A4, $\neg\mathsf{K}(E_j)$ is a fixed point of K, for all $j \in J$. Since the set of fixed points is closed under arbitrary unions and negations, $\cap_{j\in J}\mathsf{K}(E_j)$ is also a fixed point of K. It follows that $\cap_{j\in J}\mathsf{K}(E_j) \subseteq \mathsf{K}(\cap_{j\in J} E_j)$, as desired. ∎

Each of A1$'$, A2, and A4 is necessary for Proposition 2.2. If we drop any of them, then A5 no longer necessarily holds, as the following examples show.

Example 2.3: Let $W = \{1, 2, 3\}$ and define $\mathsf{K}_0(\{1\}) = \mathsf{K}_0(\{2, 3\}) = \emptyset$, and $\mathsf{K}_0(E) = E$ for $E \neq \{1\}, \{2, 3\}$. It is easy to see that K_0 satisfies A2 and A4, but not A1$'$ (since $\{3\} \subseteq \{2, 3\}$ but $\mathsf{K}_0(\{3\}) = \{3\} \not\subseteq \emptyset = \mathsf{K}_0(\{2, 3\})$). Since K_0 does not satisfy A1$'$, *a fortiori*, it does not satisfy A1 or A5. ∎

Example 2.4: Let $W = \{1, 2, 3, \ldots\}$. Define $\mathsf{K}_1(E) = E$ if E is cofinite (that is, if the complement of E is finite) and $\mathsf{K}_1(E) = \emptyset$ otherwise. It is easy to see that K_1 satisfies A1 (and hence A1$'$) and A2, but does not satisfy A4 (since, for example, $\neg\mathsf{K}_1(\neg\{1\}) = \{1\} \neq \emptyset = \mathsf{K}(\neg\mathsf{K}(\neg\{1\}))$. K_1 does not satisfy A5, since if $E_j = \neg\{j\}$, then $\mathsf{K}_1(\cap_{j\geq 1} E_j) = \mathsf{K}_1(\{1\}) = \emptyset \neq \{1\} = \cap_{j\geq 1}\mathsf{K}_1(E_j)$. ∎

Example 2.5: Let $W = \{1, 2, 3, \ldots\}$. Define $\mathsf{K}_2(E) = W$ if E is cofinite and $\mathsf{K}_2(E) = \emptyset$ otherwise. Again, it is easy to see that K_2 satisfies A1 and A4, but does not satisfy A2 (since, for example, $\mathsf{K}_2(\neg\{1\}) = W \not\subseteq \neg\{1\}$). Taking $E_j = \neg\{j\}$ as in the previous example, note that K_2 does not satisfy A5 since $\mathsf{K}_2(\cap_j E_j) = \emptyset \neq W = \cap_j\mathsf{K}_2(E_j)$. ∎

The following theorem is the set-theoretic analogue of Theorem 2.1. Aumann [1989] proved it for the case that $M \in \mathcal{M}^{rst}$ and K satisfies A1$'$, A2–A4; this is a generalization of his result. (In light of Proposition 2.2, we can replace A1$'$ by A1.) It is just the result we would expect (modulo, perhaps, the need for A5 if we do not have all of A1, A2, and A4).

Theorem 2.6: *The K operator in the frame $(W, \mathcal{K})$ satisfies A5. Moreover, if $\mathcal{K}$ is reflexive (resp., reflexive and transitive; Euclidean and transitive; an equivalence relation) then K satisfies A2 (resp., A2 and A3; A3 and A4; A1–A4). Conversely, if K' is an operator on events satisfying A5, then there is a binary relation $\mathcal{K}$ on W such that K' is the K relation in the frame $(W, \mathcal{K})$. Moreover, if K' satisfies A2 (resp., A2 and A3; A3 and A4; A1–A4), then $\mathcal{K}$ is reflexive (resp., reflexive and transitive; Euclidean and transitive; an equivalence relation).*

Proof: The first part is straightforward and left to the reader. For the second part, given an operator K' on W, define $\mathcal{K}$ so that $\mathcal{K}(w) = \cap\{E : w \in \mathsf{K}'(E)\}$. Using the fact that in all cases K' satisfies A5, it is easy to check that K and K' agree. And just as in the standard canonical model constructions of completeness in the syntactic case (see, e.g., [Fagin, Halpern, Moses, and Vardi 1995]), we can show that A2 forces $\mathcal{K}$ to be reflexive, A3 forces it to be transitive, and A4 forces it to be Euclidean. The result follows. (Note that a reflexive, Euclidean, and transitive relation is an equivalence relation.) ∎

This theorem is a typical example of a set-theoretic soundness and completeness result. The first part can be viewed as a soundness statement, while the second part gives us completeness.

The reader familiar with the standard syntactic completeness proofs using canonical model constructions should find it instructive to compare the set-theoretic completeness proofs with those involving canonical models. The set-theoretic proofs works for an arbitrary set W of worlds; we do not have to construct a special set where each world corresponds to a maximal consistent set of formulas. The definition of the $\mathcal{K}$ relation above is very similar in spirit to that in the canonical model construction, as are the arguments that A2, A3, and A4 force the $\mathcal{K}$ relation to be reflexive, transitive, and Euclidean, respectively. However, the proof that $\mathsf{K} = \mathsf{K}'$ is simpler than the proof that a formula is true at a world in the canonical model if and only if it is an element of that world (viewed as a maximal consistent set of formulas). As we shall see, in the case of conditional logic, set-theoretic completeness proofs are also relatively simpler than syntactic ones.

I have said we can view Theorem 2.6 as a set-theoretic soundness and completeness result. The standard soundness and completeness results in logic involve a language, a proof theory, and a semantics. Is there a way we can view Theorem 2.6 as a more standard soundness and completeness result? I briefly sketch here an argument showing that we can.

Fix a finite set of worlds W_0. The set of *event descriptions (for W_0)* is the least set formed as follows: We have a symbol A for each subset A of W_0, and close off under union, complementation, and applications of the K operator. Of course, we take $\mathsf{E}_1 \cap \mathsf{E}_2$ to be an abbreviation for $\neg(\neg\mathsf{E}_1 \cup \neg\mathsf{E}_2)$. A *basic event formula (for W_0)* has the form $\mathsf{E}_1 = \mathsf{E}_2$, where E_1 and E_2 are event descriptions. Note that $\mathsf{E}_1 \subseteq \mathsf{E}_2$ can be viewed as an abbreviation for the basic event formula $\mathsf{E}_1 \cap \mathsf{E}_2 = \mathsf{E}_1$. An *event formula (for W_0)* is a Boolean combination of basic event formulas. Our language consists of event formulas. Note that the language is relative to the particular domain W_0 about which we wish to reason.

A semantic model for this language consists of an interpretation of K as an operator on subsets of W_0. This allows us to associate with each event description E a subset $v(\mathsf{E})$ of W_0 in the obvious way. Of course, each symbol A is interpreted as the corresponding subset A of W_0 (that is, $v(\mathsf{A}) = A$). Union and complementation get there standard interpretation, and the interpretation of $\mathsf{K}(\mathsf{E})$ is determined by the interpretation of K. Finally, a basic event formula $\mathsf{E}_1 = \mathsf{E}_2$ is true relative to an interpretation if $v(\mathsf{E}_1) = v(\mathsf{E}_2)$. Boolean combinations of event formulas are interpreted in the obvious way.

As for the axiom system, besides considering A1–A4 (or some subset of these axioms), we need Prop and MP from the axiom system **K** for propositional reasoning, an axiom that says $\mathsf{K}(\mathsf{E})$ is equal to some subset of W_0, axioms describing the relationship between subsets of W_0, and axioms and inference rules for dealing with equality. The axiom that says $\mathsf{K}(\mathsf{E})$ is equal to some subset of W_0 is simply

A6. $\vee_{A\subseteq W_0} \mathsf{K}(\mathsf{E}) = \mathsf{A}$.

The axioms describing the relationship between subsets of W_0 have the following form: for all subsets $A, B, C \subseteq W_0$, if $A \cup B = C$, then we have an axiom $\mathsf{a} \cup \mathsf{B} = \mathsf{C}$; similarly, if $B = \neg A$, we have the axiom $\mathsf{B} = \neg\mathsf{A}$. Call this collection of axioms Rel.

Finally, we need axioms that say equality is an equivalence relation (reflexive, symmetric, and transitive), and an inference rule that allows us to substitute equals for equals. Call these three axioms and inference rule Eq.

Let $\mathcal{A}$ be the axiom system consisting of A1, A6, Prop, MP, Rel, and Eq.

Corollary 2.7: *Let $\mathcal{A}'$ be a subset of A2, A3, A4. Then $\mathcal{A} \cup \mathcal{A}'$ is a sound and complete axiomatization for the set of frames of the form $(W_0, \mathcal{K})$ where $\mathcal{K}$ satisfies the subset of $\{$reflexive, transitive, Euclidean$\}$ corresponding to $\mathcal{A}'$.*

Proof: Soundness is obvious. For completeness, it suffices to show that if φ is an event formula that is consistent with $\mathcal{A} \cup \mathcal{A}'$, then φ is satisfied in a frame $(W_0, \mathcal{K})$ where $\mathcal{K}$ satisfies the appropriate properties. But this follows almost immediately from Theorem 2.6. Extend φ to a maximal complete subset of event formulas. This set of formulas defines an operator K on events satisfying A1 (and hence A5, since W_0 is finite). Thus, by Theorem 2.6, there is a binary relation $\mathcal{K}$ on W_0 with the right properties. ∎

If W_0 is infinite rather than finite, it seems that to prove a result like Corollary 2.7, we need to allow arbitrary unions rather than just finite unions and arbitrary disjunctions (to express the analogue of A6). I conjecture that a completeness proof exists even if we restrict the language to finite disjunctions, although I have not checked details. In any case, Corollary 2.7 does show that it is legitimate to view Theorem 2.6 as a semantic counterpart of the more usual soundness and completeness results.[6]

3 Conditional logic

As I said earlier, conditional logic has been used to capture both counterfactual reasoning and default reasoning. Stalnaker [1992] gives a short and readable survey of the philosophical issues involved. I briefly review the standard syntax and semantics here.

3.1 The Syntactic Approach: Selection Functions

As suggested above, the syntax for conditional logic is straightforward. We start with a set Φ of primitive propositions, and close it off under conjunction, negation, and applications of the binary modal operator $\rightarrow$. Thus, if φ and ψ are formulas, then so is $\varphi \rightarrow \psi$. The formula $\varphi \rightarrow \psi$, can be read as "if φ were the case, then ψ would be true" (if we want to give $\rightarrow$ a counterfactual reading) or "typically/normally/by default, if φ is the case then ψ is the case" (if we want to give it a reading more appropriate for nonmonotonic reasoning). Let $\mathcal{L}^C$ be the set of formulas that can be built up this way.

The original approach for capturing $\rightarrow$, due to Stalnaker and Thomason [Stalnaker 1968; Stalnaker and Thomason 1970], proceeds as follows: Given a language $\mathcal{L}^C$, they assume that there is a *selection function* $f : W \times \mathcal{L}^C \mapsto W$. For counterfactual reasoning, we can think of $f(w, \varphi)$ as the world "closest" to w that satisfies φ. For default reasoning, $f(w, \varphi)$ should be thought of as the most normal world (relative to w) satisfying φ. These interpretations implicitly assume that there is a unique world closest to w (or most normal relative to w) that satisfies E. Many later authors argued that there is not in general a unique closest world; ties should be allowed. I follow this interpretation here. Thus, I take a *counterfactual structure* to be a tuple (W, f, π), where $f : W \times \mathcal{L}^C \mapsto 2^W$ and π is an interpretation, as before.

The definition of $\models$ in counterfactual structures is the same as that in epistemic structures, except for the clause for $\rightarrow$, which is

$$(M, w) \models \varphi \rightarrow \psi \text{ iff } f(w, \varphi) \subseteq [\![\psi]\!]_M$$

This captures the intuition that the closest worlds to w, as defined by f, satisfy φ.

There are various restrictions we can consider placing on the selection function.

[6] I thank Giacomo Bonanno for raising this issue.

S1. $f(w,\varphi) \subseteq [\![\varphi]\!]_M$: the worlds closest to w satisfying φ are in fact φ-worlds (where w is a *φ-world* if $w \in [\![\varphi]\!]_M$).

S2. If $f(w,\varphi) \subseteq [\![\psi]\!]_M$ and $f(w,\psi) \subseteq [\![\varphi]\!]_M$, then $f(w,\varphi) = f(w,\psi)$. Stalnaker and Thomason [1970] view this as a *uniformity* condition. If the closest φ-worlds all satisfy ψ and the closest ψ-worlds all satisfy φ, then the closest φ-worlds and the closest ψ-worlds must be the same. Note that S1 and S2 together force f to be a semantic function: if $[\![\varphi]\!]_M = [\![\psi]\!]_M$, then, in the presence of S1, the antecedent to S2 will hold, so $f(w,\varphi) = f(w,\psi)$.

S3. If $w \in [\![\varphi]\!]_M$, then $f(w,\varphi) = \{w\}$: if w is a φ-world, then it is the closest φ-world to w. This restriction is particularly appropriate for counterfactual reasoning. It is not necessarily appropriate for nonmonotonic reasoning. The most normal φ-world may not be w, even if w is a φ-world.

S4. $f(w,\varphi)$ is either empty or a singleton. This captures Stalnaker's original assumption that there is a unique closest world, if there is a closest world at all.

S5. $f(w,\varphi_1 \vee \varphi_2) \subseteq f(w,\varphi_1) \cup f(w,\varphi_2)$: if w' is one of the $(\varphi_1 \vee \varphi_2)$-worlds closest to w, then it must be one of the φ_1-worlds closest to w or one of the φ_2-worlds closest to w.

S6. If $f(w,\varphi) \subseteq [\![\psi]\!]_M$ then $f(w,\varphi \wedge \psi) \subseteq f(w,\varphi)$: if the closest φ-worlds to w all satisfy ψ, then the closest $\varphi \wedge \psi$-worlds are all among the closest φ-worlds to w.[7]

S7. If $f(w,\varphi) \cap [\![\psi]\!]_M \neq \emptyset$, then $f(w,\varphi \wedge \psi) \subseteq f(w,\varphi) \cap [\![\psi]\!]_M$: this is a strengthening of S5 (at least, when $f(w,\varphi) \neq \emptyset$), which says that the closest $\varphi \wedge \psi$-worlds to w are among the closest φ-worlds that are also in $[\![\psi]\!]_M$ (provided there are any).[8]

S8. If $[\![\varphi]\!]_M \neq \emptyset$ then $f(w,\varphi) \neq \emptyset$: this says that there always is some φ-world closest to w if there are any φ-worlds.

As we shall see in Section 3.3, if we introduce an ordering on worlds and define $f(w,\varphi)$ as the φ-worlds closest to w, then these restrictions arise by taking some very natural restrictions on the ordering (indeed, restrictions S1, S2, S5, and S6 are forced on us).

There is a well-known axiom corresponding to each of these conditions except for S8. Let $\Box\varphi$ be an abbreviation for $\neg\varphi \to$ *false* and let $\Diamond\varphi$ be an abbreviation of $\neg\Box\neg\varphi$.

[7]It may seem even more reasonable to replace $\subseteq$ by $=$ here. This would say that if the closest φ-worlds to w satisfy ψ, then they are the closest $\varphi \wedge \psi$-worlds. In fact, the stronger version of S6 already follows from S1, S2, and S5. (Proof: By S1, S2, and S5, $f(w,\varphi) = f(w,(\varphi \wedge \psi) \vee (\varphi \wedge \neg\psi)) \subseteq f(w,\varphi\wedge\psi) \cup f(w,\varphi\wedge\neg\psi)$. Moreover, if $f(w,\varphi) \subseteq [\![\psi]\!]$, it follows from S1 that $f(w,\varphi) \cap f(w,\varphi\wedge\neg\psi) = \emptyset$. Thus, $f(w,\varphi) \subseteq f(\varphi \wedge \psi)$.) Note that S2 can easily be obtained from S1 and the stronger version of S6. I use the weaker version of S6 here because it allows us to make technical connections to some known results in conditional logic.

[8]Again, we may want to strengthen $\subseteq$ to $=$, and again, the stronger version follows from the weaker version, in the presence of S1, S2, and S5.

Thus, $(M,w) \models \Box\varphi$ iff $f(w,\neg\varphi) = \emptyset$ and $(M,w) \models \Diamond\varphi$ iff $f(w,\varphi) \neq \emptyset$. Consider the following axioms:

Prop. All substitution instances of propositional tautologies

C0. $((\varphi \rightarrow \psi_1) \wedge (\varphi \rightarrow \psi_2)) \Rightarrow (\varphi \rightarrow (\psi_1 \wedge \psi_2))$

C1. $\varphi \rightarrow \varphi$

C2. $((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)) \Rightarrow ((\varphi \rightarrow \sigma) \Rightarrow (\psi \rightarrow \sigma))$

C3. $\varphi \Rightarrow (\psi \Leftrightarrow (\varphi \rightarrow \psi))$

C4. $(\varphi \rightarrow \psi) \vee (\varphi \rightarrow \neg\psi)$

C5. $((\varphi_1 \rightarrow \psi) \wedge (\varphi_2 \rightarrow \psi)) \Rightarrow ((\varphi_1 \vee \varphi_2) \rightarrow \psi)$

C6. $((\varphi_1 \rightarrow \varphi_2) \wedge (\varphi_1 \rightarrow \psi)) \Rightarrow ((\varphi_1 \wedge \varphi_2) \rightarrow \psi)$

C7. $(\neg(\varphi_1 \rightarrow \neg\varphi_2) \wedge (\varphi_1 \rightarrow \psi)) \Rightarrow ((\varphi_1 \wedge \varphi_2) \rightarrow \psi)$

C8. (a) $\Box\varphi \Rightarrow (\varphi \wedge (\varphi' \rightarrow \Box\varphi))$
(b) $\Diamond\varphi \Rightarrow (\varphi' \rightarrow \Diamond\varphi)$

MP. From φ and $\varphi \Rightarrow \psi$ infer ψ

RC1. From $\psi \Rightarrow \psi'$ infer $(\varphi \rightarrow \psi) \Rightarrow (\varphi \rightarrow \psi')$

All of the axioms other than C8 above are familiar from the literature. In the language of [Kraus, Lehmann, and Magidor 1990], C0 is the And Rule, C1 is Reflexivity, C5 is the Or Rule, C6 is Cautious Mononotonicity, C7 is Rational Monotonicity, and RC1 is Right Weakening. In [Kraus, Lehmann, and Magidor 1990] the focus was default reasoning, for which C0, C1, C5, C6, C7, and RC1 are considered appropriate. C2, C3, and C4 come from the literature on counterfactual reasoning; they are A6, t4.9, and t4.7 in [Stalnaker and Thomason 1970], respectively. There is one other rule considered in [Kraus, Lehmann, and Magidor 1990] called Left Logical Equivalence, which says

LLE. From $\varphi \Leftrightarrow \varphi'$ infer $(\varphi \rightarrow \psi) \Rightarrow (\varphi' \rightarrow \psi)$

This is true if f is a semantic notion, which depends only on $[\![\varphi]\!]_M$, and not the syntactic form of φ. As we have observed, this follows from S1 and S2. Not surprisingly, LLE follows readily from C1, C2, and RC1. We can also obtain C2 from C0, C1, C5, C6, RC1, and LLE, with a little effort, thus C2 holds in the framework of [Kraus, Lehmann, and Magidor 1990].

C8 is intended to characterize S8. Basically, we want to say that if $[\![\varphi]\!]_M \neq \emptyset$ (i.e., φ is satisfiable somewhere in structure M) then $\Diamond\varphi$ must be valid in M (that is, true at every world in M). Equivalently, if $M \models \Box\varphi$, then $\neg\varphi$ must be unsatisfiable in M. The following discussion may help explain how C8 captures this.

Given a structure $M = (W, f, \pi)$, define w' to be *reachable from w via φ* if $w' \in f(w,\varphi)$. We can then inductively define reachability via a sequence $\varphi_1; \ldots; \varphi_k$. If $k > 1$,

w' is reachable from w via $\varphi_1; \ldots; \varphi_k$ if there exists w'' such that w' is reachable from w'' via φ_k and w'' is reachable from w via $\varphi_1; \ldots; \varphi_{k-1}$. Finally, we say that w' is *reachable from w* if it is reachable via some sequence of formulas. Notice that $(M, w) \models \varphi_1 \rightarrow (\varphi_2 \rightarrow (\ldots(\varphi_k \rightarrow \varphi)\ldots))$ if φ is true at every world w' reachable from w via $\varphi_1; \ldots; \varphi_k$.

Lemma 3.1: *Suppose C8 is valid in M.*

(a) If $(M, w) \models \Box\varphi$, then $(M, w') \models \varphi \wedge \Box\varphi$ for every w' reachable from w via any sequence of formulas.

(b) If $(M, w) \models \Diamond\varphi$, then $(M, w') \models \Diamond\varphi$ for every w' reachable from w via any sequence of formulas.

Proof: For part (a), note that part (a) of C8 says that if $(M, w) \models \Box\varphi$, then $(M, w') \models \Box\varphi$ for every world w' reachable from w via φ'. Inductively, it follows that $(M, w') \models \Box\varphi$ for every world w' reachable from w via any sequence of formulas. Since C8 also tells us that $\Box\varphi \Rightarrow \varphi$ is valid, this means that φ is true at all worlds reachable from w via any sequence of formulas. Part (b) follows similarly from part (b) of C8. ∎

As the following theorem shows, the observation in Lemma 3.1 is enough to essentially force S8.[9]

Theorem 3.2: *Let $\mathcal{S}$ be a (possibly empty) subset of $\{S1, \ldots, S8\}$ and let $\mathcal{C}$ be the corresponding subset of $\{C1, \ldots, C8\}$. Then $\{Prop, C0, MP, RC1\} \cup \mathcal{C}$ is a sound and complete axiomatization for the language $\mathcal{L}^C$ with respect to the class of counterfactual structures satisfying the conditions in $\mathcal{S}$.*

Proof: As usual, soundness is straightforward. While the basic ideas of the completeness proof are standard (and go back to [Stalnaker and Thomason 1970]), I sketch some of the details here, since this result is somewhat more general than the ones that appear in the literature. I assume that the reader is familiar with the standard canonical model arguments from the literature.

For completeness, we must show that every consistent formula is satisfiable. Suppose that φ_0 is consistent. We first consider the case that C8 $\notin \mathcal{C}$.

Let W consist of all the maximal consistent sets of formulas in $\mathcal{L}^C$. Let $g : W \times \mathcal{L}^C \mapsto 2^{\mathcal{L}^C}$ be defined via $g(w, \varphi) = \{\psi : \varphi \rightarrow \psi \in w\}$. Define a selection f on W so that

[9] In the presence of C1, C5, C6, RC1, and LLE, C8 can be expressed in more familiar terms. It is easy to see that in the presence of C1, C5, C6, RC1, and LLE, $\Box\varphi \Rightarrow (\psi \rightarrow \varphi)$ is valid. (Proof: From C1, we have $(\varphi \wedge \psi) \rightarrow (\varphi \wedge \psi)$. By RC1, we have $(\varphi \wedge \psi) \rightarrow \varphi$. From $\neg\varphi \rightarrow$ *false* $(\Box\varphi)$ and RC1 we have $\neg\varphi \rightarrow \varphi$ and $\neg\varphi \rightarrow \psi$. From C6, we get that $\neg\varphi \wedge \psi \rightarrow \varphi$. Now by C5, we get $((\neg\varphi \wedge \psi) \vee (\varphi \wedge \psi)) \rightarrow \varphi$. Finally, by LLE, we get $\psi \rightarrow \varphi$.) Thus, C8 follows from $(\Box\varphi \Rightarrow (\Box\Box\varphi \wedge \varphi)) \wedge (\Diamond\varphi \Rightarrow \Box\Diamond\varphi)$. This is a conjunction of two axioms denoted R (for reflexivity) and U (for uniformity) by Lewis [1973]. The converse holds as well; this follows from Theorem 3.2.

$f(w,\varphi) = \{w' : g(w,\varphi) \subseteq w'\}$, and define an interpretation π such that $\pi(w,p) = \mathbf{true}$ iff $p \in w$. Let $M = (W, f, \pi)$.

We can now prove the usual *Truth Lemma*: for every formula $\varphi \in \mathcal{L}^C$, we have $(M,w) \models \varphi$ iff $\varphi \in w$. We proceed by induction; the only nontrivial case comes if φ has the form $\psi_1 \rightarrow \psi_2$. If $\psi_1 \rightarrow \psi_2 \in w$, we want to show that $(M,w) \models \psi_1 \rightarrow \psi_2$. Thus, we must show that $f(w,\psi_1) \subseteq [\![\psi_2]\!]_M$. By definition, $f(w,\psi_1) = \{w' : g(w,\psi_1) \subseteq w'\}$. Since $\psi_1 \rightarrow \psi_2 \in w$, it follows that $\psi_2 \in g(w,\psi_1)$, so $f(w,\psi_1) \subseteq \{w' : \psi_2 \in w'\}$. By the induction hypothesis, $\{w' : \psi_2 \in w'\} = [\![\psi_2]\!]_M$. Thus, $(M,w) \models \psi_1 \rightarrow \psi_2$, as desired.

For the opposite direction, suppose $(M,w) \models \psi_1 \rightarrow \psi_2$. Thus, $f(w,\psi_1) \subseteq [\![\psi_2]\!]_M$. It follows that ψ_2 must be provable from the formulas in $g(w,\psi)$, for otherwise there would be a world w' containing $g(w,\psi_1)$ and $\neg\psi_2$, and w' would be in $f(w,[\![\psi_1]\!]_M) - [\![\psi_2]\!]_M$, by the induction hypothesis. Thus, there is a finite set of formulas in $g(w,\psi_1)$, say $\{\sigma_1,\ldots,\sigma_k\}$, such that $\sigma_1 \wedge \ldots \wedge \sigma_k \Rightarrow \psi_2$ is provable. Since $\psi_1 \rightarrow \sigma_j \in w$, $j = 1,\ldots,k$, it follows from C0 that $\psi_1 \rightarrow (\sigma_1 \wedge \ldots \wedge \sigma_k) \in w$. Now applying RC1, we get that $\psi_1 \rightarrow \psi_2 \in w$, as desired.

It follows that φ_0 is satisfiable in M.

Now we have to show that f satisfies all the properties in $\mathcal{S}$. The arguments are straightforward. I consider three representative cases here.

C1: Suppose that C1 $\in \mathcal{C}$. We want to show that S1 holds. Thus, we must show that $f(w,\varphi) \subseteq [\![\varphi]\!]_M$. By definition, if $w' \in f(w,\varphi)$, then $g(w,\varphi) \subseteq w'$. If C1 holds, then $\varphi \in g(w,\varphi)$. Thus, if $w' \in f(w,\varphi)$, then $\varphi \in w'$. By the Truth Lemma, $w' \in [\![\varphi]\!]_M$.

C4: Suppose that C4 $\in \mathcal{C}$ and $f(w,\varphi) \neq \emptyset$. We want to show that $f(w,\varphi)$ is a singleton. Suppose that $w_1 \neq w_2$ are both in $f(w,\varphi)$. Thus $g(w,\varphi) \subseteq w_1 \cap w_2$. There must be some formula ψ such that $\psi \in w_1$ and $\neg\psi \in w_2$. It follows from C4 that either $(M,w) \models \varphi \rightarrow \psi$ or $(M,w) \models \varphi \rightarrow \neg\psi$. In the former case, $\psi \in g(w,\varphi)$, while in the latter case, $\neg\psi \in g(w,\varphi)$. Either way, it follows that $g(w,\varphi) - w_1 \cap w_2 \neq \emptyset$, a contradiction.

C6: Suppose C6 $\in \mathcal{C}$ and $f(w,\varphi) \subseteq [\![\psi]\!]_M$. We want to show that $f(w,\varphi \wedge \psi) \subseteq f(w,\varphi)$. Suppose that $w' \in f(w,\varphi \wedge \psi)$. This means that $g(w,\varphi \wedge \psi) \subseteq w'$. To show that $w' \in f(w,\varphi)$, we need to show that $g(w,\varphi) \subseteq w'$. But if $\sigma \in g(w,\varphi)$, then $\varphi \rightarrow \sigma \in w$. By C6, it follows that $(\varphi \wedge \psi) \rightarrow \sigma \in w$, so $\sigma \in g(w,\varphi \wedge \psi)$. Since $g(w,\varphi \wedge \psi) \subseteq w'$, it follows that $\sigma \in w'$. Thus, $g(w,\varphi) \subseteq w'$, as desired.

Finally, we must deal with the case that C8 $\in \mathcal{C}$. Let w_0 be a world in M such that $\varphi_0 \in w_0$. (Recall that φ_0 is the formula which we are trying to show is satisfiable.) Let W' consist of all worlds in W reachable from w_0. It is almost immediate from the definitions that if $w' \in W'$, then $f(w',\varphi) \subseteq W'$ for all formulas φ. Let f' and π' be the restrictions of f and π to W', respectively, and let $M(w_0) = (W', f', \pi')$. The Truth Lemma holds for W'; the proof above goes through without change. Similarly the arguments that all the properties in $\mathcal{S}$ other than S8 hold in $M(w_0)$ go through without change. Thus, φ_0 is satisfiable in $M(w_0)$, and all the properties in $\mathcal{S}$ other than $\mathcal{S}$ also hold in $M(w_0)$.

To see that S8 also holds in $M(w_0)$, suppose that $[\![\varphi]\!]_{M(w_0)} \neq \emptyset$. We want to show that $f(w, \varphi) \neq \emptyset$ for all $w \in W'$, or equivalently, that $(M(w_0), w) \models \Diamond\varphi$. Since all worlds in W' are reachable from w_0, there must be two sequences of worlds, $w_1, \ldots, w_k$ and $w'_0, w'_1, \ldots, w'_m$, and two sequences of formulas, $\varphi_1, \ldots, \varphi_k$ and $\varphi'_1, \ldots, \varphi'_m$ such that $w_k = w$, $w'_0 = w_0$, $(M(w_0), w'_m) \models \varphi$, w_j is reachable from w_{j-1} via φ_j, for $j = 1, \ldots, k$, and w'_j is reachable from w'_{j-1} via φ'_j from $j = 1, \ldots, m$. We must have $(M(w_0), w'_0) \models \Diamond\varphi$, for otherwise, by definition, we have $(M(w_0), w'_0) \models \Box\neg\varphi$, and it would follow by part (a) of Lemma 3.1 that $(M(w_0), w'_m) \models \neg\varphi$, a contradiction. Since $w'_0 = w_0$, we have that $(M(w_0), w_0) \models \Diamond\varphi$. Now aplying part (b) of Lemma 3.1, it follows that $(M(w_0), w_k) \models \Diamond\varphi$. Since $w_k = w$¡ we get the desired result. ∎

3.2 The Set-Theoretic Approach: Selection Functions

In the set-theoretic approach, rather than having a syntactic operator $\rightarrow$, we have a binary operator $\stackrel{e}{\rightarrow}$ on events (the superscript e stands for *event*); that is $\stackrel{e}{\rightarrow}: 2^W \times 2^W \rightarrow 2^W$. For ease of exposition, I write $H \stackrel{e}{\rightarrow} E$ instead of $\stackrel{e}{\rightarrow} (H, E)$. Intuitively, $w \in H \stackrel{e}{\rightarrow} E$ if, at world w, if H were to hold, then so would E. Again, we can give semantics to $\stackrel{e}{\rightarrow}$ using selection functions, but since we no longer have formulas, we replace the formula that is the second argument in the syntactic case by a set of worlds (which we can think of as the intension of a formula). Thus, a *(set-theoretic) selection function* maps $W \times 2^W$ to 2^W.

A *(set-theoretic) counterfactual structure* is then a pair $M = (W, \mathsf{f})$, where W is a set of worlds and f is a set-theoretic selection function.

We can then define the binary operator $\stackrel{e}{\rightarrow}: 2^W \times 2^W \mapsto 2^W$ in M as follows:

$$H \stackrel{e}{\rightarrow} E = \{w : \mathsf{f}(w, H) \subseteq E\}. \tag{5}$$

We can define restrictions on f completely analogous to those defined in the syntactic case. These are listed below, along with one additional restriction, S9$'$.

S1$'$. $\mathsf{f}(w, H) \subseteq H$

S2$'$. If $\mathsf{f}(w, H) \subseteq H'$ and $\mathsf{f}(w, H') \subseteq H$, then $\mathsf{f}(w, H) = \mathsf{f}(w, H')$

S3$'$. If $w \in H$, then $\mathsf{f}(w, H) = \{w\}$

S4$'$. $\mathsf{f}(w, H)$ is either empty or a singleton

S5$'$. $\mathsf{f}(w, H_1 \cup H_2) \subseteq \mathsf{f}(w, H_1) \cup \mathsf{f}(w, H_2)$

S6′. If $\mathsf{f}(w,H) \subseteq E$ then $\mathsf{f}(w,H\cap E) \subseteq \mathsf{f}(w,H)$[10]

S7′. If $\mathsf{f}(w,H)\cap E \neq \emptyset$, then $\mathsf{f}(w,H\cap E) \subseteq \mathsf{f}(w,H)\cap E$

S8′. If $H \neq \emptyset$ then $\mathsf{f}(w,H) \neq \emptyset$

S9′. If $\mathsf{f}(w,H) \subseteq E_1 \cup E_2$, then there exist H_1, H_2 such that $H_1 \cup H_2 = H$, $\mathsf{f}(w,H_1) \subseteq E_1$, and $\mathsf{f}(w,H_2) \subseteq E_2$

S9′ has no analogue in the syntactic case. Since S5′ is easily seen to be equivalent to "if there exist H_1, H_2 such that $H_1 \cup H_2 = H$, $\mathsf{f}(w,H_1) \subseteq E_1$, and $\mathsf{f}(w,H_2) \subseteq E_2$, then $\mathsf{f}(w,H) \subseteq E_1 \cup E_2$", S9′ can be viewed as a converse to S5′. It is also not hard to see that S9′ follows from S1′ and S7′. (Proof: if $H\cap E_1 \neq \emptyset$ and $H \cap E_2 \neq \emptyset$, then we can take $H_1 = H\cap E_1$ and $H_2 = H \cap (E_2 \cup \neg H_1)$; if $H \cap E_1 = \emptyset$, we can take $H_1 = \emptyset$, $H_2 = E$, and similarly if $H \cap E_2 = \emptyset$.) As we shall see, S9′ arises (along with S1′, S2′, S5′, and S6′) when the selection function f is induced by an ordering on worlds.

Each of these restrictions corresponds to an axiom, completely analogous to C0–C8. Consider the following axioms:

C0′. $\cap_{j\in J}(H \xrightarrow{e} E_j) = H \xrightarrow{e} \cap_{j\in J} E_J$ for any index set J and events $E_j, j \in J$

C1′. $(H \xrightarrow{e} H) = W$

C2′. $(H \xrightarrow{e} H') \cap (H' \xrightarrow{e} H) \cap (H \xrightarrow{e} E) \subseteq (H' \xrightarrow{e} E)$

C3′. $H \cap (H \xrightarrow{e} E) = H \cap E$

C4′. $(H \xrightarrow{e} E) \cup (H \xrightarrow{e} \neg E) = W$

C5′. $(H_1 \xrightarrow{e} E) \cap (H_2 \xrightarrow{e} E) \subseteq (H_1 \cup H_2) \xrightarrow{e} E$

C6′. $(H \xrightarrow{e} E_1) \cap (H \xrightarrow{e} E_2) \subseteq (H \cap E_1) \xrightarrow{e} E_2$

C7′. $\neg(H \xrightarrow{e} \neg E_1) \cap (H \xrightarrow{e} E_2) \subseteq (H \cap E_1) \xrightarrow{e} E_2$

C8′. $H \xrightarrow{e} \emptyset = \emptyset$ if $H \neq \emptyset$

C9′. $H \xrightarrow{e} (E_1 \cup E_2) \subseteq \cup_{\{H_1,H_2 : H_1 \cup H_2 = H\}}((H_1 \xrightarrow{e} E_1) \cap (H_2 \xrightarrow{e} E_2))$

We can now state and prove the semantic soundness and completeness result.

[10] Again, the stronger version, with $\subseteq$ replaced by equality, follows from S1′, S2′, and S5′, using arguments almost identical to the earlier ones. Similarly, a stronger version of S7′ follows from S1′, S2′, and S5′.

Theorem 3.3: *Let $\mathcal{S}$ be a (possibly empty) subset of $\{S1', \ldots, S9'\}$, let $\mathcal{C}$ be the corresponding subset of $\{C1', \ldots, C9'\}$, and let W be a set of worlds. If f is a set-theoretic selection function on W that satisfies the conditions in $\mathcal{S}$ and $\stackrel{e}{\rightarrow}$ is defined in (W, f) by (5), then $\stackrel{e}{\rightarrow}$ satisfies C0′ and the axioms in $\mathcal{C}$. Conversely, if the function $\rightsquigarrow: 2^W \times 2^W \mapsto 2^W$ satisfies C0′ and the axioms in $\mathcal{C}$, then there is a selection function f on W satisfying $\mathcal{S}$ such that $\rightsquigarrow$ is the counterfactual operator $\stackrel{e}{\rightarrow}$ in (W, f).*

Proof: It is easy to check that if f satisfies the conditions in $\mathcal{S}$, then $\stackrel{e}{\rightarrow}$ satisfies C0′ and all the conditions in $\mathcal{C}$. For the second half, given a function $\rightsquigarrow: 2^W \times 2^W \mapsto 2^W$ that satisfies C0′ and the axioms in $\mathcal{C}$, define $\mathsf{f}(w, H) = \cap\{E : w \in H \rightsquigarrow E\}$. I leave it to the reader to check that $\rightsquigarrow$ is the counterfactual operator $\stackrel{e}{\rightarrow}$ in (W, f) and satisfies $\mathcal{S}$. ∎

Again, a number of points are worth making. First, observe how C8′ captures S8′ far more directly than C8 captures S8. We did not have to struggle to find an axiom corresponding to this condition. Next, observe that the completeness proof proceeds in somewhat the same spirit as the syntactic completeness proof, but avoids the construction of a canonical model. It works whichever model we start with. An analogue to Corollary 2.7 can also be proved, showing that, again, we are entitled to view Theorem 3.3 as a soundness and completeness result. Finally, as I now show, we need the full strength of C0′ for completeness.

Consider the obvious finitary analogue of C0′:

C10′. $(H \stackrel{e}{\rightarrow} E_1) \cap (H \stackrel{e}{\rightarrow} E_2) = H \stackrel{e}{\rightarrow} (E_1 \cap E_2)$

I actually give two examples showing that we cannot in general replace C0′ with C10′. The first is a simple example that satisfies C1′–C3′, C5′–C10′, but not C0′ (or C4′). Then I give a somewhat more sophisticated example that satisfies all of C1′–C10′, but not C0′. Since $\stackrel{e}{\rightarrow}$ must satisfy C0′, this shows that C10′ does not suffice, even in the presence of all the other properties.

Example 3.4: Let $W = \{1, 2, 3, \ldots\}$. Define $\rightsquigarrow$ so that if $w \in H$, then $w \in H \rightsquigarrow E$ iff $w \in E$; if $w \notin H$, then $w \in H \rightsquigarrow E$ iff (a) $H \cap \neg E = \emptyset$ or (b) H is infinite and $H \cap \neg E$ is finite. I leave to the reader the somewhat tedious (but straightforward) task of checking the $\rightsquigarrow$ satisfies C1′–C3′, C5′–C10′. It does not satisfy C0′, since if $E_j = \neg\{j\}$, then $E_1 \rightsquigarrow E_1 = W$ and $E_1 \rightsquigarrow E_j = E_j$ for for $j = 2, 2, 3, \ldots$. But $\cap_j E_j = \emptyset$, $E_1 \rightsquigarrow \emptyset = \emptyset$, and $\cap_j (E_1 \rightsquigarrow E_j) = \{1\}$. Thus, C0′ does not hold. Note that $\rightsquigarrow$ also does not satisfy C4′ since, for example $(\{1, 2\} \rightsquigarrow \{1\}) \cup (\{1, 2\} \rightsquigarrow \neg\{1\}) = \{1, 2\} \neq W$. ∎

Example 3.5: For this example, we need to review some material on filters and ultrafilters [Bell and Slomson 1974]. A *filter* on W is a nonempty set $\mathcal{U}$ of subsets of W that is closed under supersets (i.e., if $E \in \mathcal{U}$ and $E \subset E'$, then $E' \in \mathcal{U}$) and finite intersections, and does not contain the empty set. An *ultrafilter* is a maximal filter, that is, a filter that is not a subset of any other filter. A *principal ultrafilter* is an ultrafilter which consists of all the supersets of a particular element of W. (It is easy to check that this is indeed an ultrafilter.) A *nonprincipal ultrafilter* is an ultrafilter that is not a principal ultrafilter. Note that if $\mathcal{U}$ is an ultrafilter, then for any set H, either H or $\neg H$ must be in $\mathcal{U}$. From this it follows that if $H \in \mathcal{U}$, then for any set E, we must have that one of $H \cap E$ or $H \cap \neg E$ is in $\mathcal{U}$. (Proof: If neither is in $\mathcal{U}$, then H, $\neg H \cup E$, and $\neg H \cup \neg E$ are all in $\mathcal{U}$. Since $\mathcal{U}$ is closed under finite intersections, it follows that $\emptyset \in \mathcal{U}$. This contradicts the fact that $\mathcal{U}$ is a filter.)

Let $\mathcal{U}$ be a nonprincipal ultrafilter on $W = \{1, 2, 3, \ldots\}$. (Nonprincipal ultrafilters on W can be shown to exist using Zorn's Lemma.) Define $\rightsquigarrow'$ as follows: $w \in H \rightsquigarrow' E$ if and only if the following conditions hold:

(a) if $w \in H$, then $w \in H \cap E$,

(b) if $w \notin H$ and $H \in \mathcal{U}$, then $H \cap E \in \mathcal{U}$,

(c) if $w \notin H$ and $H \notin \mathcal{U}$, then $H = \emptyset$ or $\min(H \cap E) = \min(H)$ (where $\min(F)$ denotes the minimal element of F).

Again, it is straightforward (but tedious) to show that $\rightsquigarrow'$ satisfies C1′–C10′. For example, to see that C4′ holds, we must show that, for all E, we have $(H \rightsquigarrow' E) \cup (H \rightsquigarrow' \neg E) = W$. We do a case-by-case analysis. If $w \in H$, then, by clause (a), then if $w \in E$, we also have $w \in H \rightsquigarrow' E$, by definition; otherwise, $w \in H \rightsquigarrow' \neg E$. If $w \notin H$ and $H \in \mathcal{U}$, then one of $H \cap E$ or $H \cap \neg E$ must be in $\mathcal{U}$, since otherwise, since $\mathcal{U}$ is an ultrafilter, we would have both $\neg H \cup E \in \mathcal{U}$ and $\neg H \cup \neg E \in \mathcal{U}$, from which it follows that $\neg H \in U$, contradicting the assumption that $H \in \mathcal{U}$. Finally, if $w \notin H$ and $H \notin \mathcal{U}$, then either $\min(H) = \min(H \cap E)$ or $\min(H) = \min(H \cap \neg E)$. In all cases, we have $w \in (H \rightsquigarrow' E) \cup (H \rightsquigarrow' \neg E)$.

To see that $\rightsquigarrow'$ does not satisfy C0′, note that a nonprincipal ultrafilter can contain only infinite sets, and so must contain all cofinite sets. Taking the sets E_i as in Example 3.4, it follows that E_i and $E_i \cap E_j$ are both in $\mathcal{U}$, for all i, j (since these are all cofinite sets). It then follows from the definition that $E_1 \rightsquigarrow' E_1 = W$ and $E_1 \rightsquigarrow' E_j = E_j$. Thus, we get a violation of C0′ just as in Example 3.4. ∎

3.3 Preferential Orders

I have described $f(w, \varphi)$ as the "closest" worlds to w satisfying φ. This suggests that there is an underlying ordering on worlds. Lewis [1973] made this intuition explicit as follows. A *preferential frame* is a pair (W, R), where R is a ternary *preferential relation* on a set W of possible worlds. For technical reasons that will be discussed at the end of this subsection, I assume that W is finite in this subsection. We typically write $w' \preceq^w w''$ rather than $R(w, w', w'')$. This should be thought of as saying that w' is at least as close to w as w''; thus, $\preceq^w$ represents the "at least as close to w" relation. Let $W_w = \{w' : \exists w''(w' \preceq^w w'')\}$. We can think of W_w as the domain of $\preceq^w$. Intuitively, the worlds not in W_w are so far away from w that they cannot even be discussed. As would be expected from the intuition, we require that $\preceq^w$ be a partial preorder on W_w, that is, a reflexive, transitive relation.[11] We define the relation $\prec^w$ by taking $w' \prec^w w''$ if $w' \preceq^w w''$ and not$(w'' \preceq^w w')$.

Given our intuitions regarding closeness, the following requirements seem reasonable:

P1. $w \in W_w$ and is the minimal element with respect to $\preceq^w$ (so that w is closer to itself than any other element): formally, for all $w' \in W_w$, we have $w \preceq^w w$ and $w \prec^w w'$.

P2. $\preceq^w$ is a total preorder on W_w: that is, for all $w', w'' \in W_w$, either $w' \preceq^w w''$ or $w'' \preceq^w w'$.

P3. $\preceq^w$ is a linear order on W_w: that is, for all $w', w'' \in W_w$, if $w' \neq w''$, then either $w' \prec^w w'$ or $w'' \prec^w w'$.

P4. $W_w = W$ for all $w \in W$

In a preferential frame, we can define a (set-theoretic) selection function $\mathsf{f}_\preceq$ such that $\mathsf{f}_\preceq(w, H)$ are the worlds closest to w, according to $\preceq^w$, that are in H. Formally, we have

$$\mathsf{f}_\preceq(w, H) = \{w' \in H \cap W_w : \text{if } w'' \prec^w w' \text{ then } w'' \notin H\}.$$

This gives us a way of defining a binary operator $\xrightarrow{e}$ in preferential frames, by an immediate appeal to Definition (5).

We can also define a syntactic version (which is actually what Lewis did). A *preferential structure* is a tuple $M = (W, R, \pi)$, where (W, R) is a preferential frame and π is an interpretation. Roughly speaking, we would now like to define a syntactic selection function $f_\preceq$ in M by taking

$$f_\preceq(w, \varphi) = \{w' \in [\![\varphi]\!]_M \cap W_w : \text{if } w'' \prec^w w' \text{ then } w'' \notin [\![\varphi]\!]_M\}, \tag{6}$$

and then use this selection function to give semantics to conditional formulas. The only problem is that we have not yet defined $[\![\varphi]\!]_M$. The formal definition does not use the selection function directly. Nevertheless, it is easy to check that the formal definition is consistent with this intuition.

Formally, the definition of $\models$ in preferential structures is the same as that in counterfactual structures, except for the clause for $\rightarrow$, which is

$(M, w) \models \varphi \rightarrow \psi$ if $\mathsf{f}_\preceq(w, [\![\varphi]\!]_M) \subseteq [\![\psi]\!]_M$

This is well-defined, since the inductive definition of $\models$ guarantees that the sets $[\![\varphi]\!]_M$ and $[\![\psi]\!]_M$ have already been defined. This definition makes precise the intuition that

[11]Note that $\preceq^w$ is not necessarily anti-symmetric. That is why it is a preorder, not an order. Nor is it necessarily total; totality is forced by condition P2, which will be defined shortly.

$\varphi \rightarrow \psi$ holds at world w if the φ-worlds closest to w (according to the ordering $\preceq^w$ used at w) all satisfy ψ.

The following syntactic completeness result is well known. It says that in preferential structures, $\rightarrow$ satisfies C0, C1, C2, C5, C6, and RC1; moreover, P1, P2, P3, and P4 give us C3, C7, C4, and C8, respectively.

Theorem 3.6: [Burgess 1981; Friedman and Halpern 1994; Lewis 1973] *Let $\mathcal{P}$ be any (possibly empty) subset of {P1,P2,P3,P4} let $\mathcal{C}$ be the corresponding subset of {C3,C7,C4,C8}, and let W be a finite set of worlds. Let $\mathcal{M}^{\mathcal{P}}$ be the class of preferential structures where the ternary relation R satisfies the conditions in $\mathcal{P}$. Then $\mathcal{C}\cup${Prop, C0, C1, C2, C5, C6, MP, RC1} is a sound and complete axiomatization for the language $\mathcal{L}^C$ for the class of preferential structures in $\mathcal{M}^{\mathcal{P}}$.*[12]

What about set-theoretic completeness? Not surprisingly in the light of the previous theorem, it turns out that $f_\preceq$ satisfies S1′, S2′, S5′, S6′; moreover, P1, P2, P3, and P4 correspond to S3′, S7′, S4′, and S8′, respectively. The interesting thing is that we also get S9′.

Lemma 3.7: *Let $\mathcal{P}$ be a subset of {P1,P2,P3,P4}, and let $\mathcal{S}$ be the corresponding subset of {S4′,S7′,S4′,S8′}. If (W, R) is a preferential frame satisfying the properties in $\mathcal{P}$, then $\mathsf{f}_\preceq$ satisfies S1′, S2′, S5′, S6′, S9′ and all the properties in $\mathcal{S}$.*

Proof: Proving that $\mathsf{f}_\preceq$ satisfies all the properties is straightforward. I consider only S9′ here. To see that $f_\preceq$ satisfies S9′, suppose $\mathsf{f}_\preceq(w,H) \subseteq E_1 \cup E_2$. Let $E_i' = E_i \cap f(w,H)$, for $j = 1,2$. Let E_j^H consist of all the elements in H that are at least as far from w as some element in E_j. That is, $E_j^H = \{w' \in H : \exists w'' \in E_j'(w'' \preceq^w w')\}$ Then define $H_1 = (E_1^H - (E_2' - E_1')) \cup (H - W_w)$ and $H_2 = (E_2^H - (E_1' - E_2')) \cup (H - W_w)$. It is not hard to show that $H = H_1 \cup H_2$. (Proof: By construction, $A^H \subseteq H$, so we must have $H_1 \cup H_2 \subseteq H$. For the opposite containment, note that if $v \in H - W_w$, then the construction guarantees that $v \in H_1 \cap H_2$. Also note that $E_j' \cap W_w \subseteq E_j^H$, since if $v \in E_j'$ then $v \in E_j \cap H$ and if $v \in W_w$, then $v \preceq^w v$. Thus, $E_j' \subseteq H_j$ for $j = 1,2$. Finally, if $v \in H \cap W_w - (E_1' \cup E_2')$, then choose $v' \in \mathsf{f}_\preceq(w,H)$ such that $v' \preceq^w v$. Such a v' must exist by definition, $v' \in H$ by construction, and $v' \in E_1 \cup E_2$ since $\mathsf{f}_\preceq(w,H) \subseteq E_1 \cup E_2$. If $v' \in E_j$, then we must have $v \in E_j^H$ and hence $v \in H_j$.)

It remains to show that $\mathsf{f}_\preceq(w,H_j) \subseteq E_j$, $j = 1,2$. I consider the case that $j = 1$. (The proof for $j = 2$ is almost identical.) Suppose that $v \in \mathsf{f}_\preceq(w,H_1)$. Thus, $v \in E_1^H - (E_2' - E_1')$, since $\mathsf{f}_\preceq(w,H_1) \subseteq W_w$. Since $v \in E_1^H$, there exists some $v' \in E_1'$ such that $v' \preceq^w v$. Since, as we observed above, $E_1' \subseteq H_1$, we have $v' \in H_1$. Thus we cannot have $v' \prec^w v$, for otherwise $v \notin \mathsf{f}_\preceq(w,H_j)$. Thus, $v \preceq^w v'$. It follows that $v \in \mathsf{f}_\preceq(w,H)$, for otherwise there would be some $v'' \in E_1' \cup E_2'$ such that $v'' \prec^w v$. This would mean

[12]Note that since C1, C5, C6, RC1, and LLE all hold (recall that LLE follows from C1, C2, and RC1), we can replace C8 by the simpler $(\Box\varphi \Rightarrow (\Box\Box\varphi \wedge \varphi)) \wedge (\Diamond\varphi \Rightarrow \Box\Diamond\varphi)$.

that $v'' \prec^w v'$, contradicting the fact that $v' \in E_1'$. Thus, it follows that $v \in E_1' \cup E_2'$. Since $v \in H_1$, we cannot have $v \in E_2' - E_1'$. Hence $v \in E_1'$, as desired. ∎

We then get the following set-theoretic soundness and completeness result. Note that we can use C10$'$ instead of C0$'$, since we are restricting to finite sets of worlds.

Theorem 3.8: *Let $\mathcal{P}$ be any (possibly empty) subset of $\{P1,P2,P3,P4\}$, let $\mathcal{C}$ be the corresponding subset of $\{C3',C7',C4',C8'\}$, and let W be a finite set of worlds. If R is a ternary relation on W that satisfies the conditions in $\mathcal{P}$ then $\xrightarrow{e}$ satisfies C1$'$, C2$'$, C5$'$, C6$'$, C9$'$, C10$'$, and all the axioms in $\mathcal{C}$. Conversely, if $\rightsquigarrow: 2^W \times 2^W \mapsto 2^W$ and satisfies C1$'$, C2$'$, C5$'$, C6$'$, C9$'$, C10$'$, and the axioms in $\mathcal{C}$, then there is a ternary relation R on W satisfying the conditions in $\mathcal{P}$ such that $\rightsquigarrow$ is the counterfactual operator $\xrightarrow{e}$ in (W, R).*

Proof: The first half (soundness) follow immediately from Lemma 3.7 and Theorem 3.3. For the second half, suppose that $\rightsquigarrow$ satisfies C1$'$, C2$'$, C5$'$, C6$'$, C9$'$, C10$'$, and the axioms in $\mathcal{C}$. First assume that C7$' \notin \mathcal{C}$. Define $w' \preceq^w w''$ iff $w \in \{w', w''\} \rightsquigarrow \{w'\}$ and $w \notin \{w', w''\} \rightsquigarrow \emptyset$. (This way of defining the ordering is essentially due to [Kraus, Lehmann, and Magidor 1990]; a slightly different ordering for the case that C7$'$ is in $\mathcal{C}$ is described at the end of the proof.) We must show that $\preceq^w$ is a partial order on W_w and that the operator $\xrightarrow{e}$ determined by this relation agrees with $\rightsquigarrow$. The following lemma will prove useful:

Lemma 3.9: $W_w = \{w' : w \notin \{w'\} \rightsquigarrow \emptyset\}$.

Proof: Let $W' = \{w' : w \notin \{w'\} \rightsquigarrow \emptyset\}$. If $w' \in W'$, then we have $w' \preceq^w w'$ (since, by C1$'$, we must have $w \in \{w'\} \rightsquigarrow \{w'\}$) and thus $w' \in W_w$. Conversely, if $w' \in W_w$, then there must be some w'' such that $w' \preceq^w w''$. Thus,

$$w \in \{w', w''\} \rightsquigarrow \{w'\} \tag{7}$$

and

$$w \notin \{w', w''\} \rightsquigarrow \emptyset. \tag{8}$$

Suppose, by way of contradiction, that

$$w \in \{w'\} \rightsquigarrow \emptyset. \tag{9}$$

From C0$'$ and (9), it follows that

$$w \in \{w'\} \rightsquigarrow \{w''\}. \tag{10}$$

By C1$'$, we have

$$w \in \{w''\} \rightsquigarrow \{w''\}. \tag{11}$$

By C5$'$ applied to (10) and (11), we have

$$w \in \{w', w''\} \rightsquigarrow \{w''\}. \tag{12}$$

Finally, from C10$'$ applied to (7) and (12), we get $w \in \{w', w''\} \rightsquigarrow \emptyset$. But this contradicts (8). ∎

It follows immediately from C1$'$ and Lemma 3.9 that $\preceq^w$ is reflexive on W_w. To show that it is transitive, it is easy to see that it suffices to show that if E_1, E_2, E_3 are disjoint sets, then

$$((E_1 \cup E_2) \rightsquigarrow E_1) \cap ((E_2 \cup E_3) \rightsquigarrow E_2) \subseteq ((E_1 \cup E_3) \rightsquigarrow E_1).$$

(Transitivity follows easily from the special case that $E_j = \{w_j\}$, $j = 1, 2, 3$, where w_1, w_2, and w_3 are different worlds.)

For the proof, it is useful to have two preliminary lemmas.

Lemma 3.10: *If $E \subseteq E'$ then $H \rightsquigarrow E \subseteq H \rightsquigarrow E'$.*

Proof: If $E \subseteq E'$, then by C10$'$ we have

$$(H \rightsquigarrow E) \cap (H \rightsquigarrow E') = H \rightsquigarrow (E \cap E') = H \rightsquigarrow E. \quad ∎$$

Lemma 3.11: $(H_1 \rightsquigarrow E_1) \cap (H_2 \rightsquigarrow E_2) \subseteq (H_1 \cup H_2) \rightsquigarrow (E_1 \cup E_2)$.

Proof: Using Lemma 3.10 and C5$'$, we have

$$(H_1 \rightsquigarrow E_1) \cap (H_2 \rightsquigarrow E_2) \subseteq (H_1 \rightsquigarrow (E_1 \cup E_2)) \cap (H_2 \rightsquigarrow (E_1 \cup E_2)) \subseteq (H_1 \cup H_2) \rightsquigarrow (E_1 \cup E_2).$$

∎

From Lemma 3.11, we have that

$$((E_1 \cup E_2) \rightsquigarrow E_1) \cap ((E_2 \cup E_3) \rightsquigarrow E_2) \subseteq (E_1 \cup E_2 \cup E_3) \rightsquigarrow (E_1 \cup E_2) \tag{13}$$

Applying C1$'$ and Lemma 3.11, we also get

$$(E_1 \cup E_2) \rightsquigarrow E_1 = ((E_1 \cup E_2) \rightsquigarrow E_1) \cap (E_3 \rightsquigarrow E_3) \subseteq (E_1 \cup E_2 \cup E_3) \rightsquigarrow (E_1 \cup E_3) \tag{14}$$

Finally, using (13), (14), C10$'$, and C6$'$, we get

$$\begin{array}{rl} & ((E_1 \cup E_2) \rightsquigarrow E_1) \cap ((E_2 \cup E_3) \rightsquigarrow E_2) \\ \subseteq & ((E_1 \cup E_2 \cup E_3) \rightsquigarrow (E_1 \cup E_2)) \cap ((E_1 \cup E_2 \cup E_3) \rightsquigarrow (E_1 \cup E_3)) \\ = & ((E_1 \cup E_2 \cup E_3) \rightsquigarrow E_1) \cap ((E_1 \cup E_2 \cup E_3) \rightsquigarrow (E_1 \cup E_3)) \\ \subseteq & (E_1 \cup E_3) \rightsquigarrow E_1. \end{array}$$

Thus $\preceq^w$ is transitive, as desired.

Next, we must show that C3$'$, C4$'$, and C8$'$ give us P1, P2, and P4, respectively. (Recall that I have deferred the case of C7$'$.)

Suppose that $\rightsquigarrow$ satisfies C3$'$. Then $\{w\} \cap (\{w\} \rightsquigarrow \emptyset) = \emptyset$, so $w \notin \{w\} \rightsquigarrow \emptyset$; thus, $w \in W_w$ by Lemma 3.9. Moreover, since $\{w, w'\} \cap (\{w, w'\} \rightsquigarrow \{w\}) = \{w\}$, we have $w \preceq^w w'$ for all $w' \in W_w$, showing that P1 holds.

If $\rightsquigarrow$ satisfies C4$'$, then (using C1$'$ and C10$'$), we have

$$\begin{aligned} W &= \{w', w''\} \rightsquigarrow \{w'\} \cup \{w', w''\} \rightsquigarrow (W - \{w'\}) \\ &= \{w', w''\} \rightsquigarrow \{w'\} \cup (\{w', w''\} \rightsquigarrow (W - \{w'\}) \cap \{w', w''\} \rightsquigarrow \{w', w''\}) \\ &= \{w', w''\} \rightsquigarrow \{w'\} \cup \{w', w''\} \rightsquigarrow \{w''\}. \end{aligned}$$

It easily follows that for all $w', w'' \in W_w$, we have either $w' \preceq^w w''$ or $w'' \preceq^w w'$. Thus, P2 holds.

Finally, it is immediate from the definitions that if $\rightsquigarrow$ satisfies C8$'$, then $W_w = W$.

It remains to check that $\xrightarrow{e}$ and $\rightsquigarrow$ agree. We prove that

$$H \xrightarrow{e} E \text{ iff } H \rightsquigarrow E. \tag{15}$$

Note that C1$'$, C2$'$, C5$'$, C6$'$, and C10$'$ hold for both $\xrightarrow{e}$ and $\rightsquigarrow$. For $\xrightarrow{e}$, this follows from the first half of the theorem; for $\rightsquigarrow$, it follows by assumption. These are the only properties used in the proof. Note that this means that Lemmas 3.10 and 3.11 apply to both $\xrightarrow{e}$ and $\rightsquigarrow$.

Using C10$'$ and C1$'$, we have that $H \xrightarrow{e} E = H \xrightarrow{e} (H \cap E)$, and similarly for $\rightsquigarrow$; thus, we can assume without loss of generality that $E \subseteq H$. We proceed by induction on $|E|$. (Here we are making heavy use of the fact that W is finite.) If $E = \emptyset$, then if $H = \emptyset$, the result follows from C1$'$. If $H \neq \emptyset$, then for each $w' \in H$, we have

$$H \rightsquigarrow \emptyset = ((H \rightsquigarrow \emptyset) \cap (H \rightsquigarrow \{w'\})) \subseteq \{w'\} \rightsquigarrow \emptyset.$$

Thus, we have

$$H \rightsquigarrow \emptyset \subseteq \cap_{w' \in H} \{w'\} \rightsquigarrow \emptyset.$$

By C5$'$, we actually have

$$H \rightsquigarrow \emptyset = \cap_{w' \in H} \{w'\} \rightsquigarrow \emptyset.$$

An identical argument works if we replace $\rightsquigarrow$ by $\xrightarrow{e}$. Thus, it suffices to show that $\{w'\} \rightsquigarrow \emptyset = \{w'\} \xrightarrow{e} \emptyset$. But, by Lemma 3.9, $w \in \{w'\} \rightsquigarrow \emptyset$ iff $w' \notin W_w$. It is also immediate from the definitions that $w' \notin W_w$ iff $w \in \{w'\} \xrightarrow{e} \emptyset$. Thus, $\{w'\} \rightsquigarrow \emptyset = \{w'\} \xrightarrow{e} \emptyset$, as desired.

If $|E| = 1$, suppose $E = \{w'\}$. We proceed by a subinduction on $|H|$. If $|H| = 1$, then the result is immediate from C1$'$, since $E \subseteq H$. If $|H| = 2$, let $H = \{w', w''\}$. If $H \rightsquigarrow E$, there are a number of cases to consider. First suppose that $w', w'' \in W_w$. Then $w \in \{w', w''\} \xrightarrow{e} \{w'\}$ iff $w' \prec^w w''$, which implies $w \in \{w', w''\} \rightsquigarrow \{w'\}$ by definition. If $w' \notin W_w$ then,by Lemma 3.9, we have $w \in \{w'\} \rightsquigarrow \emptyset$. By C1$'$, we have $w \in \{w''\} \rightsquigarrow \{w''\}$. Now by Lemma 3.11, we have $w \in \{w', w''\} \rightsquigarrow \{w''\}$. By C10$'$, it follows that $w \in \{w', w''\} \rightsquigarrow \emptyset$. By the induction hypothesis, we have $w \in \{w', w''\} \xrightarrow{e} \emptyset$, and by Lemma 3.10, we have $w \in \{w', w''\} \xrightarrow{e} \{w'\}$, as desired. Finally, if $w'' \notin W_w$, then $w \in \{w''\} \rightsquigarrow \emptyset$. By the induction hypothesis, we have $w \in \{w''\} \xrightarrow{e} \emptyset$, and by C1$'$, we have $w \in \{w'\} \xrightarrow{e} \{w'\}$. Now by Lemma 3.11, we have $w \in \{w', w''\} \xrightarrow{e} \{w'\}$, as desired.

For the converse, note that if $w \in \{w', w''\} \xrightarrow{e} \{w'\}$, then either $w' \prec^w w''$ or $w'' \notin W_w$. If $w' \prec^w w''$, then we have $w \in \{w', w''\} \rightsquigarrow \{w'\}$ by definition, while if $w'' \notin W_w$, then by Lemma 3.9, we have $w \in \{w''\} \rightsquigarrow \emptyset$. Since, by C1$'$, we also have $w \in \{w'\} \rightsquigarrow \{w'\}$, the desired result follows from Lemma 3.11.

To complete the subinduction, suppose $|H| > 2$. If $w \in H \rightsquigarrow \{w'\}$, then by Lemma 3.10, we also have $w \in H \rightsquigarrow \{w', w''\}$ for all $w'' \in H$. Thus, by C6$'$, $w \in \{w', w''\} \rightsquigarrow \{w'\}$ for all $w'' \in H$. By the induction hypothesis, we have $w \in \{w', w''\} \xrightarrow{e} \{w'\}$ for all $w'' \in H$. By C5$'$, we have $w \in H \xrightarrow{e} \{w'\}$. A symmetric argument works for the converse (replacing the roles of $\xrightarrow{e}$ and $\rightsquigarrow$).

Finally, suppose $|E| > 1$. Choose some $w' \in E$. By C9$'$, we have that if $w \in H \xrightarrow{e} E$, then $w \in (H_1 \rightsquigarrow (E - \{w'\})) \cap (H_2 \xrightarrow{e} \{w'\})$ for some H_1, H_2 such that $H_1 \cup H_2 = H$. By the induction hypothesis, $H_1 \xrightarrow{e} E - \{w'\} = H_1 \rightsquigarrow E - \{w'\}$ and $H_2 \xrightarrow{e} \{w'\} = H_2 \rightsquigarrow \{w'\}$. By Lemma 3.11, $H_1 \rightsquigarrow E - \{w'\} \cap H_2 \rightsquigarrow \{w'\} \subseteq H \rightsquigarrow E$. Thus, $w \in H \rightsquigarrow E$. The opposite containment is obtained by a symmetric argument.

Now we must deal with the case that C7$' \in \mathcal{C}$. The argument is similar in spirit to that given in [Friedman and Halpern 1998]. In this case, $\preceq^w$ is not necessarily a total order. However, we can show that $\prec^w$ is *modular*: if $w_1 \prec^w w_2$, then for all $w_3 \in W_w$, either $w_3 \prec^w w_2$ or $w_1 \prec^w w_3$. To see this, suppose $w_1 \prec^w w_2$ and it is not the case that $w_3 \prec^w w_2$. Then

$$w \in \{w_1, w_2\} \rightsquigarrow \{w_1\} \tag{16}$$

and

$$w \notin \{w_2, w_3\} \rightsquigarrow \{w_3\}. \tag{17}$$

Since $w \in \{w_3\} \rightsquigarrow \{w_3\}$, from (16) and Lemma 3.11 it follows that

$$w \in \{w_1, w_2, w_3\} \rightsquigarrow \{w_1, w_3\}. \tag{18}$$

We also must have

$$w \in \{w_1, w_2, w_3\} \rightsquigarrow \{w_1\}, \tag{19}$$

for otherwise, we have $w \in \neg(\{w_1, w_2, w_3\} \rightsquigarrow \neg\neg\{w_1\})$, so by C7$'$ and (18), we would have

$$w \in \{w_2, w_3\} \rightsquigarrow \{w_1, w_3\}. \tag{20}$$

Since we also have $w \in \{w_2, w_3\} \rightsquigarrow \{w_2, w_3\}$, an application of C10$'$ gives us $w \in \{w_2, w_3\} \xrightarrow{e} \{w_3\}$, contradicting (17). From (18), (19), and C6$'$, we immediately get

$$w \in \{w_1, w_3\} \xrightarrow{e} \{w_1\}. \tag{21}$$

We also cannot have $w \in \{w_1, w_3\} \xrightarrow{e} \{w_3\}$, for then by C10$'$ and (21), we would have $w \in \{w_1, w_3\} \xrightarrow{e} \emptyset$. From Lemma 3.10, we then would get $w \in \{w_1, w_3\} \xrightarrow{e} \{w_3\}$, and from C6$'$, that $w \in \{w_3\} \xrightarrow{e} \emptyset$, contradicting the assumption that $w_3 \in W_w$. We can therefore conclude, using (21), that $w_1 \prec^w w_3$, giving us the desired modular order.

Once we have a modular order, we can easily define a total order from it. Define $w' \leq^w w''$ either if $w' \prec^w w''$ or neither $w' \prec^w w''$ nor $w'' \prec^w w'$ hold. It is a standard result (and not hard to show) that $\leq^w$ is a total order if $\prec^w$ is modular. (See [Halpern 1997, Lemma 2.6] for a proof.) Moreover, it is easy to see $w' <^w w''$ iff $w' \prec^w w''$. Thus, $f_{\leq} = f_{\preceq}$, so that if we use $\leq^w$ to define the ternary relation, our previous argument shows that $\xrightarrow{e}$ and $\rightsquigarrow$ still agree. This completes the proof of Theorem 3.8. Although the proof is certainly nontrivial, it is still significantly simpler than the corresponding syntactic proof (see, for example, [Burgess 1981; Friedman and Halpern 1994]). ∎

I conclude with some remarks on the case that W is infinite. In this case, there may not be a "closest" world to w satisfying φ. As a result, the definition of $\rightarrow$ used here based on an ordering gives counterintuitive results. It is perhaps easier to understand this issue in semantic terms, using the selection function $\mathsf{f}_{\preceq}$. For example, let $W^\infty = \{0, 1, 2, \ldots, \infty\}$ and for all i, we have $i+1 \prec^\infty i$, that is, $i+1$ is closer to ∞ than i. Then if $H = W - \{\infty\}$, we would have $\mathsf{f}_{\preceq}(\infty, H) = \emptyset$ and $\infty \in H \xrightarrow{e} \emptyset$ according to the definition above, because there is no world "closest" to ∞ in H. This does not accord with the usual intuitions for $\xrightarrow{e}$.

This particular problem disappears if we require $\preceq^w$ to be *well-founded*, which means that there are no infinitely descending $\preceq^w$ sequences of the form $\ldots w_n \preceq^w w_{n-1} \preceq^w \cdots \preceq^w w_0$, as is essentially done by Kraus, Lehmann, and Magidor [1990]. Lewis gives a definition that seems to give us the properties we want even if $\preceq^w$ is not well-founded. Roughly speaking, his definitions says that $\varphi \rightarrow \psi$ holds at w if all worlds sufficiently close to w that satisfy φ also satisfy ψ. More precisely, Lewis defines $\rightarrow$ as follows. Given a preferential structure $M = (W, R, \pi)$, we have

> $(M, w) \models \varphi \rightarrow \psi$, if for every world $w_1 \in [\![\varphi]\!]_M$, there is a world w_2 such that (a) $w_2 \preceq^w w_1$ (so that w_2 is at least as close to w as w_1), (b) $w_2 \in [\![\varphi \wedge \psi]\!]_M$, and (c) for all worlds $w_3 \prec^w w_2$, we have $w_3 \in [\![\varphi \Rightarrow \psi]\!]_M$ (so any world closer to w than than w_2 that satisfies φ also satisfies ψ).

It is not hard to show that Lewis's definition coincides with that given here if $\prec$ is well-founded (and, in particular, if W is finite). Moreover, with this definition, Theorem 3.6 holds even if W is infinite (except that P3 has to be strengthened to require that $\preceq^w$ be well founded as well as linear in order to get it to correspond to C4). Thus, it seems that by taking Lewis's definition, we get precisely the properties we want.

Unfortunately, as results of [Friedman, Halpern, and Koller 1996] show, appearances here are somewhat deceiving. Lewis's definition is still not appropriate for counterfactual or nonmonotonic reasoning in infinite domains, once we have a rich enough language. In

[Friedman, Halpern, and Koller 1996], the language considered is first-order conditional logic, but the problems can be demonstrated using the set-theoretic approach as well.

We can easily define an operator $\stackrel{e}{\rightarrow}$ that captures Lewis's definition, although it does not correspond to a selection function. For example, C0′ does not hold with this definition in general (even though it does hold if $\stackrel{e}{\rightarrow}$ is defined in terms of a selection function). Consider the domain W^∞ above, and let $H_k = \{k, k+1, k+2, \ldots\}$. Then we have $\infty \in H_0 \stackrel{e}{\rightarrow} H_k$ for all k, but since $\cap_k H_k = \emptyset$, we have $(H_0 \stackrel{e}{\rightarrow} \cap_k H_k) = \emptyset$.

The fact that C0′ does not hold in general is not bad. It is not clear that we want it for counterfactual and nonmonotonic reasoning in the infinite case. For example, consider a lottery. If we think of E_j as corresponding to "player j wins the lottery" and J as being the set of players, then we might well want to have $\cap_{j\in J}(W \stackrel{e}{\rightarrow} \neg E_j)$, which just says that, for each player j in J, normally player j does not win the lottery (giving $\stackrel{e}{\rightarrow}$ a normality reading) and, in addition, $W \stackrel{e}{\rightarrow} (\cup_{jinJ} E_j)$, which says that normally someone wins the lottery. But this is incompatible with C0′.

While Lewis's definition does not force C0′, results of [Friedman, Halpern, and Koller 1996] show that other properties do hold with Lewis's definition that are arguably just as undesirable as C0′ in the infinite case. For example, it is easy to show that the following property holds, for any index set J:

$$((\cup_{j\in J} H_j) \stackrel{e}{\rightarrow} \neg H_1) \cap \cap_{j\in J}((H_1 \cup H_j) \stackrel{e}{\rightarrow} H_1) = \emptyset. \tag{22}$$

(22) encodes a variant of the lottery paradox. Consider a lottery with J players, where player 1 has bought more tickets than any other player. It might then seem reasonable to say that player 1 is more likely to win than any other player, but still unlikely to win. If we think of H_i as corresponding to "player i wins the lottery" and we give $\stackrel{e}{\rightarrow}$ a "typicality" reading, then this is exactly what the left-hand side of (22) says. However, the fact that the right-hand side is the empty set says that this situation cannot happen, according to Lewis's definition.[13]

As shown in [Friedman, Halpern, and Koller 1996], there is another approach that can be used for giving semantics to conditional logic that involves *plausibility measures* [Friedman and Halpern 1998], which works appropriately even for first-order conditional logic (with infinite domains). Of course, this approach too can be captured by a set-theoretic approach, but the details of that would take us well beyond the scope of this paper.

4 Conclusion

The goal of this paper is to show that we can still get the benefit of an axiomatic proof theory even if we work at the semantic level. Indeed, at the semantic level we may

[13]If J is finite, then (22) follows easily from C5′, C8′, and C10′. It would follow from an infinitary version of C5′ if J is infinite, but the infinitary version of C5′ does not follow from the other properties.

get more axioms and easier completeness proofs. This should not be interpreted as an argument to abandon the more traditional, syntactic approach. Syntactic methods have their place, particularly when we do not have one fixed model in mind about which we are reasoning. However, these results are further evidence showing that when we are working with a fixed model, semantic reasoning can be a powerful tool.

Acknowledements I would like to thank Giacomo Bonanno and the two anonymous referees for their insightful comments on the paper.

References

Aumann, R. J. (1989). Notes on interactive epistemology. Cowles Foundation for Research in Economics working paper.

Bell, J. L. and A. B. Slomson (1974). *Models and Ultraproducts: An Introduction (third revised printing)*. Amsterdam/New York: North-Holland/American Elsevier.

Boutilier, C. (1994). Conditional logics of normality: a modal approach. *Artificial Intelligence 68*, 87–154.

Burgess, J. (1981). Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic 22*, 76–84.

Clarke, E. M., E. A. Emerson, and A. P. Sistla (1986). Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Trans. on Programming Languages and Systems 8*(2), 244–263. An early version appeared in *Proc. 10th ACM Symposium on Principles of Programming Languages*, 1983.

Fagin, R., J. Y. Halpern, and N. Megiddo (1990). A logic for reasoning about probabilities. *Information and Computation 87*(1/2), 78–128.

Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi (1995). *Reasoning about Knowledge*. Cambridge, Mass.: MIT Press.

Friedman, N. and J. Y. Halpern (1994). On the complexity of conditional logics. In *Principles of Knowledge Representation and Reasoning: Proc. Fourth International Conference (KR '94)*, pp. 202–213. San Francisco, Calif.: Morgan Kaufmann.

Friedman, N. and J. Y. Halpern (1998). Plausibility measures and default reasoning. *Journal of the ACM*. Accepted for publication. Also available at http://www.cs.berkeley.edu/~nir.

Friedman, N., J. Y. Halpern, and D. Koller (1996). Conditional first-order logic revisited. In *Proceedings, Thirteenth National Conference on Artificial Intelligence (AAAI '96)*, pp. 1305–1312.

Halpern, J. Y. (1990). An analysis of first-order logics of probability. *Artificial Intelligence 46*, 311–350.

Halpern, J. Y. (1997). Defining relative likelihood in partially-ordered preferential structures. *Journal of A.I. Research 7*, 1–24.

Halpern, J. Y. and M. Y. Vardi (1991). Model checking vs. theorem proving: a manifesto. In V. Lifschitz (Ed.), *Artificial Intelligence and Mathematical Theory of Computation (Papers in Honor of John McCarthy)*, pp. 151–176. San Diego, Calif.: Academic Press.

Hintikka, J. (1962). *Knowledge and Belief.* Ithaca, N.Y.: Cornell University Press.

Kraus, S., D. Lehmann, and M. Magidor (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence 44*, 167–207.

Lewis, D. K. (1973). *Counterfactuals.* Cambridge, Mass.: Harvard University Press.

Milgrom, P. (1981). An axiomatic characterization of common knowledge. *Econometrica 49*(1), 219–222.

Pearl, J. (1995). Causation, action, and counterfactuals. In A. Gammerman (Ed.), *Computational Learning and Probabilistic Reasoning*, pp. 235–255. New York: Wiley.

Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory*, Number 2 in American Philosophical Quarterly monograph series. Blackwell, Oxford. Also appears in *Ifs* (Ed. W. L. Harper, R. C. Stalnaker and G. Pearce), Reidel, Dordrecht, Netherlands, 1981.

Stalnaker, R. C. (1992). Notes on conditional semantics. In Y. Moses (Ed.), *Theoretical Aspects of Reasoning about Knowledge: Proc. Fourth Conference*, pp. 316–328. San Francisco, Calif.: Morgan Kaufmann.

Stalnaker, R. C. and R. Thomason (1970). A semantical analysis of conditional logic. *Theoria 36*, 246–281.