



HAL
open science

A Rule-Based Modal View of Causal Reasoning

Emiliano Lorini

► **To cite this version:**

Emiliano Lorini. A Rule-Based Modal View of Causal Reasoning. 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023), Aug 2023, Macau, China. pp.3286-3295, 10.24963/ijcai.2023/366 . hal-04308847

HAL Id: hal-04308847

<https://hal.science/hal-04308847>

Submitted on 27 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Rule-Based Modal View of Causal Reasoning

Emiliano Lorini

IRIT, CNRS, Toulouse University, France

Emiliano.Lorini@irit.fr

Abstract

We present a novel rule-based semantics for causal reasoning as well as a number of modal languages interpreted over it. They enable us to represent some fundamental concepts in the theory of causality including causal necessity and possibility, interventionist conditionals and Lewisian conditionals. We provide complexity results for the satisfiability checking and model checking problem for these modal languages. Moreover, we study the relationship between our rule-based semantics and the structural equation modeling (SEM) approach to causal reasoning, as well as between our rule-based semantics for causal conditionals and the standard semantics for belief base change.

1 Introduction

Causal reasoning is a central topic for AI nowadays given the importance of causal concepts such as (causal) explanation and responsibility for AI applications. It is an area of research at the crossroad of different disciplines ranging from logic and philosophy to economics.

Two approaches to the formalization of causal reasoning should be distinguished. On the one hand, we have the *rule-based* (or syntactic) approach whereby causal laws are seen as rules expressed in a given logical language, such as propositional logic. This approach grounds on and is tightly interconnected with earlier work on non-monotonic reasoning in the area of knowledge representation [Geffner, 1990; Lifschitz, 1997; McCain and Turner, 1997; Bochman, 2003; Giunchiglia *et al.*, 2004]. The recent work by Bochman, culminated in [Bochman, 2021], is representative of this approach. It relates the rule-based approach to the theory of counterfactuals [Bochman, 2018b] and actual causality [Bochman, 2018a].

On the other hand, we have the *structural equation modeling* (SEM) approach whereby the causal connections between variables are expressed via a system of structural equations. Pearl's work on causality [Pearl, 2009] is probably the most famous example. The SEM approach has been successfully applied to formalizing a wide variety of concepts relevant to AI including actual causality [Halpern, 2008; Halpern, 2016; Beckers and Vennekens, 2017; Beckers, 2021], explanation

[Woodward, 2003; Woodward and Hitchcock, 2003; Halpern and Pearl, 2005b], responsibility and blame [Chockler and Halpern, 2004; Alechina *et al.*, 2017; Halpern and Kleiman-Weiner, 2018], discrimination [Chockler and Halpern, 2022]. The relationship between the rule-based and the SEM approach was explored in [Bochman and Lifschitz, 2015], where it was shown that structural equation models are representable in the causal calculus introduced in [Bochman, 2003; Giunchiglia *et al.*, 2004]. In the general SEM setting, for every endogenous variable and for every value assignment for the other variables (endogenous and exogenous ones), we must specify the value that the variable will take.

In the present work, we introduce a novel rule-based semantics for a number of modal languages that support reasoning about causality. It relies on the notion of causal base, i.e., a set of explicit causal information in propositional form. Our proposal is related to some recent work in the area of epistemic logic devoted to comparing the extensional semantics of epistemic logic based on Kripke models with a succinct semantics using belief bases [Lorini, 2018; Lorini, 2019; Lorini, 2020]. Here we make an analogous operation: we compare the SEM semantics for modal languages of causality with a succinct semantics using causal bases.

The general purpose of this paper is twofold. First of all, we want to show that, when variables are assumed to be Boolean, causal reasoning can be performed via classical deductive reasoning within a standard modal framework in which causal information is expressed in propositional form. We firmly believe that Boolean (or binary) causal models are interesting *per se* since they cover, or at least are closely connected with, a variety of interesting models and concepts including propositional opinion diffusion [Grandi *et al.*, 2015; Christoff and Grossi, 2017], binary neural networks and diagrams [Lewis, 1986; Hubara *et al.*, 2016; Narodytska *et al.*, 2018; Shi *et al.*, 2020], Boolean networks [Kauffman, 1969]. Secondly, we want to shed light on the intimate connection between causal reasoning and epistemic reasoning that, we believe, is difficult to grasp using the SEM approach. Our rule-based semantics, which leverages the concept of causal base analogous to that of belief base, makes this task easier.

Outline. The paper is organized as follows. In Section 2, we discuss related work in more detail. In Section 3, we introduce our rule-based semantics as well as a modal language for reasoning about the notions of causal necessity and

possibility interpreted over it. We provide complexity results for both satisfiability and model checking for this language. Section 4 is devoted to exploring the relationship between our rule-based semantics and the SEM semantics. We show that the latter corresponds to a specific instance of the former where causal information is represented in an equational form. In Section 5, we extend the basic framework by interventionist conditional operators in Halpern & Pearl’s style [Halpern, 2000; Halpern and Pearl, 2005a]. We introduce a novel semantics for interventionist conditionals relying on the concept of replacement and again compare it with the SEM semantics. Our semantics helps to better understand the relationship between intervention and belief change. Intervention is seen as a replacement operation, a kind of operation originally introduced in the theory of belief change to replace one sentence by another in a belief base [Hansson, 2009]. Finally, in Section 6, we present a second extension with Lewisian conditional modalities [Lewis, 1973]. We interpret them by means of a notion of comparative similarity that we compute from the causal base. We compare the Lewisian conditional modality with the interventionist modality and spell the conditions under which the latter can be seen as a special case of the former. Moreover, we elucidate the relationship between our rule-based semantics for counterfactual conditionals and the update semantics of belief base contraction and revision.

2 Related Work

Structural equation models have been exploited as a semantics for a number of modal languages of interventionist conditionals [Galles and Pearl, 1998; Halpern, 2000; Halpern and Pearl, 2005a]. Both the axiomatic and the complexity aspects of these modal languages have been investigated [Galles and Pearl, 1998; Halpern, 2000] as well as their connection with the logic of counterfactual conditionals [Halpern, 2013; Zhang, 2013]. The SEM semantics used in these works is more general than our ruled-based semantics, since variables are not necessarily Boolean. However, it has the disadvantage of not being succinct, thereby making model checking for these modal languages of causality not exploitable in practice. Models used there are huge: every endogenous variable is associated to a function that fully describes how the value of the variable varies depending on the values of the other (endogenous and exogenous) variables. This description of the causal system is exponential in the number of variables. For instance, in the Boolean case, if there are n endogenous variables and m exogenous ones, a table with $n \times 2^{(n+m)-1}$ entries (i.e., $2^{(n+m)-1}$ entries for each endogenous variable) is needed. Moreover, the description of the causal system becomes infinite if there is an infinite number of variables. Our rule-based semantics does not have this limitation. As we will shown in Section 4, we can fully describe a causal system with a set of equational formulas of polynomial size, even in the infinite variable case. This is appealing from the point of view of formal verification.

It is also important to underline some differences between our approach to causal reasoning and Bochman’s approach cited above. Ours is a classic work in modal logic. At the syntactic level, we start with a basic modal language for rea-

soning about causal necessity. We borrow from [Burks, 1951] the idea of modeling causal necessity by means of a normal modality $S5$. Then, we extend it with dyadic modal operators for interventionist and Lewisian conditional. At the semantic level, we use accessibility relations over possible states/worlds for interpreting the different modal operators. Bochman’s language presented in [Bochman, 2018b] is not a modal language in the usual modal logic sense. Namely, it is not an extension of the propositional language with different modal operators, in which we can write any Boolean combination of propositional formulas and modal formulas as well as formulas with nested modalities. His semantics does not use accessibility relations over possible worlds/states. His language consists of causal rules of the form $A \Rightarrow B$, where A and B are propositional formulas. He uses a nonmonotonic semantics based on the notion of *exact model* for a set of causal rules, that is, a consistent set of information that is closed under the causal rules and in which every information is explained by other information in the set. Given the significant differences between the two approaches, we do not know whether and how our concept of causal necessity, and the satisfiability and model checking problem defined in our framework can be “translated” into his framework.

3 Logical Framework

In this section, we introduce a simple modal language for reasoning about causality. Following [Lorini, 2020], the semantics of our language uses the notion of “state” consisting of (i) a causal base, including all causal information which regulates the physical world, and (ii) a valuation of propositional facts that is compatible with the causal information. At the syntactic level, our language extends propositional logic by an operator for representing explicit causal information and another operator for representing causal necessity of facts.

3.1 Semantics

Assume a countably infinite set of atomic propositions $\mathbb{P} = \{p, q, \dots\}$ and let $\mathcal{L}_{\text{PROP}}(\mathbb{P})$ be the propositional language built from \mathbb{P} . We assume the language $\mathcal{L}_{\text{PROP}}(\mathbb{P})$ contains the symbol \top (“true”) as primitive, while \perp (“false”) is defined from it as usual $\perp =_{\text{def}} \neg \top$.

Given a propositional formula $\omega \in \mathcal{L}_{\text{PROP}}(\mathbb{P})$, the set of symbols from \mathbb{P} occurring in ω is noted $\mathbb{P}(\omega)$. The definition extends to a set of propositional formulas $X \subseteq \mathcal{L}_{\text{PROP}}(\mathbb{P})$ in a straightforward manner: $\mathbb{P}(X) = \bigcup_{\omega \in X} \mathbb{P}(\omega)$. The main constituent of our semantics is the following notion of state.

Definition 1 (State). *A state is a pair $S = (C, V)$ where $C \subseteq \mathcal{L}_{\text{PROP}}(\mathbb{P})$ is a causal base and $V \subseteq \mathbb{P}$ is a propositional valuation such that*

$$\forall \omega \in C, V \models \omega,$$

with $V \models \omega$ meaning that the propositional formula ω is true under the valuation V and defined inductively, as usual.

The set of all states is noted \mathbf{S} . A state $S = (C, V)$ is said to be finite if both C and V are finite. The set of finite states is noted \mathbf{S}_{Fin} .

The valuation V represents the actual environment and is assumed to be compatible with the information of the actual

causal base. That is, if ω is included in the actual causal base (i.e., $\omega \in C$) then it should be true in the actual environment (i.e., $V \models \omega$).

We use the generic term “causal information” to indicate propositional formulas in a causal base. We use the more specific term “causal rule/law” for information in a causal base expressed through single implication $\omega_1 \rightarrow \omega_2$ or double implication $\omega_1 \leftrightarrow \omega_2$. With a slight abuse of terminology, we call our semantics a rule-based semantics. To be more precise, it should be called a semantics based on causal information in propositional form. We stick to the former naming since it is more concise. The following definition introduces the notion of causal compatibility.

Definition 2 (Causal compatibility). *We define \equiv_c to be the binary relation on the set \mathbf{S} such that, for every $S = (C, V), S' = (C', V') \in \mathbf{S}$:*

$$S \equiv_c S' \text{ if and only if } C = C'.$$

$S \equiv_c S'$ means that state S and state S' are causally compatible since they share the same causal information. It is easy to verify that \equiv_c is an equivalence relation and that the following holds for every $S = (C, V), S' = (C', V') \in \mathbf{S}$:

$$\text{if } S \equiv_c S' \text{ then } \forall \omega \in C, V' \models \omega. \quad (1)$$

The latter means that the causal information of the actual state must be true at all states that are causally compatible with it. To see this, suppose $S \equiv_c S'$. Thus, by Definition 2, $C \subseteq C'$. Moreover, by definition of \mathbf{S} , $\forall \omega \in C'$, we have $V' \models \omega$. Hence, since $C \subseteq C'$, $\forall \omega \in C, V' \models \omega$.

3.2 Language

In this section, we define a two-layer modal language for talking about explicit causal information and causal necessity. It is defined by the following grammar:

$$\begin{aligned} \mathcal{L}_{CI}(\mathbb{P}) &\stackrel{\text{def}}{=} \alpha ::= p \mid \top \mid \neg\alpha \mid \alpha \wedge \alpha \mid \Delta\omega, \\ \mathcal{L}_{CIN}(\mathbb{P}) &\stackrel{\text{def}}{=} \varphi ::= \alpha \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi, \end{aligned}$$

where p ranges over \mathbb{P} and ω ranges over $\mathcal{L}_{PROP}(\mathbb{P})$. The other Boolean connectives \vee, \rightarrow and \leftrightarrow are defined in the usual way. The first layer language $\mathcal{L}_{CI}(\mathbb{P})$ (Language for Causal Information) is the language for talking about explicit causal information. The formula $\Delta\omega$ has to be read “the information ω is causally relevant” or “information α is part of the causal description of the actual state”. The second layer language $\mathcal{L}_{CIN}(\mathbb{P})$ (Language for Causal Information and Necessity) extends the first layer by a modal operator \Box for causal necessity in Burks’ style [Burks, 1951]. In particular the formula $\Box\varphi$ has to be read “it is causally necessary that φ ” or, by adopting Burks’ reading, “ φ is true on causal grounds”. As usual, we define $\Diamond\varphi =_{\text{def}} \neg\Box\neg\varphi$, where the formula $\Diamond\varphi$ has to be read “it is causally possible that φ ”.

The language $\mathcal{L}_{CN}(\mathbb{P})$ (Language for Causal Necessity) is the fragment of $\mathcal{L}_{CIN}(\mathbb{P})$ in which we can only talk about causal necessity:

$$\mathcal{L}_{CN}(\mathbb{P}) \stackrel{\text{def}}{=} \varphi ::= p \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi,$$

where p ranges over \mathbb{P} . $\mathcal{L}_{PCN}(\mathbb{P})$ (Language for Propositional Causal Necessity) is the fragment of $\mathcal{L}_{CN}(\mathbb{P})$ in which we can

only talk about the causal necessity of propositional facts. It is defined by the following grammar:

$$\mathcal{L}_{PCN}(\mathbb{P}) \stackrel{\text{def}}{=} \varphi ::= p \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\omega,$$

where p ranges over \mathbb{P} and ω ranges over $\mathcal{L}_{PROP}(\mathbb{P})$.

To sum up, we consider four languages $\mathcal{L}_{CI}(\mathbb{P})$, $\mathcal{L}_{CIN}(\mathbb{P})$, $\mathcal{L}_{CN}(\mathbb{P})$ and $\mathcal{L}_{PCN}(\mathbb{P})$ with the following inclusion relations:

$$\mathcal{L}_{CI}(\mathbb{P}) \subset \mathcal{L}_{CIN}(\mathbb{P}) \text{ and } \mathcal{L}_{PCN}(\mathbb{P}) \subset \mathcal{L}_{CN}(\mathbb{P}) \subset \mathcal{L}_{CIN}(\mathbb{P}).$$

$\mathcal{L}_{CIN}(\mathbb{P})$ is the most general one. We interpret its formulas relative to a set of states U and a specific state S included in it. We call U *context* (or *universe*) of interpretation. (We omit semantic interpretations for the Boolean connectives \neg, \wedge and for \top since they are defined in the usual way.)

Definition 3 (Semantic interpretation). *Let $U \subseteq \mathbf{S}$ and $S = (C, V) \in U$. Then:*

$$\begin{aligned} (S, U) \models p &\iff p \in V, \\ (S, U) \models \Delta\omega &\iff \omega \in C, \\ (S, U) \models \Box\varphi &\iff \forall S' \in U : \text{if } S \equiv_c S' \\ &\text{then } (S', U) \models \varphi. \end{aligned}$$

The causal necessity modality \Box is a so-called S5 modality. According to the previous definition, it is causally necessary that φ iff, φ is true at all states in the context that are causally compatible with the actual state. The modality Δ has a set-theoretic interpretation: the information ω is causally relevant if it is included in the actual causal base. For notational convenience, we write $S \models \varphi$ instead of $(S, \mathbf{S}) \models \varphi$. \mathbf{S} is also called the *global context*.

As highlighted by the following proposition, the causal necessity of a propositional fact coincides with its deducibility from a causal base.

Proposition 1. *Let $S = (C, V) \in \mathbf{S}$ and $\omega \in \mathcal{L}_{PROP}(\mathbb{P})$. Then, $S \models \Box\omega$ iff $\omega \in Cn(C)$, with Cn the classical deductive closure operator over the propositional language $\mathcal{L}_{PROP}(\mathbb{P})$.*

As a simple example to illustrate the semantics, consider the state $S = (\{\omega_1, \omega_2\}, \emptyset)$ with $\omega_1 = (p_1 \wedge \neg p_2) \rightarrow p_3$ and $\omega_2 = p_4 \rightarrow \neg p_2$. We have that the causal information ω_1 and ω_2 are relevant at state S and this is a causal necessity, that is, $S \models (\Delta\omega_1 \wedge \Delta\omega_2) \wedge \Box(\Delta\omega_1 \wedge \Delta\omega_2)$. Furthermore, it is causally necessary that “if p_1 and p_4 then p_3 ”, that is, $S \models \Box((p_1 \wedge p_4) \rightarrow p_3)$.

3.3 Satisfiability and Model Checking

A formula $\varphi \in \mathcal{L}_{CIN}(\mathbb{P})$ is said to be satisfiable if there exists a context $U \subseteq \mathbf{S}$ and a state $S \in U$ such that $(S, U) \models \varphi$.

Satisfiability checking problem

Given: $\varphi \in \mathcal{L}_{CIN}(\mathbb{P})$.

Question: Is φ satisfiable?

We formulate model checking as the problem of verifying whether a formula is true at a given *finite* state.

Model checking problem

Given: $\varphi \in \mathcal{L}_{CIN}(\mathbb{P})$ and $S \in \mathbf{S}_{Fin}$.

Question: Do we have $S \models \varphi$?

The first result of this section highlights that satisfiability checking for the language $\mathcal{L}_{\text{CIN}}(\mathbb{P})$ as well as for its fragments $\mathcal{L}_{\text{CN}}(\mathbb{P})$ and $\mathcal{L}_{\text{PCN}}(\mathbb{P})$ has the same complexity as SAT. The proof of the theorem relies on the fact that for any satisfiable formula we can construct a model of polynomial size which satisfies it.

Theorem 1. *Let $\Lambda \in \{\text{CIN}, \text{CN}, \text{PCN}\}$. Then, satisfiability checking for $\mathcal{L}_\Lambda(\mathbb{P})$ -formulas is NP-complete.*

In our framework model checking has the same complexity as satisfiability checking. This is due to its compact formulation which makes it different from standard model checking for the modal logic S5 over Kripke models which is known to be polynomial with respect to the size of the input formula to be checked and the size of the input model [Grädel and Otto, 1999].

In order to prove NP-hardness of model checking, we use a polynomial reduction of SAT to it. In order to prove its NP-membership, we rely on the construction that we use for proving Theorem 1.

Theorem 2. *Let $\Lambda \in \{\text{CIN}, \text{CN}, \text{PCN}\}$. Then, model checking for \mathcal{L}_Λ -formulas is NP-complete.*

4 Relation with Structural Equation Models

In this section we study the connection between our rule-based semantics for causal reasoning and the SEM semantics. We first define a subclass of states in the sense of Definition 1, called *equational* states, in which an information in a causal base can only be expressed in equational form: a rule specifying the necessary and sufficient condition for making a certain atomic proposition true. Then, we define a variant of structural equation models in which variables are assumed to be binary (viz. Boolean). We call them binary causal models (BCMs). We show that BCMs are essentially equivalent to equational states as far as reasoning about propositional causal necessity is concerned.

4.1 Equational States

Let $\mathcal{L}_{\text{EQ}}(\mathbb{P})$ be the the fragment of the propositional language $\mathcal{L}_{\text{PROP}}(\mathbb{P})$ which contains equational formulas:

$$\mathcal{L}_{\text{EQ}}(\mathbb{P}) = \left\{ p \leftrightarrow \omega : p \in \mathbb{P} \text{ and } \omega \in \mathcal{L}_{\text{PROP}}(\mathbb{P} \setminus \{p\}) \right\}.$$

An equational state is finite state whose causal information is expressed in equational form and which includes at most one equational formula for each variable.

Definition 4 (Equational state). *A state $S = (C, V)$ is said to be in equational form if and only if C is finite, $C \subseteq \mathcal{L}_{\text{EQ}}(\mathbb{P})$, $V \subseteq \mathbb{P}(C)$ and*

$$\forall p \in \mathbb{P}, \forall p \leftrightarrow \omega, p \leftrightarrow \omega' \in C, \omega = \omega'.$$

The set of states in equational form, or equational states, is noted \mathbf{S}_{Eq} .

As a side note, model checking for equational states is also NP-complete. NP-membership is evident because of Theorem 2. The proof of NP-hardness uses the same argument as the proof of NP-hardness for Theorem 2: polynomial reduction of SAT to model checking.

An interesting aspect of equational states is their graphical counterpart. Specifically, given an equational state $S = (C, V)$, we can extract the causal graph $\Gamma_S = (N_S, \mathcal{P}_S)$ with $N_S = \mathbb{P}(C)$ and where the *causal parent* function $\mathcal{P}_S : N \rightarrow 2^N$ is defined as follows, for every $p \in N$:

- (i) $\mathcal{P}_S(p) = \mathbb{P}^+(\omega)$ if $\exists \omega \in \mathcal{L}_{\text{PROP}}(\mathbb{P} \setminus \{p\})$ s.t. $p \leftrightarrow \omega \in C$,
- (ii) $\mathcal{P}_S(p) = \emptyset$ otherwise,

where $\mathbb{P}^+(\omega) = \mathbb{P}(\omega) \cup \{\top\}$ if \top occurs in ω , and $\mathbb{P}^+(\omega) = \mathbb{P}(\omega)$ if \top does not occur in ω . In equational states we can partition the set of variables into two sets: the set of exogenous variables and the set of endogenous ones. Specifically, a variable p is said to be *exogenous* in state S if it has no parents in the corresponding causal graph Γ_S (i.e., $\mathcal{P}_S(p) = \emptyset$), it is said to be *endogenous* in S otherwise. As the following proposition highlights, an endogenous variable behaves functionally, i.e., its truth value is entirely determined by the truth values of the other variables.

Proposition 2. *Let $S = (C, V) \in \mathbf{S}_{\text{Eq}}$, $p \in \mathbb{P}(C)$ and X, X' such that $X \cup X' = (\mathbb{P}(C) \setminus \{p\})$ and $X \cap X' = \emptyset$. Then,*

$$\begin{aligned} \text{if } p \text{ is endogenous in } S \text{ then } S \models & \square(\text{con}_{X, X'} \rightarrow p) \vee \\ & \square(\text{con}_{X, X'} \rightarrow \neg p), \end{aligned}$$

with $\text{con}_{X, X'} =_{\text{def}} \bigwedge_{q \in X} q \wedge \bigwedge_{q \in X'} \neg q$.

Let us focus on an example of social influence in a multi-agent setting in order to illustrate the concept of equational state and its graphical counterpart.

Example 1. *Consider a set of five agents $\text{Agt} = \{\text{Ann}, \text{Bob}, \text{Cath}, \text{Mary}, \text{Ted}\}$. Each agent has an opinion about the use of nuclear energy for electricity production. In particular, she/he can be either in favor of or against it. An agent's opinion may be influenced by and depend on other agents' opinions. In order to represent agents' opinions, we suppose \mathbb{P} includes atomic propositions of type $\text{pro}(x)$ with $x \in \text{Agt}$, $\text{pro}(x)$ meaning that x is favor and $\neg \text{pro}(x)$ meaning that x is against the use of nuclear energy for electricity production. We assume we have perfect information about the dependency relations between the variables. In particular, we know that variables $\text{pro}(\text{Ann})$ and $\text{pro}(\text{Ted})$ are exogenous (i.e., Ann and Ted's opinions are independent of the other agents' opinions), while all other variables are exogenous. The causal connections between the variables are described by the following three equational formulas:*

$$\begin{aligned} \omega_1 &= \text{pro}(\text{Cath}) \leftrightarrow (\text{pro}(\text{Ann}) \wedge \text{pro}(\text{Bob})), \\ \omega_2 &= \text{pro}(\text{Mary}) \leftrightarrow (\text{pro}(\text{Cath}) \rightarrow \text{pro}(\text{Ted})), \\ \omega_3 &= \text{pro}(\text{Bob}) \leftrightarrow \text{pro}(\text{Mary}). \end{aligned}$$

The three formulas express, respectively, that (i) Cath will be in favor of the use of nuclear energy if and only if Ann and Bob are unanimously in favor (ω_1), (ii) Mary will be in favor if and only if Ted is in favor in case Cath is in favor too (ω_2), and (iii) Bob is in favor if and only if Mary is in favor (ω_3). Consider the state $S = (\{\omega_1, \omega_2, \omega_3\}, \{\text{pro}(x) : x \in \text{Agt}\})$. It is straightforward to verify that S is an equational state. The (influence) causal graph induced by it is represented in

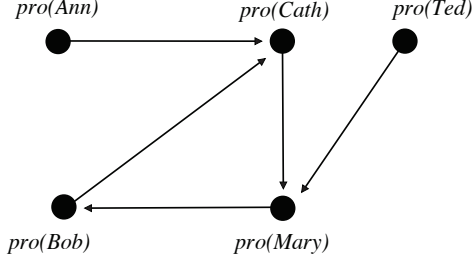


Figure 1: Influence causal graph

Figure 1. Clearly, there exists at least one state which is compatible with the causal information $\omega_1, \omega_2, \omega_3$:

$$S \models \diamond \top.$$

Moreover, at state S it is causally necessary that Bob is in favor of the use of nuclear energy. In other words, Bob has no freedom, he is trapped in the influence process:

$$S \models \Box \text{pro}(\text{Bob}).$$

Moreover, it is not causally possible that Cath is in favor while Ted is against:

$$S \models \neg \diamond (\text{pro}(\text{Cath}) \wedge \neg \text{pro}(\text{Ted})).$$

As the following theorem indicates, for every formula of the language of causal necessity, there exists a mapping from finite states to equational states such that the input state satisfies the formula if and only if the output state satisfies it too.

Theorem 3. Let $\varphi \in \mathcal{L}_{\text{CN}}(\mathbb{P})$. Then, there exists a function $g : \mathbf{S}_{\text{Fin}} \rightarrow \mathbf{S}_{\text{Eq}}$ such that

$$S \models \varphi \text{ iff } g(S) \models \varphi.$$

From Theorem 3 we can conclude that for every formula of the language of causal necessity to be checked relative to a finite state, we can transform the latter into an equational state and check the formula relative to it. In other words, the general model checking problem defined in Section 3.3 can be reduced to model checking relative to equational states.

The theorem relies on the fact that, given a formula φ , we can transform a finite causal base C into a causal base C' containing a single equational formula ω_{eq} such that the set of propositional interpretations for the atoms in $\mathbb{P}(\varphi)$ satisfying all formulas in C is the same as the set of propositional interpretations for the atoms in $\mathbb{P}(\varphi)$ satisfying ω_{eq} . This guarantees that from a finite state $S = (C, V)$ we can build an equational state $S' = (\{\omega_{\text{eq}}\}, V')$ such that $S \models \varphi$ iff $S' \models \varphi$. However, the transformation is exponential. The equational formula ω_{eq} uses the canonical DNF of the Boolean function representing the set of interpretations for the atoms in $\mathbb{P}(C)$ that satisfy all formulas in C . The size of a such a canonical DNF is exponential in the size of $\mathbb{P}(C)$. We believe it is not possible to find a polynomial transformation of a finite state into an equational state that guarantees equisatisfaction of φ .

Note that Theorem 3 cannot be generalized to the more general language $\mathcal{L}_{\text{CIN}}(\mathbb{P})$ since states in \mathbf{S}_{Eq} only allow causal information to be expressed in an equational form.

4.2 Binary Causal Model

Before defining the notion of binary causal model some preliminary notions are needed. A binary assignment for the set of variables $\mathbb{X} \subseteq \mathbb{P}$ is a function

$$\mathcal{I}_{\mathbb{X}} : \mathbb{X} \rightarrow \{0, 1\}.$$

The set of binary assignments for \mathbb{X} is noted $\text{Asg}_{\mathbb{X}}$. Interpretation $|\cdot|$ of propositional formulas in $\mathcal{L}_{\text{PROP}}(\mathbb{X})$ relative to assignments in $\text{Asg}_{\mathbb{X}}$ is as usual: $|p|_{\text{Asg}_{\mathbb{X}}} = \{\mathcal{I}_{\mathbb{X}} \in \text{Asg}_{\mathbb{X}} : \mathcal{I}_{\mathbb{X}}(p) = 1\}$ for $p \in \mathbb{X}$; $|\top|_{\text{Asg}_{\mathbb{X}}} = \text{Asg}_{\mathbb{X}}$; $|\neg \omega|_{\text{Asg}_{\mathbb{X}}} = \text{Asg}_{\mathbb{X}} \setminus |\omega|_{\text{Asg}_{\mathbb{X}}}$; $|\omega \wedge \omega'|_{\text{Asg}_{\mathbb{X}}} = |\omega|_{\text{Asg}_{\mathbb{X}}} \cap |\omega'|_{\text{Asg}_{\mathbb{X}}}$.

Definition 5 (Binary causal model). A binary causal model (BCM) is a tuple $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ where \mathbb{V}_{exo} is a finite set of exogenous variables, \mathbb{V}_{end} is a finite set of endogenous variables such that $\mathbb{V}_{\text{end}}, \mathbb{V}_{\text{exo}} \subseteq \mathbb{P}$ and $\mathbb{V}_{\text{end}} \cap \mathbb{V}_{\text{exo}} = \emptyset$, and, for every $p \in \mathbb{V}_{\text{end}}$:

$$\mathcal{F}_p : \text{Asg}_{\mathbb{V} \setminus \{p\}} \rightarrow \{0, 1\}$$

with $\mathbb{V} = \mathbb{V}_{\text{exo}} \cup \mathbb{V}_{\text{end}}$.

For notational convenience, we sometimes abbreviate $\mathbb{V} \setminus \{p\}$ as \bar{p} . Moreover, when the context is unambiguous, we write Asg instead of $\text{Asg}_{\mathbb{V}}$ and note $\mathcal{I}, \mathcal{I}', \dots$ its elements.

Definition 6 (Solutions of a BCM). Let $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ be a BCM and let $\mathcal{I} \in \text{Asg}$. We say that \mathcal{I} is a solution of Θ if and only if, for every $p \in \mathbb{V}_{\text{end}}$:

$$\mathcal{I}(p) = \mathcal{F}_p(\mathcal{I}|_{\bar{p}}),$$

where $\mathcal{I}|_{\bar{p}}$ is the restriction of function \mathcal{I} to \bar{p} . The set of solutions of Θ is noted $\text{Sol}(\Theta)$.

Definition 7 (Pointed BCM). A pointed BCM is a pair (Θ, \mathcal{I}) with $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ a BCM and $\mathcal{I} \in \text{Sol}(\Theta)$. The class of pointed BCMs is noted \mathbf{B} .

Pointed BCMs provide a natural semantics for interpreting formulas of the language of propositional causal necessity. In particular, given a pointed BCM (Θ, \mathcal{I}) with $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ we can interpret formulas of the language $\mathcal{L}_{\text{PCN}}(\mathbb{V})$, as follows. (Boolean cases and \top are again omitted since they are defined in the standard way.)

$$(\Theta, \mathcal{I}) \models p \iff \mathcal{I}(p) = 1,$$

$$(\Theta, \mathcal{I}) \models \Box \omega \iff \text{Sol}(\Theta) \subseteq |\omega|_{\text{Asg}}.$$

Notice in particular the interpretation of the causal necessity modality \Box : it is causally necessary that ω if ω is true for all solutions of the actual causal model.

The following theorem highlights the tight similarity between equational states and pointed BCMs. Every equational state can be mapped into a pointed BCM and, conversely, every pointed BCM can be mapped into an equational state such that the truth of the formulas in the language of propositional causal necessity is invariant between them.

Theorem 4. There exists a surjection $f : \mathbf{S}_{\text{Eq}} \rightarrow \mathbf{B}$ such that if $f(S) = (\Theta, \mathcal{I})$ with $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ then

$$\forall \varphi \in \mathcal{L}_{\text{PCN}}(\mathbb{V}), S \models \varphi \text{ iff } (\Theta, \mathcal{I}) \models \varphi.$$

Note that Theorem 4 is conceptually stronger than Theorem 3 since the function f works for all formulas in the language $\mathcal{L}_{\text{PCN}}(\mathbb{V})$, while the function g in Theorem 3 only works for a single formula in the language $\mathcal{L}_{\text{CN}}(\mathbb{P})$.

5 Interventionist Conditionals

In this section, we are going to define a semantics for interventions which relies on a replacement operation on causal bases in Hansson's style [Hansson, 2009]. We will show that, when restricting to equational states, it corresponds to the SEM semantics for interventions [Halpern, 2000; Halpern and Pearl, 2005a]. Before introducing the semantics, we first define the syntax of interventionist conditional modalities.

5.1 Syntax

The antecedent of an interventionist conditional is an event, namely, a set of basic interventions on propositional variables of the form $p=1$ (the variable p is set to 'true') or $p=0$ (the variable p is set to 'false') with at most one intervention per variable. The set of basic interventions is defined as follows:

$$Int = \{p=\tau : p \in \mathbb{P} \text{ and } \tau \in \{0, 1\}\},$$

while the set of events is defined as follows:

$$Evt = \{\{p_1=\tau_1, \dots, p_k=\tau_k\} \subset Int\}$$

$$\forall k', k'' \in \{0, \dots, k\}, \text{ if } k' \neq k'' \text{ then } p_{k'} \neq p_{k''}.$$

Elements of Evt are noted E, E', \dots . For notational convenience, given $p=\tau \in Int$, we define its corresponding equational formula as follows: $eq(p=1) = p \leftrightarrow \top$ and $eq(p=0) = p \leftrightarrow \perp$. We extend function eq to sets in the expected way: $eq(\{p_1=\tau_1, \dots, p_k=\tau_k\}) = \{eq(p_1=\tau_1), \dots, eq(p_k=\tau_k)\}$. Moreover, for each $E \in Evt$, we define the partial function val_E with domain \mathbb{P} and codomain $\{0, 1\}$: $val_E(p) = 1$ if $p=1 \in E$, $val_E(p) = 0$ if $p=0 \in E$, and undefined otherwise.

We consider the following extension of the language $\mathcal{L}_{PCN}(\mathbb{P})$ by interventionist conditional operators of the form $[E]$ and their negation:

$$\mathcal{L}_{PCN-Int}(\mathbb{P}) \stackrel{\text{def}}{=} \gamma ::= \varphi \mid [E]\omega \mid \neg[E]\omega \mid \gamma \wedge \gamma,$$

with φ ranging over $\mathcal{L}_{PCN}(\mathbb{P})$, ω ranging over $\mathcal{L}_{PROP}(\mathbb{P})$ and $E \in Evt$. $\mathcal{L}_{PCN-Int}$ stands for "Language for Propositional Causal Necessity plus Interventions". The new formula $[E]\omega$ has to be read "it will be causally necessary that ω , after the occurrence of the event E " or "if the event E occurred, ω would be necessarily true". As usual, we define $\langle E \rangle \omega =_{\text{def}} \neg[E]\neg\omega$. Note we only allow negation over arbitrary formulas in $\mathcal{L}_{PCN}(\mathbb{P})$ and over the modality $[E]$. Indeed, we are merely interested in talking about causal necessity (i.e., $[E]$) and causal possibility (i.e., $\langle E \rangle$) of propositional facts after intervention. Negation over arbitrary formulas in $\mathcal{L}_{PCN-Int}(\mathbb{P})$ (i.e., $\neg\gamma$) would add unnecessary complications that we prefer to leave aside for the moment.

Note also we could have built the extension with interventionist conditionals on the top of the language $\mathcal{L}_{CIN}(\mathbb{P})$ instead of the more restricted language $\mathcal{L}_{PCN}(\mathbb{P})$. We did not do this since the language $\mathcal{L}_{PCN-Int}(\mathbb{P})$ is sufficient to achieve our goal of comparing the interpretation of the concept of intervention in the rule-based semantics with its interpretation in the SEM semantics. The complexity upper bound for model checking we will provide in the next section generalizes straightforwardly to the more general language for interventionist conditionals built on the top of $\mathcal{L}_{CIN}(\mathbb{P})$.

5.2 Semantics

In order to define interventionist conditionals semantically, some preliminary notions are needed. The effect of an event (a set of interventions) is to add information about the truth values of certain propositional variables to the causal base and to remove from the causal base all equational formulas about those variables. The latter is called the revocable part of the causal base in the light of the event. In other words, the revocable part of a causal base C in the light of an event E , noted $Rev(C, E)$, includes all causal information that is superseded by some intervention in E .

Definition 8 (Revocable part of a causal base). *Let $C \subseteq \mathcal{L}_{PROP}(\mathbb{P})$ and $E \in Evt$. We define*

$$Rev(C, E) = \{p \leftrightarrow \omega \in (C \cap \mathcal{L}_{EQ}(\mathbb{P})) : p=1 \in E \text{ or } p=0 \in E\}.$$

The following definition introduces the notion of causal compatibility post intervention.

Definition 9 (Causal compatibility post intervention). *Let $E \in Evt$. We define \Rightarrow^E to be the binary relation on the set of states \mathbf{S} such that, for every $S = (C, V), S' = (C', V') \in \mathbf{S}$:*

$$S \Rightarrow^E S' \text{ if and only if } C' = (C \setminus Rev(C, E)) \cup eq(E).$$

$S \Rightarrow^E S'$ means that state $S' = (C', V')$ is compatible with state $S = (C, V)$ after the occurrence of event E . According to previous definition, the latter is the case if the causal base C' is the result of the following replacement operation applied to the causal base C : (i) remove from the causal base C all formulas which are revocable in the light of the event E , and then (ii) add to the resulting base the equational formulas corresponding to the interventions in E .

By means of the relation \Rightarrow^E , we can provide a semantic interpretation of the interventionist modalities $[E]$. They, too, are interpreted relative to a set of states (a context) and a specific state S included in it. Let $U \subseteq \mathbf{S}$ and $S \in U$. Then,

$$(S, U) \models [E]\omega \iff \forall S' \in U : \text{if } S \Rightarrow^E S' \text{ then } (S', U) \models \omega.$$

The other constructs of the language $\mathcal{L}_{PCN-Int}(\mathbb{P})$ (i.e., φ , $\neg[E]\omega$ and $\gamma \wedge \gamma$) are interpreted in the usual way. Again, to simplify notation, we write $S \models \gamma$ instead of $(S, \mathbf{S}) \models \gamma$. Note that the interventionist modality with the empty event coincides with the causal necessity operator, i.e., the formula $[\emptyset]\omega \leftrightarrow \Box\omega$ is valid.

The following theorem provides a complexity upper bound of model checking for the language extended with interventionist modalities. Its proof relies on two facts: (i) the formula to be checked is a conjunction of the form $\gamma_1 \wedge \dots \wedge \gamma_m$ where each conjunct is of the form φ , $[E]\omega$ or $\neg[E]\omega$ with $\varphi \in \mathcal{L}_{PCN}(\mathbb{P})$, (ii) model checking the formula $[E]\omega$ relative to a finite state is reducible to model checking the formula $\Box(\bigwedge_{p=\tau \in E} eq(p=\tau) \rightarrow \omega) \in \mathcal{L}_{PCN}(\mathbb{P})$ relative to the finite state obtained from the input state through the removal of all causal information which is revocable in the light of the event E . Thus, thanks to Theorem 2, model checking for $\mathcal{L}_{PCN-Int}(\mathbb{P})$ -formulas is solvable by a poly-time deterministic Turing machine querying an NP-oracle.

Theorem 5. *Model checking for $\mathcal{L}_{\text{PCN-Int}}(\mathbb{P})$ -formulas is in $P^{NP} = \Delta_2^P$.*

Let us go back to our running example.

Example 2 (cont.). *Let us consider again the equational state $S = (\{\omega_1, \omega_2, \omega_3\}, \{\text{pro}(x) : x \in \text{Agt}\})$ as defined in Example 1. We have shown that in this state Bob cannot be against the use of nuclear energy and that Cath cannot be in favor while Ted is against. It is interesting to observe that, by forcing Cath to be in favor of the use of nuclear energy through intervention on her opinion we cut the influence of Ann and Bob's opinions on her opinion. Consequently, (i) it becomes possible for Bob to be against the use of nuclear energy, and (ii) it becomes possible for Cath to be in favor while Ted is against. Formally, we have:*

$$S \models \langle \{\text{pro}(\text{Cath})=1\} \rangle \neg \text{pro}(\text{Bob}) \wedge \langle \{\text{pro}(\text{Cath})=1\} \rangle (\text{pro}(\text{Cath}) \wedge \neg \text{pro}(\text{Ted})).$$

However, the intervention $\text{pro}(\text{Mary})=1$ does not open the possibility for Bob to be against when Ted is in favor:

$$S \not\models \langle \{\text{pro}(\text{Cath})=1\} \rangle (\neg \text{pro}(\text{Bob}) \wedge \text{pro}(\text{Ted})).$$

5.3 Back to Structural Equational Models

Before concluding, it is worthwhile to explore the connection between the rule-based semantics for interventionist conditionals defined in the previous section and the standard SEM semantics. The following definition specifies the way a BCM is updated through an event. Our definition is slightly different from the update semantics used in the SEM literature whereby interventions only apply to endogenous variables, and the partition of the set of variables into exogenous and endogenous ones does not change under intervention. We assume interventions can also apply to exogenous variables thereby transforming them into endogenous ones. The justification is that we take exogenous and endogenous variables as part of the BCM relative to which formulas are evaluated.

Definition 10 (BCM after intervention). *Let $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ be a BCM and $E = \{p_1=\tau_1, \dots, p_k=\tau_k\} \in \text{Evt}$. We define the BCM resulting from the update of Θ by E as the tuple $\Theta^E = (\mathbb{V}_{\text{exo}}^E, \mathbb{V}_{\text{end}}^E, (\mathcal{F}_p^E)_{p \in \mathbb{V}_{\text{end}}^E})$ such that*

$$\begin{aligned} \mathbb{V}_{\text{exo}}^E &= \mathbb{V}_{\text{exo}} \setminus \{p_1, \dots, p_k\}, \\ \mathbb{V}_{\text{end}}^E &= \mathbb{V}_{\text{end}} \cup \{p_1, \dots, p_k\}, \\ \forall p \in (\mathbb{V}_{\text{end}}^E \setminus \{p_1, \dots, p_k\}), \mathcal{F}_p^E &= \mathcal{F}_p, \text{ and} \\ \forall p \in \{p_1, \dots, p_k\}, \text{Range}(\mathcal{F}_p^E) &= \{\text{val}_E(p)\}. \end{aligned}$$

The condition $\text{Range}(\mathcal{F}_p^E) = \{\text{val}_E(p)\}$ just means that $\mathcal{F}_p^E(\mathcal{I}_{\bar{p}}) = \text{val}_E(p)$ for every $\mathcal{I}_{\bar{p}} \in \text{Asg}_{\bar{p}}$.

The modality $[E]$ has the following interpretation relative to a pointed BCM:

$$(\Theta, \mathcal{I}) \models [E]\omega \iff \text{Sol}(\Theta^E) \subseteq |\omega|_{\text{Asg}}.$$

This means that the conditional expression $[E]\omega$ holds if the consequent ω is true for all solutions of the actual BCM updated by the event E .

The following theorem is a generalization of Theorem 4 to the language $\mathcal{L}_{\text{PCN-Int}}(\mathbb{V})$: there exists a truth preserving onto relation between equational states and BCMs for the language extended by interventionist modalities.

Theorem 6. *There exists a surjection $f : \mathbf{S}_{\text{Eq}} \rightarrow \mathbf{B}$ such that if $f(S) = (\Theta, \mathcal{I})$ with $\Theta = (\mathbb{V}_{\text{exo}}, \mathbb{V}_{\text{end}}, (\mathcal{F}_p)_{p \in \mathbb{V}_{\text{end}}})$ then*

$$\forall \varphi \in \mathcal{L}_{\text{PCN-Int}}(\mathbb{V}), S \models \varphi \text{ iff } (\Theta, \mathcal{I}) \models \varphi.$$

6 Lewisian Conditionals

In this section, we introduce a notion of counterfactual conditional interpreted by means of a comparative similarity relation. In the standard analysis of counterfactual conditionals the notion of similarity is given as a primitive [Lewis, 1973] or replaced by an abstract selection function [Stalnaker, 1968] which selects for every world and formula φ the set of most similar φ -worlds to it. In line with Ginsberg's earlier work on the computationally grounded semantics for counterfactuals [Ginsberg, 1986], we compute the comparative similarity relation from causal bases.

6.1 Semantics

A straightforward way to compute the comparative similarity relation between states relative to a causal base is by comparing the causal information of the causal base that is relevant in each state. This idea is made precise by the following definition.

Definition 11 (Similarity relation between states). *Let $S = (C, V), S' = (C', V'), S'' = (C'', V'') \in \mathbf{S}$. We say that state S' is at least as similar to state S as state S'' is, noted $S'' \preceq_S S'$, if*

$$(C \cap C'') \subseteq (C \cap C').$$

According to the previous definition, state S' is at least as similar to state S as state S'' is, if the set of causal information shared by S and S'' is included in the set of causal information shared by S and S' . In other words, state S' is at least as causally compatible with state S as state S'' is. This criterion is qualitative since it relies on set inclusion. A quantitative criterion, which is not studied here, would be based on comparing the two cardinalities $|(C \cap C'')|$ and $|(C \cap C')|$.

Following Lewis, we use the similarity relation to define the set of closest (or most similar) states to a given state satisfying a given property φ .

Definition 12 (Set of φ -closest states). *Let $U \subseteq \mathbf{S}$, $S = (C, V) \in U$ and $\varphi \in \mathcal{L}_{\text{CIN}}(\mathbb{P})$. We define*

$$\begin{aligned} \text{Closest}(\varphi, S, U) &= \{S' \in U : S' \models \varphi \text{ and } \nexists S'' \in U \\ &\text{s.t. } S'' \models \varphi \text{ and } S' \prec_S S''\}, \end{aligned}$$

where $S' \prec_S S''$ iff $S' \preceq_S S''$ and $S'' \not\preceq_S S'$.

$\text{Closest}(\varphi, S, U)$ is the set of closest φ -states to S in U , according to the similarity relation \preceq_S . For notational convenience, we write $\text{Closest}(\varphi, S)$ instead of $\text{Closest}(\varphi, S, \mathbf{S})$.

On the syntactic side, we extend the modal language $\mathcal{L}_{\text{CIN}}(\mathbb{P})$ introduced in Section 3.2 by counterfactual conditional modalities in Lewis' style of the form $\varphi_1 \Box \rightarrow \varphi_2$, with

$\varphi_1, \varphi_2 \in \mathcal{L}_{\text{CIN}}(\mathbb{P})$, and interpret them relative to a context $U \subseteq \mathbf{S}$ and a state $S \in U$:

$$(S, U) \models \varphi_1 \Box \varphi_2 \iff \forall S' \in \text{Closest}(\varphi_1, S, U) : (S', U) \models \varphi_2.$$

According to the previous semantic interpretation, the conditional $\varphi_1 \Box \varphi_2$ (read “if φ_1 were true, φ_2 would be true”) holds at a given state S relative to the context U in case φ_2 is true at all the closest φ_1 -states to S in U .

6.2 Relationship with Belief Change

There is a tight correspondence between counterfactual conditionals whose antecedents and consequents are expressed by propositional formulas and causal base change operations. In order to elucidate this correspondence the notion of ω -remainder, borrowed from formal models of theory change [Alchourrón *et al.*, 1985] and belief base change [Hansson, 1993], is needed.

Definition 13 (ω -remainder). *Let $C \subseteq \mathcal{L}_{\text{PROP}}(\mathbb{P})$, $\omega \in \mathcal{L}_{\text{PROP}}(\mathbb{P})$ and $C' \subseteq C$. We say that C' is a ω -remainder of C if*

- (i) $\omega \notin \text{Cn}(C')$,
- (ii) $\forall C'' \subseteq C$, if $C' \subset C''$ then $\omega \in \text{Cn}(C'')$,

where we recall that Cn is the classical deductive closure operator over the propositional language $\mathcal{L}_{\text{PROP}}(\mathbb{P})$. We note $C \perp \omega$ the set of all ω -remainders of C .

According to the previous definition, an ω -remainder of a causal base C is a maximal subset of the original causal base from which the propositional fact ω is not deducible. The following theorem is the first key result of this section. It highlights the strong connection between counterfactual conditionals and the notions of base contraction and revision.

Theorem 7. *Let $S = (C, V) \in \mathbf{S}$ and $\omega_1, \omega_2 \in \mathcal{L}_{\text{PROP}}(\mathbb{P})$. Then, the following three items are equivalent:*

1. $S \models \omega_1 \Box \omega_2$,
2. $\forall C' \in C \perp \neg \omega_1, (C', V) \models \Box(\omega_1 \rightarrow \omega_2)$,
3. $\omega_2 \in \bigcap_{C' \in C \perp \neg \omega_1} \text{Cn}(C' \cup \{\omega_1\})$.

According to the previous theorem, declaring that a counterfactual statement $\omega_1 \Box \omega_2$ with propositional antecedent ω_1 and consequent ω_2 be true at a certain state is the same as stating the causal necessity of the fact that the antecedent implies the consequent relative to every ω_1 -remainder of the actual causal base. Moreover, by virtue of Proposition 1, the latter is equivalent to declaring that the consequent ω_2 is included in the deductive closure of each ω_1 -remainder of the actual causal base after expanding it by ω_1 . This operation is called *full meet belief base revision scheme* or *simple base revision* in the belief revision literature [Fagin *et al.*, 1983; Ginsberg, 1986; Nebel, 1989; Nebel, 1991; del Val, 1992; Rott, 1993].

The equivalence between items 1 and 3 in Theorem 7 is in line with the so-called Ramsey test [Stalnaker, 1968]. According to the latter, the validity of an epistemic conditional can be verified by adding the antecedent to a stock of beliefs, revising it if necessary, and checking whether the consequent is deducible from the resulting epistemic state.

6.3 Relationship with Interventionist Conditionals

The second key result of this section relates interventionist conditionals to Lewisian conditionals. It pinpoints a condition under which an interventionist conditional can be reduced to a Lewisian conditional.

Theorem 8. *Let $S \in \mathbf{S}_{E_q}$. Then,*

if $(S, \mathbf{S}_{E_q}) \models \langle E \rangle \top$ then

$$\left((S, \mathbf{S}_{E_q}) \models [E]\omega \text{ iff } (S, \mathbf{S}_{E_q}) \models \left(\bigwedge_{p=\tau \in E} \Delta eq(p=\tau) \right) \Box \omega \right).$$

According to the previous theorem, when considering equational states, if after the intervention there exists at least one state which is compatible with the causal base, then declaring that “ ω will be causally necessary after intervention” is the same as declaring that “if all information corresponding to the intervention holds, ω would be true”. More generally, if we restrict to equational states — that as shown in Sections 4.2 and 5.3 correspond to BCMs — the existence of a solution after intervention guarantees the reducibility of the interventionist conditional to a Lewisian conditional.

7 Conclusion

We hope we have successfully shown that our rule-based semantics offers a natural framework (i) for modeling some crucial concepts in the theory of causality including causal necessity and possibility, interventionist and counterfactual conditional, and (ii) for elucidating the connection between causal reasoning and epistemic reasoning, with special emphasis on belief base change. We have mainly focused on the complexity of satisfiability and model checking.

Future work will be devoted to studying the proof-theoretic aspects of the different modal languages introduced in the paper. We also intend to explore the practical implications of our work. We have shown that reasoning in our framework is in NP when dealing with causal necessity and possibility and in P^{NP} in the extension by interventionist conditionals. We plan to come up with reductions of our satisfiability and model checking problems to SAT with the aim of automating causal reasoning.

Last but not least, we plan to use our framework for formalizing some interesting causal concepts with special attention to actual causality and causal explanation. As recently shown in [Beckers, 2021], the SEM semantics can successfully formalize different notions of actual cause including the Halpern & Pearl’s definition [Halpern and Pearl, 2005a], the famous NESS (*Necessary Element of a Sufficient Set*) definition [Wright, 1988] and a counterfactual variant of the NESS definition. The correspondence results between equational states and binary SEMs we proved (Theorems 4 and 6) guarantee that all these definitions can be formalized in our semantics as well, provided that the variables are binary. In order to model causal explanation, we will extend our modal languages for causal reasoning with epistemic operators. Indeed, we conceive an explanation as a causal attribution by an explainer, namely, as a belief of the explainer about the actual cause of a given fact or event.

Acknowledgments

Support from the ANR project ALoRS “Action, Logical Reasoning and Spiking networks” (grant number ANR-21-CE23-0018-01) and the ANR-3IA Artificial and Natural Intelligence Toulouse Institute (ANITI) is gratefully acknowledged.

References

- [Alchourrón *et al.*, 1985] C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50:510–530, 1985.
- [Alechina *et al.*, 2017] N. Alechina, J. Y. Halpern, and B. Logan. Causality, responsibility and blame in team plans. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2017)*, pages 1091–1099. ACM, 2017.
- [Beckers and Vennekens, 2017] S. Beckers and J. Vennekens. The transitivity and asymmetry of actual causation. *Ergo*, 4(1), 2017.
- [Beckers, 2021] S. Beckers. The counterfactual NESS definition of causation. In *Proceedings of the Thirty-Fifth AAI Conference on Artificial Intelligence (AAAI-21)*, pages 6210–6217. AAAI Press, 2021.
- [Bochman and Lifschitz, 2015] A. Bochman and V. Lifschitz. Pearl’s causality in a logical setting. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*, pages 1446–1452. AAAI Press, 2015.
- [Bochman, 2003] A. Bochman. A logic for causal reasoning. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*, pages 141–146. Morgan Kaufmann, 2003.
- [Bochman, 2018a] A. Bochman. Actual causality in a logical setting. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*, pages 1730–1736. ijcai.org, 2018.
- [Bochman, 2018b] A. Bochman. On laws and counterfactuals in causal reasoning. In *Proceedings of the Sixteenth International Conference on Principles of Knowledge Representation and Reasoning (KR 2018)*, pages 494–503. AAAI Press, 2018.
- [Bochman, 2021] A. Bochman. *A Logical Theory of Causality*. MIT Press, 2021.
- [Burks, 1951] A. W. Burks. The logic of causal propositions. *Mind*, 60(239):363–382, 1951.
- [Chockler and Halpern, 2004] H. Chockler and J. Y. Halpern. Responsibility and blame: A structural-model approach. *Journal of Artificial Intelligence Research*, 22:93–115, 2004.
- [Chockler and Halpern, 2022] H. Chockler and J. Y. Halpern. On testing for discrimination using causal models. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22)*, pages 5548–5555. AAAI Press, 2022.
- [Christoff and Grossi, 2017] Z. Christoff and D. Grossi. Stability in binary opinion diffusion. In *Proceedings of the 6th International Workshop on Logic, Rationality, and Interaction (LORI 2017)*, volume 10455 of *Lecture Notes in Computer Science*, pages 166–180. Springer, 2017.
- [del Val, 1992] A. del Val. Computing knowledge base updates. In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR’92)*, pages 740–750. Morgan Kaufmann, 1992.
- [Fagin *et al.*, 1983] R. Fagin, J. D. Ullman, and M. Y. Vardi. On the semantics of updates in databases. In *Proceedings of the Second ACM SIGACT-SIGMOD Symposium on Principles of Database Systems, March 21-23, 1983, Colony Square Hotel, Atlanta, Georgia, USA*, pages 352–365. ACM, 1983.
- [Galles and Pearl, 1998] D. Galles and J. Pearl. An axiomatic characterization of causal counterfactuals. *Foundation of Science*, 3(1):151–182, 1998.
- [Geffner, 1990] H. Geffner. Causal theories for nonmonotonic reasoning. In H. E. Shrobe, T. G. Dietterich, and W. R. Swartout, editors, *Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-90)*, pages 524–530. AAAI Press / The MIT Press, 1990.
- [Ginsberg, 1986] M. L. Ginsberg. Counterfactuals. *Artificial Intelligence*, 30(1):35–79, 1986.
- [Giunchiglia *et al.*, 2004] E. Giunchiglia, J. Lee, V. Lifschitz, N. McCain, and H. Turner. Nonmonotonic causal theories. *Artificial Intelligence*, 153(1-2):49–104, 2004.
- [Grädel and Otto, 1999] E. Grädel and M. Otto. On logics with two variables. *Theoretical Computer Science*, 224:73–113, 1999.
- [Grandi *et al.*, 2015] U. Grandi, E. Lorini, and L. Perrusset. Propositional opinion diffusion. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, pages 989–997. ACM, 2015.
- [Halpern and Kleiman-Weiner, 2018] J. Y. Halpern and M. Kleiman-Weiner. Towards formal definitions of blameworthiness, intention, and moral responsibility. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, pages 1853–1860. AAAI Press, 2018.
- [Halpern and Pearl, 2005a] J. Y. Halpern and J. Pearl. Causes and explanations: a structural-model approach. Part I: Causes. *British Journal for Philosophy of Science*, 56(4):843–887, 2005.
- [Halpern and Pearl, 2005b] J. Y. Halpern and J. Pearl. Causes and explanations: a structural-model approach. Part II: Explanations. *British Journal for Philosophy of Science*, 56(4):889–911, 2005.
- [Halpern, 2000] J. Y. Halpern. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research*, 12:317–337, 2000.
- [Halpern, 2008] J. Y. Halpern. Defaults and normality in causal structures. In G. Brewka and J. Lang, editors,

- Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference (KR 2008)*, pages 198–208. AAAI Press, 2008.
- [Halpern, 2013] J. Y. Halpern. From causal models to counterfactual structures. *Review of Symbolic Logic*, 6(2):305–322, 2013.
- [Halpern, 2016] J. Y. Halpern. *Actual causality*. MIT Press, 2016.
- [Hansson, 1993] S. O. Hansson. Theory contraction and base contraction unified. *Journal of Symbolic Logic*, 58(2):602–625, 1993.
- [Hansson, 2009] S. O. Hansson. Replacement—a Sheffer stroke for belief change. *Journal of Philosophical Logic*, 38(2):127–149, 2009.
- [Hubara *et al.*, 2016] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio. Binarized neural networks. In *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016*, pages 4107–4115, 2016.
- [Kauffman, 1969] S. Kauffman. Homeostasis and differentiation in random genetic control networks. *Nature*, 224(5215):177–178, 1969.
- [Lewis, 1973] D. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, 1973.
- [Lewis, 1986] D. Lewis. Postscripts to “causation”. In D. Lewis, editor, *Philosophical Papers*, volume 2, pages 172–213. Oxford University Press, 1986.
- [Lifschitz, 1997] V. Lifschitz. On the logic of causal explanation (research note). *Artificial Intelligence*, 96(2):451–465, 1997.
- [Lorini, 2018] E. Lorini. In praise of belief bases: Doing epistemic logic without possible worlds. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, pages 1915–1922. AAAI Press, 2018.
- [Lorini, 2019] E. Lorini. Exploiting belief bases for building rich epistemic structures. In *Proceedings of the Seventeenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2019)*, volume 297 of *EPTCS*, pages 332–353, 2019.
- [Lorini, 2020] E. Lorini. Rethinking epistemic logic with belief bases. *Artificial Intelligence*, 282, 2020.
- [McCain and Turner, 1997] N. McCain and H. Turner. Causal theories of action and change. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*, pages 460–465. AAAI Press / The MIT Press, 1997.
- [Narodytska *et al.*, 2018] N. Narodytska, S. Prasad Kaviswanathan, L. Ryzhyk, M. Sagiv, and T. Walsh. Verifying properties of binarized deep neural networks. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18)*, pages 6615–6624. AAAI Press, 2018.
- [Nebel, 1989] B. Nebel. A knowledge level analysis of belief revision. In *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR’89)*. Toronto, Canada, May 15-18 1989, pages 301–311. Morgan Kaufmann, 1989.
- [Nebel, 1991] B. Nebel. Belief revision and default reasoning: Syntax-based approaches. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR’91)*, pages 417–428. Morgan Kaufmann, 1991.
- [Pearl, 2009] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2009.
- [Rott, 1993] H. Rott. Belief contraction in the context of the general theory of rational choice. *Journal of Symbolic Logic*, 58(4):1426–1450, 1993.
- [Shi *et al.*, 2020] W. Shi, A. Shih, A. Darwiche, and A. Choi. On tractable representations of binary neural networks. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020)*, pages 882–892, 2020.
- [Stalnaker, 1968] R. Stalnaker. A theory of conditionals. In N. Rescher, editor, *Studies in Logical Theory*, pages 28–45. Oxford University Press, 1968.
- [Woodward and Hitchcock, 2003] J. Woodward and C. Hitchcock. Explanatory generalizations, part i: A counterfactual account. *Noûs*, 37(1):1–24, 2003.
- [Woodward, 2003] J. Woodward. *Making Things Happen: a Theory of Causal Explanation*. Oxford University Press, 2003.
- [Wright, 1988] R. W. Wright. Causation, responsibility, risk, probability, naked statistics, and proof: Pruning the bramble bush by clarifying the concepts. *Iowa Law Review*, 73:1001–1077, 1988.
- [Zhang, 2013] J. Zhang. A Lewisian logic of causal counterfactuals. *Minds and Machines*, 23(1):77–93, 2013.