Strong Equivalence for Causal Theories

Hudson Turner *

Computer Science Department
University of Minnesota, Duluth
Duluth, MN 55812 USA
hudson@d.umn.edu

Abstract. Strong equivalence is an important property for nonmonotonic formalisms, allowing safe local changes to a nonmonotonic theory. This paper considers strong equivalence for nonmonotonic causal theories of the kind introduced by McCain and Turner. Causal theories T_1 and T_2 are strongly equivalent if, for every causal theory T, $T_1 \cup T$ and $T_2 \cup T$ are equivalent (that is, have the same causal models). The paper introduces a convenient characterization of this property in terms of so-called SE-models, much as was done previously for answer set programs and default theories. A similar result is provided for the nonmonotonic modal logic UCL. The paper also introduces a reduction from the problem of deciding strong equivalence of two causal theories to the problem of deciding equivalence of two sets of propositional formulas.

1 Introduction

Strong equivalence is an important property for nonmonotonic formalisms; it allows one to safely make local changes to a nonmonotonic theory (without having to consider the whole theory). The generic property can be defined thus: theories P and Q (in some nonmonotonic formalism) are *strongly equivalent* if, for every theory R, theories $P \cup R$ and $Q \cup R$ are equivalent. Such P can be safely replaced with Q, with no need to consider the context in which P occurs.

Lifschitz, Pearce and Valverde [13] introduced the notion of strong equivalence and used Heyting's logic of here-and-there to characterize strong equivalence for logic programs with nested expressions [14] (and, more generally, for equilibrium logic [19]). A closely related characterization in terms of so-called SE-models was introduced in [24]. In that paper, a variant of SE-models was also used to characterize strong equivalence for default logic [21, 7]. In [25], SE-models were used to characterize strong equivalence for the weight constraint programs of Niemelä and Simons [18]. The current paper introduces another variant of SE-models to characterize strong equivalence for nonmonotonic causal theories.

Nonmonotonic causal theories were introduced by McCain and Turner in [17]. The current paper considers the extension of causal theories, introduced in [8], in which the atomic parts of formulas are equalities of a kind that may be found in constraint satisfaction problems. This is convenient when formulas are used to

* Partially supported by NSF Career Grant #0091773.

talk about states of a system. For instance, to describe the location of a person in an apartment, we can use equalities like

$$Loc = Kitchen, Loc = LivingRoom, Loc = Bathroom, Loc = Bedroom.$$

The direct effect of walking to the kitchen can then be described simply by saying that there is a cause for the first of these atomic formulas to be true (from which it will follow that there is a cause for the others to be false). By comparison, in a strictly Boolean language, the relationship between these atomic formulas must be expressed by additional formulas; in a case like this, one must express, so to speak, not only the logical relationship between them (exactly one is true at each moment) but also the causal relationship.

An implementation of causal theories—the Causal Calculator (CCALC)[1]—has been applied to several challenge problems in the theory of commonsense knowledge [12, 11, 1, 5], to wire-routing problems [6], and to formalization of multi-agent computational systems [2, 3]. Lee and Lifschitz [9] have shown how to use causal theories to represent "additive" fluents—fluents with numerical values that may be affected by concurrent actions whose interacting effects must be suitably (often additively) combined. In working out these various applications, our understanding of the causal theories formalism has begun to mature.

The mathematical results of the current paper contribute to this maturation process. In addition to providing a mathematically simple characterization of strong equivalence for causal theories, the paper describes an easy reduction of the problem of deciding strong equivalence of two causal theories to the problem of deciding equivalence of two comparably-sized sets formulas of propositional logic. This makes it straightforward to check strong equivalence of finite causal theories automatically, via standard satisfiability solvers.

The paper also includes a characterization of strong equivalence for the nonmonotonic modal logic UCL [23], which can be understood as an extension of the causal theories formalism. One of the initial motivations for the introduction of UCL was to provide a more adequate semantic foundation for causal theories. UCL is obtained by imposing a simple fixpoint condition on standard S5 modal logic. Thus, standard S5 provides a basis in UCL for the kinds of replacement properties guaranteed by strong equivalence. That is, in UCL one can safely replace any subtheory with an S5-equivalent, without changing the set of causal models. What we will see is that the SE-models for UCL are actually a slight specialization of the S5-models. (So S5-equivalence implies strong equivalence, but the converse does not hold.)

We proceed as follows. Section 2 reviews the class of propositional formulas from which causal theories and UCL will be built. Section 3 reviews the syntax and semantics of causal theories. Section 4 introduces and discusses the SE-model characterization of strong equivalence for causal theories. Section 5 presents the reduction from strong equivalence of causal theories to equivalence of sets of propositional formulas. Section 6 introduces the syntax and semantics of UCL, and shows that causal theories are subsumed by UCL. Section 7 presents the

[1] `http://www.cs.utexas.edu/users/tag/cc/`

SE-model characterization of strong equivalence for UCL. Section 8 consists of concluding remarks.

2 Formulas

Following [8], the class of formulas defined here is similar to the class of propositional formulas, but a little bit more general: we will allow atomic parts of a formula to be equalities of the kind found in constraint satisfaction problems.

A (multi-valued propositional) *signature* is a set σ of symbols called *constants*, along with a nonempty finite set $Dom(c)$ of symbols, disjoint from σ, assigned to each constant c. We call $Dom(c)$ the *domain* of c. An *atom* of a signature σ is an expression of the form

$$c = v$$

("the value of c is v") where $c \in \sigma$ and $v \in Dom(c)$. A *formula* over σ is a propositional combination of atoms.

An *interpretation* of σ is a function that maps every element of σ to an element of its domain. An interpretation I *satisfies* an atom $c = v$ if $I(c) = v$. The satisfaction relation is extended from atoms to arbitrary formulas according to the usual truth tables for the propositional connectives.

As usual, the symbol $\models$ denotes the satisfaction relation. An interpretation *satisfies* a set of formulas if it satisfies each formula in the set. An interpretation that satisfies a formula, or a set of formulas, is said to be one of its *models*. Formulas or sets of formulas are *equivalent* if they have the same models.

A *Boolean* constant is one whose domain is the set $\{\mathbf{f}, \mathbf{t}\}$ of truth values. A *Boolean* signature is one whose constants are Boolean. If c is a Boolean constant, we will sometimes write c as shorthand for the atom $c = \mathbf{t}$. Under this convention, when the syntax and semantics defined above are restricted to Boolean signatures and to formulas that do not contain $\mathbf{f}$, they turn into the usual syntax and semantics of classical propositional formulas.

3 Causal Theories

3.1 Syntax

Begin with a multi-valued propositional signature σ. By a (causal) *rule* we mean an expression of the form

$$F \Leftarrow G$$

("F is caused if G is true") where F and G are formulas over σ, called the *head* and the *body* of the rule. A *causal theory* is a set of causal rules.

3.2 Semantics

For any causal theory T and interpretation I of its signature, let

$$T^I = \{ F : F \Leftarrow G \in T, I \models G \} .$$

We say that I is a *causal model* of T if I is the unique model of T^I.

3.3 Examples, Remarks, an Auxiliary Definition and Two Facts

The following example makes use of the convention that for a Boolean constant c, we can write c as shorthand for the atom $c=\mathbf{t}$.

Example 1. Let T be the causal theory over the Boolean signature $\{p, q\}$ whose rules are

$$\begin{aligned} p &\Leftarrow q \ , \\ q &\Leftarrow q \ , \\ \neg q &\Leftarrow \neg q \ . \end{aligned}$$

Of the four interpretations of this signature, only the interpretation I_1 that maps both p and q to $\mathbf{t}$ is a causal model. Indeed, $T^{I_1} = \{p, q\}$, whose unique model is I_1. Let I_2 be such that $I_2(p) = \mathbf{f}$ and $I_2(q) = \mathbf{t}$. Then $T^{I_2} = \{p, q\}$, which is not satisfied by I_2, which is therefore not a causal model of T. Finally, if I maps q to $\mathbf{f}$, then $T^I = \{\neg q\}$, which has two models. Consequently, such an I cannot be a causal model of T.

For technical convenience, we extend the definition of satisfaction to causal rules and thus to causal theories; to this end, $I \models F \Leftarrow G$ if $I \models G \supset F$.

In defining satisfaction of causal theories thus, we certainly do not mean to suggest that the causal connective $\Leftarrow$ can be identified with the material conditional $\supset$. To the contrary, notice for instance that the latter two rules in Example 1 are satisfied by every interpretation of the signature, and yet each describes a distinct condition under which a distinct state of affairs has a cause.

The following observation is easily verified.

Fact 1 *For any causal theory T and interpretation I, $I \models T$ iff $I \models T^I$.*

The following easy corollary can also be useful.

Fact 2 *For any causal theory T and interpretation I, if I is a causal model of T, then $I \models T$.*

Next we consider an example that makes use of non-Boolean constants.

Example 2. Take $\sigma = \{a, b\}$ with $Dom(a) = \{0, 1\}$ and $Dom(b) = \{0, 1, 2\}$. Let

$$T = \left\{ \begin{array}{c} a=0 \not\equiv b=0 \Leftarrow \top \ , \\ a=1 \Leftarrow a=1 \ , \\ \neg b=1 \Leftarrow b=2 \ , \\ \bot \Leftarrow a=0 \wedge b=1 \end{array} \right\} .$$

For each $j \in \{0, 1\}$ and $k \in \{0, 1, 2\}$, let I_{jk} be the interpretation of σ such that $I_{jk}(a) = j$ and $I_{jk}(b) = k$. Notice that $I_{00} \not\models a=0 \not\equiv b=0 \Leftarrow \top$, so $I_{00} \not\models T$, and we can conclude by Fact 2 that I_{00} is not a causal model of T. For the same reason, I_{11} and I_{12} are not causal models of T. And since $I_{01} \not\models \bot \Leftarrow a=0 \wedge b=1$, Fact 2 also shows that I_{01} is not a causal model of T. For the remaining interpretations, we have

$$T^{I_{02}} = \{\, a{=}0 \not\equiv b{=}0,\ \neg b{=}1 \,\} \ ,$$
$$T^{I_{10}} = \{\, a{=}0 \not\equiv b{=}0,\ a{=}1 \,\} \ .$$

$T^{I_{02}}$ has two models (namely, I_{02} and I_{10}), so I_{02} is not a causal model of T. The unique model of $T^{I_{10}}$ is I_{10}, so I_{10} is a causal model of T.

The definition of causal models reflects the following intuitions (with respect to those features of worlds that are described by the constants in the signature of causal theory T).

- For each interpretation I, T^I is a description of exactly what is caused in worlds like I.
- The "causally explained" worlds are those in which (i) everything that is caused obtains and, moreover, (ii) everything that obtains is caused.

Accordingly, I is a causal model of T iff $I \models T^I$ and, so to speak, everything about I has a cause according to T (that is, no interpretation different from I satisfies T^I). Requirement (i) is not surprising. Requirement (ii) is more interesting, and since it is the key to the mathematically simple fixpoint definition, we go so far as to name it "the principle of universal causation." In practice, it is often accomodated in part by writing rules of the form

$$F \Leftarrow F$$

which can be understood to stipulate that F has a cause whenever F is the case.

Similarly, in applications of causal theories to reasoning about action, the commonsense law of inertia is most typically expressed by rules of the form

$$t{+}1\!:\!c{=}v \Leftarrow t{+}1\!:\!c{=}v \wedge t\!:\!c{=}v$$

which can be understood to stipulate that if c has value v at time t and keeps that value at time $t+1$ then there is a cause for $c{=}v$ at time $t+1$. (Intuitively, the cause is inertia.) As explored elsewhere, such rules provide a robust solution to the frame problem in the context of causal theories.

For further discussion, see [17, 23, 8]. Additional examples can also be found in the application papers cited in the introduction.

4 Strong Equivalence for Causal Theories

Causal theories T_1 and T_2 are *equivalent* if they have the same causal models.

Causal theories T_1 and T_2 are *strongly equivalent* if, for every causal theory T, $T_1 \cup T$ and $T_2 \cup T$ are equivalent.

An *SE-structure* for causal theories over signature σ is simply a pair of interpretations of σ. An SE-structure (I, J) is an *SE-model* of causal theory T if $J \models T$ and $I \models T^J$.

Theorem 1. *Causal theories are strongly equivalent iff they have the same SE-models.*

Hence, strongly equivalent causal theories not only have the same satisfying interpretations but also agree on what is caused with respect to each of them.

It is clear that strong equivalence is maintained under substitution of equivalent formulas in heads and bodies of rules. What is more interesting is determining when one set of rules can be safely replaced with another.

The following result appears as Proposition 4 in [8].

Fact 3 (*i*) *Replacing a rule*

$$F \wedge G \Leftarrow H$$

in a causal theory by the rules

$$F \Leftarrow H,\ G \Leftarrow H$$

does not affect its causal models. (*ii*) *Replacing a rule*

$$F \Leftarrow G \vee H$$

in a causal theory by the rules

$$F \Leftarrow G,\ F \Leftarrow H$$

does not affect its causal models.

These are essentially claims about strong equivalence, and they are easily checked using Theorem 1. For instance, to check part (*i*), consider any SE-structure (I, J). Notice first that $J \models F \wedge G \Leftarrow H$ iff $J \models F \Leftarrow H$ and $J \models G \Leftarrow H$. Then notice that $I \models F \wedge G$ iff $I \models F$ and $I \models G$.

The following result shows that replacement of a formula F by a formula G is safe in the presence of the rule $F \equiv G \Leftarrow \top$.

Proposition 1. *Let T_1 and T_2 be causal theories, with T_2 obtained from T_1 by replacing any number of occurrences of formula F by formula G. Let*

$$E = \{ F \equiv G \Leftarrow \top \}\ .$$

Then $T_1 \cup E$ and $T_2 \cup E$ are strongly equivalent.

This proposition follows easily from Theorem 1, given that the replacement property holds for formulas.[2]

The choice of the rule $F \equiv G \Leftarrow \top$ in the statement of Proposition 1 is optimal in the sense that adding it to T_1 "eliminates" all and only those SE-models (I, J) of T_1 for which either $I \not\models F \equiv G$ or $J \not\models F \equiv G$.

Before presenting the proof of Theorem 1, we consider one more example of its use.

To this end, let us agree to call a causal theory *conjunctive* if the head of each of its rules is a conjunction (possibly empty) of atoms.

---

[2] That is, if F_1 and F_2 are formulas, with F_2 obtained from F_1 by replacing any number of occurrences of formula F by formula G, and $I \models F \equiv G$, then $I \models F_1 \equiv F_2$.

Proposition 2. *Let $c_1 = v_1$ and $c_2 = v_2$ be distinct atoms that can be jointly falsified. (Constants c_1 and c_2 need not be distinct.) The causal theory*

$$\{ c_1 = v_1 \vee c_2 = v_2 \Leftarrow \top \}$$

is not strongly equivalent to any conjunctive causal theory.[3]

Proof. Let T be the causal theory above, and suppose that T_c is a strongly equivalent conjunctive causal theory. (We will derive a contradiction.) Take interpretations I and J such that I satisfies only the first of the atoms in T and J satisfies only the second. (This is possible since the atoms are distinct and each can be falsified.) Both (I, J) and (J, J) are SE-models of T. By Theorem 1, (I, J) and (J, J) are SE-models of T_c. Hence, $I \models T_c^J$ and $J \models T_c^J$. Since each element of T_c^J is a conjunction of atoms, we can conclude that c_1 does not appear in T_c^J; indeed, any atom involving constant c_1 would be falsified by either I or J, since they were chosen to disagree on c_1. Similarly, c_2 cannot appear in T_c^J. And since there are no more constants in the signature, and T_c^J is a satisfiable set of conjunctions of atoms, we conclude that $T_c^J = \emptyset$. Let I_0 be an interpretation that falsifies both $c_1 = v_1$ and $c_2 = v_2$. (By assumption they are jointly falsifiable.) Then (I_0, J) is an SE-model of T_c but not of T, which by Theorem 1 yields the contradiction we seek. $\square$

Now let us consider the proof of Theorem 1, which is quite similar to the proofs of the corresponding strong equivalence theorems from [24, 25] for logic programs with nested expressions, default theories, and weight constraint programs.

The following lemma is easily verified.

Lemma 1. *Causal theories with the same SE-models have the same causal models.*

Proof (of Theorem 1). Assume that causal theories T_1 and T_2 have the same SE-models. We need to show that they are strongly equivalent. So consider an arbitrary causal theory T. We need to show that $T_1 \cup T$ and $T_2 \cup T$ have the same causal models. Since T_1 and T_2 have the same SE-models, so do $T_1 \cup T$ and $T_2 \cup T$, and it follows by Lemma 1 that $T_1 \cup T$ and $T_2 \cup T$ also have the same causal models.

Now assume that T_1 and T_2 have different SE-models. Without loss of generality, assume that (I, J) is an SE-model of T_1 and not of T_2. We need to show that T_1 and T_2 are not strongly equivalent. Consider two cases.

Case 1: $J \not\models T_2$. Take

$$T = \{ c = v \Leftarrow \top \ : \ J(c) = v \} \ .$$

Because (I, J) is an SE-model of T_1, $J \models T_1$, and by Fact 1, $J \models T_1^J$. Meanwhile, it is clear that J is the unique model of T^J. Consequently, J is the unique model

[3] A similar result for logic programs appears in [25].

of $T_1^J \cup T^J = (T_1 \cup T)^J$. That is, J is a causal model of $T_1 \cup T$. On the other hand, since $J \not\models T_2$, $J \not\models T_2 \cup T$, and it follows by Fact 2 that J is not a causal model of $T_2 \cup T$. Hence, in this case, T_1 and T_2 are not strongly equivalent.

Case 2: $J \models T_2$. It follows that $I \not\models T_2^J$, since (I, J) is not an SE-model of T_2. On the other hand, it follows by Fact 1 that $J \models T_2^J$. We may conclude that $I \neq J$. Take

$$T = \{\, \neg c_1 = v_1 \supset c_2 = v_2 \Leftarrow \top \,:\, I(c_1) = v_1,\, J(c_2) = v_2 \,\} \ .$$

Notice that the only models of T^J are I and J. Since $I \not\models T_2^J$ while $J \models T_2^J$, we may conclude that J is the unique model of $T_2^J \cup T^J = (T_2 \cup T)^J$. That is, J is a causal model of $T_2 \cup T$. On the other hand, since (I, J) is an SE model of T_1, $I \models T_1^J$. We may conclude that $I \models T_1^J \cup T^J = (T_1 \cup T)^J$, and since $I \neq J$, it follows that J is not a causal model of $T_1 \cup T$. Hence, in this case too, T_1 and T_2 are not strongly equivalent. $\square$

5 Strong Equivalence of Causal Theories as Equivalence of Sets of Propositional Formulas

As we will show, strong equivalence of two causal theories can be determined by checking the equivalence of two corresponding, comparably-sized sets of propositional formulas. Consequently, if the causal theories are finite, we can decide strong equivalence by deciding the unsatisfiability of a corresponding propositional formula. Of course this can in turn be decided by a standard satisfiability solver, once we reduce the multi-valued propositional formula to a Boolean propositional formula (which is straightforward).

The idea of the encoding is first to make a "copy" of the signature, since we are interested in SE-structures, which are pairs of interpretations.

For any signature σ, let signature $\sigma' = \{c' : c \in \sigma\}$, where c' is unique for each $c \in \sigma$, and σ' is disjoint from σ, with $Dom(c') = Dom(c)$ for each $c \in \sigma$. For any interpretation I of σ, let I' be the interpretation of σ' such that $I'(c') = I(c)$ for all $c \in \sigma$.

Thus, the SE-structures for signature σ are in one-to-one correspondence with the interpretations of $\sigma' \cup \sigma$, each of which can be represented in the form $I' \cup J$ for some SE-structure (I, J).

Of course there is no problem encoding the first SE-model condition, $J \models T$, and all that is needed for the second condition, $I \models T^J$, is to make a "copy" of the head of each rule as in the following.

For any causal theory T over signature σ,

$$se(T) = \{\, G \supset F \wedge F' \,:\, F \Leftarrow G \in T \,\}$$

where F' is the formula obtained from F by replacing each occurrence of each constant $c \in \sigma$ with the corresponding constant $c' \in \sigma'$. So $se(T)$ is a set of formulas over $\sigma' \cup \sigma$.

Lemma 2. *For any causal theory T, an SE-structure (I, J) is an SE-model of T iff $I' \cup J \models se(T)$.*

Proof. Observe that the following two conditions are equivalent.

- $J \models F \Leftarrow G$, and
  if $J \models G$ then $I \models F$.
- $I' \cup J \models G \supset F \wedge F'$.

Consequently, $J \models T$ and $I \models T^J$ iff $I' \cup J \models se(T)$. □

Theorem 2. *Causal theories T_1 and T_2 are strongly equivalent iff $se(T_1)$ and $se(T_2)$ are equivalent.*

Proof. Follows easily from Lemma 2, given that the mapping $(I, J) \mapsto I' \cup J$ is a bijection from the set of SE-structures for σ—the signature of T_1 and T_2—to the set of interpretations of $\sigma' \cup \sigma$—the signature of $se(T_1)$ and $se(T_2)$. □

Notice that if T is finite, so is $se(T)$. Abusing notation slightly, let us understand $se(T)$ to stand for the conjunction of its elements in the case that T is finite. Then we have the following corollary to Theorem 2.

Corollary 1. *Finite causal theories T_1 and T_2 are strongly equivalent iff the formula $se(T_1) \not\equiv se(T_2)$ is unsatisfiable.*

All that remains in order to automate the decision process is to reduce the formula $se(T_1) \not\equiv se(T_2)$ to a corresponding Boolean propositional formula, which is straightforward. (See Appendix A in [8], for instance.)

The encoding above closely resembles previous encodings [20, 15] for reducing strong equivalence of logic programs to equivalence in classical propositional logic. It is also related to the reduction in [25] of strong equivalence of two weight constraint programs to inconsistency of a single corresponding weight constraint program.

6 UCL

UCL is a modal nonmonotonic logic obtained from standard S5 modal logic by imposing a simple fixpoint condition that reflects the "principle of universal causation" (discussed in Sect. 3 in connection with causal theories). In [23], UCL was defined not only in the (Boolean) propositional case, but also for nonpropositional languages that include first and second-order quantifiers. In the current paper, we consider a different extension of (Boolean) propositional UCL, built from the multi-valued propositional formulas defined in Sect. 2.

The fundamental distinction in UCL—between propositions that have a cause and propositions that (merely) obtain—is expressed by means of the modal operator C, read as "caused." For example, one can write

$$G \supset \mathsf{C}F \tag{1}$$

to say that F is caused whenever G obtains. If we assume that F and G are formulas of the kind defined in Sect. 2, then formula (1) corresponds to the causal rule $F \Leftarrow G$. This claim is made precise in Theorem 3 below.

6.1 Syntax

UCL formulas are obtained by extending the (implicit) recursive definition of formulas from Sect. 2 with an additional case for the modal operator C, in the usual way for modal logic:

- If F is a UCL formula, then so is $\mathsf{C}F$.

A *UCL theory* is a set of UCL formulas.

6.2 Semantics

An *S5-structure* is a pair (I, S) such that I is an interpretation and S is a set of interpretations (all of the same signature) to which I belongs. *Satisfaction* of a UCL formula by an S5-structure is defined by the standard recursions over the propositional connectives, plus the following two conditions:

- if p is an atom, $(I, S) \models p$ iff $I \models p$,
- $(I, S) \models \mathsf{C}F$ iff, for all $J \in S$, $(J, S) \models F$.

We call the S5-structures that satisfy a UCL formula its *S5-models*. UCL formulas are *S5-equivalent* if they have the same S5-models. We extend these notions to UCL theories (that is, sets of UCL formulas) in the usual way.

For a UCL theory T, if $(I, S) \models T$, we say that (I, S) is an *I-model* of T, thus emphasizing the distinguished interpretation I.

We say that I is *causal model* of T if $(I, \{I\})$ is the unique I-model of T.

6.3 UCL Subsumes Causal Theories

Theorem 3. *For any causal theory T, the causal models of T are precisely the causal models of the corresponding UCL theory*

$$\{ G \supset \mathsf{C}F \; : \; F \Leftarrow G \in T \} \; .$$

The proof presented here is essentially the same as the proof given in [23] for the (Boolean) propositional case, despite the fact that the current paper uses a different, more general definition of a propositional formula.

We begin the proof with a lemma that follows easily from the definitions.

Lemma 3. *For every causal rule $F \Leftarrow G$ and S5-structure (I, S), the following two conditions are equivalent.*

- $(I, S) \models G \supset \mathsf{C}F$.
- *If $I \models G$, then, for all $J \in S$, $J \models F$.*

Proof (of Theorem 3). Let T' be the UCL theory corresponding to causal theory T.

Assume that I is the unique model of T^I. By Lemma 3, $(I, \{I\}) \models T'$. Let S be a superset of $\{I\}$ such that $(I, S) \models T'$. By Lemma 3, for all $J \in S$, $J \models T^I$. It follows that $S = \{I\}$, so $(I, \{I\})$ is the unique I-model of T'.

Assume that $(I, \{I\})$ is the unique I-model of T'. By Lemma 3, $I \models T^I$. Assume that $J \models T^I$. By Lemma 3, $(I, \{I, J\}) \models T$. It follows that $I = J$, so I is the unique model of T^I. $\square$

A similar result in [23] shows that UCL with first and second-order quantification subsumes the nonpropositional causal theories of [10], which in turn subsume the (Boolean) propositional causal theories of [17].

As mentioned previously, one of the original motivations for UCL was to provide a more adequate semantic foundation for causal theories. Another was to unify two closely related approaches to action representation, one based on causal theories and the other based essentially on default logic [16, 22]. See [23] for more on UCL, including its close relationship to disjunctive default logic [7] (an extension of Reiter's default logic).

7 Strong Equivalence for UCL

It is clear that S5-equivalence implies strong equivalence for UCL theories, but the converse does not hold. For example, let $T = \{\neg \mathsf{C}p, \neg \mathsf{C}\neg p\}$. No superset of T has a causal model, so all supersets of T are strongly equivalent, yet T is not S5-equivalent to $T \cup \{p\}$, for instance.

To capture strong equivalence, we slightly strengthen the notion of S5-model.

An S5-model (I, S) of UCL theory T is an *SE-model* of T if $(I, \{I\}) \models T$.

That is, (I, S) is an SE-model iff both $(I, \{I\})$ and (I, S) are S5-models. Notice, for instance, that the example UCL theory T above has no SE-models, which is consistent with the observation that no superset of T has a causal model.

Theorem 4. *UCL theories are strongly equivalent iff they have the same SE-models.*

The proof of this theorem is similar to the proofs of SE-model characterizations of strong equivalence for other nonmonotonic formalisms. In particular, it is very much like the earlier proof of Theorem 1 (for causal theories), and so is given a sketchier presentation.

We begin with the usual easily verified lemma.

Lemma 4. *UCL theories with the same SE-models have the same causal models.*

Proof (of Theorem 4). The right-to-left direction is just as in the proof of Theorem 1, except that now we are interested in UCL theories (instead of causal theories) and we use Lemma 4 (instead of Lemma 1).

For the other direction, assume without loss of generality that (I, S) is an SE-model of T_1 but not of T_2.

Case 1: $(I, \{I\}) \not\models T_2$. Take

$$T = \{\, \mathsf{C}c{=}v \;:\; I(c) = v \,\}$$

and the rest is easy. (See the corresponding case in the proof of Theorem 1).

Case 2: $(I, \{I\}) \models T_2$. So $(I, S) \not\models T_2$, and thus $S \neq \{I\}$. Take

$$T = \{ F \supset \mathsf{C}c{=}v \,:\, F \in T_2,\, (I,S) \not\models F,\, I(c) = v \} \ .$$

Notice first that $(I, S) \models T_1 \cup T$, and since $S \neq \{I\}$, I is not a causal model of $T_1 \cup T$. Notice next that $(I, \{I\}) \models T_2 \cup T$. Assume that $(I, S') \models T_2 \cup T$. Since $(I, S) \not\models T_2$ and $(I, S') \models T_2$, there is an $F \in T_2$ such that $(I, S) \not\models F$ and $(I, S') \models F$. And since $(I, S') \models T$, we may conclude that $S' = \{I\}$. So I is a causal model of $T_2 \cup T$. □

8 Concluding Remarks

This paper introduces mathematically simple characterizations of strong equivalence for causal theories and for the nonmonotonic modal logic UCL. (It also shows that even when multi-valued propositional atoms are incorporated, UCL subsumes causal theories, as expected.)

One of the original motivations for UCL was to provide a more adequate semantic foundation for causal theories, thus securing replacement properties like those guaranteed by strong equivalence. The results of this paper imply that, from this point of view, (propositional) UCL was quite successful. Indeed, from Lemma 3 it follows that causal theories have the same SE-models iff the corresponding UCL theories have the same S5-models. (To see this, let T' be the UCL theory corresponding to causal theory T, and use Lemma 3 to verify that $(J, S) \models T'$ iff, for all $I \in S$, (I, J) is an SE-model of T.) Consequently, S5 modal logic, on which the fixpoint semantics of UCL is based, is perfectly adequate for characterizing strong equivalence of causal theories. [4]

At any rate, it is nice to have the simpler SE-models used to characterize strong equivalence of causal theories in Theorem 1. Moreover, they lead naturally to the result of Theorem 2, which reduces strong equivalence of causal theories to a question of equivalence of comparably-sized sets of propositional formulas.

References

[1] Varol Akman, Selim Erdoğan, Joohyung Lee, Vladimir Lifschitz, and Hudson Turner. Representing the Zoo World and the Traffic World in the language of the Causal Calculator. *Artificial Intelligence*, 2003. To appear.

[2] A. Artikis, M. Sergot, and J. Pitt. An executable specification of an argumentation protocol. In *Proc. of Artificial Intelligence and Law (ICAIL)*, pages 1–11, 2003.

[3] A. Artikis, M. Sergot, and J. Pitt. Specifying electronic societies with the Causal Calculator. In *Proc. of Workshop on Agent-Oriented Software III (AOSE)*. Springer LNCS 2585, pages 1–15, 2003.

[4] Alex Bochman. A logic for causal reasoning. In *Proc. IJCAI'03*, pages 141–146, 2003.

[4] Bochman [4] independently establishes this fact (along with many others), expressed rather differently (in terms of smallest sets of rules closed under an inference system).

[5] J. Campbell and V. Lifschitz. Reinforcing a claim in commonsense reasoning. In *Logical Formalizations of Commonsense Reasoning: Papers from 2003 AAAI Spring Symposium*, pages 51–56, 2003.

[6] Esra Erdem, Vladimir Lifschitz, and Martin Wong. Wire routing and satisfiability planning. In *Proc. CL-2000*, pages 822–836, 2000.

[7] Michael Gelfond, Vladimir Lifschitz, Halina Przymusińska, and Miroslaw Truszczyński. Disjunctive defaults. In *Proc. KR'91*, pages 230–237, 1991.

[8] Enrico Giunchiglia, Joohyung Lee, Vladimir Lifschitz, Norman McCain, and Hudson Turner. Nonmonotonic causal theories. *Artificial Intelligence*, 2003. To appear.

[9] Joohyung Lee and Vladimir Lifschitz. Describing additive fluents in action language C+. In *Proc. IJCAI'03*, pages 1079–1084, 2003.

[10] Vladimir Lifschitz. On the logic of causal explanation. *Artificial Intelligence*, 96:451–465, 1997.

[11] Vladimir Lifschitz. Missionaries and cannibals in the Causal Calculator. In *Proc. KR'00*, pages 85–96, 2000.

[12] Vladimir Lifschitz, Norman McCain, Emilio Remolina, and Armando Tacchella. Getting to the airport: the oldest planning problem in AI. In Jack Minker, editor, *Logic-Based Artificial Intelligence*, pages 147–165. Kluwer, 2000.

[13] Vladimir Lifschitz, David Pearce, and Agustín Valverde. Strongly equivalent logic programs. *ACM Transactions on Computational Logic*, 2:526–541, 2001.

[14] Vladimir Lifschitz, L.R. Tang, and Hudson Turner. Nested expressions in logic programs. *Annals of Mathematics and Artificial Intelligence*, 25(2–3):369–390, 1999.

[15] Fangzhen Lin. Reducing strong equivalence of logic programs to entailment in classical logic. In *Proc. KR'02*, pages 170–176, 2002.

[16] Norman McCain and Hudson Turner. A causal theory of ramifications and qualifications. In *Proc. of IJCAI-95*, pages 1978–1984, 1995.

[17] Norman McCain and Hudson Turner. Causal theories of action and change. In *Proc. AAAI-97*, pages 460–465, 1997.

[18] Ilkka Niemelä and Patrik Simons. Extending the SMODELS system with cardinality and weight constraints. In Jack Minker, editor, *Logic-Based Artificial Intelligence*, pages 491–521. Kluwer, 2000.

[19] David Pearce A new logical characterization of stable models and answer sets. In *Non-Monotonic Extensions of Logic Programming (Lecture Notes in Artificial Intelligence 1216)*, pages 57–70, 1997.

[20] David Pearce, Hans Tompits, and Stefan Woltran. Encodings for equilibrium logic and logic programs with nested expressions. In *Proc. of the 10th Portuguese Conf. on AI (Lecture Notes in Artificial Intelligence 2258)*, pages 306–320, 2001.

[21] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1,2):81–132, 1980.

[22] Hudson Turner. Representing actions in logic programs and default theories: A situation calculus approach. *Journal of Logic Programming*, 31(1–3):245–298, 1997.

[23] Hudson Turner. A logic of universal causation. *Artificial Intelligence*, 113:87–123, 1999.

[24] Hudson Turner. Strong equivalence for logic programs and default theories (made easy). In *Logic Programming and Nonmonotonic Reasoning: Proc. of Sixth Int'l Conf. (Lecture Notes in Artificial Intelligence 2173)*, pages 81–92, 2001.

[25] Hudson Turner. Strong equivalence made easy: Nested expressions and weight constraints. *Theory and Practice of Logic Programming*, 3(4&5):609–622, 2003.