A Rule-based Modal Framework for Causal Reasoning

EMILIANO LORINI*, IRIT, CNRS, Toulouse University, France

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, causal necessity post-intervention and causal counterfactuals. We provide complexity results for the satisfiability checking and model checking problems for these modal languages. Moreover, we study the relationship between our rule-based semantics and the structural equation modeling (SEM) approach to causal reasoning. Finally, we use our semantics to elucidate the relationship between causal counterfactuals and belief change.

JAIR Associate Editor: Davide Grossi

JAIR Reference Format:
Emiliano Lorini. 2026. A Rule-based Modal Framework for Causal Reasoning . *Journal of Artificial Intelligence Research* 85, Article 13 (February 2026), 42 pages. DOI: 10.1613/jair.1.18960

1 Introduction

Causal reasoning is a central topic for AI nowadays given the relevance of causal concepts such as causal explanation for machine learning and causal responsibility for multi-agent theories and 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 (Bochman 2003; Geffner 1990; Giunchiglia et al. 2004; Lifschitz 1997; McCain and Turner 1997). 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 (Beckers 2021; Beckers and Vennekens 2017; Halpern 2016, 2008), explanation (Halpern and Pearl 2005b; Woodward 2003; Woodward and Hitchcock 2003), responsibility and blame (Alechina et al. 2017; Chockler and Halpern 2004; Halpern and Kleiman-Weiner 2018), discrimination (Chockler and Halpern 2022) and harm (Beckers, Chockler, et al. 2024). 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.

*Corresponding Author.

Author's Contact Information: Emiliano Lorini, ORCID: 0000-0002-7014-6756, Emiliano.Lorini@irit.fr, IRIT, CNRS, Toulouse University, Toulouse, France.

DOI: 10.1613/jair.1.18960

In the present work, we introduce a novel rule-based semantics for representing causal information through Boolean (or binary) variables, and use it to interpret a number of modal languages that support reasoning about causality. We believe that restricting to causal reasoning with Boolean variables is interesting *per se* since it covers, or at least is closely connected to, a variety of interesting models and concepts including propositional opinion diffusion (Christoff and Grossi 2017; Grandi et al. 2015), binary neural networks and diagrams (Hubara et al. 2016; Lewis 1986; Narodytska et al. 2018; Shi et al. 2020), and Boolean networks (Kauffman 1969). There is a crucial difference between the standard semantics for modal logic relying on multi-relational Kripke structures (Blackburn et al. 2001) and our semantics. In the standard Kripke semantics, the notion of possible state (or world) in a model is undecomposed and accessibility relations between states are given as primitives. In our semantics, a state is decomposed into two elements: i) a propositional valuation that represents the ontic component of the state (i.e., how the state is), and ii) a set of rules in propositional form, viz. a causal base, that represents the causal theory regulating the state (i.e., how the state could be). Moreover, various types of accessibility relation between states can be directly computed from the causal bases. These include a causal compatibility relation (i.e., whether two states share the same causal base), a causal compatibility post-intervention relation (i.e., whether two states shares the same causal base after the execution of an intervention), and a causal comparative similarity relation (i.e., whether a first state is at least as causally similar to a reference state as a second state). In a related work, we explored a similar idea by replacing the extensional semantics of epistemic logic based on multi-relational Kripke models with a succinct semantics using belief bases (Lorini 2019, 2018, 2020).

We will show that our rule-based semantics is useful both computionally and conceptually. On the computational side, we will use it to provide a succinct formulation of model checking in which causal properties, expressed by means of formulas of the modal languages we will introduce, are verified with respect to a finite state with no need to specify the set of possible states and the accessibility relations. Indeed, as pointed out above, in our semantics a state contains the sufficient information to compute *ex post* the set of possible states and the accessibility relations that are needed to perform causal reasoning. This is particularly useful for automatic verification of causal properties in concrete applications. On the conceptual side, by leveraging the concept of causal base which is analogous to that of belief base, our rule-based semantics offers the right level of generality to explore the intimate connection between causal reasoning and epistemic reasoning that, we believe, is difficult to grasp using the structural equation modeling approach. In particular, we will use our semantics to elucidate the relationship between causal counterfactual conditionals, belief revision and interventions.

The central concept underlying the modal languages introduced in this paper, and interpreted through the rule-based semantics, is *causal necessity*. This concept traces back to Burks (Burks 1951), and earlier to Kant[1] and Hume[2]. Burks appears to be the first to offer a modal-logical analysis of this notion, though the idea of treating causal necessity as a modal concept was already suggested by Russell (Russell 1913). He distinguishes causal from logical necessity, taking the latter to be stronger (that is, all logically necessary facts are also causally necessary), a distinction that mirrors that between nomic and metaphysical necessity in philosophy (Kment 2014, Chapter 7). Following Burks, we model causal necessity using an S5 modal operator. This operator identifies the facts that hold true in all states causally compatible with the actual state, where a state S' is considered causally compatible with the actual state S if they share the same causal base.

We distinguish plain causal necessity in Burks' sense from causal necessity *post-intervention*. While plain causal necessity captures causal necessity under the assumption that the causal information does not change, causal necessity post-intervention captures causal necessity after an intervention has taken place, where an intervention is a causal change operation (or event) that consists in fixing the value of a certain variable thereby making it causally independent of the other variables. In this sense, causal necessity post-intervention has a conditional

[1]See (Friedman 2017).
[2]See (Stroud 1978).

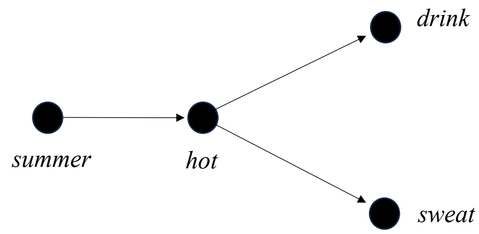

Fig. 1. Hot summer causal graph

reading: a given fact φ is said to be causally necessary post-intervention E if and only if "if the intervention E occurred, φ would be necessarily true".

To better understand the difference between plain causal necessity and causal necessity post-intervention, consider a scenario whose causal aspects are represented by a propositional causal base that specifies the direct causal connections between four variables: it is a non-rainy day in the middle of the summer (*summer*), the physical environment is very hot (*hot*), the person drinks a lot (*drink*), the person sweats a lot (*sweat*). Specifically, suppose the causal base includes three rules stating that "the physical environment is very hot if and only if it is a non-rainy day in the middle of summer", "the person drinks a lot if and only if the physical environment is very hot", and "the person sweats a lot if and only if the physical environment is very hot". In formal terms,

$$hot \leftrightarrow summer,$$
$$drink \leftrightarrow hot,$$
$$sweat \leftrightarrow hot.$$

The causal graph induced by the three causal rules is given in Figure 1. In this example, the following two facts are causally necessary: i) if the physical environment is very hot then the person will sweat a lot, and ii) if the person drinks a lot then it will sweat a lot. Thus, plain causal necessity does not allow us to identify the cause of the fact that the person sweats a lot since it only captures the positive correlations between the variables and, from a correlational point of view, there is no difference between the fact that the physical environment is very hot and the fact that the person drinks a lot with respect to the fact that the person sweats a lot. In order to identify the cause of a person sweating a lot we have to rely on causal necessity post-intervention. In particular, we have to perform the following interventionist test. We first intervene on the variable *drink* (e.g., by the person getting into the habit of drinking a lot even on non-hot days) so that the causal influence of the variable *hot* on it is interrupted (left side of Figure 2). After this intervention, it is no longer causally necessary that if the person drinks a lot then it will sweat a lot. Indeed, after this intervention, it is possible that the person drinks a lot (on a non-hot day) without sweating a lot. Then, we intervene on the variable *hot* (e.g., by the person getting into the habit of taking a sauna regularly on non-summer days) (right side of Figure 2). After this intervention, it is still causally necessary that if the physical environment is very hot then the person will sweat a lot. So, from the point of view of causal necessity post-intervention, there is an asymmetry between the fact that the physical environment is very hot and the fact that the person drinks a lot with respect to the fact that the person sweats a lot. Such an asymmetry leads us to conclude that the person's sweating depends on the physical environment being very hot but not on the person drinking a lot.

Outline. The paper is organized as follows. In Section 2, we discuss some related work in more detail. In Section 3, we introduce our rule-based semantics as well as a modal language for reasoning about plain causal necessity. We show how to interpret the language by the semantics. In Section 4 we focus on the satisfiability checking

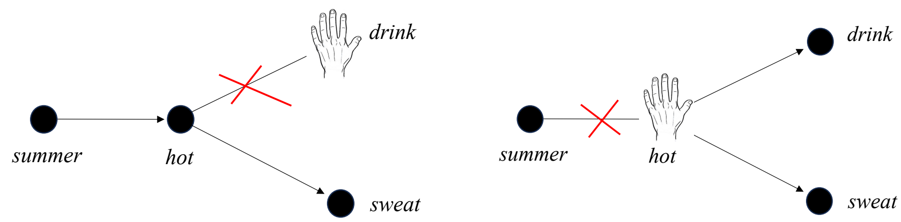

Fig. 2. Intervention on variable *drink* (left), intervention on variable *hot* (right).

problem for our modal language and prove its NP-completeness. In Section 5, we provide a succinct formulation of model checking in our rule-based semantics and prove its Θ_2^p-completeness. With 'succinct' we mean that the model with respect to which a formula has to be checked is not given explicitly with its set of possible states and its causal compatibility relations, but it is given in a compact form. Our complexity analysis of model checking relies on two polynomial embeddings: of Carnap's modal logic into model checking and of model checking into Carnap's modal logic. Our results highlight that in our semantics model checking is intrinsically more complex than satisfiability checking given its succinctness. Section 6 is devoted to exploring the relationship between our rule-based semantics and the SEM semantics for causal reasoning. We show that the latter corresponds to a specific instance of the former where causal information is represented in an equational form. In Section 7, we extend our modal framework with modal operators of causal necessity post-intervention. We introduce a novel semantics for interventions 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). We show that Θ_2^p-completeness for model checking generalizes to the language of causal necessity post-intervention. Finally, in Section 8, we present a second extension with causal counterfactual conditionals. We interpret them by means of a notion of comparative similarity that we compute from causal bases. We compare causal counterfactual conditionals to post-intervention causal necessity and spell out 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. To make the paper more readable, the proofs are given in the technical Appendix A at the end of the paper.

To sum up, the major contributions of the present paper are:

- a rule-based semantics for Boolean causality and a family of modal languages for reasoning about causal information, plain causal necessity and causal necessity post-intervention interpreted through this semantics;
- a succinct formulation of the model-checking problem for the modal languages under the rule-based semantics, together with tight complexity results for it;
- an analysis of the relationship between the rule-based semantics and the SEM semantics;
- an analysis, with the support of the rule-based semantics, of the relationship between causal counterfactual conditionals, belief revision and causal necessity post-intervention.

We believe that this is the first attempt to provide a logical framework for causal reasoning that is shown to be sufficiently general to account for a rich variety of causal notions (e.g., causal information, plain causal necessity and causal necessity post-intervention), in which model checking can be formulated succinctly, and which allows to link causal reasoning to reasoning about belief change. This work is an extended and improved version of a paper appeared in the proceedings of the IJCAI 2023 conference (Lorini 2023).

2 Related Work

We organize the discussion of related work along four axes: the structural equation modeling approach to causal reasoning, the modeling of causality in non-deterministic domains, team semantics for causal reasoning, and Bochman's nonmonotonic rule-based approach to causal reasoning.

2.1 SEM Semantics for Logics of Causal Reasoning

Structural equation models (SEM) 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 Lewis' logic of counterfactual conditionals (Halpern 2013; Zhang 2013).

A disadvantage of the SEM semantics used by Halpern & Pearl compared to our rule-based semantics is the fact that it is not succinct, thereby making model checking for modal languages of causality not exploitable in practice. Models used by Halpern & Pearl in (Galles and Pearl 1998; Halpern 2000; Halpern and Pearl 2005a) 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 6, 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.

There is also a key conceptual difference between our modal approach to causal reasoning and Halpern & Pearl's. As emphasized in the introduction, the concept underlying our modal languages is causal necessity in its plain and post-intervention form. Halpern & Pearl's modal language is designed as a dynamic language which supports reasoning about the effects of interventions on the values of the variables. As explicitly stated by Halpern (Halpern 2000, p. 320), his language "...borrows ideas from dynamic logic..." (Harel and Tiuryn 2000). It does not involve any concept of causal necessity.

From a certain point of view, the SEM semantics used by Halpern & Pearl can be considered more general than our ruled-based semantics, since variables are not necessarily Boolean. Binary causal models, as a special case of causal models with Boolean variables, have been studied in depth in (Aleksandrowicz et al. 2017; Eiter and Lukasiewicz 2002). But from another point of view, their semantics can be considered less general than ours for at least two reasons. First of all, unlike SEMs, in our rule-based semantics there is no rigid distinction between endogenous and exogenous variables which is fixed at the meta-level. We represent causal information by a set of propositional formulas in which all propositional variables have the same status. Secondly, our semantics allows for incomplete information about the underlying causal model. However, as we will show in Section 6, when we restrict to equational causal bases, i.e., causal bases in which propositional formulas are in equational form, we have complete information about the causal model and we can extract the exogenous and endogenous variables from it.

2.2 Non-Deterministic Causal Models

The notion of incomplete information about the underlying causal model, which can be represented in our semantics, has some aspects in common with that of non-deterministic causal model, recently introduced in (Beckers 2025) as a generalization of a standard causal model with deterministic effects. Contrarily to a deterministic causal model, in a non-deterministic causal model various values of an endogenous variable may be possible, once the values of the other (exogenous and endogenous) variables are specified. Causation in non-deterministic domains has also been recently studied in the action-theoretic framework of situation calculus (Khan, Lespérance, et al. 2025) following previous work on the situation calculus semantics for causal reasoning (Batusov and Soutchanski 2018; Khan and Soutchanski 2020).

2.3 Causal Team Semantics

Another work to be mentioned is Barbero and coll.'s recent application of the team semantics (Hodges 1997) to causal reasoning (Barbero and Sandu 2021; Barbero and Yang 2020). They show that the notion of interventionist conditional à la Halpern & Pearl and of functional dependency can be naturally expressed by the notion of causal team: a set of assignments for the variables sharing a common domain, called a team, extended with a functional component capturing the causal dependencies between the variables. The causal team semantics share with the SEM semantics the idea of representing causal information in a functional form. It shares with our rule-based semantics the idea of considering a set of assignments (resp. a set of possible states in our setting) that are compatible with the available causal information specified in functional form (resp. in propositional form in our setting).

Like the SEM semantics and unlike our rule-based semantics, the causal team semantics distinguishes exogenous from endogenous variables at the meta-level. Moreover, like the SEM semantics, it was recently generalized in (Barbero 2024) in order to represent causal models with non-deterministic effects.

2.4 Bochman's Nonmonotonic Approach to Causal Reasoning

We conclude this section by underlining some differences between our approach to causal reasoning and Bochman's approach mentioned in the introduction. Ours is a classic work in modal logic. At the syntactic level, we start with a basic modal language for reasoning about plain 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 modal operators for causal necessity post-intervention and Lewisian counterfactual conditionals. At the semantic level, we use accessibility relations over possible states/worlds for interpreting the different modal operators. Bochman'sf language presented in (Bochman 2021, 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, it is not clear whether and how our concepts of causal necessity in plain and post-intervention form, 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 in propositional form 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 $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ be the propositional language built from $\mathbb{P}$. We assume the language $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ contains the symbol $\top$ ("true") as primitive, while $\bot$ ("false") is defined from it as usual $\bot =_{def} \neg\top$.

Given a propositional formula $\omega \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, the set of symbols from $\mathbb{P}$ occurring in ω is denoted by $\mathbb{P}(\omega)$. The definition extends to a set of propositional formulas $X \subseteq \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ 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 $C \subseteq \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ is a causal base and $V \subseteq \mathbb{P}$ is a propositional valuation such that*

$$\forall \omega \in C,\ V \models \omega, \qquad \textbf{(CausalCons)}$$

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

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

The causal base C of the state S represents the causal theory that regulates it. We call this the state S's *causal-theoretical* (or simply *causal*) component. The propositional valuation V represents how the state S is, namely, the set of atomic propositions that are true (resp. false) at it. We call this the state S's *ontic* component. According to the Condition (**CausalCons**), a state's ontic component must be consistent with the state's causal component. In other words, how the state is must be in agreement with the causal theory that specifies how the state could be.

We use the generic term "causal information" to indicate the propositional formulas in the causal base C. 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 $\equiv$ to be the binary relation on the set $\mathbf{S}$ such that, for every $S = (C, V), S' = (C', V') \in \mathbf{S}$:*

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

$S \equiv S'$ means that state S and state S' are causally compatible since they share the same causal information. We define $\equiv(S) = \{S' \in \mathbf{S} : S \equiv S'\}$ to be the set of states that are causally compatible with the state S.

It is easy to verify that $\equiv$ is an equivalence relation and that the following holds for every $S = \big(C, V\big), S' = \big(C', V'\big) \in \mathbf{S}$:

$$\text{if } S \equiv S' \text{ then } \forall \omega \in C, V' \models \omega. \tag{1}$$

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 $S \equiv S'$. 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 plain causal necessity. It is defined by the following grammar:

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

where p ranges over $\mathbb{P}$ and ω ranges over $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$. The other Boolean connectives $\vee$ (or), $\oplus$ (exclusive or), $\rightarrow$ (implication) and $\leftrightarrow$ (double implication) are defined in the usual way. The first layer language $\mathcal{L}_{\mathsf{CI}}(\mathbb{P})$ (Language for Causal Information) is the language for talking about explicit causal information. The formula $\triangle\omega$ has to be read "ω is part of the causal theory regulating the actual state" or, more shortly, "ω is a causally relevant information". The second layer language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ (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 $\Diamond\varphi =_{def} \neg\Box\neg\varphi$, where the formula $\Diamond\varphi$ has to be read "it is causally possible that φ".

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

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

where p ranges over $\mathbb{P}$.

The language $\mathcal{L}_{\mathsf{PCN}}(\mathbb{P})$ (Language for *Propositional* Causal Necessity) is the fragment of $\mathcal{L}_{\mathsf{CN}}(\mathbb{P})$ in which we can only talk about the causal necessity of propositional facts. It is defined by the following grammar:

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

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

It is worth to consider these fragments for both conceptual and technical reasons. Indeed, each fragment has its conceptual counterpart. Moreover, they will be used to build a connection between the rule-based semantics for causal reasoning and the SEM semantics.

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

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

$\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is the most general one.

We interpret the formulas of the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ relative to a set of states U and a specific state S included in it. We call U *context* (or *universe*) of interpretation. Specifically, given $S \in U \subseteq \mathbf{S}$ and $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$, we write $(S, U) \models \varphi$ to mean that φ is true at (S, U) and define it inductively, as follows.

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

$$\begin{aligned}
(S, U) &\models \top, \\
(S, U) &\models p \iff p \in V, \\
(S, U) &\models \triangle\omega \iff \omega \in C, \\
(S, U) &\models \neg\varphi \iff (S, U) \not\models \varphi, \\
(S, U) &\models \varphi \wedge \psi \iff (S, U) \models \varphi \text{ and } (S, U) \models \psi, \\
(S, U) &\models \Box\varphi \iff \forall S' \in U, \text{ if } S \equiv S' \text{ then } (S', U) \models \varphi.
\end{aligned}$$

The pair (S, U) with $U \subseteq \mathbf{S}$ and $S \in U$ with respect to which formulas are interpreted is also called a *model*.

The causal necessity modality $\Box$ is a so-called S5 modality. According to the previous definition, it is causally necessary that φ if and only if, φ is true at all states in the context that are causally compatible with the actual state. The modality $\triangle$ 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*. Moreover, given $\omega_{\mathrm{c}} \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, we define

$$\mathbf{S}(\omega_{\mathrm{c}}) = \left\{S = (C, V) \in \mathbf{S} : V \models \omega_{\mathrm{c}}\right\}.$$

$\mathbf{S}(\omega_c)$ is the context induced by the hard (or integrity) constraint ω_c. We moreover define $\mathbf{S}_{Fin}(\omega_c) = \mathbf{S}_{Fin} \cap \mathbf{S}(\omega_c)$, where we recall $\mathbf{S}_{Fin}$ is the set of finite states (see Definition 1).

As highlighted by the following proposition, the causal necessity of a propositional fact ω relative to the context induced by a hard constraint ω_c coincides with its deducibility from the actual causal base extended with the hard constraint. Its proof is given in Section A.1 in the technical appendix at the end of the paper.

Proposition 1. *Let $S = (C, V) \in \mathbf{S}$ and $\omega_c, \omega \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$. Then,*

$$\big(S, \mathbf{S}(\omega_c)\big) \models \Box\omega \text{ iff } \omega \in Cn\big(C \cup \{\omega_c\}\big),$$

with Cn the classical deductive closure operator over the propositional language $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$.

Let us introduce a simple example to illustrate the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ and its semantics interpretation.

Example 1 (Video game). *Consider a video game character that is controlled via keyboard by a human video gamer. The human can choose the configuration of the keyboard controls that regulate the character's actions in the video game. Thus, each configuration corresponds to a causal base that specifies how the character will behave depending on the keyboard key activated by the human. Suppose the character can only execute five actions: 'move up' (up), 'move down' (do), 'move right' (ri), 'move left' (le) and 'execute an attack with the sword' (at).*[3] *Moreover, there are five keyboard keys that the human can activate during the game: $ke_1, \dots, ke_5$.*

In the actual situation the controls are set in such a way that i) if key 1 is activated (ke_1), then the character will move up; if key 2 is activated (ke_2), then it will move down; if key 3 is activated (ke_3), then it will move right; if key 4 is activated (ke_4), then it will move left; it will execute an attack with its sword if and only if key 5 is activated (ke_5). Moreover, in the actual situation ii) no key is activated and the character takes no action. Finally, iii) there is a hard constraint in the game that prevents the human from activating two different keys at the same time. (We presuppose, without including this aspect in the formalization, that if the human presses two or more keys, then none of them will be activated.) So, according to the hypotheses i), ii) and iii), we are at the state $S_0 = (C_0, V_0)$ and at the context $\mathbf{S}(\omega_0)$ induced by the hard constraint ω_c with:

$$\begin{aligned} C_0 &= \big\{ke_1 \to up, ke_2 \to do, ke_3 \to ri, ke_4 \to le, ke_5 \leftrightarrow at\big\},\\ V_0 &= \emptyset,\\ \omega_c &= \bigwedge_{\substack{x,y \in \{1,2,3,4,5\}:\\ x \neq y}} (ke_x \to \neg ke_y). \end{aligned}$$

It is easy to verify that at $\big(S_0, \mathbf{S}(\omega_c)\big)$ it is causally relevant that if key 1/2/3/4 is activated then the character will move up/down/right/left, and that key 5 is activated if and only if the character attacks with its sword. Moreover, it is causally necessary that if key 1 is activated then the character will not attack with its sword. In formal terms, we have

$$\begin{aligned}\big(S_0, \mathbf{S}(\omega_c)\big) \models & \triangle(ke_1 \to up) \wedge \triangle(ke_2 \to do) \wedge \triangle(ke_3 \to ri) \wedge \triangle(ke_4 \to le) \wedge \triangle(ke_5 \leftrightarrow at) \wedge \\ & \Box(ke_1 \to \neg at).\end{aligned}$$

In the previous example there is incomplete information about the underlying causal model. Indeed, for every proposition in $\{up, do, ri, le\}$, there is just a rule in the form of an implication specifying a sufficient condition for the proposition being true, and there is no rule in the form of a double implication specifying the necessary and sufficient condition for the proposition being true. This means that the activation of a key 1, 2, 3 or 4 is a sufficient condition for the character to move in the corresponding direction (up, down, right, left), but it is not a necessary condition. The movement of the character could be determined by an exogenous event that cannot be identified

[3] An example of this game is the Nintendo classic "The Legend of Zelda" (1986):
https://zelda.fandom.com/wiki/The_Legend_of_Zelda.

in advance due to the lack of information about the game and that is not included in the representation of the causal model (e.g., the character could be accidentally moved by an enemy). As emphasized in Section 2, this is an important difference between our semantics and the standard SEM semantics in which it is assumed that the value of an endogenous variable is univocally determined once the values of the other variables are specified.

4 Satisfiability Checking

A formula φ of the general language of causal information and necessity $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is said to be satisfiable if there exists a model (S, U) (i.e., a context $U \subseteq \mathbf{S}$ and a state $S \in U$) such that $(S, U) \models \varphi$. Validity checking is the dual of satisfiability checking: $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is said to be valid if $\neg\varphi$ is not satisfiable.

The following proposition provides a list of interesting validities. Its proof is given in Section A.2 in the technical appendix.

Proposition 2. *The following formulas are valid:*

$$\big(\Box\varphi \wedge \Box(\varphi \rightarrow \psi)\big) \rightarrow \Box\psi \qquad (\mathbf{K}_{\Box})$$

$$\Box\varphi \rightarrow \varphi \qquad (\mathbf{T}_{\Box})$$

$$\Box\varphi \rightarrow \Box\Box\varphi \qquad (\mathbf{4}_{\Box})$$

$$\neg\Box\varphi \rightarrow \Box\neg\Box\varphi \qquad (\mathbf{5}_{\Box})$$

$$\triangle\omega \rightarrow \Box\omega \qquad (\mathbf{Mix1}_{\triangle,\Box})$$

$$\triangle\omega \rightarrow \Box\triangle\omega \qquad (\mathbf{Mix2}_{\triangle,\Box})$$

$$\neg\triangle\omega \rightarrow \Box\neg\triangle\omega \qquad (\mathbf{Mix3}_{\triangle,\Box})$$

Moreover,

$$\textit{if } \varphi \textit{ is valid then } \Box\varphi \textit{ is valid} \qquad (\mathbf{Nec}_{\Box})$$

Validities ($\mathbf{K}_{\Box}$), ($\mathbf{T}_{\Box}$), ($\mathbf{4}_{\Box}$) and ($\mathbf{5}_{\Box}$) together with the so-called necessitation property ($\mathbf{Nec}_{\Box}$) highlight that the causal necessity modality $\Box$ is a normal S5 modality. Validities ($\mathbf{Mix1}_{\triangle,\Box}$), ($\mathbf{Mix2}_{\triangle,\Box}$) and ($\mathbf{Mix3}_{\triangle,\Box}$) capture some interaction properties between causal information and causal necessity: if ω is a causally relevant information, then it is causally necessary that ω; if ω is a causally relevant information, then it is causally necessary that ω is a causally relevant information; if ω is not a causally relevant information, then it is causally necessary that ω is not a causally relevant information.

As the following theorem highlights, satisfiability checking for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ has the same complexity as SAT. NP-hardness follows from the NP-hardness of SAT that is clearly poly-time reducible to satisfiability checking for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ due to the fact that $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P}) \subset \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$. NP-membership relies on the fact that for any satisfiable formula of the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ we can construct a model of polynomial size which satisfies it. The proof of the theorem is given in Section A.3 in the technical appendix.

Theorem 1. *Let $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$. Then, checking whether φ is satisfiable is NP-complete.*

The previous NP-completeness results for satisfiability checking naturally extends to the languages $\mathcal{L}_{\mathsf{CN}}(\mathbb{P})$ and $\mathcal{L}_{\mathsf{PCN}}(\mathbb{P})$ since $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P}) \subset \mathcal{L}_{\mathsf{CN}}(\mathbb{P})$ and $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P}) \subset \mathcal{L}_{\mathsf{PCN}}(\mathbb{P})$.

In the next section we move to model checking.

5 Model Checking

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

Model checking for formulas in $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$
Given: $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$, $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $S \in \mathbf{S}_{Fin}(\omega_c)$.
Question: Do we have $\big(S, \mathbf{S}(\omega_c)\big) \models \varphi$?

Note that when the hard constraint ω_c equals to $\top$ model checking consists in verifying whether the formula φ is true at the finite state S relative to the global context $\mathbf{S}$, i.e., in verifying whether $S \models \varphi$.

In the rest of this section, we are going to show that in our framework model checking has a higher complexity than satisfiability checking. In particular, it is Θ_2^p-complete. This is due to its succinct 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 the following section, we introduce Carnap's modal logic, a logic which is representative of the Θ_2^p class.

5.1 The Θ_2^p Class and Carnap's Modal Logic

Before establishing the complexity of the model checking problem it is useful to remind some complexity classes that are relevant to our analysis. The complexity class Δ_2^p (also denoted P^{NP}) consists of those decision problems that can be solved by a deterministic polynomial time Turing machine with access to an NP oracle. The class Θ_2^p (also denoted $\mathrm{P}_{||}^{\mathrm{NP}}$) is the subclass of Δ_2^p consisting of those problems for which the queries to the NP oracle are *non-adaptive*, i.e., they can be performed in parallel. The latter problems are opposed to those problems in Δ_2^p for which the queries to the NP oracle are *adaptive*, i.e., the way a query to the NP oracle is formulated may depend on the answers to the previous queries. The class Θ_2^p has been studied in (Jenner and Torán 1995; Wagner 1990). As shown in (Gottlob 1995)[4], an interesting logic characterizing this class is Carnap's modal logic (Carnap 1947).

The language of Carnap's modal logic is a standard mono-modal language $\mathcal{L}_{\mathfrak{C}}(\mathbb{P})$ defined by the following grammar:

$$\mathcal{L}_{\mathfrak{C}}(\mathbb{P}) \quad \stackrel{\text{def}}{=} \quad \varphi \quad ::= \quad p \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi \mid \blacksquare\varphi \mid \blacklozenge\varphi,$$

where p ranges over $\mathbb{P}$. The operator $\blacksquare$ captures the notion of strong provability: the formula $\blacksquare\varphi$ can be read "φ is provable". Its dual $\blacklozenge$ captures admissibility: the formula $\blacklozenge\varphi$ can be read "φ is admissible".

As shown in (Gottlob 1995), Carnap's modal logic has a natural possible worlds semantics based on the notion of *universal structure with finite valuations* for the set of atomic propositions $\mathbb{P}$. It consists of all propositional valuations that assign positive truth values to a *finite* a number of atomic propositions from $\mathbb{P}$, that is,

$$\mathbf{W}_{\mathbb{P}}^{fin} = \{w \in \mathbf{W}_{\mathbb{P}} : w \text{ is finite}\},$$

where $\mathbf{W}_{\mathbb{P}} = 2^{\mathbb{P}}$ is the *universal structure* for the set of atomic propositions $\mathbb{P}$.

In particular, a formula of Carnap's modal logic can be interpreted relative to a specific valuation (or world) $w \in \mathbf{W}_{\mathbb{P}}^{fin}$, as follows. (We omit the interpretation of the Boolean connectives and of $\top$ since they are defined in the usual way.)

$$\begin{aligned} w \models^{fin} p &\iff p \in w, \\ w \models^{fin} \blacksquare\varphi &\iff \forall v \in \mathbf{W}_{\mathbb{P}}^{fin}, v \models^{fin} \varphi, \\ w \models^{fin} \blacklozenge\varphi &\iff \exists v \in \mathbf{W}_{\mathbb{P}}^{fin} \text{ such that } v \models^{fin} \varphi. \end{aligned}$$

We are going to show that the restriction to finite valuations is not necessary. Specifically, the formulas of the language $\mathcal{L}_{\mathfrak{C}}(\mathbb{P})$ can be interpreted relative to a world $w \in \mathbf{W}_{\mathbb{P}}$ as follows, without affecting the resulting set of

[4]See also Eiter and Gottlob (1997).

validities:

$$w \models p \iff p \in w,$$
$$w \models \blacksquare\varphi \iff \forall v \in \mathbf{W}_{\mathbb{P}}, v \models \varphi,$$
$$w \models \blacklozenge\varphi \iff \exists v \in \mathbf{W}_{\mathbb{P}} \text{ such that } v \models \varphi.$$

To show the equivalence between the semantics based on the universal structure with finite valuations $\mathbf{W}_{\mathbb{P}}^{fin}$ and the one based on the universal structure $\mathbf{W}_{\mathbb{P}}$, we need to introduce two different notions of of validity for Carnap's modal logic.

A formula φ of the language $\mathcal{L}_{\mathfrak{C}}(\mathbb{P})$ is said to be Carnap valid, denoted $\models_{\mathfrak{C}} \varphi$, if $v \models \varphi$ for all $v \in \mathbf{W}_{\mathbb{P}}$. The formula φ is said to be *finite* Carnap valid, denoted $\models_{\mathfrak{C}}^{fin} \varphi$, if $v \models^{fin} \varphi$ for all $v \in \mathbf{W}_{\mathbb{P}}^{fin}$. As the following proposition highlights, Carnap validity and finite Carnap validity coincide. Its proof is given in Section A.4 in the technical appendix.

Proposition 3. *Let $\varphi \in \mathcal{L}_{\mathfrak{C}}(\mathbb{P})$. Then,*

$$\models_{\mathfrak{C}} \varphi \textit{ iff } \models_{\mathfrak{C}}^{fin} \varphi.$$

It is known that finite Carnap validity checking is Θ_2^p-complete (Gottlob 1995, Theorem 5.1.2). Thus, because of Proposition 3, Carnap validity checking is also Θ_2^p-complete. In the next section, we are going to provide two poly-time reductions: of our model checking problem into Carnap validity checking, and vice-versa. Thanks to these two polynomial embeddings, we will state a Θ_2^p-completeness result for our model checking problem.

5.2 Poly-time Reductions and Complexity

We first prove the Θ_2^p lower bound for model checking and then we turn to the Θ_2^p upper bound.

5.2.1 Lower Bound. Proving that our model checking problem is Θ_2^p-hard is the easiest part. Indeed, we can find a polynomial embedding of Carnap validity checking into it. In particular, let

$$tr : \mathcal{L}_{\mathfrak{C}}(\mathbb{P}) \longrightarrow \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$$

be the translation from the language of Carnap modal logic to the language of causal information and necessity such that

$$tr(p) = p,$$
$$tr(\neg\varphi) = \neg tr(\varphi),$$
$$tr(\varphi \wedge \psi) = tr(\varphi) \wedge tr(\psi),$$
$$tr(\blacksquare\varphi) = \Box tr(\varphi),$$
$$tr(\blacklozenge\varphi) = \Diamond tr(\varphi).$$

As the following theorem highlights, a formula of Carnap's modal logic is Carnap valid if and only if at the empty state $(\emptyset, \emptyset)$ relative to the global context $\mathbf{S}$ (we recall that $S \models \varphi$ abbreviates $(S, \mathbf{S}) \models \varphi$) it is causally necessary that its translation is true. Its proof is given in Section A.5 in the technical appendix.

Theorem 2. *Let $\varphi \in \mathcal{L}_{\mathfrak{C}}(\mathbb{P})$. Then,*

$$\models_{\mathfrak{C}} \varphi \textit{ iff } (\emptyset, \emptyset) \models \Box tr(\varphi).$$

The following theorem is a direct corollary of Theorem 2, the fact that the size of $tr(\varphi)$ is polynomial in the size of the input formula φ and, as pointed out above, the fact that Carnap validity checking is Θ_2^p-hard.

Theorem 3. *Model checking for formulas in $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is Θ_2^p-hard.*

In the next section, we focus on the complexity upper bound of our model checking problem.

5.2.2 Upper Bound. We define a translation from our language of causal information and necessity to the language of Carnap's modal logic built over the following set of atomic formulas $\mathbb{P}_\triangle$ that includes the basic atomic propositions as well as one special atomic formula for every "triangle" formula $\triangle\omega$:

$$\mathbb{P}_\triangle = \mathbb{P} \cup \{p_{\triangle\omega} : \triangle\omega \in \mathcal{L}_{\mathsf{CI}}(\mathbb{P})\}.$$

Specifically, for each finite state $S = (C, V) \in \mathbf{S}_{Fin}$ and hard constraint $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, we define the translation

$$tr_{S,\omega_c} : \mathcal{L}_{\mathsf{CIN}}(\mathbb{P}) \longrightarrow \mathcal{L}_{\mathfrak{C}}(\mathbb{P}_\triangle)$$

such that:

$$\begin{aligned}
tr_{S,\omega_c}(p) &= \begin{cases} \top, & \text{if } p \in V, \\ \bot, & \text{otherwise,} \end{cases} \\
tr_{S,\omega_c}(\top) &= \top, \\
tr_{S,\omega_c}(\triangle\omega) &= \begin{cases} \top, & \text{if } \omega \in C, \\ \bot, & \text{otherwise,} \end{cases} \\
tr_{S,\omega_c}(\neg\varphi) &= \neg tr_{S,\omega_c}(\varphi), \\
tr_{S,\omega_c}(\varphi \wedge \psi) &= tr_{S,\omega_c}(\varphi) \wedge tr_{S,\omega_c}(\psi), \\
tr_{S,\omega_c}(\Box\varphi) &= \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C} \omega \wedge \bigwedge_{\omega \in C} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\varphi)\setminus C} \neg p_{\triangle\omega}\big) \to tr^*_{C,\omega_c}(\varphi)\Big),
\end{aligned}$$

and

$$\begin{aligned}
tr^*_{C,\omega_c}(p) &= p \\
tr^*_{C,\omega_c}(\top) &= \top, \\
tr^*_{C,\omega_c}(\triangle\omega) &= p_{\triangle\omega}, \\
tr^*_{C,\omega_c}(\neg\varphi) &= \neg tr^*_{C,\omega_c}(\varphi), \\
tr^*_{C,\omega_c}(\varphi \wedge \psi) &= tr^*_{C,\omega_c}(\varphi) \wedge tr^*_{C,\omega_c}(\psi), \\
tr^*_{C,\omega_c}(\Box\varphi) &= \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C} \omega \wedge \bigwedge_{\omega \in C} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\varphi)\setminus C} \neg p_{\triangle\omega}\big) \to tr^*_{C,\omega_c}(\varphi)\Big),
\end{aligned}$$

with

$$\begin{aligned}
\Gamma(p) &= \emptyset, \\
\Gamma(\triangle\omega) &= \{\omega\}, \\
\Gamma(\neg\varphi) &= \Gamma(\varphi), \\
\Gamma(\psi_1 \wedge \psi_2) &= \Gamma(\psi_1) \cup \Gamma(\psi_2), \\
\Gamma(\Box\psi) &= \Gamma(\psi).
\end{aligned}$$

For each atomic proposition p (resp. "triangle" formula $\triangle\omega$), the translation tr_{S,ω_c} returns the symbol $\top$ if p (resp. ω) belongs to the valuation V (resp. to the causal base C) of the state S. Otherwise, it returns the symbol $\bot$. The causal necessity modality $\Box$ is translated into the strong provability modality $\blacksquare$ by restricting to the possible worlds that i) satisfy the hard constraint ω_c as well as the causal information of the state S (i.e., $\omega_c \wedge \bigwedge_{\omega \in C} \omega$), ii) include all causal information that is present at S (i.e., $\bigwedge_{\omega \in C} p_{\triangle\omega}$), and iii) do not include the φ-relevant causal information (i.e., $\Gamma(\varphi)$) that is not present at S (i.e., $\bigwedge_{\omega\in\Gamma(\varphi)\setminus C}\neg p_{\triangle\omega}$). Note that the translation tr^{*}_{C,ω_c} is applied after the application of the translation tr_{S,ω_c} to the modality $\Box$. The only difference between tr_{S,ω_c} and tr^{*}_{C,ω_c} is in the translation of the atomic propositions and of the "triangle" formulas. Unlike tr_{S,ω_c}, tr^{*}_{C,ω_c} translates them into their direct counterparts on the side of the language $\mathcal{L}_{\mathfrak{C}}(\mathbb{P}_{\triangle})$ and does not need to evaluate them in relation to the state S. This is the reason why it needs to record the causal base C for handling the modality $\Box$, but it does not need to record the full state S.

The following lemma will turn out to be useful for proving that the previous translation provides a polynomial embedding of our model checking problem into Carnap's modal logic. It states that, for every formula of our language of causal information and necessity, either its translation into the language of Carnap's modal logic or the negation of its translation is Carnap valid. Its proof is given in Section A.6 in the technical appendix.

Lemma 1. *Let $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$. Then,*

$$\models_{\mathfrak{C}} tr_{S,\omega_c}(\varphi) \text{ or } \models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\varphi).$$

The following theorem is the main result of this section. It states that a formula of the language of causal information and necessity is true at a finite state S under the hard constraint ω_c if and only if its translation $tr_{S,\omega_c}(\varphi)$ is Carnap valid. Its proof is given in Section A.7 in the technical appendix.

Theorem 4. *Let $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$, $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $S = (C, V) \in \mathbf{S}_{Fin}(\omega_c)$. Then,*

$$\big(S, \mathbf{S}(\omega_c)\big) \models \varphi \text{ iff } \models_{\mathfrak{C}} tr_{S,\omega_c}(\varphi).$$

The following theorem is a direct corollary of i) Theorem 4, ii) the fact that the size of $tr_{S,\omega_c}(\varphi)$ is polynomial in the size of φ, and iii) the Θ^p_2-membership of the problem of checking whether a formula of the language $\mathcal{L}_{\mathfrak{C}}(\mathbb{P})$ is Carnap valid.

Theorem 5. *Model checking for formulas in $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is in Θ^p_2.*

We have proved that our model checking problem for the language of causal information and necessity $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is Θ^p_2-complete (Theorems 3 and 5). This result is in contrast to Theorem 1 proved in Section 4 that satisfiability checking for the same language is NP-complete. Thus, model checking for this language is intrinsically more complex that satisfiability checking. This is due to its succinct representation.

6 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.

6.1 Equational States

In order to represent causal information, we consider propositional formulas in equational form. An equational formula for a proposition p is a propositional formula of the form $p \leftrightarrow \omega$ which unambiguously specifies the truth value of p using a propositional formula ω made of propositions other than p. The set of equational formulas built from the set of atomic propositions $\mathbb{P}$ is denoted by $\mathcal{L}_{\mathsf{EQ}}(\mathbb{P})$, that is:

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

For every $p \in \mathbb{P}$, $\mathcal{L}_{\text{EQ}}(p)$ is the set of equational formulas for the atomic proposition p.

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

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*

$$\forall p \in \mathbb{P}, \forall p \leftrightarrow \omega, p \leftrightarrow \omega' \in C, \omega = \omega'. \qquad \textbf{(SingleEq)}$$

The set of states in equational form, or equational states, is denoted by S_{Eq}.

Note that the Condition (**SingleEq**) is equivalent to $\forall p \in \mathbb{P}, |\mathcal{L}_{\text{EQ}}(p)| \leq 1$.

From an equational state, it is straightforward to extract a set of endogenous variables and a set of exogenous ones. A variable is endogenous if there is an equational formula for it in the actual causal base, it is exogenous if it appears in the actual causal base but there is no equational formula for it.

Definition 5 (Exogenous and endogenous variables). *Let $S = (C, V) \in \mathrm{S}_{Eq}$. Its set of exogenous variables $exo(S)$ and its set of endogenous variables $end(S)$ are defined, as follows:*

$$end(S) = \{p \in \mathbb{P}(C) : \exists \omega \in \mathcal{L}_{\text{PROP}}(\mathbb{P} \setminus \{p\}) \text{ such that } p \leftrightarrow \omega \in C\},$$
$$exo(S) = \mathbb{P}(C) \setminus end(S).$$

As the following proposition highlights, an endogenous variable behaves functionally, i.e., its truth value is univocally determined once the truth values of the other variables are specified.

Proposition 4. *Let $S = (C, V) \in \mathrm{S}_{Eq}$, $p \in \mathbb{P}(C)$ and X, X' such that $X \cup X' = (\mathbb{P}(C) \setminus \{p\})$ and $X \cap X' = \emptyset$. Then,*

$$\text{if } p \in end(S) \text{ then } S \models \Box(\mathsf{con}_{X,X'} \to p) \vee \Box(\mathsf{con}_{X,X'} \to \neg p),$$

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

From an equational state it is also possible to extract its graphical counterpart. To show how to do this, it is useful to introduce some additional notation: for a propositional formula ω, we define $\mathbb{P}^+(\omega) = \mathbb{P}(\omega) \cup \{\top\}$ if $\top$ occurs in ω, and $\mathbb{P}^+(\omega) = \mathbb{P}(\omega)$ if $\top$ does not occur in ω. Moreover, for a causal base C, we define $\mathbb{P}^+(C) = \bigcup_{\omega \in C} \mathbb{P}^+(\omega)$.

Definition 6 (Induced causal graph). *Let $S = (C, V) \in \mathrm{S}_{Eq}$. Then, the causal graph induced by S is the pair $\Gamma_S = (N_S, \mathcal{P}_S)$ with $N_S = \mathbb{P}^+(C)$ and where the causal parent function $\mathcal{P}_S : N_S \longrightarrow 2^N$ is defined as follows: if $\top \in N_S$, then*

$$\mathcal{P}_S(\top) = \emptyset,$$

and, for every $p \in N_S \cap \mathbb{P}$,

$$\mathcal{P}_S(p) = \begin{cases} \emptyset, & \text{if } \mathcal{L}_{\text{EQ}}(p) \cap C = \emptyset, \\ \mathbb{P}^+(\omega), & \text{if } p \leftrightarrow \omega \in C. \end{cases}$$

Note that the causal parent function $\mathcal{P}_S$ is well-defined in virtue of the Condition (**SingleEq**) in Definition 4. Since S is an equational state, this condition guarantees that there cannot be more that one equational formula for the same atomic proposition.

The following fact is evident.

Fact 1. *We have $N_S \cap \mathbb{P} = end(S) \cup exo(S)$. Moreover, if $p \in N_S \cap \mathbb{P}$ then, $\mathcal{P}_S(p) = \emptyset$ iff $p \in exo(S)$.*

Let us go back to the hot summer example we informally presented in Section 1 to illustrate the concept of equational state and its graphical counterpart.

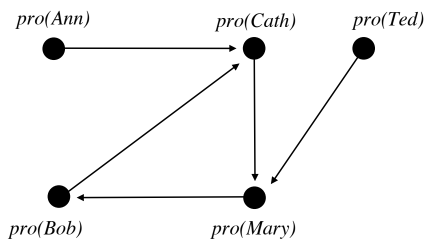

Fig. 3. Social influence causal graph

Example 2 (Hot summer). *We consider the equational state $S_0 = (C_0, V_0)$ whose causal base and propositional valuation are*

$$C_0 = \{hot \leftrightarrow summer, drink \leftrightarrow hot, sweat \leftrightarrow hot\},$$
$$V_0 = \{summer, hot, drink, sweat\}.$$

The causal graph induced by the causal base C_0 is the one depicted in Figure 1 in Section 1. Notice that it is a directed acyclic graph (DAG) with no causal loops. It is easy to verify that at state S_0, it is causally necessary that if the person drinks a lot then it will sweat a lot, and that if the physical environment is very hot then the person will sweat a lot. In formal terms,

$$S_0 \models \Box(drink \rightarrow sweat) \wedge \Box(hot \rightarrow sweat).$$

Let us consider a more complex example involving causal loops.

Example 3 (Social influence). *Consider a set of five agents $Agt = \{Ann, Bob, Cath, Mary, Ted\}$. Each agent has an opinion about the use of nuclear energy for electricity production. In particular, it 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 $pro(x)$ with $x \in Agt$, $pro(x)$ meaning that x is favor and $\neg pro(x)$ 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 $pro(Ann)$ and $pro(Ted)$ are exogenous (i.e., Ann and Ted's opinions are independent of the other agents' opinions), while all other variables are endogenous. The causal connections between the variables are described by the following three equational formulas:*

$$\omega_1 = pro(Cath) \leftrightarrow \big(pro(Ann) \wedge pro(Bob)\big),$$
$$\omega_2 = pro(Mary) \leftrightarrow \big(pro(Cath) \rightarrow pro(Ted)\big),$$
$$\omega_3 = pro(Bob) \leftrightarrow pro(Mary).$$

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 $S_0 = \big(\{\omega_1, \omega_2, \omega_3\}, \{pro(x) : x \in Agt\}\big)$. It is straightforward to verify that S_0 is an equational state. The (influence) causal graph induced by it is represented in Figure 3. Clearly, there exists at least one state which is compatible with the causal information $\omega_1, \omega_2, \omega_3$:

$$S_0 \models \Diamond\top.$$

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:

$$S_0 \models \Box pro(Bob).$$

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

$$S_0 \models \neg\Diamond\big(pro(Cath) \wedge \neg pro(Ted)\big).$$

6.2 Binary Causal Model

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

$$\mathcal{I}_Z : Z \longrightarrow \{0, 1\}.$$

The set of binary assignments for Z is denoted by Asg_Z. The interpretation $|.|$ of propositional formulas in $\mathcal{L}_{\text{PROP}}(Z)$ relative to the assignments in Asg_Z is defined as usual:

$$\begin{aligned}
|p|_{Asg_Z} &= \big\{\mathcal{I}_Z \in Asg_Z : \mathcal{I}_Z(p) = 1\big\} \text{ for } p \in Z,\\
|\top|_{Asg_Z} &= Asg_Z,\\
|\neg\omega|_{Asg_Z} &= Asg_Z \setminus |\omega|_{Asg_Z},\\
|\omega_1 \wedge \omega_2|_{Asg_Z} &= |\omega_1|_{Asg_Z} \cap |\omega_2|_{Asg_Z}.
\end{aligned}$$

Definition 7 (Binary causal model). *A binary causal model (BCM) is a tuple* $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p\in\mathbb{V}_{end}}\big)$ *where* $\mathbb{V}_{exo}$ *is a finite set of* exogenous *variables,* $\mathbb{V}_{end}$ *is a finite set of* endogenous *variables such that* $\mathbb{V}_{end}, \mathbb{V}_{exo} \subseteq \mathbb{P}$ *and* $\mathbb{V}_{end} \cap \mathbb{V}_{exo} = \emptyset$*, and, for every* $p \in \mathbb{V}_{end}$*:*

$$\mathcal{F}_p : Asg_{\mathbb{V}\setminus\{p\}} \longrightarrow \{0, 1\}$$

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

For notational convenience, we sometimes abbreviate $\mathbb{V} \setminus \{p\}$ as $\overline{p}$. Moreover, when the context is unambiguous, we write Asg instead of $Asg_{\mathbb{V}}$ and denote its elements by $\mathcal{I}, \mathcal{I}', \ldots$

Definition 8 (Solutions of a BCM). *Let* $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p\in\mathbb{V}_{end}}\big)$ *be a BCM and let* $\mathcal{I} \in Asg$*. We say that* $\mathcal{I}$ *is a solution of* Θ *if and only if, for every* $p \in \mathbb{V}_{end}$*:*

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

where $\mathcal{I}|_{\overline{p}}$ *is the restriction of function* $\mathcal{I}$ *to* $\overline{p}$*. The set of solutions of* Θ *is denoted by* $Sol(\Theta)$.

Definition 9 (Pointed BCM). *A pointed BCM is a pair* $(\Theta, \mathcal{I})$ *with* $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p\in\mathbb{V}_{end}}\big)$ *a BCM and* $\mathcal{I} \in Sol(\Theta)$*. The class of pointed BCMs is denoted by* $\mathbf{B}$.

The following example is given to illustrate Definitions 7, 8 and 9.

Example 4 (BCM for three atomic propositions). *An example of BCM for three atomic propositions, one exogenous and two endogenous, is the tuple* $\Theta_0 = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p\in\mathbb{V}_{end}}\big)$ *with:*

$$\begin{aligned}
&\mathbb{V}_{exo} = \{p\}, && \mathcal{F}_q(p \mapsto 1, r \mapsto 1) = 1,\\
&\mathbb{V}_{end} = \{q, r\}, && \mathcal{F}_r(p \mapsto 0, q \mapsto 0) = 1,\\
&\mathcal{F}_q(p \mapsto 0, r \mapsto 0) = 0, && \mathcal{F}_r(p \mapsto 0, q \mapsto 1) = 0,\\
&\mathcal{F}_q(p \mapsto 0, r \mapsto 1) = 0, && \mathcal{F}_r(p \mapsto 1, q \mapsto 0) = 1,\\
&\mathcal{F}_q(p \mapsto 1, r \mapsto 0) = 1, && \mathcal{F}_r(p \mapsto 1, q \mapsto 1) = 0.
\end{aligned}$$

It is easy to verify that the set of solutions of Θ_0 *is:*

$$Sol(\Theta_0) = \{(p \mapsto 0, q \mapsto 0, r \mapsto 1), (p \mapsto 1, q \mapsto 1, r \mapsto 0)\}$$

Thus, the pair $(\Theta_0, (p \mapsto 0, q \mapsto 0, r \mapsto 1))$ *is a pointed BCM built from* Θ_0 *since the point of evaluation* $(p \mapsto 0, q \mapsto 0, r \mapsto 1)$ *is a solution of* Θ_0*.*

Pointed BCMs provide a natural semantics for interpreting formulas of the language of propositional causal necessity $\mathcal{L}_{\mathsf{PCN}}(\mathbb{P})$. The interpretation of the more general language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ through BCMs would require a more general semantics based on sets of BCMs instead of single BCMs. We do not give this semantics here since the goal of this section is simply to elucidate the connection between the rule-based semantics and the SEM semantics for a language of causal reasoning in which nested modalities are not allowed. Notice that this is a common assumption in existing languages for causal reasoning, including the language of interventionist conditionals presented in (Halpern 2000) in which nested conditionals are disallowed.

Given a pointed BCM $(\Theta, \mathcal{I})$, we can interpret formulas of the language $\mathcal{L}_{\mathsf{PCN}}(\mathbb{V})$, as follows. (Boolean cases and $\top$ are again omitted since they are defined in the standard way.)

$$(\Theta, \mathcal{I}) \models p \iff \mathcal{I}(p) = 1,$$
$$(\Theta, \mathcal{I}) \models \Box\omega \iff Sol(\Theta) \subseteq |\omega|_{Asg}.$$

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

Example 5 (BCM for three atomic propositions (cont.)). *It is easy to verify that at the pointed BCM introduced in Example 4 both* p *and* q *are false while* r *is true, and moreover it is causally necessary that* q *is true iff* r *is false. Formally,*

$$(\Theta_0, (p \mapsto 0, q \mapsto 0, r \mapsto 1)) \models \neg p \wedge \neg q \wedge r \wedge \Box(q \leftrightarrow \neg r).$$

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. The proof of the theorem is given in Section A.8 in the technical appendix.

Theorem 6. *There exists a surjection* $f : \mathbf{S}_{Eq} \longrightarrow \mathbf{B}$ *such that if* $f(S) = (\Theta, \mathcal{I})$ *then*

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

The reason why in the statement of the previous Theorem 6 we have a surjection instead of a bijection is that there might be two (or more) equational states whose causal bases are syntactically different but co-extensional, i.e., they lead to the same set of causally compatible states.

7 Extension with Causal Necessity Post-Intervention

In this section, we extend the modal language presented in Section 3 with modal operators of causal necessity post-intervention. In order to interpret the resulting language, we 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, our semantics corresponds to the SEM semantics for interventions. We also show that the Θ_2^p-completeness result for model checking given in Section 5 naturally generalizes to the extended language.

7.1 Language

We conceive an intervention as a possibly empty finite set of equational formulas of type $p \leftrightarrow \top$ or $p \leftrightarrow \bot$ with at most one equational formula for each atomic proposition. To make presentation more concise, we sometimes write $p \leftrightarrow \tau$ with $\tau \in \{\top, \bot\}$. The set of interventions is defined as follows:

$$Int = \big\{E \subseteq \mathcal{L}_{\mathsf{EQ}}(\mathbb{P}) : \forall p \in \mathbb{P}, E \cap \mathcal{L}_{\mathsf{EQ}}(p) = \emptyset \text{ or } E \cap \mathcal{L}_{\mathsf{EQ}}(p) = \{p \leftrightarrow \top\} \text{ or } E \cap \mathcal{L}_{\mathsf{EQ}}(p) = \{p \leftrightarrow \bot\}\big\}.$$

Elements of *Int* are denoted by $E, E', \dots$ since we conceive an intervention as a special kind of event.

For every finite set of atomic propositions $Z \subseteq \mathbb{P}$, Int_Z denotes the set of interventions for Z, that is,

$$Int_Z = \big\{E \in Int : (\forall p \in Z, p \leftrightarrow \top \in E \text{ or } p \leftrightarrow \bot \in E) \text{ and } (\forall p \notin Z, p \leftrightarrow \top \notin E \text{ and } p \leftrightarrow \bot \notin E)\big\}.$$

We consider the following extension of the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ defined in Section 3.2 with modal operators of causal necessity post-intervention of the form $[E]$:

$$\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P}) \quad \stackrel{\text{def}}{=} \quad \varphi \quad ::= \quad \alpha \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi \mid [E]\varphi,$$

with α ranging over $\mathcal{L}_{\mathsf{CI}}(\mathbb{P})$ and $E \in Int$. $\mathcal{L}_{\mathsf{CIN-Int}}$ stands for "Language for Causal Information and Necessity *plus* Interventions". The new formula $[E]\varphi$ has to be read "it will be causally necessary that φ, after the occurrence of the intervention E" or "if the intervention E occurred, φ would be necessarily true". As usual, we define $\langle E\rangle\varphi =_{def} \neg[E]\neg\varphi$.

We also consider the following fragment of the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ in which we can talk about plain causal necessity and causal necessity post-intervention about propositional facts (i.e., without nesting modal operators):

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

where p ranges over $\mathbb{P}$, ω ranges over $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $E \in Int$. $\mathcal{L}_{\mathsf{PCN-Int}}$ stands for "Language for Propositional Causal Necessity *plus* Interventions". We introduce this fragment since later in the paper we will use it to compare our rule-based semantics and the SEM semantics with respect to the concept of intervention.

7.2 Semantics

In order to semantically interpret the new operators of causal necessity post-intervention, some preliminary notions are needed. The effect of an intervention 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 intervention. In other words, the revocable part of a causal base C in the light of an intervention E, denoted by $Rev(C,E)$, includes all causal information that is superseded by some element of E.

Definition 10 (Revocable part of a causal base). *Let $C \subseteq \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $E \in Int$. We define*

$$Rev(C,E) = \Big\{p \leftrightarrow \omega \in \big(C \cap \mathcal{L}_{\mathsf{EQ}}(\mathbb{P})\big) : p \leftrightarrow \top \in E \text{ or } p \leftrightarrow \bot \in E\Big\}.$$

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

Definition 11 (Causal compatibility post-intervention). *Let $E \in Int$. We define $\Rightarrow^E$ to be the binary relation on the set of states* S *such that, for every $S = (C, V), S' = (C', V') \in$* S*:*

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

$S \Rightarrow^E S'$ means that state $S' = (C', V')$ is compatible with state $S = (C, V)$ after the occurrence of the intervention 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 equational formulas which are revocable in the light of the intervention E, and then (ii) add to the resulting base the equational formulas in the intervention E.

The following fact is evident: if a state is causally compatible with an equational state after an intervention, then it is equational. Indeed, a replacement operation for a given intervention can only replace an equational formula for a proposition by another equational formula for the same proposition. Formally:

Fact 2. *If $S \in \mathbf{S}_{Eq}$ and $S \Rightarrow^E S'$ then $S' \in \mathbf{S}_{Eq}$.*

By means of the relation $\Rightarrow^E$, we can provide a semantic interpretation of the causal necessity post-intervention modalities $[E]$. They, too, are interpreted relative to a model, namely, a set of states U and a specific state S included in it. The other constructs of the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ are interpreted as in Definition 3. So, we do not need to repeat their semantic interpretations.

Definition 12 (Semantic interpretation (cont.)). *For every $U \subseteq \mathbf{S}$ and $S \in U$:*

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

Thus, it is causally necessary that φ after the occurrence of the intervention E, if and only if φ is true at all states that are compatible with the actual state after the occurrence of the intervention E. Again, to simplify notation, we write $S \models \varphi$ instead of $(S, \mathbf{S}) \models \varphi$.

Note that the causal necessity post-intervention modality with the empty intervention coincides with the plain causal necessity modality, i.e., the formula $[\emptyset]\omega \leftrightarrow \Box\omega$ is valid.

Let us go back to the Example 2 introduced in Section 6.1 to illustrate the causal necessity post-intervention modality.

Example 6 (Hot summer (cont.)). *It is routine to verify that, from the point of view of causal necessity post-intervention, there is an asymmetry between the fact that the physical environment is very hot and the fact that the person drinks a lot with respect to the fact that the person sweats a lot. Indeed, we have*

$$\begin{aligned} S_0 \models &\neg[\{drink \leftrightarrow \top\}](drink \rightarrow sweat) \wedge \\ &[\{hot \leftrightarrow \top\}](hot \rightarrow sweat). \end{aligned}$$

This means that by positively intervening on the proposition "drink", the positive correlation between this proposition and the proposition "sweat" is broken. On the contrary, the positive correlation between the proposition "hot" and the proposition "sweat" is preserved after a positive intervention on "hot".

For the sake of exhaustiveness, let us also illustrate the notion of intervention in the context of the Example 3 about social influence.

Example 7 (Social influence (cont.)). *We have shown that in the equational state $S_0 = \big(\{\omega_1, \omega_2, \omega_3\}, \{pro(x) : x \in Agt\}\big)$ defined in Example 3, Bob cannot be against the use of nuclear energy and 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:*

$$\begin{aligned} S_0 \models &\langle\{pro(Cath) \leftrightarrow \top\}\rangle\neg pro(Bob) \wedge \\ &\langle\{pro(Cath) \leftrightarrow \top\}\rangle\big(pro(Cath) \wedge \neg pro(Ted)\big). \end{aligned}$$

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

$$S_0 \models \neg\langle\{pro(Cath) \leftrightarrow \top\}\rangle\big(\neg pro(Bob) \wedge pro(Ted)\big).$$

In his analysis of causal reasoning (Halpern 2000) and of actual causality (Halpern 2015, 2016), Halpern considers *acyclic* causal models, namely, causal models in which dependency relations between variables contain no cycles. Following Halpern, it is worth to consider equational states whose induced causal graphs are DAGs, as the one in the Example 2. As the following proposition highlights, for such states once the values of the exogenous variables and of the irrelevant variables are fixed, there is a unique causally compatible state after an intervention. Moreover, in virtue of the previous Fact 2, such a unique state must be equational. The proof of the proposition is given in Section A.9 in the technical appendix.

Proposition 5. *Let* $S = (C, V) \in \mathbf{S}_{Eq}$ *be an equational state such that its causal graph* Γ_S *is a DAG and let* $E \in Int_Z$ *for some* $Z \subseteq end(S)$*. Then, there is a unique* $S' = (C', V') \in \mathbf{S}$ *such that*

$$S \Rightarrow^E S' \text{ and } V \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big) = V' \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big).$$

We denote with S^E *such a unique state.*

In the following section we focus again on the model checking problem by generalizing it to the language of causal necessity post-intervention.

7.3 Model Checking

The model checking problem for the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ is analogous to the one for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ defined in Section 5. We verify whether a formula φ of the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ is true at a finite state S under a hard constraint ω_c:

Model checking for formulas in $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$
Given: $\varphi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$, $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $S \in \mathbf{S}_{Fin}(\omega_c)$.
Question: Do we have $\big(S, \mathbf{S}(\omega_c)\big) \models \varphi$?

Clearly, by virtue of Theorem 3 and of the fact that $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P}) \subset \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$, model checking for formulas in $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ is Θ_2^p-hard. We are going to show that the Θ_2^p-membership result given in Section 5.2.2 for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ generalizes to the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ in a natural way.

It is sufficient to generalize the translations tr_{S,ω_c} and tr^*_{C,ω_c} given in Section 5.2.2 to the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$, by adding the following case for the modality $[E]$:

$$tr_{S,\omega_c}([E]\varphi) = \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C^E} \omega \wedge \bigwedge_{\omega \in C^E} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\varphi) \setminus C^E} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C^E,\omega_c}(\varphi)\Big),$$

and

$$tr^*_{C,\omega_c}([E]\varphi) = \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C^E} \omega \wedge \bigwedge_{\omega \in C^E} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\varphi) \setminus C^E} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C^E,\omega_c}(\varphi)\Big),$$

with

$$C^E = \big(C \setminus Rev(C,E)\big) \cup E.$$

The following theorem is the analogue of Theorem 4 for the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$. Its proof is given in Section A.10 in the technical appendix.

Theorem 7. *Let* $\varphi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$*,* $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ *and* $S = (C, V) \in \mathrm{S}_{Fin}(\omega_c)$*. Then,*

$$\big(S, \mathrm{S}(\omega_c)\big) \models \varphi \text{ iff } \models_{\mathfrak{C}} tr_{S,\omega_c}(\varphi).$$

The following Theorem 8 follows from Theorem 7 and from the Θ_2^p-membership of Carnap validity checking similarly to how Theorem 5 followed from Theorem 4 and from the Θ_2^p-membership of Carnap validity checking in Section 5.2.2.

Theorem 8. *Model checking for formulas in* $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ *is in* Θ_2^p*.*

So, in the light of Theorem 8, we can affirm that extending the language of plain causal necessity with modalities of causal necessity post-intervention does not increase the complexity of the model checking problem.

Before concluding this section, we would like to consider a variant of the model checking problem for equational states. Similar to the model checking problems for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ defined above and for the language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ defined in Section 5, it consists in verifying whether $\big(S, \mathrm{S}_{Eq}(\omega_c)\big) \models \varphi$, where φ is a formula of the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ (resp. $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$), the context $\mathrm{S}_{Eq}(\omega)$ is the set of equational states satisfying the hard constraint $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, i.e., $\mathrm{S}_{Eq}(\omega_c) = \big\{S \in \mathrm{S}_{Eq} : S \models \omega_c\big\}$, and S is an equational state in $\mathrm{S}_{Eq}(\omega_c)$. This "equational" variant of the model checking problem is also Θ_2^p-complete both for $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ and for $\varphi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$. On the one hand, Θ_2^p-membership follows from Theorem 8. Indeed, checking whether $\big(S, \mathrm{S}_{Eq}(\omega_c)\big) \models \varphi$ with $S \in \mathrm{S}_{Eq}(\omega_c)$ and $\varphi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ is equivalent to checking whether $\big(S, \mathrm{S}(\omega_c)\big) \models \varphi$, because of Fact 2 in Section 7.2 that guarantees that if S is an equational state then any state reachable from it by means of a composition of the accessibility relations $\equiv$ and $\Rightarrow^E$ is also an equational state. On the other hand, Θ_2^p-hardness follows from i) Theorem 2 stating that that Carnap validity checking is poly-time reducible to model checking for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ relative to the model (S_0, S) with $S_0 = (\emptyset, \emptyset)$, ii) the fact that the empty state S_0 is an equational state, and iii) the fact that checking whether $(S_0, \mathrm{S}) \models \varphi$ is equivalent to checking $(S_0, \mathrm{S}_{Eq}) \models \varphi$, because of Fact 2.

7.4 Back to Structural Equational Models

In this section we are going to explore the connection between the rule-based semantics for causal necessity post-intervention defined in Section 7.2 and the standard SEM semantics. We specify the way in which a BCM is updated through an intervention. Our definition differs slightly from the update semantics used in the SEM literature, where interventions apply only to endogenous variables and the partition of variables into exogenous and endogenous sets remains unchanged 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, while in the standard SEM approach the partition between exogenous and endogenous variables is given at the meta-level and is causal model-invariant.

To define the notion of intervention-based BCM update, it is useful to introduce some additional notation. For each intervention $E \in Int$, we define val_E to be the partial function with domain $\mathbb{P}$ and codomain $\{0, 1\}$ such that, for all $p \in \mathbb{P}$,

$$val_E(p) = \begin{cases} 1, & \text{if } p \leftrightarrow \top \in E, \\ 0, & \text{if } p \leftrightarrow \bot \in E, \\ \text{undefined, otherwise.} \end{cases}$$

It is the value function of the intervention E.

Definition 13 (Intervention-based BCM update). *Let* $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p \in \mathbb{V}_{end}}\big)$ *be a BCM and* $E \in Int$*. The BCM resulting from the update of* Θ *by the intervention* E *is the tuple* $\Theta^E = \big(\mathbb{V}^E_{exo}, \mathbb{V}^E_{end}, (\mathcal{F}^E_p)_{p \in \mathbb{V}^E_{end}}\big)$ *such that,* *if* $E = \emptyset$ *then* $\Theta^E = \Theta$, *and if* $E = \{p_1 \leftrightarrow \tau_1, \ldots, p_k \leftrightarrow \tau_k\}$ *then:*

$$\begin{aligned}
&\mathbb{V}^E_{exo} = \mathbb{V}_{exo} \setminus \{p_1, \ldots, p_k\},\\
&\mathbb{V}^E_{end} = \mathbb{V}_{end} \cup \{p_1, \ldots, p_k\},\\
&\forall p \in \big(\mathbb{V}^E_{end} \setminus \{p_1, \ldots, p_k\}\big), \mathcal{F}^E_p = \mathcal{F}_p, \text{ and}\\
&\forall p \in \{p_1, \ldots, p_k\}, Range(\mathcal{F}^E_p) = \big\{val_E(p)\big\}.
\end{aligned}$$

The condition $Range(\mathcal{F}^E_p) = \{val_E(p)\}$ just means that $\mathcal{F}^E_p(\mathcal{I}_{\overline{p}}) = val_E(p)$ for every $\mathcal{I}_{\overline{p}} \in Asg_{\overline{p}}$.

Let us go back to the Example 4 of BCM with three atomic propositions we provided in Section 6.2 to illustrate Definition 13.

Example 8 (BCM for three atomic propositions (cont.)). *The result of updating the BCM* $\Theta_0 = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p \in \mathbb{V}_{end}}\big)$ *defined in Example 4 with the intervention* $E_1 = \{q \leftrightarrow \top\}$ *is the BCM* $\Theta_0^{E_1} = \big(\mathbb{V}^{E_1}_{exo}, \mathbb{V}^{E_1}_{end}, (\mathcal{F}^{E_1}_p)_{p \in \mathbb{V}^{E_1}_{end}}\big)$ *such that:*

$$\begin{array}{ll}
\mathbb{V}^{E_1}_{exo} = \{p\}, & \mathcal{F}^{E_1}_q(p \mapsto 1, r \mapsto 1) = 1,\\
\mathbb{V}^{E_1}_{end} = \{q, r\}, & \mathcal{F}^{E_1}_r(p \mapsto 0, q \mapsto 0) = 1,\\
\mathcal{F}^{E_1}_q(p \mapsto 0, r \mapsto 0) = 1, & \mathcal{F}^{E_1}_r(p \mapsto 0, q \mapsto 1) = 0,\\
\mathcal{F}^{E_1}_q(p \mapsto 0, r \mapsto 1) = 1, & \mathcal{F}^{E_1}_r(p \mapsto 1, q \mapsto 0) = 1,\\
\mathcal{F}^{E_1}_q(p \mapsto 1, r \mapsto 0) = 1, & \mathcal{F}^{E_1}_r(p \mapsto 1, q \mapsto 1) = 0.
\end{array}$$

Let us now consider the intervention $E_2 = \{p \leftrightarrow \bot, q \leftrightarrow \top\}$. *The result of updating the BCM* Θ_0 *by this intervention is the BCM* $\Theta_0^{E_2} = \big(\mathbb{V}^{E_2}_{exo}, \mathbb{V}^{E_2}_{end}, (\mathcal{F}^{E_2}_p)_{p \in \mathbb{V}^{E_2}_{end}}\big)$ *such that:*

$$\begin{array}{ll}
\mathbb{V}^{E_2}_{exo} = \emptyset, & \mathcal{F}^{E_2}_r(p \mapsto 0, q \mapsto 1) = 0,\\
\mathbb{V}^{E_2}_{end} = \{p, q, r\}, & \mathcal{F}^{E_2}_r(p \mapsto 1, q \mapsto 0) = 1,\\
\mathcal{F}^{E_2}_q(p \mapsto 0, r \mapsto 0) = 1, & \mathcal{F}^{E_2}_r(p \mapsto 1, q \mapsto 1) = 0,\\
\mathcal{F}^{E_2}_q(p \mapsto 0, r \mapsto 1) = 1, & \mathcal{F}^{E_2}_p(q \mapsto 1, r \mapsto 1) = 0,\\
\mathcal{F}^{E_2}_q(p \mapsto 1, r \mapsto 0) = 1, & \mathcal{F}^{E_2}_p(q \mapsto 1, r \mapsto 1) = 0,\\
\mathcal{F}^{E_2}_q(p \mapsto 1, r \mapsto 1) = 1, & \mathcal{F}^{E_2}_p(q \mapsto 1, r \mapsto 1) = 0,\\
\mathcal{F}^{E_2}_r(p \mapsto 0, q \mapsto 0) = 1, & \mathcal{F}^{E_2}_p(q \mapsto 1, r \mapsto 1) = 0.
\end{array}$$

Note that the intervention E_1 *does not modify the set of exogenous variables or the set of endogenous ones. On the contrary, the intervention* E_2 *transforms the variable* p *from exogenous to endogenous since it fixes its value to* $\bot$.

The notion of intervention-based BCM update of Definition 13 can be used to provide a semantic interpretation of the formulas of the language $\mathcal{L}_{\mathsf{PCN-Int}}(\mathbb{V})$, namely the language for propositional causal necessity *plus* interventions. We just need to extend the semantic interpretation of the formulas in the language $\mathcal{L}_{\mathsf{PCN}}(\mathbb{V})$ we provided in Section 7.1 with the following rule for the formula $[E]\omega$, with $\omega \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{V})$. Given a pointed BCM $(\Theta, \mathcal{I})$ with $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p \in \mathbb{V}_{end}}\big)$ and $\mathbb{V} = \mathbb{V}_{exo} \cup \mathbb{V}_{end}$:

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

Thus, it will be causally necessary that ω after the occurrence of the intervention E, if and only if ω holds at all solutions of the actual BCM updated by the intervention E.

Example 9 (BCM for three atomic propositions (cont.)). *Consider again the pointed BCM* $\big(\Theta_0, (p \mapsto 0, q \mapsto 0, r \mapsto 1)\big)$ *introduced in Example 4. It is easy to verify that at this pointed BCM, after the occurrence of the intervention* $E_1 = \{q \leftrightarrow \top\}$ *it is causally necessary that* $q \wedge \neg r$:

$$\big(\Theta_0, (p \mapsto 0, q \mapsto 0, r \mapsto 1)\big) \models [E_1](q \wedge \neg r).$$

The following theorem is a generalization of Theorem 6 to the language $\mathcal{L}_{\mathsf{PCN-Int}}(\mathbb{V})$: there exists a truth preserving onto relation between equational states and BCMs for the language extended by with causal necessity post-intervention modalities. Its proof is given in Section A.11 in the technical appendix.

Theorem 9. *There exists a surjection* $f : \mathbf{S}_{Eq} \longrightarrow \mathbf{B}$ *such that if* $f(S) = (\Theta, \mathcal{I})$ *then*

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

8 Causal Counterfactuals

This section concludes our analysis of causal reasoning by bringing causal counterfactual conditionals to the fore. We use our rule-based semantics to define a comparative similarity relation between states. 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.

There are two main motivations for supplementing our analysis with causal counterfactuals in Lewis' style. First of all, we want to show that our rule-based semantics offers a natural and flexible framework for elucidating the connection between causal counterfactuals and belief base revision and, more generally, between causal reasoning and epistemic reasoning. Secondly, we want to compare the notion of causal counterfactual conditional with the notion of causal necessity post-intervention introduced in Section 7 and to elucidate the conditions under which the latter can be seen as a special case of the former.

8.1 Semantics and Language

A natural way to compute the comparative similarity of two states with respect to a reference state is to consider the causal information that each state shares with the reference state. This idea is made precise by the following definition.

Definition 14 (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, denoted by* $S'' \preceq_S S'$, *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'', if the causal information that S'' shares with S is included in the causal information that S' shares with S or, said differently, the shared causal component of S'' and S is included in the shared causal component of S' and S. This means that state S' is at least as causally compatible with state S as state S''.

The relation $\preceq_S$ captures a form of *causal* comparative similarity insofar as it is determined exclusively by the causal components C, C' and C'' of the three states S, S' and S''. An alternative to it would be a form of *ontic* comparative similarity which is determined exclusively by the states' ontic components V, V' and V'': state S' is at least as similar to state S as state S'' if the ontic component of S'' differs from that of S at least as much as the ontic component of S' differs from that of S (i.e., $V \Delta V' \subseteq V \Delta V''$, with Δ standing for symmetric difference), under the assumption that the three states are regulated by the same causal theory (i.e., $S', S'' \in \equiv(S)$). A second alternative would be a combination of the previous two variants, a form of *mixed* comparative similarity, according to which state S' is at least as similar to state S as state S'' if i) the ontic component of S'' differs from that of S at least as much as the ontic component of S' differs from that of S (i.e., $V \Delta V' \subseteq V \Delta V''$), and ii) the shared causal component of S'' and S is included in the shared causal component of S' and S (i.e., $(C \cap C'') \subseteq (C \cap C')$).[5]

[5]The idea that comparative similarity between states can be computed by considering the states' two components is in line with Lewis' distinction between backtracking and non-backtracking counterfactuals (Lewis 1979). Roughly speaking, according to Lewis, in a backtracking

In this section, we focus on the causal variant of comparative similarity given in Definition 14 and do not consider the ontic and mixed variant since our primary aim is to elucidate the connection between causal reasoning and belief change. In belief change only the beliefs (which correspond to the causal information in a causal base) matter.[6] An in-depth analysis of the the mixed variant can be found in (Aguilera-Ventura et al. 2025).

On the syntactic side, we extend the modal language $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ introduced in Section 7.1 by counterfactual conditional modalities in Lewis' style of the form $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ which is read "if φ were true, ψ would be true", and whose dual $\varphi \mathrel{\Diamond\!\!\rightarrow} \psi =_{def} \neg(\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi)$ is read "if φ were true, ψ might be true". The resulting language is defined by the following grammar:

$$\varphi \quad ::= \quad \alpha \mid \top \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi \mid [E]\varphi \mid \varphi \mathrel{\Box\!\!\rightarrow} \varphi,$$

where α ranges over $\mathcal{L}_{\mathsf{CI}}(\mathbb{P})$ and $E \in \mathit{Int}$. The counterfactual conditional modalities are interpreted relative to a context $U \subseteq \mathbf{S}$ and a state $S \in U$, as follows:

$$(S, U) \models \varphi \mathrel{\Box\!\!\rightarrow} \psi \iff \forall S' \in U, \text{ if } S' \in \mathit{Closest}(\varphi, S, U) \text{ then } (S', U) \models \psi,$$

with

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

and ($S' \prec_S S''$ iff $S' \preceq_S S''$ and $S'' \npreceq_S S'$). $\mathit{Closest}(\varphi,S,U)$ is the set of closest φ-states to S in U, according to the similarity relation $\preceq_S$. According to the previous semantic interpretation, the conditional $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ holds at a given state S relative to the context U in case ψ is true at all the closest φ-states to S in U. For notational convenience, we write $\mathit{Closest}(\varphi,S)$ instead of $\mathit{Closest}(\varphi,S,\mathbf{S})$.

8.2 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 15 (ω-remainder). *Let $C \subseteq \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, $\omega \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $C' \subseteq C$. We say that C' is a ω-remainder of C if*

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

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

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 causal counterfactual conditionals and the notions of base contraction and revision. Its proof is given in Section A.12 in the technical appendix.

Theorem 10. *Let $S = (C, V) \in \mathbf{S}$ and $\omega_1, \omega_2 \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$. Then, the following three items are equivalent:*

(1) $S \models \omega_1 \mathrel{\Box\!\!\rightarrow} \omega_2$,

counterfactual only the propositional atoms representing the initial conditions can be changed to satisfy the antecedent of the conditional, while the causal laws are kept fixed. This corresponds to the *ontic* variant of comparative similarity. On the contrary, in a non-backtracking counterfactual, the causal laws can also be changed. This corresponds to the *causal* variant of comparative similarity given in Definition 14 and the *mixed* variant.

[6] The ontic component would become relevant if we compared causal reasoning to knowledge change. Unlike beliefs, knowledge is necessarily true in the objective world.

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

According to the previous theorem, declaring that a counterfactual statement $\omega_1 \Box\!\!\rightarrow \omega_2$ with propositional antecendent ω_1 and consequent ω_2 is true at a certain state is equivalent to saying that it is causally necessary, relative to every ω_1-remainder of the actual causal base, that ω_1 implies the ω_2. 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, 1991; Rott 1993; Val 1992). It is distinguished from slightly different operations such as WIDTIO ("When In Doubt Throw It Out") (Winslett 1990) which takes the intersection of the ω_1-remainders of the actual base first and then, after having added ω_1 to the resulting base, compute its deductive closure.[7]

The equivalence between items 1 and 3 in Theorem 10 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.

8.3 Relationship with Causal Necessity Post-Intervention

The second key result of this section is the following theorem that elucidates the connection between causal necessity post-intervention and causal counterfactual conditional. It pinpoints a condition under which the former can be reduced to the latter. Its proof is given in Section A.13 in the technical appendix.

Theorem 11. *Let $S \in \mathrm{S}_{Eq}$ and $\omega \in \mathcal{L}_{\mathrm{PROP}}(\mathbb{P})$. Then,*

$$(S, \mathrm{S}_{Eq}) \models \langle E \rangle \top \rightarrow \Big([E]\omega \leftrightarrow \big(\bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau) \Box\!\!\rightarrow \omega \big) \Big).$$

According to the previous theorem, when considering equational states, if there exists at least one state which is compatible with the actual state after an intervention E, then declaring that "it will be causally necessary that the propositional fact ω is true after the intervention E" is equivalent to declaring that "if the information pertaining to the intervention E were causally relevant, ω would be true". More generally, if we restrict to equational states — that as shown in Sections 6.2 and 7.4 correspond to BCMs — the existence of a solution after the intervention guarantees the reducibility of causal necessity post-intervention about a propositional fact to a causal counterfactual conditional. We go back to the social influence scenario of the Examples 3 and 7 to illustrate the previous theorem.

Example 10 (Social influence (cont.)). *In the Example 7 we have shown that*

$$S_0 \models \langle \{pro(Cath) \leftrightarrow \top\} \rangle \neg pro(Bob),$$

with $S_0 = \big(\{\omega_1, \omega_2, \omega_3\}, \{pro(x) : x \in Agt\}\big)$, which is equivalent to

$$S_0 \models \neg [\{pro(Cath) \leftrightarrow \top\}] pro(Bob).$$

Therefore, we also have

$$S_0 \models \langle \{pro(Cath) \leftrightarrow \top\} \rangle \top.$$

[7] The difference between a belief change operation and a belief change scheme was pointed out by (Nebel 1998). The main difference between the two is that the result of a belief base change operation is a belief base, whereas the result of a belief base change scheme is a theory (i.e., a deductively closed belief base).

Thus, thanks to Theorem 11, we have

$$S_0 \models \neg\Big(\triangle\big(pro(Cath) \leftrightarrow \top\big) \mathrel{\Box\!\!\rightarrow} pro(Bob)\Big),$$

which is equivalent to

$$S_0 \models \triangle\big(pro(Cath) \leftrightarrow \top\big) \mathrel{\Diamond\!\!\rightarrow} \neg pro(Bob).$$

The latter means that, at state S_0, if the information that Cath is in favor of the use of nuclear energy were causally relevant, Bob might be against the use of nuclear energy.

9 Conclusion

We hope we have successfully shown that our rule-based semantics offers a natural framework for (i) modeling some crucial concepts in the theory of causality including plain causal necessity, causal necessity post-intervention and causal counterfactuals, (ii) model checking causal properties, and (iii) elucidating the connection between causal reasoning and epistemic reasoning and, in particular, between causal counterfactuals and belief change. We have provided a number of complexity results for satisfiability and model checking for the different modal languages we presented. We have shown that in our framework model checking is intrinsically more complex than satisfiability checking because of its succinctness. For instance, while satisfiability checking for the language of causal information and necessity is NP-complete, model checking for this language and for its extension with causal necessity post-intervention modalities is Θ_2^p-complete. Directions of future research are manifold.

Proof-theory and model checking algorithms. We plan to study the proof-theoretic aspects of the two modal languages $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ and $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$. We expect the principles given in Proposition 2 to provide a sound and complete axiomatization for the language $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$. We also intend to explore the practical implications of our work. In particular, we plan to design model-checking algorithms for our modal languages, with the goal of automating causal reasoning. We have already taken some steps in this direction in a recent work (de Lima and Lorini 2024), where we proposed a QBF-based model-checking algorithm for a modal language of conditional causal necessity, in which the condition is a causal change operation. This language slightly generalizes the language of causal necessity post-intervention presented in Section 7. We have applied the algorithm to the verification of causality in legal cases (Liepina et al. 2025). In the light of the Θ_2^p-complete complexity result we presented, we believe this algorithm is non-optimal since TQBF (true quantified boolean formulas) is a PSPACE-complete problem.

Causal concepts. We also 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 6 and 9) guarantee that all these definitions can be formalized in our semantics as well, provided that the variables are binary. We have started to explore this direction in (de Lima and Lorini 2024) by showing that through our semantics we can express the notion of actual cause given in (Halpern 2015). In order to model causal explanation, we will extend our modal languages for causal reasoning with epistemic operators. Indeed, in agreement with (Halpern 2016; Halpern and Pearl 2005b), 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. This will require an integration of the semantics for causal reasoning we presented in Section 3 with the belief base semantics for epistemic logic we presented in (Lorini 2020).

Agency. As it is clear from the video game Example 1 given in Section 3.2, causality is closely connected to the notion of agency, namely, the fact that an agent (e.g., the human video gamer), causes a certain event to occur by choosing a certain action (e.g., the action of activating a key). We plan to explore the connection between our approach to causal reasoning and the logic STIT (the logic of "seeing to it that"), one of the most popular logics of agency (Belnap et al. 2001). We believe it is possible to interpret the atemporal version of STIT (the version of STIT with no temporal operators involved) (Balbiani et al. 2008; Lorini and Schwarzentruber 2011) through our semantics. Future work will be devoted to identifying the subclass of models of our semantics that correspond to atemporal STIT models.

Time. Last but not least, we plan to develop a temporal variant of our semantics. The problem of defining a semantics of causal reasoning in which time is explicit has been recently tackled in (Gladyshev et al. 2023, 2025). We will consider states in which causal information is expressed through formulas of linear temporal logic LTL instead of propositional logic and whose ontic component is a finite trace instead of a propositional valuation. In particular, we will consider states of the form $S = (C, \lambda)$ such that $C \subseteq \mathcal{L}_{\mathsf{LTL}}(\mathbb{P})$ and $\lambda = V_1 \ldots V_n$, where $V_k \subseteq \mathbb{P}$ for every $1 \leq k \leq n$ and $\mathcal{L}_{\mathsf{LTL}}(\mathbb{P})$ is the language of LTL defined by the following grammar:

$$\mathcal{L}_{\mathsf{LTL}}(\mathbb{P}) \quad \stackrel{\text{def}}{=} \quad \omega \quad ::= \quad \top \mid p \mid \neg\omega \mid \omega \wedge \omega \mid \mathsf{X}\omega \mid \omega \mathsf{U} \omega.$$

X and U are, respectively, the temporal operators 'next' and 'until'. States will have to satisfy the following condition which is analogous to the Condition (**CausalCons**) of Definition 1:

$$\forall \omega \in C,\ \lambda \models \omega, \qquad (\textbf{CausalCons}^*)$$

where $\lambda \models \omega$ is the usual interpretation of LTL formulas relative to finite traces, as given in (De Giacomo and Vardi 2013). The language of causal information $\mathcal{L}_{\mathsf{CI}}(\mathbb{P})$ defined in Section 3.2 will have to be redefined by assuming that ω ranges over $\mathcal{L}_{\mathsf{LTL}}(\mathbb{P})$ instead of $\mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$. Future work will be devoted to studying the complexity of model checking for the modal languages $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ and $\mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$ in this alternative version of the semantics with explicit time. We expect the problem to be PSPACE-complete.

Acknowledgments

This work is supported by the ANR project DesiRes "Designing Artificial Reasoners for Communication".

References

C. Aguilera-Ventura, X. Liu, E. Lorini, and D. Rozplokhas. 2025. "A Non-Interventionist Approach to Causal Reasoning Based on Lewisian Counterfactuals." In: *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2025)*. ijcai.org, 4301–4310.

C. Alchourrón, P. Gärdenfors, and D. Makinson. 1985. "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions." *The Journal of Symbolic Logic*, 50, 510–530.

N. Alechina, J. Y. Halpern, and B. Logan. 2017. "Causality, Responsibility and Blame in Team Plans." In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2017)*. ACM, 1091–1099.

G. Aleksandrowicz, H. Chockler, J. Y. Halpern, and A. Ivrii. 2017. "The Computational Complexity of Structure-Based Causality." *Journal of Artificial Intelligence Research*, 58, 431–451.

P. Balbiani, A. Herzig, and N. Troquard. 2008. "Alternative axiomatics and complexity of deliberative STIT theories." *Journal of Philosophical Logic*, 37(4), 387–406.

F. Barbero. 2024. "On the Logic of Interventionist Counterfactuals Under Indeterministic Causal Laws." In: *Proceedings of the 13th International Symposium on the Foundations of Information and Knowledge Systems (FoIKS 2024)* (Lecture Notes in Computer Science). Vol. 14589. Springer, 203–221.

F. Barbero and G. Sandu. 2021. "Team Semantics for Interventionist Counterfactuals: Observations vs. Interventions." *Journal of Philosophical Logic*, 50, 3, 471–521.

F. Barbero and F. Yang. 2020. "Counterfactuals and Dependencies on Causal Teams: Expressive Power and Deduction Systems." In: *Proceedings of the 13th Conference on Advances in Modal Logic (AiML 2020)*. College Publications, 73–93.

V. Batusov and M. Soutchanski. 2018. "Situation Calculus Semantics for Actual Causality." In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI 2018)*. AAAI Press, 1744–1752.

S. Beckers. 2025. "Nondeterministic Causal Models." In: *Proceedings of the 4th Conference on Causal Learning and Reasoning (CLeaR 2025)* (Proceedings of Machine Learning Research). Vol. 275. PMLR, 1532–1554.

S. Beckers. 2021. "The Counterfactual NESS Definition of Causation." In: *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*. AAAI Press, 6210–6217.

S. Beckers, H. Chockler, and J. Y. Halpern. 2024. "A Causal Analysis of Harm." *Minds and Machines*, 34, 3, 34.

S. Beckers and J. Vennekens. 2017. "The Transitivity and Asymmetry of Actual Causation." *Ergo*, 4, 1.

N. Belnap, M. Perloff, and M. Xu. 2001. *Facing the future: agents and choices in our indeterminist world*. Oxford University Press, New York.

P. Blackburn, M. de Rijke, and Y. Venema. 2001. *Modal Logic*. Cambridge University Press, Cambridge.

A. Bochman. 2003. "A Logic For Causal Reasoning." In: *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*. Morgan Kaufmann, 141–146.

A. Bochman. 2021. *A Logical Theory of Causality*. MIT Press.

A. Bochman. 2018a. "Actual Causality in a Logical Setting." In: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI 2018)*. ijcai.org, 1730–1736.

A. Bochman. 2018b. "On Laws and Counterfactuals in Causal Reasoning." In: *Proceedings of the Sixteenth International Conference on Principles of Knowledge Representation and Reasoning (KR 2018)*. AAAI Press, 494–503.

A. Bochman and V. Lifschitz. 2015. "Pearl's Causality in a Logical Setting." In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*. AAAI Press, 1446–1452.

A. W. Burks. 1951. "The Logic of Causal Propositions." *Mind*, 60, 239, 363–382.

R. Carnap. 1947. *Meaning and Necessity*. The University of Chicago Press.

H. Chockler and J. Y. Halpern. 2022. "On Testing for Discrimination Using Causal Models." In: *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22)*. AAAI Press, 5548–5555.

H. Chockler and J. Y. Halpern. 2004. "Responsibility and Blame: A Structural-Model Approach." *Journal of Artificial Intelligence Research*, 22, 93–115.

Z. Christoff and D. Grossi. 2017. "Stability in Binary Opinion Diffusion." In: *Proceedings of the 6th International Workshop on Logic, Rationality, and Interaction (LORI 2017)* (Lecture Notes in Computer Science). Vol. 10455. Springer, 166–180.

G. De Giacomo and M. Y. Vardi. 2013. "Linear Temporal Logic and Linear Dynamic Logic on Finite Traces." In: *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013)*. IJCAI/AAAI, 854–860.

T. de Lima and E. Lorini. 2024. "Model Checking Causality." In: *Proceedings of the Thirty-Thirth International Joint Conference on Artificial Intelligence, (IJCAI 2024)*. ijcai.org.

T. Eiter and G. Gottlob. 1997. "The Complexity Class Theta$_2$P: Recent Results and Applications in AI and Modal Logic." In: *Proceedings of the 11th International Symposium on Fundamentals of Computation Theory (FCT'97)* (LNCS). Vol. 1279. Springer, 1–18.

T. Eiter and T. Lukasiewicz. 2002. "Complexity results for structure-based causality." *Artificial Intelligence*, 142, 1, 53–89.

R. Fagin, J. D. Ullman, and M. Y. Vardi. 1983. "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*. ACM, 352–365.

M. Friedman. 2017. "Kant's Conception of Causal Necessity and Its Legacy." In: *Kant and the Laws of Nature*. Ed. by M. Massimi and A. Breitenbach. Cambridge University Press, 195–213.

D. Galles and J. Pearl. 1998. "An axiomatic characterization of causal counterfactuals." *Foundation of Science*, 3, 1, 151–182.

H. Geffner. 1990. "Causal Theories for Nonmonotonic Reasoning." In: *Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-90)*. AAAI Press / The MIT Press, 524–530.

M. L. Ginsberg. 1986. "Counterfactuals." *Artificial Intelligence*, 30, 1, 35–79.

E. Giunchiglia, J. Lee, V. Lifschitz, N. McCain, and H. Turner. 2004. "Nonmonotonic causal theories." *Artificial Intelligence*, 153, 1-2, 49–104.

M. Gladyshev, N. Alechina, M. Dastani, D. Doder, and B. Logan. 2023. "Dynamic Causality." In: *Proceedings of the 26th European Conference on Artificial Intelligence (ECAI 2023)* (Frontiers in Artificial Intelligence and Applications). Vol. 372. IOS Press, 867–874.

M. Gladyshev, N. Alechina, M. Dastani, D. Doder, and B. Logan. 2025. "Temporal Causal Reasoning with (Non-Recursive) Structural Equation Models." In: *Proceedings of the 39th Annual AAAI Conference on Artificial Intelligence (AAAI 2025)*. AAAI, 14949–14957.

G. Gottlob. 1995. "NP Trees and Carnap's Modal Logic." *Journal of the ACM*, 42, 2, 421–457.

E. Grädel and M. Otto. 1999. "On logics with two variables." *Theoretical Computer Science*, 224, 73–113.

U. Grandi, E. Lorini, and L. Perrussel. 2015. "Propositional Opinion Diffusion." In: *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*. ACM, 989–997.

J. Y. Halpern. 2015. "A Modification of the Halpern-Pearl Definition of Causality." In: *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)*. AAAI Press, 3022–3033.

J. Y. Halpern. 2016. *Actual causality*. MIT Press.

J. Y. Halpern. 2000. "Axiomatizing causal reasoning." *Journal of Artificial Intelligence Research*, 12, 317–337.

J. Y. Halpern. 2008. "Defaults and Normality in Causal Structures." In: *Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference (KR 2008)*. Ed. by G. Brewka and J. Lang. AAAI Press, 198–208.

J. Y. Halpern. 2013. "From Causal Models to Counterfactual Structures." *Review of Symbolic Logic*, 6, 2, 305–322.
J. Y. Halpern and M. Kleiman-Weiner. 2018. "Towards Formal Definitions of Blameworthiness, Intention, and Moral Responsibility." In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18)*. AAAI Press, 1853–1860.
J. Y. Halpern and J. Pearl. 2005a. "Causes and explanations: a structural-model approach. Part I: Causes." *British Journal for Philosophy of Science*, 56, 4, 843–887.
J. Y. Halpern and J. Pearl. 2005b. "Causes and explanations: a structural-model approach. Part II: Explanations." *British Journal for Philosophy of Science*, 56, 4, 889–911.
S. O. Hansson. 2009. "Replacement—A Sheffer Stroke for Belief Change." *Journal of Philosophical Logic*, 38, 2, 127–149.
S. O. Hansson. 1993. "Theory Contraction and Base Contraction Unified." *Journal of Symbolic Logic*, 58, 2, 602–625.
D. Harel D. ADN Kozen and J. Tiuryn. 2000. *Dynamic Logic*. MIT Press, Cambridge.
W. Hodges. 1997. "Compositional semantics for a language of imperfect information." *Logic Journal of the IGPL*, 5, 4, 539–563.
I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio. 2016. "Binarized Neural Networks." In: *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016*, 4107–4115.
B. Jenner and J. Torán. 1995. "Computing Functions with Parallel Queries to NP." *Theoretical Computer Science*, 141, 1&2, 175–193.
S. Kauffman. 1969. "Homeostasis and differentiation in random genetic control networks." *Nature*, 224, 5215, 177–178.
S. M. Khan, Y. Lespérance, and M. Rostamigiv. 2025. "Reasoning about Actual Causes in Nondeterministic Domains." In: *Proceedings of the 39th Annual AAAI Conference on Artificial Intelligence (AAAI 2025)*. AAAI.
S. M. Khan and M. Soutchanski. 2020. "Necessary and Sufficient Conditions for Actual Root Causes." In: *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*. Vol. 325. IOS Press, 800–808.
B. Kment. 2014. *Modality and Explanatory Reasoning*. Oxford University Press, Oxford.
D. K. Lewis. 1979. "Counterfactual Dependence and Time's Arrow." *Noûs*, 13, 4, 455–476.
D. K. Lewis. 1973. *Counterfactuals*. Harvard University Press, Cambridge.
D. K. Lewis. 1986. "Postscripts to "Causation"." In: *Philosophical Papers*. Vol. 2. Ed. by D. K. Lewis. Oxford University Press, 172–213.
R. Liepina, T. de Lima, E. Lorini, G. Pisano, and G. Sartor. 2025. "A Causal Model Checker for Legal Cases." In: *Proceedings of the 20th International Conference on Artificial Intelligence and Law (ICAIL 2025)*. ACM.
V. Lifschitz. 1997. "On the Logic of Causal Explanation (Research Note)." *Artificial Intelligence*, 96, 2, 451–465.
E. Lorini. 2023. "A Rule-Based Modal View of Causal Reasoning." In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI 2023)*. ijcai.org, 3286–3295.
E. Lorini. 2019. "Exploiting Belief Bases for Building Rich Epistemic Structures." In: *Proceedings of the Seventeenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2019)* (EPTCS). Vol. 297, 332–353.
E. Lorini. 2018. "In Praise of Belief Bases: Doing Epistemic Logic Without Possible Worlds." In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*. AAAI Press, 1915–1922.
E. Lorini. 2020. "Rethinking epistemic logic with belief bases." *Artificial Intelligence*, 282.
E. Lorini and F. Schwarzentruber. 2011. "A logic for reasoning about counterfactual emotions." *Artificial Intelligence*, 175, 3-4, 814–847.
N. McCain and H. Turner. 1997. "Causal Theories of Action and Change." In: *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI-97)*. AAAI Press / The MIT Press, 460–465.
N. Narodytska, S. P. Kasiviswanathan, L. Ryzhyk, M. Sagiv, and T. Walsh. 2018. "Verifying Properties of Binarized Deep Neural Networks." In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18)*. AAAI Press, 6615–6624.
B. Nebel. 1989. "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*. Morgan Kaufmann, 301–311.
B. Nebel. 1991. "Belief Revision and Default Reasoning: Syntax-Based Approaches." In: *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*. Morgan Kaufmann, 417–428.
B. Nebel. 1998. "How hard is it to revise a belief base?" In: *Handbook of Defeasible Reasoning and Uncertainty Management Systems*. Kluwer Academic. Vol. 3: Belief Change. Ed. by D. Dubois and H. Prade, 77–145.
J. Pearl. 2009. *Causality: Models, Reasoning and Inference*. Cambridge University Press.
H. Rott. 1993. "Belief Contraction in the Context of the General Theory of Rational Choice." *Journal of Symbolic Logic*, 58, 4, 1426–1450.
B. Russell. 1913. "On the Notion of Cause." *Proceedings of the Aristotelian Society*, 13, 1–26.
W. Shi, A. Shih, A. Darwiche, and A. Choi. 2020. "On Tractable Representations of Binary Neural Networks." In: *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020)*, 882–892.
R. Stalnaker. 1968. "A theory of conditionals." In: *Studies in Logical Theory*. Ed. by N. Rescher. Oxford University Press, 28–45.
B. Stroud. 1978. "Hume and the Idea of Causal Necessity." *Philosophical Studies*, 33, 1, 39–59.
A. del Val. 1992. "Computing Knowledge Base Updates." In: *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*. Morgan Kaufmann, 740–750.
K. W. Wagner. 1990. "Bounded Query Classes." *SIAM Journal on Computing*, 19, 5, 833–846.
M. Winslett. 1990. *Updating logical databases*. Cambridge University Press.

J. Woodward. 2003. *Making Things Happen: a Theory of Causal Explanation.* Oxford University Press.

J. Woodward and C. Hitchcock. 2003. "Explanatory generalizations, part I: A counterfactual account." *Noûs*, 37, 1, 1–24.

R. W. Wright. 1988. "Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof: Pruning the Bramble Bush by Clarifying the Concepts." *Iowa Law Review*, 73, 1001–1077.

J. Zhang. 2013. "A Lewisian Logic of Causal Counterfactuals." *Minds and Machines*, 23, 1, 77–93.

A Proofs

In this technical annex we present the proofs of the results presented in the paper.

A.1 Proof of Proposition 1

Proof. We are going to prove the equivalent statement $\big(S, \mathrm{S}(\omega_{\mathrm{c}})\big) \models \Diamond\neg\omega$ iff $\omega \notin Cn\big(C \cup \{\omega_{\mathrm{c}}\}\big)$. The latter is equivalent to say that $C \cup \{\omega_{\mathrm{c}}, \neg\omega\}$ is propositionally consistent.

In virtue of the strong completeness of propositional logic, propositional consistency of $C \cup \{\omega_{\mathrm{c}}, \neg\omega\}$ means that we can find a propositional valuation $V' \subseteq \mathbb{P}$ which satisfies all formulas in $C \cup \{\omega_{\mathrm{c}}, \neg\omega\}$. The latter implies that we can find a state $S' = (C', V') \in \mathbf{S}$ such that $C' = C \cup \{\omega_{\mathrm{c}}, \neg\omega\}$ and (i) $V' \models \omega_3$ for all $\omega_3 \in C'$. Let $S'' = (C'', V'')$ with $C'' = C$ and $V'' = V'$. We have (ii) $S'' \in \mathrm{S}(\omega_{\mathrm{c}})$ and $V'' \models \neg\omega$ because of (i) and the fact that the truth value of a propositional formula only depends on the component V'' of S'', which is equal to V'. Thus, there exists $S'' = (C'', V'') \in \mathrm{S}(\omega_{\mathrm{c}})$ such that $S \equiv S''$ and $V'' \models \neg\omega$. The latter implies that $\big(S, \mathrm{S}(\omega_{\mathrm{c}})\big) \models \Diamond\neg\omega$.

Let us prove the left-to-right direction. Suppose $\big(S, \mathrm{S}(\omega_{\mathrm{c}})\big) \models \Diamond\neg\omega$ with $S = (C, V)$. Thus, there exists $S'' = (C'', V'') \in \mathrm{S}(\omega_{\mathrm{c}})$ such that $S \equiv S''$ and $V'' \models \neg\omega$. The latter implies that there exists $V'' \subseteq \mathbb{P}$ such that $V'' \models \omega_3$ for all $\omega_3 \in C \cup \{\omega_{\mathrm{c}}, \neg\omega\}$. The latter means that $C \cup \{\omega_{\mathrm{c}}, \neg\omega\}$ is propositionally consistent. □

A.2 Proof of Proposition 2

Proof. We only prove validities (**T**$_\Box$), (**4**$_\Box$), (**5**$_\Box$), (**Mix1**$_{\triangle,\Box}$), (**Mix2**$_{\triangle,\Box}$) and (**Mix3**$_{\triangle,\Box}$). The proofs of validity (**T**$_\Box$) and property (**Nec**$_\Box$) are just straightforward and we do not give them.

Proof of validity (**T**$_\Box$). Let (S, U) be a model such that $(S, U) \models \Box\varphi$. The latter is equivalent to $\forall S' \in U,$ if $S \equiv S'$ then $(S', U) \models \varphi$. Thus, since $S \in U$ and $S \equiv S$, we have that $(S, U) \models \varphi$.

Proof of validity (**4**$_\Box$). Let (S, U) be a model such that $(S, U) \models \Box\varphi$. The latter is equivalent to $\forall S' \in U,$ if $S \equiv S'$ then $(S', U) \models \varphi$. We have that if $S \equiv S'$ then $\equiv(S) = \equiv(S')$, since $\equiv$ is an equivalence relation. Thus, $\forall S' \in U,$ if $S \equiv S'$ then $\big(\forall S'' \in U,$ if $S' \equiv S''$ then $(S'', U) \models \varphi\big)$. The latter is equivalent to $\forall S' \in U,$ if $S \equiv S'$ then $(S', U) \models \Box\varphi$, which is equivalent to $(S, U) \models \Box\Box\varphi$.

Proof of validity (**5**$_\Box$). Let (S, U) be a model such that $(S, U) \models \neg\Box\varphi$. The latter is equivalent to $\exists S' \in U$ such that $S \equiv S'$ and $(S', U) \models \neg\varphi$. We have that if $S \equiv S'$ then $\equiv(S) = \equiv(S')$, since $\equiv$ is an equivalence relation. Thus, $\forall S' \in U,$ if $S \equiv S'$ then $\big(\exists S'' \in U$ such that $S' \equiv S''$ and $(S'', U) \models \neg\varphi\big)$. The latter is equivalent to $\forall S' \in U,$ if $S \equiv S'$ then $(S', U) \models \neg\Box\varphi$, which is equivalent to $(S, U) \models \Box\neg\Box\varphi$.

Proof of validity (**Mix1**$_{\triangle,\Box}$). Let (S, U) be a model with $S = (C, V)$ such that $(S, U) \models \triangle\omega$. The latter is equivalent to $\omega \in C$. We have that if $S \equiv S'$ with $S' = (C', V')$ then $C = C'$. Thus, $\forall S' = (C', V') \in U,$ if $S \equiv S'$ then $\omega \in C'$. The latter implies that $\forall S' = (C', V') \in U,$ if $S \equiv S'$ then $V' \models \omega$, which is equivalent to $\forall S' = (C', V') \in U,$ if $S \equiv S'$ then $(S', U) \models \omega$. The latter is equivalent to $(S, U) \models \Box\omega$.

Proof of validity (**Mix2**$_{\triangle,\Box}$). Let (S, U) be a model with $S = (C, V)$ such that $(S, U) \models \triangle\omega$. The latter is equivalent to $\omega \in C$. We have that if $S \equiv S'$ with $S' = (C', V')$ then $C = C'$, since $\equiv$ is an equivalence relation. Thus, $\forall S' \in U,$ if $S \equiv S'$ then $(S', U) \models \triangle\omega$. The latter is equivalent to $(S, U) \models \Box\triangle\omega$.

Proof of validity (**Mix3**$_{\triangle,\Box}$). Let (S, U) be a model with $S = (C, V)$ such that $(S, U) \models \neg\triangle\omega$. The latter is equivalent to $\omega \notin C$. We have that if $S \equiv S'$ with $S' = (C', V')$ then $C = C'$, since $\equiv$ is an equivalence relation. Thus, $\forall S' \in U,$ if $S \equiv S'$ then $(S', U) \models \neg\triangle\omega$. The latter is equivalent to $(S, U) \models \Box\neg\triangle\omega$. □

A.3 Proof of Theorem 1

Proof. NP-hardness is evident.

In order to prove NP-membership, we show that if φ is satisfiable, then there exists $U \subseteq \mathbf{S}$ and $S = (C, V) \in U$ such that (S, U) satisfies φ and U has at most $|\varphi|$ states, where $|\varphi|$ is the *size* of φ, i.e., the number of symbols used to write it.

Let $U \subseteq \mathbf{S}$ and $S = (C, V) \in U$ such that $(S, U) \models \varphi$. We define (S, U') where $U' = \big(\equiv(S) \cap U\big)$.

Observe that $\big(\equiv(S') \cap U'\big) = U'$ for every $S' \in U'$.

By induction on the structure of ψ, it is easy to check that, for every $\psi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ and for every $S' \in U'$, $(S', U') \models \psi$ iff $(S', U) \models \psi$. Therefore, $(S, U') \models \varphi$ since $(S, U) \models \varphi$.

Let F_φ be the set of subformulas of φ of the form $\Box\psi$ for which $(S', U') \models \neg\Box\psi$ for all $S' \in U'$. Since $U' \neq \emptyset$, for each formula $\Box\psi \in F_\varphi$, there should be S_ψ in U' such that $(S_\psi, U') \models \neg\psi$.

We define (S, U'') where $U'' = \{S\} \cup \{S_\psi : \Box\psi \in F_\varphi\}$.

It is easy to check that U'' has at most $|\varphi|$ states since $|F_\varphi| \leq |sub(\varphi)| \leq |\varphi|$. By induction on the structure of φ, it is routine exercise to prove that, for all $S' \in U''$ and for all $\psi \in sub(\varphi)$, $(S', U') \models \psi$ iff $(S', U'') \models \psi$. The only non-trivial case is when ψ is of the form $\Box\chi$. We only prove the right-to-left direction, as the left-to-right one is straightforward. Suppose $(S', U') \models \neg\Box\chi$. Thus, $(S', U') \models \Box\neg\Box\chi$. Consequently, by definition of U', $(S'', U') \models \neg\Box\chi$ for all $S'' \in U'$. Hence, $\Box\chi \in F_\varphi$ and $(S_\chi, U') \models \neg\chi$. By construction, $S_\chi \in U''$ and, by induction hypothesis, $(S_\chi, U'') \models \neg\chi$. Since $S_\chi \in \big(\equiv(S') \cap U''\big)$, it follows that $(S', U'') \models \neg\Box\chi$.

Thus, we have $(S, U'') \models \varphi$, since $(S, U') \models \varphi$.

Now, let us prove that the satisfiability problem is in NP. We have shown that every satisfiable formula is satisfiable in a model which is polysize in $|\varphi|$.[8] Here is an non-deterministic algorithm to check if a given formula φ is satisfiable:

- guess non-deterministically a pair (S, U) such that $U \subseteq \mathbf{S}$, $S = (C, V) \in U$ and the size of U is bounded by $|\varphi|$,
- check whether $(S, U) \models \varphi$.

This algorithm non-deterministically runs in polynomial time. So, checking satisfiability of formulas in $\mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ is in NP. □

A.4 Proof of Proposition 3

Proof. We consider the function $g_\varphi : \mathbf{W}_\mathbb{P} \longrightarrow \mathbf{W}_\mathbb{P}^{fin}$ such that i) $g_\varphi(w) = w$ if $v \in \mathbf{W}_\mathbb{P}^{fin}$, and ii) $g_\varphi(w) = w \cap \mathbb{P}(\varphi)$ if $w \in (\mathbf{W}_\mathbb{P} \setminus \mathbf{W}_\mathbb{P}^{fin})$, where $\mathbb{P}(\varphi)$ is the set of atomic propositions in $\mathbb{P}$ occurring in φ. The function g_φ is clearly surjective. By induction on the structure of φ, we prove that, for every $w \in \mathbf{W}_\mathbb{P}$, $w \models \varphi$ iff $g_\varphi(w) \models^{fin} \varphi$. The case $\varphi = p$ and the Boolean cases are routine exercises. The following is the proof for the case $\varphi = \blacksquare\psi$:

$$
\begin{aligned}
w \models \blacksquare\psi \quad & \text{iff} \quad \forall v \in \mathbf{W}_\mathbb{P}, v \models \psi, \\
& \text{iff} \quad \forall v \in \mathbf{W}_\mathbb{P}, g_{\blacksquare\psi}(v) \models^{fin} \psi \text{ (by induction hypothesis)}, \\
& \text{iff} \quad \forall v \in \mathbf{W}_\mathbb{P}^{fin}, v \models^{fin} \psi \text{ (since } g_{\blacksquare\psi} \text{ is surjective)}, \\
& \text{iff} \quad g_\varphi(w) \models^{fin} \blacksquare\psi.
\end{aligned}
$$

[8]A model (S, U) is said to be polysize in $k \in \mathbb{N}$ if there is polynomial δ such that U contains at most $\delta(k)$ states (Blackburn et al. 2001, Definition 6.6).

The following is the proof for the case $\varphi = \blacklozenge\psi$:

$$\begin{aligned} w \models \blacklozenge\psi \quad & \text{iff} \quad \exists v \in \mathbf{W}_{\mathbb{P}} \text{ s.t. } v \models \psi, \\ & \text{iff} \quad \exists v \in \mathbf{W}_{\mathbb{P}} \text{ s.t. } g_{\blacksquare\psi}(v) \models^{fin} \psi \text{ (by induction hypothesis)}, \\ & \text{iff} \quad \exists v \in \mathbf{W}^{fin}_{\mathbb{P}} \text{ s.t. } v \models^{fin} \psi \text{ (since } g_{\blacksquare\psi} \text{ is surjective)}, \\ & \text{iff} \quad g_{\varphi}(w) \models^{fin} \blacklozenge\psi. \end{aligned}$$

We can use the function g_φ to show that "$\models_{\mathfrak{C}} \varphi$ iff $\models^{fin}_{\mathfrak{C}} \varphi$":

$$\begin{aligned} \models_{\mathfrak{C}} \varphi \quad & \text{iff} \quad \forall v \in \mathbf{W}_{\mathbb{P}}, v \models \varphi, \\ & \text{iff} \quad \forall v \in \mathbf{W}_{\mathbb{P}}, g_{\varphi}(v) \models^{fin} \varphi, \\ & \text{iff} \quad \forall v \in \mathbf{W}^{fin}_{\mathbb{P}}, v \models^{fin} \varphi \text{ (since } g_{\varphi} \text{ is surjective)}, \\ & \text{iff} \quad \models^{fin}_{\mathfrak{C}} \varphi. \end{aligned}$$

□

A.5 Proof of Theorem 2

Proof. We consider the function $f_0 : \mathbf{W}_{\mathbb{P}} \longrightarrow \equiv(S_0)$ where $S_0 = (\emptyset, \emptyset)$, $\equiv(S_0) = \{S \in \mathbf{S} : S_0 \equiv S\}$ and such that, for every $v \in \mathbf{W}_{\mathbb{P}}$ and for every $S = (C, V) \in \equiv(S_0)$, $f(v) = S$ iff $v = V$. It is straightforward to show that i) f_0 is a bijection. Moreover, by induction on the structure of φ, it is routine to show that ii) for every $v \in \mathbf{W}_{\mathbb{P}}$, we have "$v \models \varphi$ iff $f_0(v) \models tr(\varphi)$". Thus,

$$\begin{aligned} \models_{\mathfrak{C}} \varphi \quad & \text{iff} \quad \forall v \in \mathbf{W}_{\mathbb{P}}, v \models \varphi, \\ & \text{iff} \quad \forall v \in \mathbf{W}_{\mathbb{P}}, f_0(v) \models tr(\varphi) \text{ (by item ii)}, \\ & \text{iff} \quad \forall S \in \equiv(S_0), S \models tr(\varphi) \text{ (since } f_0 \text{ is a bijection)}, \\ & \text{iff} \quad S_0 \models \Box tr(\varphi). \end{aligned}$$

□

A.6 Proof of Lemma 1

Proof. The proof is by induction on the structure of φ.

Case $\varphi = p$, $\varphi = \top$ and $\varphi = \triangle\omega$. They are clear since $\top$ and $\neg\bot$ are Carnap valid.

Case $\varphi = \neg\psi$. By induction hypothesis, we have $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi)$ or $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi)$. The latter is equivalent to $\models_{\mathfrak{C}} \neg\neg tr_{S,\omega_c}(\psi)$ or $\models_{\mathfrak{C}} tr_{S,\omega_c}(\neg\psi)$ which in turn is equivalent to $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\neg\psi)$ or $\models_{\mathfrak{C}} tr_{S,\omega_c}(\neg\psi)$.

Case $\varphi = \psi_1 \wedge \psi_2$. By induction hypothesis, we have $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_1)$ or $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_1)$, and $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_2)$ or $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_2)$. Four sub-cases are possible.

Sub-case $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_1)$ and $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_2)$. The latter implies $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_1 \wedge \psi_2)$.

Sub-case $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_1)$ and $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_2)$. The latter implies $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_1 \wedge \psi_2)$.

Sub-case $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_1)$ and $\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi_2)$. The latter implies $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_1 \wedge \psi_2)$.

Sub-case $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_1)$ and $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_2)$. The latter implies $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi_1 \wedge \psi_2)$.

Case $\varphi = \Box\psi$. To prove this case it is sufficient to show that for every $\chi \in \mathcal{L}_{\mathfrak{C}}$, we have $\models_{\mathfrak{C}} \blacksquare\chi$ or $\models_{\mathfrak{C}} \neg\blacksquare\chi$. Suppose toward a contradiction that $\not\models_{\mathfrak{C}} \blacksquare\chi$ and $\not\models_{\mathfrak{C}} \neg\blacksquare\chi$. The latter means that $\exists v, u \in \mathbf{W}_{\mathbb{P}_\triangle}$ such that $v \models \blacklozenge\neg\chi$ and $u \models \blacksquare\chi$. The latter implies that $\exists v' \in \mathbf{W}_{\mathbb{P}_\triangle}$ such that $v' \models \neg\chi$ and $\forall u' \in \mathbf{W}_{\mathbb{P}_\triangle}, u' \models \chi$. The latter leads to a contradiction. □

A.7 Proof of Theorem 4

Proof. Let $\varphi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$, $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and $S = (C, V) \in \mathbf{S}_{Fin}(\omega_c)$. The proof of "$\big(S, \mathbf{S}(\omega_c)\big) \models \varphi$ iff $\models_{\mathfrak{C}} tr_{S,\omega_c}(\varphi)$" is by induction on the structure of φ.

Case $\varphi = p$. We distinguish two sub-cases: $p \in V$ and $p \notin V$. Suppose $p \in V$. Clearly, $\big(S, \mathbf{S}(\omega_c)\big) \models p$. Moreover, $tr_{S,\omega_c}(p) = \top$ and, clearly $\models_{\mathfrak{C}} \top$. Thus, $\models_{\mathfrak{C}} tr_{S,\omega_c}(p)$. In an analogous way, it can be shown that if $p \notin V$ then $\big(S, \mathbf{S}(\omega_c)\big) \not\models p$ and $\not\models_{\mathfrak{C}} tr_{S,\omega_c}(p)$.

Case $\varphi = \triangle\omega$. The proof is analogous to the proof of the previous case.

Case $\varphi = \neg\psi$. To prove this case we rely on Lemma 1 we stated above. We first prove the left-to-right direction. Suppose $\big(S, \mathbf{S}(\omega_c)\big) \models \neg\psi$. The latter is equivalent to $\big(S, \mathbf{S}(\omega_c)\big) \not\models \psi$ that, by induction hypothesis, is equivalent to $\not\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi)$. By Lemma 1, the latter implies $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi)$ which is equivalent to $\models_{\mathfrak{C}} tr_{S,\omega_c}(\neg\psi)$. As for the right-to-left direction, suppose $\models_{\mathfrak{C}} tr_{S,\omega_c}(\neg\psi)$. The latter is equivalent to $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\psi)$ which in turn implies $\not\models_{\mathfrak{C}} tr_{S,\omega_c}(\psi)$. By induction hypothesis, the latter is equivalent to $\big(S, \mathbf{S}(\omega_c)\big) \not\models \psi$ which in turn is equivalent to $\big(S, \mathbf{S}(\omega_c)\big) \models \neg\psi$.

Case $\varphi = \Box\psi$. To prove this case some preliminary notions and results are needed. Let

$$\equiv(S, \omega_c) = \{S' \in \mathbf{S}(\omega_c) : S \equiv S'\}.$$

Moreover, given a state $S = (C, V) \in \mathbf{S}$, a formula $\psi \in \mathcal{L}_{\mathsf{CIN}}(\mathbb{P})$ and a hard constraint $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, we define the following set of (C, ψ, ω_c)-relevant worlds from $\mathbf{W}_{\mathbb{P}_\triangle}$:

$$\begin{aligned}\mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c) = \Big\{v \in \mathbf{W}_{\mathbb{P}_\triangle} : \ & v \models \omega_c,\\ & \forall \omega \in C, v \models \omega,\\ & \{p_{\triangle\omega} : \omega \in C\} \subseteq v,\\ & \text{and } \big\{p_{\triangle\omega} : \omega \in (\Gamma(\psi) \setminus C)\big\} \cap v = \emptyset\Big\},\end{aligned}$$

and the following function

$$f_S : \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c) \longrightarrow \equiv(S, \omega_c)$$

such that, for all $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c)$ and for all $S' = (C', V') \in \equiv(S, \omega_c)$,

$$f_S(v) = S' \text{ iff } V' = \big(v \cap \mathbb{P}\big).$$

The function f_S maps (C, ψ, ω_c)-relevant worlds into S-causally compatible states satisfying the hard constraint ω_c.

Lemma 2. *Let $S = (C, V) \in \mathbf{S}$. Then, the function f_S is total and surjective.*

Proof. Consider an arbitrary $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c)$. We have that $v \models \omega$ for all $\omega \in C \cup \{\omega_c\}$. The latter guarantees that there exists $S' = (C', V') \in \mathbf{S}$ such that $C' = C$ and $V' = \big(v \cap \mathbb{P}\big)$. Thus, there exists $S' = (C', V') \in \equiv(S, \omega_c)$ such that $V' = \big(v \cap \mathbb{P}\big)$. Hence, there exists $S' = (C', V') \in \equiv(S, \omega_c)$ such that $f_S(v) = S'$. This means that f_S is total.

Let us prove surjectivity. Consider an arbitrary $S' = (C', V') \in \equiv(S, \omega_c)$. We have that $\forall \omega \in C \cup \{\omega_c\}$, $V' \models \omega$ since $C' = C$ and $S' \in \mathbf{S}(\omega_c)$. This implies that there exists $v \in \mathbf{W}_{\mathbb{P}_\triangle}$ such that $\forall \omega \in C \cup \{\omega_c\}, v \models \omega$ and $V' = \big(v \cap \mathbb{P}\big)$. (The fact that in virtue of Proposition 3 we consider $\mathbf{W}_{\mathbb{P}_\triangle}$ instead of $\mathbf{W}^{fin}_{\mathbb{P}}$ is essential here to guarantee the existence of such v.) The latter implies that there exists $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c)$ such that $V' = \big(v \cap \mathbb{P}\big)$. Hence, there exists $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c)$ such that $f_S(v) = S'$. □

The following lemma is crucial to prove our last case. It relies on Lemma 2.

Lemma 3. *Let* $S = (C, V) \in \mathrm{S}_{Fin}(\omega_c)$ *and* $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c)$. *Then,*

$$v \models tr^*_{C,\omega_c}(\psi) \text{ iff } \big(f_S(v), \mathrm{S}(\omega_c)\big) \models \psi.$$

Proof. Suppose $S = (C, V) \in \mathrm{S}_{Fin}(\omega_c)$ and $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c)$. We prove that "$v \models tr^*_{C,\omega_c}(\psi)$ iff $\big(f_S(v), \mathrm{S}(\omega_c)\big) \models \psi$" by induction on the structure of ψ.

Case $\psi = p$. Suppose $f_S(v) = S' = (C', V')$. Thus, $V' = \big(v \cap \mathbb{P}\big)$. Hence, "$v \models tr^*_{C,\omega_c}(p)$ iff $\big(f_S(v), \mathrm{S}(\omega_c)\big) \models p$" since $tr^*_{C,\omega_c}(p) = p$.

Case $\psi = \triangle\omega$. Suppose $f_S(v) = S' = (C', V')$. Since $f_S(v) \in \equiv (S, \omega_c)$, we have $C' = C$. Thus, since $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \triangle\omega, \omega_c)$, we have "$v \models p_{\triangle\omega}$ iff $\omega \in C'$". Therefore, we have "$v \models p_{\triangle\omega}$ iff $f_S(v) \models \triangle\omega$ iff $\big(f_S(v), \mathrm{S}(\omega_c)\big) \models \triangle\omega$". Thus, since $tr^*_{C,\omega_c}(\triangle\omega) = p_{\triangle\omega}$, we have "$v \models tr^*_{C,\omega_c}(\triangle\omega)$ iff $\big(f_S(v), \mathrm{S}(\omega_c)\big) \models \triangle\omega$".

Cases $\psi = \neg\chi$ and $\psi = \chi_1 \wedge \chi_2$. Their proof is straightforward and we can simply omit it.

Case $\psi = \Box\chi$. The case is proved as follows:

$$\begin{aligned}
v \models tr^*_{C,\omega_c}(\Box\chi) \quad & \text{iff} \quad v \models \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C} \omega \wedge \bigwedge_{\omega \in C} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\chi)\setminus C} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C,\omega_c}(\chi)\Big),\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \chi, \omega_c), u \models tr^*_{C,\omega_c}(\chi),\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \chi, \omega_c), \big(f_S(u), \mathrm{S}(\omega_c)\big) \models \chi \text{ (by induction hypothesis)},\\
& \text{iff} \quad \forall S' \in \equiv (S, \omega_c), \big(S', \mathrm{S}(\omega_c)\big) \models \chi \text{ (since, according to Lemma 2,}\\
& \qquad\quad f_S \text{ is surjective)},\\
& \text{iff} \quad \forall S' \in \equiv\big(f_S(v), \omega_c\big), \big(S', \mathrm{S}(\omega_c)\big) \models \chi\\
& \qquad\quad \text{(since } \equiv\big(S, \omega_c\big) = \equiv\big(f_S(v), \omega_c\big) \text{ due to the fact that } f_S(v) \in \equiv\big(S, \omega_c\big)\text{)},\\
& \text{iff} \quad \big(f_S(v), \mathrm{S}(\omega_c)\big) \models \Box\chi.
\end{aligned}$$

□

Now we have everything we need to prove the case $\varphi = \Box\psi$:

$$\begin{aligned}
\big(S, \mathrm{S}(\omega_c)\big) \models \Box\psi \quad & \text{iff} \quad \forall S' \in \equiv (S, \omega_c), \big(S', \mathrm{S}(\omega_c)\big) \models \psi,\\
& \text{iff} \quad \forall v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c), v \models tr^*_{C,\omega_c}(\psi) \text{ (by Lemma 3 and since, according}\\
& \qquad\quad \text{to Lemma 2, } f_S \text{ is surjective)},\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle},\\
& \qquad\quad u \models \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C} \omega \wedge \bigwedge_{\omega \in C} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\psi)\setminus C} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C,\omega_c}(\psi)\Big),\\
& \text{iff} \quad \models_{\mathfrak{C}} tr_{S,\omega_c}(\Box\psi).
\end{aligned}$$

□

A.8 Proof of Theorem 6

Proof. Let $S = (C, V)$ be an equational state. We define the pointed BCM $f(S) = (\Theta, \mathcal{I})$ with $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p \in \mathbb{V}_{end}}\big)$ as follows:

- $\mathbb{V}_{exo} = exo(S)$,
- $\mathbb{V}_{end} = end(S)$,
- $\forall p \in \mathbb{V}_{end}, \forall \mathcal{I}_{\overline{p}} \in Asg_{\overline{p}}$,

$$\mathcal{F}_p(\mathcal{I}_{\overline{p}}) = 1 \text{ iff } \exists p \leftrightarrow \omega \in C \text{ such that } \mathcal{I}_{\overline{p}} \in |\omega|_{Asg_{\overline{p}}}.$$

- $\forall q \in \mathbb{V}, \mathcal{I}(q) = 1$ iff $q \in V$.

We are going to show that f is a surjection from $\mathbf{S}_{Eq}$ to $\mathbf{B}$. In order to prove this, we consider an arbitrary $(\Theta, \mathcal{I}) \in \mathbf{B}$ with $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p \in \mathbb{V}_{end}}\big)$. We show that there exists $S^{(\Theta,\mathcal{I})} = (C^{(\Theta,\mathcal{I})}, V^{(\Theta,\mathcal{I})}) \in \mathbf{S}_{Eq}$ such that $f(S^{(\Theta,\mathcal{I})}) = (\Theta, \mathcal{I})$.

For every $p \in \mathbb{V}_{end}$, we take the canonical DNF (CDNF) expression of the Boolean function $\mathcal{F}_p$ built from the set of variables $\mathbb{V} \setminus \{p\}$ and note $cdnf_{\mathcal{F}_p}$ it. We define

$$C^{(\Theta,\mathcal{I})} = \{p \leftrightarrow cdnf_{\mathcal{F}_p} : p \in \mathbb{V}_{end}\},$$
$$V^{(\Theta,\mathcal{I})} = \{q \in \mathbb{P} : \mathcal{I}(q) = 1\}.$$

Since $|cdnf_{\mathcal{F}_p}|_{Asg_{\overline{p}}} = \big\{\mathcal{I}_{\overline{p}} \in Asg_{\overline{p}} : \mathcal{F}_p(\mathcal{I}_{\overline{p}}) = 1\big\}$, we have that $f(S^{(\Theta,\mathcal{I})}) = (\Theta, \mathcal{I})$.

Now, suppose that $f(S) = (\Theta, \mathcal{I})$ with $S = (C, V)$ an equational state and $\Theta = \big(\mathbb{V}_{exo}, \mathbb{V}_{end}, (\mathcal{F}_p)_{p \in \mathbb{V}_{end}}\big)$. By induction on the structure of φ, we are going to verify that $\forall \varphi \in \mathcal{L}_{\mathsf{PCN}}(\mathbb{V}), S \models \varphi$ iff $(\Theta, \mathcal{I}) \models \varphi$.

The following proposition will be useful for rest of the proof.

Proposition 6. *Let $S \in \mathbf{S}_{Eq}$ and $f(S) = (\Theta, \mathcal{I})$. Then,*

- *if $\mathcal{I}' \in Sol(\Theta)$ then $\exists S' = (C', V') \in \mathbf{S}$ s.t. $S \equiv S'$ and $V' \cap \mathbb{V} = \big\{q \in \mathbb{V} : \mathcal{I}'(q) = 1\big\}$,*
- *if $S \equiv S' = (C', V')$ then $\exists \mathcal{I}' \in Sol(\Theta)$ s.t. $V' \cap \mathbb{V} = \big\{q \in \mathbb{V} : \mathcal{I}'(q) = 1\big\}$.*

Suppose $\varphi = q$. Then, $S \models q$ iff $q \in V$ iff $\mathcal{I}(q) = 1$ iff $(\Theta, \mathcal{I}) \models q$. Boolean cases are just routine. Let us prove the case $\varphi = \Box\omega$.

We first prove the left-to-right direction. Suppose, towards a contradiction, that $S \models \Box\omega$ and $(\Theta, \mathcal{I}) \models \Diamond\neg\omega$. The latter means that $Sol(\Theta) \cap |\neg\omega|_{Asg} \neq \emptyset$. By the first item of Proposition 6, the latter implies that $\exists \mathcal{I}' \in Sol(\Theta) \cap |\neg\omega|_{Asg}, \exists S' = (C', V') \in \mathbf{S}$ such that $S \equiv S'$ and $V' \cap \mathbb{V} = \big\{q \in \mathbb{V} : \mathcal{I}'(q) = 1\big\}$. Since $\omega \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{V})$, the latter implies that $\exists S' \in \mathbf{S}$ such that $S \equiv S'$ and $S' \models \neg\omega$. The latter means that $S \models \Diamond\neg\omega$ which leads to a contradiction.

Let us now prove the right-to-left direction. Suppose, towards a contradiction, that $(\Theta, \mathcal{I}) \models \Box\omega$ and $S \models \Diamond\neg\omega$. The latter means that $\exists S' = (C', V') \in \mathbf{S}$ such that $S \equiv S'$ and $S' \models \neg\omega$. By the second item of Proposition 6, the latter implies that $\exists \mathcal{I}' \in Sol(\Theta), \exists S' = (C', V') \in \mathbf{S}$ such that $S \equiv S', S' \models \neg\omega$ and $V' \cap \mathbb{V} = \big\{q \in \mathbb{V} : \mathcal{I}'(q) = 1\big\}$. Since $\omega \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{V})$, the latter implies that $Sol(\Theta) \cap |\neg\omega|_{Asg} \neq \emptyset$. The latter means that $(\Theta, \mathcal{I}) \models \Diamond\neg\omega$ which leads to a contradiction. □

A.9 Proof of Proposition 5

Proof. The proof of the "at most" part is the most convoluted one and we prove it first. Let $S = (C, V)$ be an equational state such that Γ_S is a DAG and let $E \in Int_Z$ for some $Z \subseteq end(S)$. Moreover, let $S' = (C', V'), S'' = (C'', V'') \in \mathbf{S}$ such that $S \Rightarrow^E S', S \Rightarrow^E S''$ and

$$(1)\ V \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big) = V' \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big) = V'' \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big).$$

Since S is an equational state, by Fact 2, S' and S'' are equational states too. From the fact that Γ_S is a DAG, $S \Rightarrow^E S'$ and $S \Rightarrow^E S''$ we have that

$$(2)\ C' = C'' \text{ and } \Gamma_{S'} = \Gamma_{S''} \text{ is a DAG too.}$$

Moreover:

$$(3)\ \mathbb{P}(C) = \mathbb{P}(C') = \mathbb{P}(C''),$$
$$exo(S) = exo(S') = exo(S''),$$
$$end(S) = end(S') = end(S'').$$

Since $\Gamma_{S'}$ is a DAG we can build the following ranking of the elements of $N_{S'}$:

$$Rank^0_{S'} = \Big\{p \in N_{S'} : \nexists q \in N_{S'} \text{ s.t. } p \neq q \text{ and } q \in \mathcal{P}_{S'}(p)\Big\},$$
$$Rank^{k+1}_{S'} = \Big\{p \in N_{S'} : \exists q \in Rank^k_{S'} \text{ s.t. } q \neq p \text{ and } q \in \mathcal{P}_{S'}(p)$$
$$\text{and } \forall q' \in \big(N_{S'} \setminus \bigcup_{k' \leq k} Rank^{k'}_{S'}\big), q' \notin \mathcal{P}_{S'}(p)\Big\}.$$

It is easy to verify that rank 0 contains only and all the exogenous variables in S' *plus* eventually the symbol $\top$ if $\top \in N_{S'}$. So, once we know the values of the endogenous variables in S' (i.e., $V' \cap exo(S')$), we can directly compute the values of each variable p of rank 1 using the equational formula $p \leftrightarrow \omega$ in C' and knowing that ω only contains variables of rank 0. Then, we can compute values of each variable of rank 2: we use the equational formula $p \leftrightarrow \omega$ in C', knowing that ω only contains variables of rank strictly higher than 2 for which we computed the values at the previous steps. And so on so forth for the ranks higher than 2. At the end of the process, which is finite since $N_{S'}$ is finite, the values for all variables in $\mathbb{P}(C')$ will be computed.

Because of (1) and (3), we have that S' and S'' have the same exogenous variables and assign the same values to them. Moreover, (2) guarantees that the same process can be applied to S'' so as to obtain the same values for the variables in $\mathbb{P}(C'') = \mathbb{P}(C')$. Item (1) guarantee that S' and S'' assign the same truth values to the variables that are not in $\mathbb{P}(C)$. Thus, again by (3), we can conclude that $V' = V''$.

The latter together with (2) guarantee that $S' = S''$.

As for the "at least" part, we just need to show that given an equational state $S = (C, V)$ whose causal graph is a DAG and $E \in Int_Z$ for some $Z \subseteq end(S)$, there exists at least one state $S' = (C', V')$ such that $S \Rightarrow^E S'$ and

$$V \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big) = V' \cap \Big(exo(S) \cup \big(\mathbb{P} \setminus \mathbb{P}(C)\big)\Big).$$

The intervention E transforms the equational causal base C of the state S into a new equational causal base C' whose endogenous and exogenous variables are the same as in C. It is routine to verify that since Γ_S is a DAG, the causal graph induced by C' is a DAG too. In the rest of the proof, we just need to use the same method as the one used in the proof of the "at most" part. We construct the ranking of the elements of the DAG induced by C'. Then, we assign to the exogenous variables the values they take in C and, step after step, we compute the values for the variables in all ranks. At the end of the process, the values for all variables in $\mathbb{P}(C') = \mathbb{P}(C)$ have been computed. We assign to the variables in $\mathbb{P} \setminus \mathbb{P}(C')$ the values they take at S. We note V' the resulting valuation. The pair (C', V') is the state we looked for.

□

A.10 Proof of Theorem 7

Proof. Given $S = (C, V) \in \mathrm{S}_{Fin}(\omega_c)$ and $\varphi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$, we have to prove that "$\big(S, \mathrm{S}(\omega_c)\big) \models \varphi$ iff $\models_{\mathfrak{C}} tr_{S,\omega_c}(\varphi)$". The proof is again by induction on the structure of φ.

The proofs of the cases $\varphi = p$, $\varphi = \top$, $\varphi = \triangle\omega$ and $\varphi = \Box\psi$ are identical to those given in the proof of Theorem 4.

Case $\varphi = \neg\psi$. The proof of this case relies on the fact that for every $\varphi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$, we have $\models_{\mathfrak{C}} tr_{S,\omega_c}(\varphi)$ or $\models_{\mathfrak{C}} \neg tr_{S,\omega_c}(\varphi)$. This fact can be proved in a way analogous to Lemma 1 in Section 5.2.2.

Case $\varphi = [E]\psi$. The proof of this case follows the lines of the proof of the case $\varphi = \Box\psi$ in the proof of Theorem 4. Some preliminary notions and results are needed. Given a state $S = (C, V) \in \mathbf{S}$, a formula $\psi \in \mathcal{L}_{\mathsf{CIN-Int}}(\mathbb{P})$, a hard constraint $\omega_c \in \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$ and a sequence of interventions $\pi = E_1 \dots E_k$ from *Int*, we define the following set of (C, ψ, ω_c, π)-relevant worlds from $\mathbf{W}_{\mathbb{P}_\triangle}$:

$$\begin{aligned}\mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi) = \Big\{ v \in \mathbf{W}_{\mathbb{P}_\triangle} : {} & v \models \omega_c, \\ & \forall \omega \in C^\pi, v \models \omega, \\ & \{p_{\triangle\omega} : \omega \in C^\pi\} \subseteq v, \\ & \text{and } \big\{p_{\triangle\omega} : \omega \in (\Gamma(\psi) \setminus C^\pi)\big\} \cap v = \emptyset \Big\},\end{aligned}$$

where

$$\begin{aligned}C^E &= \big(C \setminus Rev(C,E)\big) \cup E, \\ C^{E\pi} &= (C^E)^\pi.\end{aligned}$$

Moreover, we define the following function

$$f_S^\pi : \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi) \longrightarrow \Rightarrow^\pi(S, \omega_c)$$

with

$$\Rightarrow^\pi(S, \omega_c) = \big\{S' = (C', V') \in \mathbf{S}(\omega_c) : C' = C^\pi\big\},$$

such that, for all $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)$ and for all $S' = (C', V') \in \Rightarrow^\pi(S, \omega_c)$,

$$f_S^\pi(v) = S' \text{ iff } V' = \big(v \cap \mathbb{P}\big).$$

The function f_S^π maps (C, ψ, ω_c, π)-relevant worlds into S-causally compatible states after the occurrence of the sequence of interventions π satisfying the hard constraint ω_c. The following Lemma 4 is the analogue of Lemma 2 in the proof of Theorem 4.

Lemma 4. *Let $S = (C, V) \in \mathbf{S}$ and $\pi = E_1 \dots E_k$ a sequence of interventions from Int. Then, the function f_S^π is total and surjective.*

Proof. Consider an arbitrary $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)$. We have that $v \models \omega$ for all $\omega \in C^\pi \cup \{\omega_c\}$. The latter guarantees that there exists $S' = (C', V') \in \mathbf{S}(\omega_c)$ such that $C' = C^\pi$ and $V' = \big(v \cap \mathbb{P}\big)$. Thus, there exists $S' = (C', V') \in\Rightarrow^\pi(S, \omega_c)$ such that $V' = \big(v \cap \mathbb{P}\big)$. Hence, there exists $S' = (C', V') \in\Rightarrow^\pi(S, \omega_c)$ such that $f_S^\pi(v) = S'$. This means that f_S^π is total.

Let us prove surjectivity. Consider an arbitrary $S' = (C', V') \in\Rightarrow^\pi(S, \omega_c)$. We have that $\forall \omega \in C^\pi \cup \{\omega_c\}, V' \models \omega$ since $C' = C^\pi$ and $S' \in \mathbf{S}(\omega_c)$. This implies that there exists $v \in \mathbf{W}_{\mathbb{P}_\triangle}$ such that $\forall \omega \in C^\pi \cup \{\omega_c\}, v \models \omega$ and $V' = \big(v \cap \mathbb{P}\big)$. (The fact that in virtue of Proposition 3 we consider $\mathbf{W}_{\mathbb{P}_\triangle}$ instead of $\mathbf{W}_{\mathbb{P}}^{fin}$ is essential here to guarantee the existence of such v.) The latter implies that there exists $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)$ such that $V' = \big(v \cap \mathbb{P}\big)$. Hence, there exists $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)$ such that $f_S^\pi(v) = S'$. □

The following lemma, which is provable by means of Lemma 4, is crucial for the proof of our last case. It is the analogue of Lemma 3 in the proof of Theorem 4.

Lemma 5. *Let $S = (C, V) \in \mathbf{S}_{Fin}(\omega_c)$, $\pi = E_1 \dots E_k$ a sequence of interventions from Int, and $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)$. Then,*

$$v \models tr^*_{C^\pi, \omega_c}(\psi) \textit{ iff } \big(f_S^\pi(v), \mathbf{S}(\omega_c)\big) \models \psi.$$

Proof. Suppose $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)$. We prove that "$v \models tr^*_{C^\pi,\omega_c}(\psi)$ iff $\big(f^\pi_S(v), \mathrm{S}(\omega_c)\big) \models \psi$" by induction on the structure of ψ.

The proofs of the cases $\psi = p$, $\psi = \neg\chi$ and $\psi = \chi_1 \wedge \chi_2$ are identical to those given in the proof of Lemma 3 inside the proof of Theorem 4.

Case $\psi = \triangle\omega$. Suppose $f^\pi_S(v) = S' = (C', V')$. Since $f^\pi_S(v) \in \Rightarrow^\pi (S, \omega_c)$, we have $C' = C^\pi$. Thus, since $v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \triangle\omega, \omega_c, \pi)$, we have "$v \models p_{\triangle\omega}$ iff $\omega \in C'$". Therefore, we have "$v \models p_{\triangle\omega}$ iff $f^\pi_S(v) \models \triangle\omega$ iff $\big(f^\pi_S(v), \mathrm{S}(\omega_c)\big) \models \triangle\omega$". Thus, since $tr^*_{C^\pi,\omega_c}(\triangle\omega) = p_{\triangle\omega}$, we have "$v \models tr^*_{C^\pi,\omega_c}(\triangle\omega)$ iff $\big(f^\pi_S(v), \mathrm{S}(\omega_c)\big) \models \triangle\omega$".

Case $\psi = \Box\chi$. The case is proved as follows:

$$
\begin{aligned}
v \models tr^*_{C^\pi,\omega_c}(\Box\chi) \quad & \text{iff} \quad v \models \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C^\pi} \omega \wedge \bigwedge_{\omega \in C^\pi} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\chi)\setminus C^\pi} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C^\pi,\omega_c}(\chi)\Big),\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \chi, \omega_c, \pi), u \models tr^*_{C^\pi,\omega_c}(\chi),\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \chi, \omega_c, \pi), \big(f^\pi_S(u), \mathrm{S}(\omega_c)\big) \models \chi \text{ (by induction hypothesis)},\\
& \text{iff} \quad \forall S' \in \Rightarrow^\pi(S, \omega_c), S' \models \chi \text{ (since, according to Lemma 4, } f^\pi_S \text{ is surjective)},\\
& \text{iff} \quad \big(f^\pi_S(v), \mathrm{S}(\omega_c)\big) \models \Box\chi \text{ (since we have } \Rightarrow^\pi(S, \omega_c) = \equiv\big(f^\pi_S(v), \omega_c\big) \text{ as a}\\
& \qquad \text{consequence of } f^\pi_S(v) \in \Rightarrow^\pi(S, \omega_c) \text{ which is implied by } v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)).
\end{aligned}
$$

Case $\psi = [E]\chi$. The case is proved as follows:

$$
\begin{aligned}
v \models tr^*_{C^\pi,\omega_c}([E]\chi) \quad & \text{iff} \quad v \models \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C^{\pi E}} \omega \wedge \bigwedge_{\omega \in C^{\pi E}} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\chi)\setminus C^{\pi E}} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C^{\pi E},\omega_c}(\chi)\Big),\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \chi, \omega_c, \pi E), u \models tr^*_{C^{\pi E},\omega_c}(\chi),\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \chi, \omega_c, \pi E), \big(f^{\pi E}_S(u), \mathrm{S}(\omega_c)\big) \models \chi \text{ (by induction hypothesis)},\\
& \text{iff} \quad \forall S' \in \Rightarrow^{\pi E}(S, \omega_c), \big(S', \mathrm{S}(\omega_c)\big) \models \chi \text{ (since, according to Lemma 4,}\\
& \qquad f^{\pi E}_S \text{ is surjective)},\\
& \text{iff} \quad \big(f^\pi_S(v), \mathrm{S}(\omega_c)\big) \models [E]\chi \text{ (since we have } \Rightarrow^{\pi E}(S, \omega_c) = \Rightarrow^E\big(f^\pi_S(v), \omega_c\big)\\
& \qquad \text{as a consequence of } f^\pi_S(v) \in \Rightarrow^\pi(S, \omega_c)\\
& \qquad \text{which is implied by } v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, \omega_c, \pi)).
\end{aligned}
$$

□

Thanks to Lemma 5, we can prove the case $\varphi = [E]\psi$:

$$
\begin{aligned}
\big(S, \mathrm{S}(\omega_c)\big) \models [E]\psi \quad & \text{iff} \quad \forall S' \in \Rightarrow^E(S, \omega_c), \big(S', \mathrm{S}(\omega_c)\big) \models \psi,\\
& \text{iff} \quad \forall v \in \mathbf{W}_{\mathbb{P}_\triangle}(C, \psi, E), v \models tr^*_{C^E,\omega_c}(\psi) \text{ (by Lemma 5 and since, according}\\
& \qquad \text{to Lemma 4, } f^E_S \text{ is surjective)},\\
& \text{iff} \quad \forall u \in \mathbf{W}_{\mathbb{P}_\triangle},\\
& \qquad u \models \blacksquare\Big(\big(\omega_c \wedge \bigwedge_{\omega \in C^E} \omega \wedge \bigwedge_{\omega \in C^E} p_{\triangle\omega} \wedge \bigwedge_{\omega \in \Gamma(\psi)\setminus C^E} \neg p_{\triangle\omega}\big) \rightarrow tr^*_{C^E,\omega_c}(\psi)\Big),\\
& \text{iff} \quad \models_{\mathfrak{C}} tr_{S,\omega_c}([E]\psi).
\end{aligned}
$$

□

A.11 Proof of Theorem 9

Proof. We define the function f in the same way as in the proof of Theorem 6 and prove that it is a surjection from $\mathbf{S}_{Eq}$ to $\mathbf{B}$ in an analogous way. The only fact to be proved is that if $f(S) = (\Theta, \mathcal{I})$ then $S \models [E]\omega$ iff $(\Theta, \mathcal{I}) \models [E]\omega$.

The following proposition will turn out to be useful in the rest of the proof.

Proposition 7. *Let $S \in \mathbf{S}_{Eq}$ and $f(S) = (\Theta, \mathcal{I})$. Then,*

- *if $\mathcal{I}' \in Sol(\Theta^E)$ then $\exists S' = (C', V') \in \mathbf{S}$ s.t. $S \Rightarrow^E S'$ and $V' \cap \mathbb{V} = \{q \in \mathbb{V} : \mathcal{I}'(q) = 1\}$,*
- *if $S \Rightarrow^E S' = (C', V')$ then $\exists \mathcal{I}' \in Sol(\Theta^E)$ s.t. $V' \cap \mathbb{V} = \{q \in \mathbb{V} : \mathcal{I}'(q) = 1\}$.*

We first prove the left-to-right direction. Suppose, towards a contradiction, that $S \models [E]\omega$ and $(\Theta, \mathcal{I}) \models \langle E \rangle \neg\omega$. The latter means that $Sol(\Theta^E) \cap |\neg\omega|_{Asg} \neq \emptyset$. By the first item of Proposition 7, the latter implies that $\exists \mathcal{I}' \in Sol(\Theta^E) \cap |\neg\omega|_{Asg}, \exists S' = (C', V') \in \mathbf{S}$ such that $S \Rightarrow^E S'$ and $V' \cap \mathbb{V} = \{q \in \mathbb{V} : \mathcal{I}'(q) = 1\}$. Since $\omega \in \mathcal{L}_{\text{PROP}}(\mathbb{V})$, the latter implies that $\exists S' \in \mathbf{S}$ such that $S \Rightarrow^E S'$ and $S' \models \neg\omega$. The latter means that $S \models \langle E \rangle \neg\omega$ which leads to a contradiction.

Let us now prove the right-to-left direction. Suppose, towards a contradiction, that $(\Theta, \mathcal{I}) \models [E]\omega$ and $S \models \langle E \rangle \neg\omega$. The latter means that $\exists S' = (C', V') \in \mathbf{S}$ such that $S \Rightarrow^E S'$ and $S' \models \neg\omega$. By the second item of Proposition 7, the latter implies that $\exists \mathcal{I}' \in Sol(\Theta^E), \exists S' = (C', V') \in \mathbf{S}$ such that $S \Rightarrow^E S'$, $S' \models \neg\omega$ and $V' \cap \mathbb{V} = \{q \in \mathbb{V} : \mathcal{I}'(q) = 1\}$. Since $\omega \in \mathcal{L}_{\text{PROP}}(\mathbb{V})$, the latter implies that $Sol(\Theta^E) \cap |\neg\omega|_{Asg} \neq \emptyset$. The latter means that $(\Theta, \mathcal{I}) \models \langle E \rangle \neg\omega$ which leads to a contradiction.

□

A.12 Proof of Theorem 10

Proof. The following lemma is propaedeutic for the proof of the main theorem.

Lemma 6. *Let $\omega \in \mathcal{L}_{\text{PROP}}(\mathbb{P})$ and $S = (C, V) \in \mathbf{S}$. Then,*

$$ClosestSub(\omega, S) \subseteq Closest(\omega, S),$$

with

$$ClosestSub(\omega, S) = \{S' \in SubSt(S) : S' \models \omega \text{ and } \nexists S'' \in SubSt(S) \text{ s.t. } S'' \models \omega \text{ and } S' \prec_S S''\},$$

and $SubSt(S) = \{S' = (C', V') \in \mathbf{S} : C' \subseteq C\}$.

Proof. Let $S = (C, V) \in \mathbf{S}$. Suppose that $S' = (C', V') \in ClosestSub(\omega, S)$. The latter means that (i) $S' = (C', V') \in SubSt(S)$ and (ii) $\nexists S'' \in SubSt(S)$ such that $S'' \models \omega$ and $S' \prec_S S''$. Since $SubSt(S) \subseteq \mathbf{S}$, we have (iii) $S' = (C', V') \in \mathbf{S}$. Moreover, from item (ii) by the definition of the similarity relation $\prec_S$, we have (iv) $\nexists S'' \in \mathbf{S}$ such that $S'' \models \omega$ and $S' \prec_S S''$. For otherwise, we could find $S'' = (C'', V'') \in \mathbf{S}$ such that $S'' \models \omega$ and $(C \cap C') \subsetneq (C \cap C'')$. As a consequence, we could find $S''' = (C''', V''') \in \mathbf{S}$ such that $C''' = (C \cap C'')$ and $S''' \models \omega$, since, for every propositional formula ω and for all $C''', C'' \subseteq \mathcal{L}_{\text{PROP}}(\mathbb{P})$, if $C''' \subseteq C''$ and there exists $S'' = (C'', V'') \in \mathbf{S}$ such that $S'' \models \omega$ then there exists $S''' = (C''', V''') \in \mathbf{S}$ such that $S''' \models \omega$. Thus, we could find $S''' = (C''', V''') \in \mathbf{S}$ such that $C''' \subseteq C$, $S''' \models \omega$ and $(C \cap C') \subsetneq (C \cap C''')$ which would contradict (ii). Therefore, from (iii) and (iv), we can conclude that $S' \in Closest(\omega, S)$. □

Now, let us move to the proof of the main theorem.

The equivalence between item 2 and item 3 is a consequence of Proposition 1.

Let $S = (C, V) \in \mathbf{S}$. We are going to prove the equivalence between item 1 and item 2. We first prove 1⇒2.

Suppose that (i) $S \models \varphi_1 \;\Box\!\!\rightarrow\; \varphi_2$ and, towards a contradiction, (ii) $(C', V) \models \Diamond(\varphi_1 \wedge \neg\varphi_2)$ for an arbitrary $C' \in C\bot\neg\varphi_1$.

Item (ii) means that $C' \subseteq C$, $\neg\varphi_1 \notin Cn(C')$, $\forall C'' \subseteq C$, if $C' \subset C''$ then $\neg\varphi_1 \in Cn(C'')$, and $(C', V) \models \Diamond(\varphi_1 \wedge \neg\varphi_2)$. The latter implies that (iii) $C' \subseteq C$, (iv) $(C', V) \models \Diamond(\varphi_1 \wedge \neg\varphi_2)$, and (v) $\forall C'' \subseteq C$, if $C' \subset C''$ then $(C'', V) \models \Box\neg\varphi_1$.

Items (iii) and (iv) together imply that (vi) we can find $S' = (C', V') \in SubSt(S)$ such that $S' \models \varphi_1 \wedge \neg\varphi_2$.

Items (v) and (vi) together imply that (vii) we can find $S' = (C', V') \in SubSt(S)$ such that $S' \models \varphi_1 \wedge \neg\varphi_2$ and $\nexists S'' \in SubSt(S)$ such that $S'' \models \varphi_1$ and $S' \prec_S S''$.

Item (vii) implies that (viii) we can find $S' = (C', V') \in ClosestSub(\varphi_1, S)$ such that $S' \models \varphi_1 \wedge \neg\varphi_2$.

Thus, by Lemma 6, from item (viii) we can conclude that there exists $S' = (C', V') \in Closest(\varphi_1, S)$ such that $S' \models \varphi_1 \wedge \neg\varphi_2$. The latter contradicts the initial assumption (i).

We are going to prove 2⇒1.

Suppose that (i) $(C', V) \models \Box(\varphi_1 \rightarrow \varphi_2)$ for all $C' \in C\bot\neg\varphi_1$ and, towards a contradiction, that (ii) $S \not\models \varphi_1 \;\Box\!\!\rightarrow\; \varphi_2$. The latter means that there exists $S'' = (C'', V'') \in Closest(\varphi_1, S)$ such that $S' \models \varphi_1 \wedge \neg\varphi_2$. Thus, by Lemma 6, (iii) there exists $S'' = (C'', V'') \in ClosestSub(\varphi_1, S)$ such that $S' \models \varphi_1 \wedge \neg\varphi_2$.

By Proposition 1, it is easy to verify that (iv) if $S'' = (C'', V'') \in ClosestSub(\varphi_1, S)$ then $C'' \in C\bot\neg\varphi_1$.

Thus, from items (iii) and (iv), we can conclude that there exists $C'' \in C\bot\neg\varphi_1$ such that $(C'', V'') \models \Diamond(\varphi_1 \wedge \neg\varphi_2)$ which contradicts the initial assumption (i). □

A.13 Proof of Theorem 11

Proof. We are going to prove the following equivalent result supposing that $S \in \mathbf{S}_{Eq}$:

$$\text{if } (S, \mathbf{S}_{Eq}) \models \langle E \rangle \top \text{ then}$$
$$\Big((S, \mathbf{S}_{Eq}) \models [E]\omega \text{ iff } (S, \mathbf{S}_{Eq}) \models \bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau) \;\Box\!\!\rightarrow\; \omega\Big).$$

Suppose $(S, \mathbf{S}_{Eq}) \models \langle E \rangle \top$. Hence, (i) $\exists S' = (C', V') \in \mathbf{S}_{Eq}$ such that $C' = \big(C \setminus Rev(C,E)\big) \cup E$. From the latter, it is routine to show that

$$\text{(ii) } \forall S'' = (C'', V'') \in Closest\Big(\bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau), S, \mathbf{S}_{Eq}\Big),$$
$$\Big(\big(C \setminus Rev(C,E)\big) \cup E\Big) \subseteq C'',$$

and

$$\text{(iii) } \big(\Rightarrow^E(S) \cap \mathbf{S}_{Eq}\big) \subseteq Closest\Big(\bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau), S, \mathbf{S}_{Eq}\Big),$$

with $\Rightarrow^E(S) = \{S' \in \mathbf{S} : S \Rightarrow^E S'\}$.

From (iii), we have that if $(S, \mathbf{S}_{Eq}) \models \big(\bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau)\big) \;\Box\!\!\rightarrow\; \omega$ then $(S, \mathbf{S}_{Eq}) \models [E]\omega$.

We prove the other direction of the equivalence by reductio ad absurdum. Let us suppose (iv) $(S, \mathbf{S}_{Eq}) \models [E]\omega$ and (v) $(S, \mathbf{S}_{Eq}) \not\models \big(\bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau)\big) \;\Box\!\!\rightarrow\; \omega$. From (v) we have that (vi) $\exists S'' = (C'', V'') \in Closest\big(\bigwedge_{p \leftrightarrow \tau \in E} \triangle(p \leftrightarrow \tau), S, \mathbf{S}_{Eq}\big)$ such that $(S'', \mathbf{S}_{Eq}) \models \neg\omega$.

It useful to observe that (vii) for every propositional formula ω and for all $C''', C'' \subseteq \mathcal{L}_{\mathsf{PROP}}(\mathbb{P})$, if $C''' \subseteq C''$ and there exists $S'' = (C'', V'') \in \mathbf{S}$ such that $S'' \models \omega$ then there exists $S''' = (C''', V''') \in \mathbf{S}$ such that $S''' \models \omega$.

Thus, from (vi), by (ii) and (vii), we can conclude that $\exists S''' = (C''', V''') \in \mathbf{S}_{Eq}$ such that $C''' = \big(C \setminus Rev(C,E)\big) \cup E$ and $(S''', \mathbf{S}_{Eq}) \models \neg\omega$. The latter contradicts (iv). □

Received 24 April 2025; accepted 18 December 2025