The Underdeterministic Framework

Tomasz Wysocki

Philosophy and statistics have studied two causal species, deterministic and probabilistic. There's a third species, however, hitherto unanalysed: underdeterministic causal phenomena, which are non-deterministic yet non-probabilistic. Here, I formulate a framework for modelling them. Consider a simple case. If I go out, I may stumble into you but also may miss you. If I don't go out, we won't meet. I go out. We meet. My going out is a cause of our encounter even if there was no determinate probability of us meeting conditional on my going out. The cause is neither deterministic (it didn't necessitate the effect) nor probabilistic (the relevant conditional probabilities are undefined). Rather, it's underdeterministic: it raises the modal status of the effect from causally impossible to possible. Here, I won't offer a theory of such token causes but develop the prerequisite for any such theory: the underdeterministic framework. The framework is like the deterministic structural-equations framework but with one consequential difference—an equation can return multiple values. This change allows me to define causal possibility and necessity, and corresponding notions of interventionist might- and would-counterfactuals. I also define conditional independence, which obeys the graphoid axioms, and prove that underdeterministic models satisfy the causal Markov condition. The framework can causally model situations that other frameworks cannot: decision-making under bounded uncertainty, games with multiple equilibria, infinite fair lotteries, and any other non-deterministic situation where indeterminacies are essentially non-probabilistic, or where we have a reason not to use probabilities.

1. Introduction

Not all that happens is bound to happen. Sometimes probabilities can help represent such non-deterministic events, but sometimes all there is to know is that some events may, some will, and some won't happen depending on what you decide to do. No formal framework can yet model such situations.

Here, I formulate one. I'll dub it the underdeterministic framework and build it drawing on the resources of the deterministic interventionist framework (Halpern [2016]; Pearl [2000]; Hitchcock [2001]; Spirtes et al. [2001]). I begin (sec. 2) by motivating the framework. It can represent indeterminacies that are essentially non-probabilistic, and it can be useful for modellers working under bounded uncertainty. Next (sec. 3), I introduce underdeterministic models, which are like their deterministic counterparts but with a lone exception: structural equations can return multiple values, and hence models can have multiple solutions. Consequently (sec. 4), the framework affords (non-probabilistic) modal notions and counterfactuals. An event is causally possible/necessary if and only if it occurs on some/all solutions; a might-/would-counterfactual holds in a model if the consequent is causally possible/necessary after the antecedent is brought about with an intervention.

The modal notions, in turn, underlie the notion of conditional independence (sec. 5), which satisfies the graphoid axioms (Pearl and Paz [1985]; Verma and Pearl [1990]), and hence behaves similarly to probabilistic dependence. The notion lets me formulate the causal Markov condition, which is satisfied by any underdeterministic model. I relegate all substantial proofs to the appendix (sec. 6).

2. Motivation

Underdeterministic causal phenomena are non-deterministic but non-probabilistic. My aim here is to put forward a framework that will allow for modelling such phenomena: token causation, type causation, counterfactuals, and causal decision theory, for instance. Before I start formulating the framework, I'll offer some reasons for why these phenomena deserve our attention.

The first reason for developing the framework is metaphysical: events are sometimes genuinely non-deterministic even though this non-determinism doesn't take the form of objective chances. For starters, consider a fair infinite lottery (Norton [2021b], chapters 14–15). I bet on 8 and I win; my betting on 8 was a cause of my winning, even though there's no probability distribution over lottery tickets that respects the lottery's fairness. For a more physically plausible case, take Norton's dome (Norton [2003]). A pebble left on the dome's top may slide down or stay unmoved indefinitely, and no probability distribution can govern these possibilities. Yet, if I put the pebble on the top of the dome, it slides down at night, hits the base, and wakes me up, my leaving it there caused me to wake up: had I not played with the dome, I would have remained asleep. Finally, there are genuine theories in science where non-determinism cannot take a form of objective chances. According to the hypothesis of eternal inflation, the universe began with an indefinitely long expansion, during which infinitely many pocket universes were seeded. There's no probability of a pocket universe like ours being seeded (Norton [2021]), yet the Big Bang does count as a cause of our pocket universe. The framework could also prove useful for debates about cosmic fine-tuning (Friederich [2022]), where a range of values of some physical constant are treated as possible, but ascribing probabilities to these values seems metaphysically unjustified.[1] And all physically plausible examples aside: non-deterministic situations where there's no fact of the matter how probable an event is seem like a genuine ontological possibility, and, as such, they deserve to be investigated.

[1] I owe this illustration to one of the reviewers.

The second reason is epistemic: sometimes we don't ascribe non-arbitrary probabilities to events even if these events are in fact probabilistic or even deterministic. Historians, for instance, investigate the causes of the fall of Saba, yet it would be unreasonable to demand they assess conditional probabilities. If I know the modal facts but not probabilities, I still can reason causally using the underdeterministic framework.[2]

The epistemic reason won't convince everyone. Bayesians, most notably, maintain that an agent should always ascribe probabilities to possible events, no matter how scarce her evidence is. I can't argue here against such imperialism. Rather, I'll just observe that, first, not everyone is a Bayesian imperialist, and these who are not might want to use my framework (for example, Douven [2021]; Norton [2021]; I'll return to them in a moment). Second, the argument from Bayesian imperialism equally applies to any use of non-probabilistic epistemic possibilities. For instance, this argument would banish epistemic logics for failing to model knowledge and belief as degrees. If you think that in philosophy there's a place for epistemic logics, you shouldn't deny it to underdeterministic causal models.

The third reason for developing the framework is pragmatic. People who argue against Bayesian imperialism often do it on the grounds that it's too demanding for us as finite beings, even as a normative ideal (Elqayam [2012]; Arkes et al. [2016]). They instead propose that evaluating a cognitive process as rational depends on the agent's goals, the domain of inquiry, and available resources. It seems that the underdeterministic framework, combined with the leximin rule, allows for normatively evaluating an agent's causal reasoning in situations where stakes are high but available resources (such as time or computational capacity) are scarce (Wysocki [unpublished-a]). If crossing the road now may get me to the other side but also may get me killed, while crossing the road once the traffic subsides will get me to the other side (and won't get me killed), I should wait for the traffic to subside—and I needn't employ probabilities to be justified in reaching this conclusion.

The fourth reason comes from philosophy of science. Economists do consider agents under bounded uncertainty: such agents know what states will become possible given the available decisions not how probable these states are (Knight [1921]; Cohen and Jaffray [1983]; Runde [1998]; Deœux [2019]). The underdeterministic framework comes in handy, for instance, if we want to make causal claims about such agents. Or consider any equilibrium model that specifies what equilibria are possible given some boundary conditions, but not how probable these equilibria are (see Spence [1973] for a famous example); game theory is rife with such models (Biswas [1997]). Deriving causal claims from equilibrium models requires the underdetermin-

[2] There have been other proposals to model non-deterministic causation without using probabilities. Spohn ([2012]) uses ranking functions instead of probabilities to represent an agent's grades of disbelief in a proposition. Among the many applications of Spohn's framework is a theory of causation. The framework can't capture all cases of underdeterminism, however, because Spohn's agents can compare any two grades of disbelief and thus any two potential effects. On another proposal, the agent's ignorance could be represented with imprecise probabilities. It turns out, though, that imprecise probabilities don't yield a notion of independence, and hence the causal Markov condition, which makes them unsuitable for theories of causation (Kinney [2018]).

istic framework. Biology also comes with its own notion (or notions) of possibility, especially salient in synthetic biology, but these possibilities aren't accompanied by probabilities (Ijäs and Koskinen [2021]; Knuuttila and Loettgers [2022]). It seems that underdeterministic models are naturally equipped for representing such possibilities.[3]

The final reason is (let me call it) conceptual: we just seem to possess a family of underdeterministic causal concepts, interesting enough to deserve analysis. I may but may not bump into you if I go out; I won't if I don't. I go out and bump into you. My going out was a cause of meeting you, and it seems that this causal claim holds in virtue of the counterfactual structure true of the case (and what actually happened). While this structure itself might hold in virtue of some probabilistic properties of the events involved, these properties are only indirectly relevant to the causal claim at hand. It is the relation between the counterfactual structure and the causal claim that needs to be analysed; invoking probabilities just muddles the analysis. Moreover, underdeterministic relations seem to be the weakest possible causal relations, which suggests that investigating underdeterministic causation will shed light on what's essential to causation *simpliciter*.

3. Models

I'll start with a case. Three economists are to recommend a tax for the wealthy. The first consults the oracle; whatever the oracle proposes, so will the economist, and the oracle may propose a 0.8 or 1 tax. The second consults both the oracle and the soothsayer and will recommend the average of their proposals; and the soothsayer may propose a 0.7 or 0.9 tax. The third may recommend what either the soothsayer or the oracle proposes. Held hostage to neoliberalism, the government will implement the smallest of four rates: the three recommendations and a 0.6 tax.

A model consists of variables, their ranges, and their structural equations, $\mathfrak{M} = \langle \mathcal{V}, \mathcal{R}, \mathcal{E} \rangle$, where for every X from $\mathcal{V}$, $\mathcal{D}_X$ denotes X's range, $\mathcal{D}_X \in \mathcal{R}$. Variable values denote atomic events (for example, that the soothsayer proposes 0.9). In the economists case, I'll use O, S, P, A, T, and G to denote the tax recommended by the oracle, the soothsayer, the first, second, and third economist, and the tax implemented by the government. The ranges agree: $\mathcal{D}_O = \ldots = \mathcal{D}_G = [0, 1]$.

A variable's structural equation identifies the variable's values that are causally possible given the values of other variables. A variable and its equation are exogenous if the equation's right side contains no variables (that is, the equation is nullary); otherwise, the variable and equation are endogenous. Equations that for some input return multiple values are underdeterministic; otherwise, they are deterministic. One variable parents another if the former is an argument in the latter's equation; the latter variable is then a child of the former. Being an an-

[3] Bayesian imperialists would have to claim that such non-probabilistic uses of possibility are wrong or at least wanting.

cestor is the transitive closure of being a parent, and being a descendant is the transitive closure of being a child.

Equations describe underdeterministic counterfactual relations between events. Specifically, exogenous equations encode primitive causal modals, while endogenous equations encode primitive causal counterfactuals. I'll use the economists case to explain what that means. The equations representing the case are:

$$O \leftarrow 0.8,\ 1, \quad S \leftarrow 0.7,\ 0.9, \quad P \leftarrow O, \quad A \leftarrow (S+O)/2, \quad T \leftarrow S,\ O, \quad G \leftarrow \min(P, A, T, 0.6).$$

O parents P, A, and T; S parents A and T; P, A, and T parent G. O's equation, exogenous underdeterministic, expresses the primitive causal possibility that the oracle proposes a 0.8 tax and the primitive causal possibility that the oracle proposes a 1 tax. The comma separates the variable's possible values, that is, read the equation as $O \in \{0.8, 1\}$. P's equation, endogenous deterministic, expresses a family of primitive would-counterfactuals 'if the oracle proposes rate O, the first economist will recommend rate O' for every value of O from O's range. T's equation is endogenous underdeterministic; it expresses a family of might counterfactuals 'if the oracle proposes rate O and the soothsayer rate S, the third economist may recommend rate O' and 'if the oracle proposes rate O and the soothsayer rate S, the third economist may recommend rate S' for every value of O and S from their ranges. Again, read the equation as $A \in \{O, S\}$. The remaining equations read analogously.

Not all counterfactuals are primitive. Although 'if the oracle proposes 0.7 and the soothsayer 0.2, the first economist will recommend 0.7' holds, the counterfactual isn't encoded by any equation but follows instead from the primitive counterfactual 'if the oracle proposes 0.7, the first economist will recommend 0.7'. Any equation must satisfy a constraint: the value of the parent makes a difference to the value of the child at least under some assignment of values to the remaining variables. In deterministic models, the constraint demands that X parents Y if and only if there are at least two values of X such that switching between these values changes Y's value under some assignment of values to the remaining variables $\mathcal{V} \setminus \{X, Y\}$ (Halpern [2016], p. 15; Hitchcock [2001], p. 280). In probabilistic models, the constraint demands that under some assignment to $\mathcal{V} \setminus \{X, Y\}$, switching between the two values of X changes Y's probabilistic distribution (Pearl [2000], p. 14; Fenton-Glynn [2017]). The underdeterministic version of the constraint is analogous: under some assignment to $\mathcal{V} \setminus \{X, Y\}$, switching between the two values of X changes Y's possible values. In the underdeterministic framework, therefore, to make a difference is to change (that is, to limit or to expand) what may ensue. Since on the interventionist approach to causation, to make a difference is to cause, a variable's parent is the same as a direct type cause. This definition entails a definition of underdeterministic type causation *simpliciter*: X type-causes Y if and only if switching between two values of X changes Y's possible values under some assignment of values to some subset of variables $\mathcal{V}$; this definition

Table 1. Solutions to the economists model.

	O	S	P	A	T	G		O	S	P	A	T	G
$\vec{\sigma}_1$	0.8	0.7	0.8	0.75	0.7	0.6	$\vec{\sigma}_5$	1	0.7	1	0.85	0.7	0.6
$\vec{\sigma}_2$	0.8	0.7	0.8	0.75	0.8	0.6	$\vec{\sigma}_6$	1	0.7	1	0.85	1	0.6
$\vec{\sigma}_3$	0.8	0.9	0.8	0.85	0.8	0.6	$\vec{\sigma}_7$	1	0.9	1	0.95	0.9	0.6
$\vec{\sigma}_4$	0.8	0.9	0.8	0.85	0.9	0.6	$\vec{\sigma}_8$	1	0.9	1	0.95	1	0.6

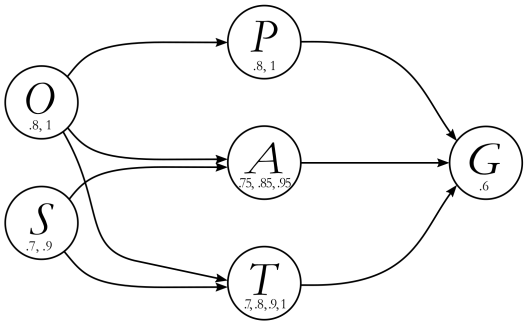

Figure 1. The economists case.

is a straightforward analogue of Woodward's ([2003], p. 75).[4]

The framework deals with recursive models only, which means that no variable is its own ancestor. In such models, variables can be ordered topologically, so that every variable comes after the variables that are arguments in its equation. I'll assume that $\mathcal{V}$ is ordered with some particular topological order—henceforth, the topological order—and every subset of $\mathcal{V}$ inherits this order. I'll denote such subsets with a capital letter topped with an arrow, for example, $\vec{X}$; I write $\vec{X}\vec{Y}$ for $\vec{X} \cup \vec{Y}$ and $X\vec{Y}$ for $\{X\} \cup \vec{Y}$.

Solutions to an underdeterministic model represent how the target situation may develop, that is, possible histories. Recursive by assumption, underdeterministic models always have at least one solution, satisfying a plausible demand that every situation may develop in at least one way. The economists case may develop in eight ways (table 1). If a model describes a situation that has already happened, you may identify one of the solutions as actual. However, what events actually transpire won't matter in the current analysis.[5] Models are best entertained using diagrams. Variables correspond to nodes (hence, I'll use the terms interchangeably), and directed edges connect a node to its children. On the diagram, a node also contains its possible values. A recursive model has an acyclic graph. For the diagram of the economists model, see figure 1.

Crucial for the framework is the notion of an assignment. An assignment over variables $\vec{X}$ maps the variables to some elements from their domains. I denote an arbitrary assignment with a lowercase letter topped with an arrow, for example, $\vec{x}$. I represent a concrete assignment as a tuple of values in the topological order, typically also indicating what variables are

[4] Unfortunately, I lack space here to investigate the underdeterministic type causation and compare it with its deterministic analogue (Hausman and Woodward [1999]).

[5] They matter when it comes to token causation (Wysocki [unpublished-c]).

being mapped, for example, assignment $\langle 1_F, 0_A\rangle$ assigns 1 to F and 0 to A. $\vec{x}[\vec{Y}]$ denotes a projection of $\vec{x}$ onto variables $\vec{Y}$, where $\vec{x}$ must assign values to all variables from $\vec{Y}$, for example, $\langle 1_F, 0_A\rangle[F] = \langle 1_F\rangle$. Any projection onto an empty list of variables produces the empty assignment, $\vec{x}[\varnothing] = \langle\rangle$. I'll also rely heavily on junction, an operation on sets of assignments. Where U is a set of assignments over $\vec{X}$, and V is a set of assignments over $\vec{Y}$, their junction $U \bowtie V$ is the maximal set of assignments over $\vec{X}\vec{Y}$ such that every assignment from the junction agrees with some assignment from U on the values of $\vec{X}$ and some assignment from V on the values of $\vec{Y}$. (For another definition of junction, see section 6.) For example, $\{\langle 0_S\rangle, \langle 1_S\rangle\} \bowtie \{\langle 0.5_P\} = \{\langle 0_S, 0.5_P\rangle, \langle 1_S, 0.5_P\rangle\}$. A solution to an underdeterministic model is an assignment over $\mathcal{V}$ that satisfies the model's equation.

Events over $\vec{X}$ can be identified with sets of assignments over $\vec{X}$.[6] A conjunctive event over $\vec{X}$ is a singleton containing an assignment over $\vec{X}$. An atomic event is a conjunctive event over a single variable. A simple event over $\vec{X}$ is a set of assignments that decomposes into a junction of sets over single variables, where for every variable from $\vec{X}$, exactly one set is over this variable. For example, event $\{\langle 0_S, 0.2_P\rangle, \langle 1_S, 0.5_P\rangle\}$ isn't simple, while $\{\langle 0_S, 0.5_P\rangle, \langle 1_S, 0.5_P\rangle\}$ is because it's equivalent to $\{\langle 0_S\rangle, \langle 1_S\rangle\} \bowtie \{\langle 0.5_P\rangle\}$. Therefore, atomic events are conjunctive, and conjunctive events are simple.

The name of an event over $\vec{X}$ is a Boolean combination of mathematical formulas over $\vec{X}$ that is satisfied by all and only assignments from the event. I won't specify the syntax of names but simply assume that for any name built with some variables and for any assignment over at least these variables, it's always possible to substitute the variables' values for the variables' names and test if the resulting sentence is true. If it is, the assignment satisfies the event name. So, a name denotes a conjunctive event over $\vec{X}$ if the name is satisfied by exactly one assignment over $\vec{X}$; a name denotes an atomic event if the name denotes a conjunctive event over one variable; if a name is a conjunction of satisfiable formulas, each over a different single variable, then the name denotes a simple event. $\top$ stands for a sentence over no variables that denotes the event $\{\langle\rangle\}$ (which conventionally counts a conjunctive simple event); for $\vec{X}{=}\varnothing$, $\vec{X}{=}\vec{x}$ is equivalent to $\top$. $\bot$ stands for a sentence over no variables that is satisfied by no assignments; thus, $\bot$ denotes the contradictory event $\varnothing$. I adopt the following conventions: I'll typically denote atomic events over X with $X{=}x$, where $x \in \mathcal{D}_X$; conjunctive events over $\vec{X}$ with $\vec{X}{=}\vec{x}$, where $\vec{x}$ is an assignment over $\vec{X}$; simple events with Greek letters topped with a dot, for example, $\dot{\varphi}$; and arbitrary events with Greek letters, for example, ψ. For instance, $\langle A, T\rangle = \langle 0.8, 0.9\rangle$ denotes the conjunctive event of the second economist recommending a 0.8 tax and the third economist recommending a 0.9 tax.

6 Therefore, events are like teams (Barbero and Sandu [2021]), which are similarly defined as sets of assignments. However, dealing with disjunctive events requires extending the formalism to sets of assignments that needn't be over the same variables (Briggs [2012]; Fine [2012]; Wysocki [unpublished-b]). There, the analogy with teams breaks.

Any simple event can be brought about in a model by an intervention.[7] An intervention represents manipulating a situation from without. In deterministic models, an intervention on the target variable changes its value directly, that is, not through affecting any other variables from the model (Woodward [2003], p. 47); this is reflected in the equation of the target variable being replaced with one that directly assigns to the variable its new value. An intervention thus makes the target variable causally independent from its original parents.[8] The intervention can change the values of the model's other variables, but these changes simply follow from the change of value in the target variable. Consequently, a post-intervention deterministic model describes a situation that would have occurred, had the intervention been performed–that is, had the target event been caused exogenously.

The idea behind interventions in the underdeterministic framework is the same—representing exogenous modifications to the target situation—although underdeterminism affords a richer apparatus. First, take a conjunctive event $\vec{X}{=}\vec{x}$. Like in deterministic models, to bring about $\vec{X}{=}\vec{x}$ in $\mathfrak{M}$, for every X from $\vec{X}$, replace X's equation with $X \leftarrow \vec{x}[X]$. Now, take a simple event $\dot{\varphi}$; say that the event is over $\vec{X} = \langle X_1, X_2, \ldots, X_k \rangle$ and it decomposes into a junction $V_1 \bowtie V_2 \bowtie \ldots \bowtie V_k$, where V_i is a non-empty set over X_i. To bring $\dot{\varphi}$ about, for every X_i from $\vec{X}$, replace X_i's equation one that assigns to X_i all values from V_i: $X_i \leftarrow V_i$. For instance, to bring about $T > 0.8 \wedge A < 0.5$, replace T's equation with $T \leftarrow (0.8, 1]$ and A's equation with $A \leftarrow [0, 0.5)$. I'll denote a post-intervention model with $\mathfrak{M}_{\dot{\varphi}}$, where $\dot{\varphi}$ is the event brought about with an intervention. The post-intervention model agrees with the pre-intervention one on variables, their ranges, and the equations for $\mathcal{V} \backslash \vec{X}$. Bringing about $\top$ leaves the model intact; you cannot bring about $\bot$, as it's not a simple event. You can also compose interventions. $\mathfrak{M}_{\dot{\varphi},\dot{\rho}}$ denotes the model produced by bringing about $\dot{\varphi}$ in $\mathfrak{M}$ and subsequently $\dot{\rho}$ in the resulting model $\mathfrak{M}_{\dot{\varphi}}$; the definition extends to any number of events. Notice that if variables in $\dot{\varphi}$ and $\dot{\rho}$ don't overlap, bringing about $\dot{\varphi}$ and then $\dot{\rho}$ is equivalent to bringing about $\dot{\varphi} \wedge \dot{\rho}$.

The underdeterministic framework, therefore, allows for a richer notion of interventions than the deterministic framework. In the latter, the post-intervention model represents the sole way that the situation, modified by the intervention, will develop (or would have developed), and an intervention can bring about only a conjunctive event. In contrast, in the underdeterministic framework, the post-intervention model can represent many ways that the modified situation may develop, and the target event itself needn't be conjunctive. So, while an intervention still corresponds to manipulating the situation, now the manipulation can be underdeterministic—it may happen in many ways. I'll reuse the example: bringing about $T > 0.8 \wedge A < 0.5$ results in a situation, represented by the post-intervention model, where the third economists may rec-

[7] The framework allows for a richer semantics of bringing about any event (Wysocki [unpublished-b]); however, this functionality requires additional formalism that I cannot present here.

[8] Probabilistic models also allow for so-called parametric (or soft) interventions, which modify, but don't entirely cut, the variable's ties to its parents (Eberhardt and Scheines [2006]). As these interventions come up in the context of causal discovery, I won't discuss them here; I will only notice they do have underdeterministic analogues (Wysocki [unpublished-b]).

ommend any rate greater than 0.8, and the second economist may recommend any rate smaller than 0.5. (If you like, as a metaphor, to think about interventions as actions of idealized agents (Menzies and Price [1993]), an underdeterministic intervention bringing about a simple event over some variables corresponds to several agents, one per variable, who act underdeterministically and independently from each other.[9] For example, $T > 0.8 \wedge A < 0.5$ is brought about by two agents; one underdeterministically chooses the third economist's recommendation from $(0.8, 1]$, and the other chooses the second economist's recommendation from $[0, 0.5)$.)

4. The Model Language: Modals And Counterfactuals

Where name ψ denotes any event, and name $\dot{\varphi}$ denotes a simple event, sentences of the model language are: modal claims $\Diamond\psi$ and $\Box\psi$, might-counterfactuals $\dot{\varphi} \Diamond\!\!\rightarrow \psi$, would-counterfactuals $\dot{\varphi} \Box\!\!\rightarrow \psi$, and any Boolean combination of modal sentences and counterfactuals. Their semantics is as follows:

The event (denoted by the event name) ψ is causally possible—henceforth, possible—in $\mathfrak{M}$ if and only if ψ happens on at least one solution to $\mathfrak{M}$:

$$\mathfrak{M} \models \Diamond\psi \quad \text{if and only if} \quad \psi \text{ is satisfied by some solution to } \mathfrak{M}. \tag{1}$$

For example, it's possible that the soothsayer proposes a 0.7 tax, and it's impossible that the second economist recommends a tax less than 0.5.

The event ψ is causally necessary—henceforth, necessary—in $\mathfrak{M}$ if and only if ψ happens on all solutions to $\mathfrak{M}$:

$$\mathfrak{M} \models \Box\psi \quad \text{if and only if} \quad \psi \text{ is satisfied by every solution to } \mathfrak{M}. \tag{2}$$

For example, it's necessary that the government implements a 0.6 tax, G=0.6, and that the recommendations of the second and the third economist differ at most by 0.15.

The operators are duals,

$$\mathfrak{M} \not\models \Diamond\psi \quad \text{if and only if} \quad \mathfrak{M} \models \Box\neg\psi, \tag{3}$$

and necessity implies possibility,

$$\text{if} \quad \mathfrak{M} \models \Box\psi \quad \text{then} \quad \mathfrak{M} \models \Diamond\psi. \tag{4}$$

In deterministic models, which have only one solution, the implication goes both ways.

Might-counterfactual 'were $\dot{\varphi}$ to happen, ψ could happen' holds in $\mathfrak{M}$ if and only if ψ is

[9] For me, like for Woodward ([2003], p. 123), this is no more than a metaphor.

possible after bringing about $\dot{\varphi}$,

$$\mathfrak{M} \vDash \dot{\varphi} \mathbin{\Diamond\!\!\rightarrow} \psi \quad \text{if and only if} \quad \mathfrak{M}_{\dot{\varphi}} \vDash \Diamond \psi. \tag{5}$$

For example, if the oracle proposes a 0.6 tax, the second economist may recommend a 0.6 tax because A=0.6 is possible on $\mathfrak{M}_{O=0.6}$.

Would-counterfactual 'were $\dot{\varphi}$ to happen, ψ would happen' holds in $\mathfrak{M}$ if and only if ψ is necessary after bringing about $\dot{\varphi}$,

$$\mathfrak{M} \vDash \dot{\varphi} \mathbin{\Box\!\!\rightarrow} \psi \quad \text{if and only if} \quad \mathfrak{M}_{\dot{\varphi}} \vDash \Box \psi. \tag{6}$$

For example, if the oracle proposes a 0.6 tax, the first economist will recommend a 0.6 tax because P=0.6 is necessary in $\mathfrak{M}_{O=0.6}$. Counterfactuals are duals,

$$\mathfrak{M} \nvDash \dot{\varphi} \mathbin{\Diamond\!\!\rightarrow} \psi \quad \text{if and only if} \quad \mathfrak{M} \vDash \dot{\varphi} \mathbin{\Box\!\!\rightarrow} \neg\psi, \tag{7}$$

which follows from equation 3. Any would-counterfactual entails a corresponding might-counterfactual,

$$\text{if} \quad \mathfrak{M} \vDash \dot{\varphi} \mathbin{\Box\!\!\rightarrow} \psi \quad \text{then} \quad \mathfrak{M} \vDash \dot{\varphi} \mathbin{\Diamond\!\!\rightarrow} \psi, \tag{8}$$

and in deterministic models, the implication goes both ways. $\top$ is necessary and thus possible, $\bot$ is impossible and thus not necessary. $\mathfrak{M} \vDash \top \mathbin{\Diamond\!\!\rightarrow} \psi$ is equivalent to $\mathfrak{M} \vDash \Diamond \psi$, and $\mathfrak{M} \vDash \top \mathbin{\Box\!\!\rightarrow} \psi$ is equivalent to $\mathfrak{M} \vDash \Diamond \psi$.

Modal claims and counterfactuals can be combined using Boolean operators. Where Φ and Ψ are sentences of the model language, $\mathfrak{M} \vDash \neg\Phi$ if and only if $\mathfrak{M} \nvDash \Phi$, $\mathfrak{M} \vDash \Phi \wedge \Psi$ if and only if $\mathfrak{M} \vDash \Phi$ and $\mathfrak{M} \vDash \Psi$, and so on for all the usual logical connectives. Neither a lone event name (for example, $\mathfrak{M} \vDash O$=0.9) nor a formula with a modal operator bound by another (for example, $\mathfrak{M} \vDash \Diamond \neg \Box \neg \Diamond O$=0.9) is a sentence of the model language.

What's causally possible and necessary is relative to a situation, and thus a model; for example, G=0.6 is causally necessary in $\mathfrak{M}$ and causally impossible in $\mathfrak{M}_{O=0.4}$. Therefore, typically some atomic events in a variable's range are causally impossible, and events in a range shouldn't be confused with causally possible events. For any event from the range, however, there's an intervention that brings it about, and so ranges can be identified with sets of atomic events that interventions can bring about; two atomic events belong to the same range if and only if they cannot be brought about simultaneously. All conjunctive events over all variables (that is, points from the junction of all ranges) form—if you let me reuse a notion from physics—the state space of the system. Typically, only a small subset of this space is causally possible: causal possibility is narrower than state-space possibility. An analogy: in the standard representation of a classical-mechanical system of particles, the state space is all the combinations of position and momentum for the particles, considered independently of the constraints provided by the system's dynamics. Causally possible states of the system—that is, the ones that may be achieved given the dynamics and initial conditions—are a minuscule subset of all combinations.[10]

As for the talk of events, it should be interpreted rather loosely. I want to let the modeller use a variable value to denote, for example, my not going out, without committing herself to a particular ontology that allows for genuinely negative events. If a model where some values denote property instantiations yields correct counterfactuals, I will welcome such uses of my framework. For example, I would be more than happy to see an underdeterministic model used in fine-tuning debates to represent the claim that life is (causally) possible under a certain values of the gravitational constant but not others; that the constant takes on a particular value is hardly a paradigmatic case of an event. If there's a sensible interpretation of the modals and counterfactuals encoded in the model, it's a sufficient justification for using it in the context at hand. In the end, I'll just say this. First, formalism is patient, and I would not want metaphysical or semantic considerations to limit the use of the framework. Second, the question of the choice and interpretation of variables and their ranges isn't any different from when asked in the context of deterministic or probabilistic models.

5. Independence and The Causal Markov Condition

The framework doesn't deal with probabilities, but it does afford a very similar—and useful—notion of independence. Take disjoint variable sets $\vec{X}$ and $\vec{Y}$. $\vec{X}$ and $\vec{Y}$ are (unconditionally) independent in a model if and only if any possible conjunctive events over $\vec{X}$ and over $\vec{Y}$ are co-possible,

$$\vec{X} \perp\!\!\!\perp \vec{Y} \mid \varnothing \qquad \text{if and only if} \qquad \begin{array}{l} \mathfrak{M} \vDash \Diamond\, \vec{X}{=}\vec{x} \text{ and } \mathfrak{M} \vDash \Diamond\, \vec{Y}{=}\vec{y} \\ \text{implies} \quad \mathfrak{M} \vDash \Diamond \left(\vec{X}{=}\vec{x} \wedge \vec{Y}{=}\vec{y}\right) \qquad \text{for any } \vec{x}, \vec{y}. \end{array} \tag{9}$$

Take any variable set $\vec{Z}$ (which is allowed to overlap with $\vec{X}\vec{Y}$). $\vec{X}$ and $\vec{Y}$ are conditionally independent given $\vec{Z}$ in a model if and only if any possible conjunctive events over $\vec{X}\vec{Z}$ and over $\vec{Y}\vec{Z}$ that agree on $\vec{Z}$ are co-possible,

$$\vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z} \qquad \text{if and only if} \qquad \begin{array}{l} \mathfrak{M} \vDash \Diamond \left(\vec{X}{=}\vec{x} \wedge \vec{Z}{=}\vec{z}\right) \text{ and } \mathfrak{M} \vDash \Diamond \left(\vec{Y}{=}\vec{y} \wedge \vec{Z}{=}\vec{z}\right) \\ \text{implies} \quad \mathfrak{M} \vDash \Diamond \left(\vec{X}{=}\vec{x} \wedge \vec{Y}{=}\vec{y} \wedge \vec{Z}{=}\vec{z}\right) \qquad \text{for any } \vec{x}, \vec{y}, \vec{z}. \end{array} \tag{10}$$

Unconditional independence is a case of conditional independence for $\vec{Z}{=}\varnothing$.

Conditional independence explicates the notion of irrelevance: $\vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z}$ means that 'knowing $\vec{z}$ [the values of $\vec{Z}$] renders $\vec{x}$ irrelevant to $\vec{y}$' (Pearl and Paz [1985], p. 5). Pearl and Paz ([1985], p. 3) illustrate the notion as follows: 'In trying to predict whether I am going to be late

[10] Again, this is an analogy. Causal models can't straightforwardly represent continuous processes.

for a meeting, it is normally a good idea to ask somebody on the street for the time. However, once I establish the precise time by listening to the radio, asking people for the time becomes superfluous and their responses would be irrelevant'.

As an example, take O and S. They are unconditionally independent. Only two conjunctive events over O and two over S are possible: O=0.8, O=1, and S=0.7, S=0.9. And it turns out (table 1) that $\langle O, S\rangle = \langle 0.8, 0.7\rangle$, $\langle O, S\rangle = \langle 0.8, 0.9\rangle$, $\langle O, S\rangle = \langle 1, 0.7\rangle$, and $\langle O, S\rangle = \langle 1, 0.9\rangle$ are also possible. The definition is satisfied. Say, you know all counterfactual relationships between the events, that is, you know the model of the case. Once you learn what the oracle in fact proposes, it won't change what you believe the soothsayer may propose. And the other way around. However, the variables are conditionally dependent on A: both $\langle O, A\rangle = \langle 0.8, 0.85\rangle$ and $\langle S, A\rangle = \langle 0.7, 0.85\rangle$ are possible, but $\langle O, S, A\rangle = \langle 0.8, 0.7, 0.85\rangle$ is impossible, and equation 10 fails. Say, you learn that the second economist recommends a 0.85 tax. That doesn't narrow down what the soothsayer may propose (0.7 or 0.9). But once you also learn that the oracle proposes 0.8, you'll infer that the soothsayer proposes 0.9.

Notice that if some conjunctive event over $\vec{X}$ is necessary in a model, $\vec{X}$ are conditionally and unconditionally independent from all variables. For example, G=0.6 is necessary in the economists model, and all other variables are independent from G.

Verma and Pearl ([1990], p. 352) notice that 'most sensible definitions of [independence] share four common properties', which they formalize as four graphoid axioms: symmetry, decomposition, weak union, and contraction. Where $\vec{W}$ is disjoint with $\vec{Y}$, the axioms are:

Symmetry

If one piece of information is irrelevant for another, the latter is also irrelevant for the former,

$$\text{if} \quad \vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z} \quad \text{then} \quad \vec{Y} \perp\!\!\!\perp \vec{X} \mid \vec{Z}. \tag{11}$$

For example, if knowing the oracle's actual proposal doesn't help you narrow down the soothsayer's possible proposals, $O \perp\!\!\!\perp S \mid \varnothing$, then knowing the soothsayer's actual proposal won't help you narrow down the oracle's possible proposals, $S \perp\!\!\!\perp O \mid \varnothing$.[11]

Decomposition

If two pieces are irrelevant for a third piece, then either piece alone is also irrelevant,

$$\text{if} \quad \vec{X}\vec{W} \perp\!\!\!\perp \vec{Y} \mid \vec{Z} \quad \text{then} \quad \vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z}. \tag{12}$$

For example, knowing the second economist's recommendation won't help you narrow down the first's if you already know what the oracle proposed, $A \perp\!\!\!\perp P \mid O$; knowing the third's recommendation won't help you narrow down the first's if you already know what the oracle

[11] For another, very appealing illustration of the axioms, see (Lauritzen [1996], p. 30).

proposed, $T \perp\!\!\!\perp P \mid O$; therefore, knowing the second's and third's recommendations won't help you narrow down the first's if you already know what the oracle proposed, $A, T \perp\!\!\!\perp P \mid O$.

Weak Union

Learning new irrelevant information won't make relevant what's already irrelevant,

$$\text{if} \quad \vec{X}\vec{W} \perp\!\!\!\perp \vec{Y} \mid \vec{Z} \quad \text{then} \quad \vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z}\vec{W}. \tag{13}$$

For example, the second's and third economist's recommendations are irrelevant for the first's given the oracle's proposal, $A, T \perp\!\!\!\perp P \mid O$; therefore, the second's is irrelevant for the first's given the oracle's proposal and the third's recommendation, $A \perp\!\!\!\perp P \mid O, T$.

Contraction

A piece that is irrelevant after you learn some other irrelevant piece was already irrelevant before you learned that other piece,

$$\text{if} \quad \vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z} \quad \text{and} \quad \vec{W} \perp\!\!\!\perp \vec{Y} \mid \vec{X}\vec{Z} \quad \text{then} \quad \vec{X}\vec{W} \perp\!\!\!\perp \vec{Y} \mid \vec{Z}. \tag{14}$$

For example, the second economist's recommendation is irrelevant for the first's recommendation given the oracle's proposal, $A \perp\!\!\!\perp P \mid O$; the third's recommendation is irrelevant for the first's recommendation given the oracle's proposal, $T \perp\!\!\!\perp P \mid O, S$; therefore, the second's and third's recommendation are irrelevant for the first's given the oracle's proposal, $A, T \perp\!\!\!\perp P \mid O$.

Underdeterministic independence satisfies the axioms (sec. 6). Therefore, I can use it to formulate the causal Markov condition. The condition states that any node is conditionally independent from its non-descendants given its parents,

$$X \perp\!\!\!\perp \vec{N} \mid \vec{P}, \qquad \text{that is,} \quad \begin{array}{l} \mathfrak{M} \vDash \Diamond\left(X{=}x \wedge \vec{P}{=}\vec{p}\right) \text{ and } \mathfrak{M} \vDash \Diamond\left(\vec{N}{=}\vec{n} \wedge \vec{P}{=}\vec{p}\right) \\ \text{implies} \quad \mathfrak{M} \vDash \Diamond\left(X{=}x \wedge \vec{N}{=}\vec{n} \wedge \vec{P}{=}\vec{p}\right) \qquad \text{for any } x, \vec{n}, \vec{p}, \end{array} \tag{15}$$

where $\vec{P}$ are X's parents, and $\vec{N}$ are all the nodes that don't descend from X; therefore, $X \notin \vec{N}$, but $\vec{P} \subseteq \vec{N}$. However, as any variable is trivially independent from some other variables conditional on these very variables, it's fine to consider non-parent non-descendants $\vec{N} \backslash \vec{P}$ instead of all non-descendants $\vec{N}$ (Hausman and Woodward [1999], p. 523).

Underdeterministic models satisfy the causal Markov condition (sec. 6); roughly, the condition follows from the fact that a variable's possible values are determined solely by the values of its parents. An illustration: what the second economist recommends is independent from the recommendations of the other two given the oracle's and the soothsayer's proposals, $A \perp\!\!\!\perp P, T \mid O, S$. Four conjunctive events over O, S, A and eight over O, S, P, T are possible;

and any conjunction of one of the four and one of the eight is possible, provided the conjoined events agree on the values of O, S. For instance, $\langle O, S, A\rangle = \langle 0.8, 0.7, 0.75\rangle$ is possible, $\langle O, S, P, T\rangle = \langle 0.8, 0.7, 0.8, .7\rangle$ is possible, and their conjunction is also possible—it corresponds to the first solution to the model ($\vec{\sigma}_1$ in table 1).

From the graphoid axioms and the causal Markov condition follows (what I'll call) the independence theorem: if some variables d-separate two node sets from each other, the sets are conditionally independent given the separating variables,

$$\text{if } \vec{Z} \text{ d-separates } \vec{X} \text{ and } \vec{Y}, \text{ then } \vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z}, \tag{16}$$

where d-separation is defined graphically (Pearl [2000], p. 17). Variables $\vec{Z}$ block an undirected path if and only if

(1) the path contains a chain of nodes $\circ \rightarrow Z \rightarrow \circ$ or a fork $\circ \leftarrow Z \rightarrow \circ$ (where $Z \in \vec{Z}$, $\circ$ stands for any node, and arrows stand for edges);

(2) or the path contains a collider $\circ \rightarrow N \leftarrow \circ$ such that neither N nor any of its descendants are in $\vec{Z}$.

$\vec{Z}$ d-separate $\vec{X}$ and $\vec{Y}$ from each other if and only if $\vec{Z}$ block every undirected path between any node from $\vec{X}$ and any node from $\vec{Y}$. Verma and Pearl ([1988], pp. 355–57) prove that if a ternary relation on nodes of a digraph satisfies symmetry, decomposition, weak union, contraction, and the Markov condition, then the relation holds between every three sets of nodes where the last set d-separates the first two. Since underdeterministic independence satisfies the graphoid axioms and the causal Markov condition, any two sets of nodes in an underdeterministic are independent given any d-separating set of nodes—that is, the independence theorem holds of underdeterministic models.

As illustrations take the relationships already analysed. O and S don't share ancestors and therefore are d-separated by $\varnothing$, and therefore $O \perp\!\!\!\perp S \mid \varnothing$. Yet, $O \not\perp\!\!\!\perp S \mid A$, which is consistent with the fact that A, a collider, doesn't d-separate O and S. $P \not\perp\!\!\!\perp T \mid \varnothing$, which is consistent with the fact that P and T share an ancestor and therefore aren't d-separated by $\varnothing$. Yet, O d-separates P and T, and therefore $P \perp\!\!\!\perp T \mid O$. Notice, though, the converse is false: for example, $O \perp\!\!\!\perp S \mid T$, although O and S aren't d-separated by T.

From the independence theorem follows the intervention corollary: intervening on proper non-ancestors[12] of some variables doesn't change what events over these variables are possible (sec. 6). In term of diagrams, the corollary expresses the obvious property that if no directed paths lead from $\vec{N}$ to $\vec{X}$, intervening on $\vec{N}$ won't change the possible values of $\vec{X}$. For example, setting S to 0.5 won't change the possible values of P—in $\mathfrak{M}_{S=0.5}$, like in $\mathfrak{M}$, the first economist may recommend a 0.8 or 1 tax but no other.

[12] Proper non-ancestors of $\vec{X}$ are variables that neither are ancestors of any node from $\vec{X}$ nor belong to $\vec{X}$.

Table 2. Solutions to the sprinklers model.

	S	P	R	W	L		S	P	R	W	L
$\vec{v}_1$	2	1	0	0	0	$\vec{v}_4$	3	0	0	0	0
$\vec{v}_2$	2	1	0	1	0	$\vec{v}_5$	3	0	1	1	0
$\vec{v}_3$	2	1	0	1	1	$\vec{v}_6$	3	0	1	1	1

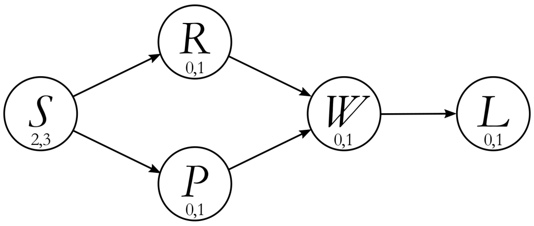

Figure 2. The sprinklers case.

Within the framework, you can also formulate the causal faithfulness condition, on which only variables that aren't d-separated are dependent (Pearl [2000]; Weinberger [2018]). The condition plays a crucial role in causal learning, that is, recovering causal models from data. The underdeterministic framework, as it stands, doesn't assume the faithfulness condition; and indeed, that the oracle's and the soothsayer's recommendations are independent conditional on the third economist's recommendation is a violation of faithfulness. However, I also haven't at all discussed causal learning of underdeterministic models. It is possible that recovering underdeterministic models would require assuming faithfulness.

Finally, I'll illustrate the entire framework with a second case—an underdeterministic version of the sprinklers case (fig. 2), which opens Pearl's ([2000], p. 15) seminal book. It may be summer or fall (S). Sprinklers are on (P) if it's spring or summer, and are off otherwise. It may rain (R) in spring or fall, but not in winter (it's too cold) or summer (because of global warming). The pavement may be wet (W) if the sprinklers have been on (because of occasional water shortages) and is wet if it has rained. The pavement may but needn't be slippery (L) if it's wet. The following equations encode the case:

$$S \leftarrow 2,3, \quad P \leftarrow S \leq 2, \quad R \leftarrow 0,(S \bmod 2), \quad W \leftarrow 1R{=}10,P, \quad L \leftarrow 0,W, \tag{17}$$

where $S = 1,2,3,4$ stands for spring, summer, fall, and winter, and for the remaining variables, 1 stands for true and 0 for false. The situation may develop in three ways (table 2).

In this model, P and R aren't independent, $P \not\perp\!\!\!\perp R \mid \varnothing$, because the sprinklers being on and it raining are possible yet not co-possible; the variables are independent conditional on the season, however, $P \perp\!\!\!\perp R \mid S$. L is independent from P and R when they are considered alone but not when they considered together, $LPR\varnothing$, because $L{=}1$ and $P{=}0 \wedge R{=}0$ are possible but not co-possible. L is independent from PR conditional on W, $L \perp\!\!\!\perp PR \mid W$, consistently with the causal Markov condition. It's causally impossible that the sprinklers are on and it's raining, $\mathfrak{M} \vDash \Box\neg(P{=}1 \wedge R{=}1)$. If the sprinklers were on and it rained, the pavement would have been wet, yet it still might have not been slippery, $\mathfrak{M} \models P{=}1 \wedge R{=}1 \;\Box\!\!\rightarrow W{=}1$ and $\mathfrak{M} \models P{=}1 \wedge R{=}1 \;\Diamond\!\!\rightarrow L{=}0$. Had been the sprinklers had been off in spring or fall, the pavement would have been neither wet nor slippery, $\mathfrak{M} \models (S{=}2 \vee S{=}4) \wedge P{=}0 \;\Box\!\!\rightarrow W{=}0 \wedge L{=}0$.

6. Conclusion

Causal underdeterminism is a genuine causal phenomenon, as witnessed by the fact that underdeterministic models can represent situations that may develop in multiple ways, that causal modals and counterfactuals have a coherent, intelligible semantics, that underdeterministic dependencies satisfy the graphoid axioms, and that underdeterministic models satisfy the causal Markov condition. Deterministic and probabilistic causal models just got a sibling.

What's next? First awaits the analysis of the underdeterministic phenomena that I already mentioned: type and token causation, a full semantics of counterfactuals, an underdeterministic decision theory, an algorithm for recovering an underdeterministic model from data, an empirical investigation of underdeterministic causation and its relation to probabilistic causation. Another task is to investigate whether the underdeterministic framework can be generalized, possibly unifying underdeterministic and probabilistic models. Imagine, for instance, a framework where structural equations returned multiple probabilistic distributions; there, underdeterministic models are edge cases where equations return many one-point distributions, and probabilistic models are edge cases where equations return single distributions.[13] Of course, it remains to be seen if such framework comes with its own notion of conditional independence that could underlie causal relations. I am hopeful.

Appendix

A1. Prerequisites

The proofs heavily rely on the notions of possibility sets and spatial concatenation.

A possibility set contains the projections onto given variables $\vec{X}$ of all solutions where a certain event φ happens:

$$\mathfrak{M}^{\varphi}[\vec{X}] = \left\{\vec{\sigma}[\vec{X}] \colon \vec{\sigma} \text{ is a solution to } \mathfrak{M} \text{ that satisfies } \varphi\right\}. \tag{18}$$

Therefore, $\vec{X}$'s possibility set conditional on φ consists of all values of $\vec{X}$ from the solutions in which φ occurs.[14] You can think of the possibility set as the sum of all conjunctive events over $\vec{X}$ that happen in the solutions that satisfy φ. Notice that φ can be over variables other

[13] I owe this suggestion to one of the reviewers.

[14] Possibility sets have an illustrative analogue in the SQL language. If you treat the set of all solutions as a table with $\mathcal{V}$ as columns, $\mathfrak{M}^{\varphi}[\vec{X}]$ reads as: select distinct $\vec{X}$ from solutions where φ. Moreover, junction ($\bowtie$) corresponds to joining tables on all shared variables.

than $\vec{X}$. $\mathfrak{M}^{\top}[\vec{X}]$ denotes the projection of all solutions onto $\vec{X}$, and therefore $\mathfrak{M}^{\top}[\mathcal{V}]$ denotes the set of all solutions. An illustration: in the economists model, $\mathfrak{M}^{O=0.8}[A] = \{\langle 0.75_A\rangle, \langle 0.85_A\rangle\}$; $\mathfrak{M}^{S>O}[P,T] = \{\langle 0.8_P, 0.8_T\rangle, \langle 0.8_P, 0.9_T\rangle\}$; $\mathfrak{M}^{P>10}[P,A,T] = \varnothing$.

Hereafter, assume that $\vec{x}$ is over $\vec{X}$, $\vec{y}$ is over $\vec{Y}$, $\vec{z}$ is over $\vec{Z}$, and $\vec{w}$ is over $\vec{W}$ unless specified otherwise. Spatial concatenation combines two assignments into one that agrees with either assignment on shared variables:

$$\vec{z} = \vec{x} \star \vec{y} \quad \text{if and only if} \quad \vec{z}[\vec{X}] = \vec{x} \text{ and } \vec{z}[\vec{Y}] = \vec{y}, \text{ where } \vec{z} \text{ is over } \vec{X}\vec{Y}. \tag{19}$$

For example, $\langle 0.8_O, 0.9_S\rangle \star \langle 0.9_S, 0.9_A\rangle = \langle 0.8_O, 0.9_S, 0.9_A\rangle$. If the assignments disagree on some shared value, the operation is indeterminate. Spatial concatenation is associative, commutative, and idempotent. The empty assignment $\langle\rangle$ is the operation's one (the neutral element). I'll typically omit brackets for assignments over single variables: $z \star \vec{x}$ means $\langle z\rangle \star \vec{x}$.

Junction is defined using spatial concatenation:

$$U \bowtie V = \bigcup_{\vec{x}\in U} \bigcup_{\vec{y}\in V} \{\vec{x} \star \vec{y}\}, \tag{20}$$

where U contains assignments over $\vec{X}$, V over $\vec{Y}$, and $U \bowtie V$ over $\vec{X}\vec{Y}$. (If $\vec{x} \star \vec{y}$ is indeterminate, just move on to the next pair of assignments.) Junction is associative and commutative, $\{\langle\rangle\}$ is its one, and $\varnothing$ is its zero (the annihilating element). If $\vec{X}$ and $\vec{Y}$ are disjoint, the combination of every element from U and every element from V ends up in $U \bowtie V$ because $\vec{x} \star \vec{y}$ is always determinate, that is, for any $\vec{x}$, $\vec{y}$,

$$\text{if} \quad \vec{X} \cap \vec{Y} = \varnothing \quad \text{then} \qquad \vec{x} \star \vec{y} \in U \bowtie V \quad \text{if and only if} \quad \vec{x} \in U \quad \text{and} \quad \vec{y} \in V. \tag{21}$$

The definition entails that

$$\bigcup_i \left(U_i \bowtie V\right) = \bigcup_i \bigcup_{\vec{x}\in U_i} \bigcup_{\vec{y}\in V} \{\vec{x} \star \vec{y}\} = \bigcup_{\vec{x}\in \bigcup_i U_i} \bigcup_{\vec{y}\in V} \{\vec{x} \star \vec{y}\} = \left(\bigcup_i U_i\right) \bowtie V, \tag{22}$$

where U_i are over the same variables.

The following identities will prove useful: For any $\vec{x}$ and $\vec{z}$,

$$\vec{x} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] \quad \text{if and only if} \quad \exists \vec{y}.\ \vec{x} \star \vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}], \tag{23}$$

which holds because $\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]$ is a projection of $\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}]$ onto $\vec{X}$.

For any $\vec{x}$, $\vec{y}$, and $\vec{z}$,

$$\vec{x} \in \mathfrak{M}^{\vec{Y}=\vec{y}\wedge\vec{Z}=\vec{z}}[\vec{X}] \quad \text{if and only if} \quad \vec{x} \star \vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}], \tag{24}$$

which holds because checking if $\vec{X}=\vec{x}$ holds in any of the solutions that satisfy $\vec{Y}=\vec{y}$ and $\vec{Z}=\vec{z}$ is equivalent to checking if $\vec{X}{=}\vec{x} \wedge \vec{Y}{=}\vec{y}$ holds in any of the solutions that satisfy $\vec{Z}{=}\vec{z}$. If $\vec{x} \star \vec{y}$ is indeterminate, assume (here and in similar cases) that the expression on the right side is false (rather than itself indeterminate).

For $\vec{x} = \vec{y}$, equation 24 implies that

$$\vec{y} \in \mathfrak{M}^{\vec{Y}=\vec{y} \wedge \vec{Z}=\vec{z}}[\vec{Y}] \quad \text{if and only if} \quad \vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}], \tag{25}$$

and for $\vec{x} = \langle\rangle$, it implies that

$$\mathfrak{M}^{\vec{Y}=\vec{y} \wedge \vec{Z}=\vec{z}}[\varnothing] \neq \varnothing \quad \text{if and only if} \quad \vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]. \tag{26}$$

Together, equation 23 and equation 24 also entail that

$$\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] = \bigcup_{\vec{y}} \mathfrak{M}^{\vec{Y}=\vec{y} \wedge \vec{Z}=\vec{z}}[\vec{X}] = \bigcup_{\vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]} \mathfrak{M}^{\vec{Y}=\vec{y} \wedge \vec{Z}=\vec{z}}[\vec{X}]. \tag{27}$$

A2. The graphoid axioms

I'll now rewrite the definition of conditional independence equation 10 as

$$\vec{X}\vec{Y} \mid \vec{Z} \quad \text{if and only if} \quad \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] = \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \quad \text{for any } \vec{z}. \tag{28}$$

In what follows, $\vec{X}\vec{W}$ and $\vec{Y}$ are disjoint. Conditional independence satisfies symmetry (equ. 11), which now reads as

$$\begin{aligned} \text{if} \quad \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] &= \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \\ \text{then} \quad \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] &= \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]. \end{aligned} \tag{29}$$

Symmetry holds because junction is commutative.

Conditional independence satisfies decomposition (equ. 12), which now reads as: for any $\vec{z}$,

$$\begin{aligned} \text{if} \quad \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}] &= \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \\ \text{then} \quad \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] &= \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]. \end{aligned} \tag{30}$$

Assume the antecedent of the axiom; decomposition holds because

$$\begin{aligned}
\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] &\overset{1}{=} \left\{\vec{x}\star\vec{y}\colon\ \exists\vec{w}.\ \vec{x}\star\vec{y}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}]\right\}\\
&\overset{2}{=} \left\{\vec{x}\star\vec{y}\colon\ \exists\vec{w}.\ \vec{x}\star\vec{y}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\bowtie\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\right\}\\
&\overset{3}{=} \left\{\vec{x}\star\vec{y}\colon\ \exists\vec{w}.\ \vec{x}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\wedge\vec{y}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\right\}\\
&\overset{4}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]\wedge\vec{y}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\right\}\\
&\overset{5}{=} \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]\bowtie\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}].
\end{aligned}$$

The first transition holds in virtue of equation 23. The second holds in virtue of the antecedent. The third holds in virtue of equation 21. The fourth holds in virtue of equation 23. The fifth holds in virtue of equation 21.

Conditional independence satisfies weak union (equ. 13), which now reads as: for any $\vec{z}$ and $\vec{w}$,

$$\begin{aligned}
&\text{if}\quad \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}] = \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\bowtie\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\\
&\quad\text{then}\quad \mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] = \mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{X}]\bowtie\mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{Y}].
\end{aligned}\tag{31}$$

Assume the antecedent of the axiom; weak union holds because

$$\begin{aligned}
\mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] &\overset{1}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\star\vec{y}\star\vec{w}\in\mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}]\right\}\\
&\overset{2}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\star\vec{y}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}]\right\}\\
&\overset{3}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\star\vec{y}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\bowtie\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\right\}\\
&\overset{4}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\wedge\vec{y}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\right\}\\
&\overset{5}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\wedge\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{W}]\wedge\vec{y}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}]\right\}\\
&\overset{6}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\wedge\vec{y}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}\vec{W}]\right\}\\
&\overset{7}{=} \left\{\vec{x}\star\vec{y}\colon\ \vec{x}\in\mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{X}]\wedge\vec{y}\in\mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}\right\}[\vec{Y}]\\
&\overset{8}{=} \mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{X}]\bowtie\mathfrak{M}^{\vec{W}=\vec{w}\wedge\vec{Z}=\vec{z}}[\vec{Y}].
\end{aligned}$$

I'll explain only the unfamiliar moves. The second transition holds in virtue of equation 24. The third holds in virtue of the antecedent. The fifth holds because $\vec{x}\star\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]$ entails $\vec{w}\in\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{W}]$, as the second set is a projection of the first. The seventh holds in virtue of equation 24.

Conditional independence satisfies contraction (equ. 14), which now reads as: for any $\vec{z}$,

$$\begin{aligned}\text{if}\quad & \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] = \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \\ \text{and}\quad & \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{Y}\vec{W}] = \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{Y}] \bowtie \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{W}] \text{ for any } \vec{x} \\ \text{then}\quad & \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}] = \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}].\end{aligned} \tag{32}$$

Assume the antecedent of the axiom; contraction holds because

$$\begin{aligned}\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}] &\overset{1}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{x}\star\vec{y}\star\vec{w} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}\vec{W}]\right\} \\ &\overset{2}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{y}\star\vec{w} \in \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{Y}\vec{W}]\right\} \\ &\overset{3}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{y}\star\vec{w} \in \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{Y}] \bowtie \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{W}]\right\} \\ &\overset{4}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{y} \in \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{Y}] \wedge \vec{w} \in \mathfrak{M}^{\vec{X}=\vec{x}\wedge\vec{Z}=\vec{z}}[\vec{W}]\right\} \\ &\overset{5}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{x}\star\vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{Y}] \wedge \vec{x}\star\vec{w} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\right\} \\ &\overset{6}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{x} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}] \wedge \vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \wedge \vec{x}\star\vec{w} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\right\} \\ &\overset{7}{=} \left\{\vec{x}\star\vec{y}\star\vec{w} : \vec{y} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}] \wedge \vec{x}\star\vec{w} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]\right\} \\ &\overset{8}{=} \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}] \bowtie \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{Y}].\end{aligned}$$

The second and fifth transition hold in virtue of (equ. 24). The third holds in virtue of the second conjunct in the axiom's antecedent. The seventh holds because $\vec{x}\star\vec{w} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}\vec{W}]$ entails $\vec{x} \in \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]$, and the latter conjunct can be omitted.

Another useful fact: variables that assume a single necessary value make no difference to dependence relations,

$$\vec{X}\vec{Y} \mid \vec{F}\vec{Z} \quad \text{if and only if} \quad \vec{X} \perp\!\!\!\perp \vec{Y} \mid \vec{Z} \qquad \text{where} \quad \mathfrak{M} \vDash \Box\, \vec{F}=\vec{f}, \tag{33}$$

which holds because $\mathfrak{M}^{\vec{F}=\vec{f}\wedge\vec{Z}=\vec{z}}[\vec{X}] = \mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]$ if all solutions satisfy $\vec{F}=\vec{f}$.

A3. The causal Markov condition

The proof requires that I first describe how a model's solutions are produced. A model has L variables, ordered topologically; the ith equation determines the values of the ith variable.

Let Σ_i be the set of all solutions to the first i equations. To produce Σ_i, $i \geq 1$, take every partial solution $\sigma_{(i-1)}$ from Σ_{i-1}, where $\Sigma_0 = \{\langle\rangle\}$. Next, take the set of all values of the ith variable that are possible given the values that $\vec{P}$, the ith variable's parents, take on in that solution: ${}_i\big(\sigma_{(i-1)}[\vec{P}]\big)$, where f_i is the ith equation. For every such possible value, append it to the partial solution it was produced from and put the new partial solution in Σ_i:[15]

$$\Sigma_i = \left\{\sigma_{(i-1)} \star v \colon\ \sigma_{(i-1)} \in \Sigma_{i-1} \wedge v \in_i \left(\sigma_{(i-1)}[\vec{P}]\right)\right\}. \tag{34}$$

The set of all solutions to $\mathfrak{M}$, $\mathfrak{M}^{\top}[\mathcal{V}]$, just is Σ_L.

Every partial solution from Σ_j, $j > i$, contains as a prefix some partial solution from Σ_i, and every partial solution from Σ_i extends to at least one partial solution from Σ_j (because equations return at least one value). Therefore, Σ_i is the projection of Σ_j onto the first i variables. For j=L, this reads as

$$\Sigma_i = \mathfrak{M}^{\top}[V_1, \ldots, V_i], \tag{35}$$

where V_i is the ith variable.

The two identities entail that the possibility set of $\vec{X}\vec{Y}$, where $\vec{X}$ are topologically prior to $\vec{Y}$ and include $\vec{Y}$'s every parent, is determined by $\vec{Y}$'s equation and the possibility set of $\vec{X}$,

$$\mathfrak{M}^{\top}[\vec{X}\vec{Y}] = \left\{\vec{x} \star \vec{y} \colon\ \vec{x} \in \mathfrak{M}^{\top}[\vec{X}] \ \wedge\ \forall i.\ \vec{y}[Y_i] \in f_i\left([\vec{P}_i]\right)\right\}, \tag{36}$$

where Y_i is that the ith variable in $\vec{Y}$, f_i is its equation, and $\vec{P}_i$ are the variable's parents, $\vec{P}_i \subseteq \vec{X}$.[16]

In virtue of the definition of the possibility set, if φ is over some variables from $V_1, \ldots, V_i$,

$$\mathfrak{M}^{\varphi}[V_1, \ldots, V_i] = \left\{\vec{\sigma}_{(i)} \colon\ \vec{\sigma}_{(i)} \in \Sigma_o \wedge \vec{\sigma}_{(i)} \text{ statisfies } \varphi\right\}. \tag{37}$$

In short, $\mathfrak{M}^{\varphi}[V_1, \ldots, V_i]$ is Σ_i filtered by φ.

Now, for the proof of the causal Markov condition. The condition states that any node is conditionally independent from its non-descendants given its parents,

$$X \perp\!\!\!\perp \vec{N} \mid \vec{P}, \qquad \text{that is,} \qquad \mathfrak{M}^{\vec{P}=\vec{p}}[X\vec{N}] = \mathfrak{M}^{\vec{P}=\vec{p}}[X] \bowtie \mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \tag{38}$$

where $\vec{P}$ are X's parents, and $\vec{N}$ are all the nodes that don't descend from $\vec{X}$; therefore, $X \notin \vec{N}$, but $\vec{P} \subseteq \vec{N}$.

Assume that in the topological order, X's non-descendants $\vec{N}$ come before X; if the order doesn't satisfy this condition, you can always rearrange the variables. Since $\vec{N}$ are X's all non-descendants, $X\vec{N}$ is closed on ancestors. Hence, if X is the ith variable, $X\vec{N}$ contains the first i

[15] For exogenous nodes, $\vec{P} = \varnothing$, and the condition on v is $v \in f_i$.

[16] The full proof would be inductive over the number of variables in $\vec{Y}$; I won't give it here, however, as it's trivial.

variables, and $\vec{N}$ contains the first $i-1$ variables. Therefore, for any $\vec{p}$ over $\vec{P}$,

$$\begin{aligned}
\mathfrak{M}^{\vec{P}=\vec{p}}[X\vec{N}] &\stackrel{1}{=} \left\{\sigma_{(i)}\colon \sigma_{(i)} \in \Sigma_i \wedge \sigma_{(i)}[\vec{P}] = \vec{p}\right\} \\
&\stackrel{2}{=} \left\{\sigma_{(i-1)} \star x\colon \sigma_{(i-1)} \in \Sigma_{i-1} \wedge x \in f_x\left(\sigma_{(i-1)}[\vec{P}]\right) \wedge \sigma_{(i-1)}[\vec{P}] = \vec{p}\right\} \\
&\stackrel{3}{=} \left\{\sigma_{(i-1)} \star x\colon \sigma_{(i-1)} \in \Sigma_{i-1} \wedge x \in f_X\left(\vec{p}\right) \wedge \sigma_{(i-1)}[\vec{P}] = \vec{p}\right\} \\
&\stackrel{4}{=} \left\{\sigma_{(i-1)} \star x\colon \sigma_{(i-1)} \in \mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \wedge x \in f_X\left(\vec{p}\right)\right\} \\
&\stackrel{5}{=} \mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \bowtie f_X\left(\vec{p}\right) \\
&\stackrel{6}{=} \mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \bowtie \mathfrak{M}^{\vec{P}=\vec{p}}[X]
\end{aligned} \tag{39}$$

The first and fourth transition hold in virtue of equation 37 and $\vec{P} \subseteq \vec{N}$. The second holds in virtue of equation 34. The fifth holds in virtue of equation 21. In the sixth transition, it's not simply the case that for any $\vec{p}$, $f_X\left(\vec{p}\right) = \mathfrak{M}^{\vec{P}=\vec{p}}[X]$ because if $\vec{P}=\vec{p}$ is impossible on $\mathfrak{M}$, $\mathfrak{M}^{\vec{P}=\vec{p}}[X] = \varnothing$, whereas a structural equation always returns a non-empty set. However, if $\vec{P}=\vec{p}$ is impossible, $\mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] = \varnothing$ too, and therefore $\mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \bowtie f_X\left(\vec{p}\right) = \varnothing$; and if $\vec{P}=\vec{p}$ is possible, $f_X\left(\vec{p}\right) = \mathfrak{M}^{\vec{P}=\vec{p}}[X]$ because all values of X returned by its equation applied to $\vec{p}$ are possible. In either case, $\mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \bowtie f_X\left(\vec{p}\right) = \mathfrak{M}^{\vec{P}=\vec{p}}[\vec{N}] \bowtie \mathfrak{M}^{\vec{P}=\vec{p}}[X]$, and the sixth transition holds.

A4. The agreement theorem and the intervention corollary

The intervention corollary states that intervening on proper non-ancestors of some variables doesn't change what events over these variables are possible. I'll prove a stronger agreement theorem: if two models agree on $\vec{X}$'s ranges and equations, $\vec{X}$ and $\vec{P}$ don't overlap, all $\vec{X}$ are screened off by $\vec{P}$ from all $\vec{X}$'s ancestors not in $\vec{X}$, and the models agree on $\vec{P}$'s possible values conditional on some event over any subset of $\vec{P}\vec{X}$, then the models agree on the possibility sets of any variables from $\vec{P}\vec{X}$ conditional on that event. That is,

$$\text{if} \quad \mathfrak{M}^{\varphi}[\vec{P}] = \mathfrak{m}^{\varphi}[\vec{P}] \quad \text{then} \quad \mathfrak{M}^{\varphi}[\vec{Y}] = \mathfrak{m}^{\varphi}[\vec{Y}], \tag{40}$$

where $\mathfrak{M}$ and $\mathfrak{m}$ agree on the ranges and equations of $\vec{X}$, $\vec{Y} \subseteq \vec{P}\vec{X}$, and φ is over at most $\vec{P}\vec{X}$. The theorem says that if you focus solely on $\vec{P}\vec{X}$ in two models, where all $\vec{X}$ ultimately descend from $\vec{P}$ alone, make sure that $\vec{P}$ agree on their possible values conditional on φ, and make sure that $\vec{X}$ agree on their equations, then the two fragments of models, when filtered by φ, fully agree with each other.

The proof is by induction on n, the longest path from $\vec{P}$ to $\vec{Y}$. For the base step, take n=0, which means that $\vec{Y} \subseteq \vec{P}$, and φ is over at most $\vec{P}$. In that case, $\mathfrak{M}^{\varphi}[\vec{P}] = \mathfrak{m}^{\varphi}[\vec{P}]$ entails $\mathfrak{M}^{\varphi}[\vec{Y}] = \mathfrak{m}^{\varphi}[\vec{Y}]$, as the latter sets are projections of the former onto $\vec{Y}$.

For the inductive step, take n=N and assume the theorem holds for any shorter path. Assume, for now, that $\vec{Y}$ is closed on all ancestors from $\vec{P}\vec{X}$; therefore, in particular, $\vec{P} \subseteq \vec{Y}$. Let $\vec{Y} = \vec{P} \uplus \vec{H} \uplus \vec{T}$, where the assumption holds of $\vec{H}$ but not of $\vec{T}$, that is, all paths from $\vec{P}$ to $\vec{H}$ are shorter than N edges, and all paths from $\vec{P}$ to $\vec{T}$ have at most N edges (and at least one edge). Per the theorem's antecedent, assume that $\mathfrak{M}^{\varphi}[\vec{P}] = \mathfrak{m}^{\varphi}[\vec{P}]$, where φ is over any subset of $\vec{P}$.

The idea behind the inductive step is to move from $\vec{P}\vec{H}\vec{T}$ to $\vec{P}\vec{H}$, apply the inductive assumption, and move back to $\vec{P}\vec{H}\vec{T}$. Since φ can be over $\vec{P}\vec{H}\vec{T}$, however, I need a different condition; denote it by φ^*. As φ^* take an event name over $\vec{P}\vec{H}$ such that $\vec{p} \star \vec{h}$ satisfies φ^* if and only if $\vec{p} \star \vec{h} \in \mathfrak{M}^{\varphi}[\vec{P}\vec{H}]$. For now, assume that you can always choose such a name; I'll prove that it is so momentarily. First, the proof:

$$
\begin{aligned}
\mathfrak{M}^{\varphi}[\vec{Y}] &\overset{1}{=} \mathfrak{M}^{\varphi}[\vec{P}\vec{H}\vec{T}] \\
&\overset{2}{=} \left\{\vec{p} \star \vec{h} \star \vec{t}\colon\ \vec{p} \star \vec{h} \star \vec{t} \in \mathfrak{M}^{\top}[\vec{P}\vec{H}\vec{T}] \ \wedge\ \vec{p} \star \vec{h} \star \vec{t} \text{ satisfies } \varphi\right\} \\
&\overset{3}{=} \left\{\vec{p} \star \vec{h} \star \vec{t}\colon\ \forall i.\ \vec{t}[T_i] \in f_i\big(\vec{p} \star \vec{h}\big) \ \wedge\ \vec{p} \star \vec{h} \in \mathfrak{M}^{\top}[\vec{P}\vec{H}] \ \wedge\ \vec{p} \star \vec{h} \star \vec{t} \text{ satisfies } \varphi\right\} \\
&\overset{4}{=} \left\{\vec{p} \star \vec{h} \star \vec{t}\colon\ \forall i.\ \vec{t}[T_i] \in f_i\big(\vec{p} \star \vec{h}\big) \ \wedge\ \vec{p} \star \vec{h} \in \mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}] \ \wedge\ \vec{p} \star \vec{h} \star \vec{t} \text{ satisfies } \varphi\right\} \\
&\overset{5}{=} \left\{\vec{p} \star \vec{h} \star \vec{t}\colon\ \forall i.\ \vec{t}[T_i] \in f_i\big(\vec{p} \star \vec{h}\big) \ \wedge\ \vec{p} \star \vec{h} \in \mathfrak{m}^{\varphi^*}[\vec{P}\vec{H}] \ \wedge\ \vec{p} \star \vec{h} \star \vec{t} \text{ satisfies } \varphi\right\} \\
&\overset{6}{=} \left\{\vec{p} \star \vec{h} \star \vec{t}\colon\ \forall i.\ \vec{t}[T_i] \in f_i\big(\vec{p} \star \vec{h}\big) \ \wedge\ \vec{p} \star \vec{h} \in \mathfrak{m}^{\top}[\vec{P}\vec{H}] \ \wedge\ \vec{p} \star \vec{h} \star \vec{t} \text{ satisfies } \varphi\right\} \\
&\overset{7}{=} \left\{\vec{p} \star \vec{h} \star \vec{t}\colon\ \vec{p} \star \vec{h} \star \vec{t} \in \mathfrak{m}^{\top}[\vec{P}\vec{H}\vec{T}] \ \wedge\ \vec{p} \star \vec{h} \star \vec{t} \text{ satisfies } \varphi\right\} \\
&\overset{8}{=} \mathfrak{m}^{\varphi}[\vec{P}\vec{H}\vec{T}] \overset{9}{=} \mathfrak{m}^{\varphi}[\vec{Y}],
\end{aligned}
$$

where $\vec{p}$, $\vec{h}$, $\vec{t}$ are over $\vec{P}$, $\vec{H}$, $\vec{T}$; T_i is the ith variable in $\vec{T}$; and f_i is this variable's equation. I abused notation: since f_i might map the values of a subset of $\vec{P}\vec{H}$ onto the values of T_i, read $f_i\big(\vec{p} \star \vec{h}\big)$ as if the function ignored all values from $\vec{p} \star \vec{h}$ that aren't the values of T_i's parents.

Crucial is the construction of φ^*. If $\mathfrak{M}^{\varphi}[\vec{P}\vec{H}]$ is finite, that's easy—just take φ^* as the disjunction $\vec{P}\vec{H} = \vec{v}_1 \vee \cdots \vee \vec{P}\vec{H} = \vec{v}_m$, where $\mathfrak{M}^{\varphi}[\vec{P}\vec{H}] = \{\vec{v}_1, \ldots, \vec{v}_m\}$. For such φ^*, the fourth and sixth transition go through in virtue of the definition of the possibility set. However, for if $\mathfrak{M}^{\varphi}[\vec{P}\vec{H}]$ is infinite, I need a different construction. Instead, as φ^* take

$$
\exists t_1, \ldots, t_k.\ \varphi[t_1, \backslash T_1, \ldots, t_k, \backslash T_k] \ \wedge\ t_1 \in f_1\big(\vec{P}\vec{H}\big) \ \wedge\ \ldots \ \wedge\ t_k \in f_k\big(\vec{P}\vec{H}\big), \tag{41}
$$

where there are k variables in $\vec{T}$, and the expression in the brackets indicates that you substitute t_1 for every occurrence of T_1 in φ, t_2 for every occurrence of T_2, and so on.[17] (I again abused notation when listing arguments in f_i.) The name is over $\vec{P}\vec{H}$, as all occurrences of $\vec{T}$ have been replaced with bound variables.

The second and eighth transition hold in virtue of equation 18, the definition of a possibility set. The third and seventh holds in virtue of equation 36 applied $\vec{T}$ and $\vec{P}\vec{H}$.

The fourth and sixth hold because for any $\vec{p} \star \vec{h} \star \vec{t}$ that satisfies φ, $\vec{p} \star \vec{h}$ is in $\mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}]$ (in $\mathfrak{m}^{\varphi^*}[\vec{P}\vec{H}]$) if and only if it's in $\mathfrak{M}^{\top}[\vec{P}\vec{H}]$ (in $\mathfrak{m}^{\top}[\vec{P}\vec{H}]$). Left-to-right: if $\vec{p} \star \vec{h} \in \mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}]$, then $\vec{p} \star \vec{h} \in \mathfrak{M}^{\top}[\vec{P}\vec{H}]$ because $\mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}] \subseteq$

[17] The syntax of the framework is flexible enough to count φ^* as a well-formed event name.

$mathfrakM^{\top}[\vec{P}\vec{H}]$. Right-to-left: if $\vec{p}\star\vec{h}\in\mathfrak{M}^{\top}[\vec{P}\vec{H}]$, then $\vec{p}\star\vec{h}\in\mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}]$ because that $\vec{p}\star\vec{h}\star\vec{t}$ satisfies φ means that $\vec{p}\star\vec{h}$ satisfies φ where values from $\vec{t}$ are substituted for the corresponding variable names from $\vec{T}$, which in turn means that $\vec{p}\star\vec{h}$ satisfies φ^*. The same considerations apply to $\mathfrak{m}$.

The fifth transition holds in virtue of the inductive assumption. But for the inductive assumption to apply, it must be that $\mathfrak{M}^{\varphi^*}[\vec{P}]=\mathfrak{m}^{\varphi^*}[\vec{P}]$. And it is. φ^* is constructed in such a way that $\vec{p}\star\vec{h}$ is possible and satisfies φ^* if and only if there's some $\vec{t}$ such that $\vec{p}\star\vec{h}\star\vec{t}$ is possible and satisfies φ. That means that $\mathfrak{M}^{\varphi}[\vec{P}\vec{H}]=\mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}]$, which in turn entails that $\mathfrak{M}^{\varphi}[\vec{P}]=\mathfrak{M}^{\varphi^*}[\vec{P}]$. Analogously for $\mathfrak{m}$. Because $\mathfrak{M}^{\varphi}[\vec{P}]=\mathfrak{m}^{\varphi}[\vec{P}]$ per the theorem's antecedent, $\mathfrak{M}^{\varphi^*}[\vec{P}]=\mathfrak{m}^{\varphi^*}[\vec{P}]$, and the inductive assumption applies: $\mathfrak{M}^{\varphi^*}[\vec{P}\vec{H}]=\mathfrak{m}^{\varphi^*}[\vec{P}\vec{H}]$. The fifth transition goes through.

This completes the proof for $\vec{Y}$ that are closed on their ancestors from $\vec{P}\vec{X}$. If $\vec{Y}$ doesn't satisfy this requirement, notice that $\vec{P}\vec{X}$ does satisfy the requirement, and therefore that $\mathfrak{M}^{\varphi}[\vec{P}]=\mathfrak{m}^{\varphi}[\vec{P}]$ entails $\mathfrak{M}^{\varphi}[\vec{P}\vec{X}]=\mathfrak{m}^{\varphi}[\vec{P}\vec{X}]$. Thus, project $\mathfrak{M}^{\varphi}[\vec{P}\vec{X}]$ and $\mathfrak{m}^{\varphi}[\vec{P}\vec{X}]$ onto $\vec{Y}$ to obtain $\mathfrak{M}^{\varphi}[\vec{Y}]=\mathfrak{m}^{\varphi}[\vec{Y}]$, which completes the proof for any $\vec{Y}$.

The intervention corollary is a consequence of equation 40. Let $\dot{\rho}$ be any simple event over $\vec{Y}$, some proper non-ancestors of $\vec{X}$. On the corollary, intervening on $\vec{Y}$ won't modify $\vec{X}$'s possible values:

$$\mathfrak{M}^{\top}_{\dot{\rho}}[\vec{X}]=\mathfrak{M}^{\top}[\vec{X}]. \tag{42}$$

Let $\vec{A}$ be $\vec{X}$'s ancestors. Rearrange the variables so $\vec{X}\vec{A}$ come before $\vec{Y}$. Bringing about $\dot{\rho}$ will now modify the equations of the variables topologically posterior to $\vec{X}$, which means that $\mathfrak{M}$ and $\mathfrak{M}_{\dot{\rho}}$ agree on $\vec{X}\vec{A}$'s equations. Since $\vec{X}\vec{A}$ are screened by $\vec{P}=\varnothing$ from their own ancestors not in $\vec{X}\vec{A}$ (as $\vec{X}\vec{A}$ is closed on ancestors), and $\mathfrak{M}^{\top}_{\varnothing}[\vec{X}]=\{\langle\rangle\}=\mathfrak{M}^{\top}[\varnothing]$, the agreement theorem applies (for $\top$ as φ), and the corollary holds.

The agreement theorem entails another corollary. If $\vec{Z}$ screen off $\vec{X}$ from $\vec{Z}$'s ancestors, deterministically intervening on $\vec{Z}$ results in the same possible values of $\vec{X}$ as conditionalizing on $\vec{Z}$, provided $\vec{Z}$'s values are possible: for any $\vec{z}$ from $\mathfrak{M}^{\top}[\vec{Z}]$,

$$\mathfrak{M}^{\vec{Z}=\vec{z}}[\vec{X}]=\mathfrak{M}^{\vec{Z}=\vec{z}}_{\vec{Z}=\vec{z}}[\vec{X}]=\mathfrak{M}^{\top}_{\vec{Z}=\vec{z}}[\vec{X}]. \tag{43}$$

The first identity holds in virtue of equation 40. The second identity holds because all solutions to $\mathfrak{M}_{\vec{Z}=\vec{z}}$ already satisfy $\vec{Z}$=$\vec{z}$.

Acknowledgements

I owe many thanks to many people. Colin Allen, Dmitri Gallow, Clark Glymour, Christopher Hitchcock, Hanti Lin, Jim Woodward, and many others gave me valuable feedback. I'm extremely grateful to Justin Shin, who read through my proofs. The suggestions of the two reviewers of this journal improved this manuscript immensely.

Department of History and Philosophy of Science

University of Pittsburgh
Pittsburgh, PA, USA
tomwysocki@pitt.edu

References

Arkes, H. R., Gigerenzer, G. and Hertwig, R. [2016]: ‘How Bad Is Incoherence?’, *Decision*, **3**, pp. 20–39.

Barbero, F. and Sandu, G. [2021]: ‘Team Semantics for Interventionist Counterfactuals: Observations vs. Interventions’, *Journal of Philosophical Logic* **50**, pp. 471–521.

Biswas, T. [1997]: *Decision-Making under Uncertainty*, London: Palgrave Macmillan.

Briggs, R. [2012]: ‘Interventionist Counterfactuals’, *Philosophical Studies*, **160**, pp. 139–66.

Cohen, M. and Jaffray, J. [1983]: ‘Rational Choice under Complete Ignorance’, *Econometrica*, **48**, pp. 1281–99.

Deœux, T. [2019]: ‘Decision-Making with Belief Functions: A Review’, *International Journal of Approximate Reasoning*, **109**, pp. 87–110.

Douven, I. [2021]: *The Art of Abduction*, Cambridge, MA: MIT Press.

Eberhardt, F. and Scheines, R. [2006]: ‘Interventions and Causal Inference’, *Philosophy of Science*, **74**, pp. 981-95.

Elqayam, S. [2012]: ‘Grounded Rationality: Descriptivism in Epistemic Context’, *Synthese*, **189**, pp. 39–49.

Fenton-Glynn, L. [2017]: ‘A Proposed Probabilistic Extension of the Halpern and Pearl Definition of “Actual Cause”’, *British Journal for the Philosophy of Science*, **68**, pp. 1061–124.

Friederich, S. [2022]: ‘Fine-Tuning’, in E. Zalta (*ed.*), *The Stanford Encyclopedia of Philosophy*, available at <plato.stanford.edu/archives/sum2022/entries/fine-tuning/>.

Fine, K. [2012]: ‘Counterfactuals without Possible Worlds’, *The Journal of Philosophy*, **109**, pp. 221–46.

Halpern, J. [2016]: *Actual Causality*, Cambridge, MA: MIT Press.

Hausman, D. and Woodward, J. [1999]: ‘Independence, Invariance, and the Causal Markov Condition’, *British Journal for the Philosophy of Science*, **50**, pp. 521–83.

Hitchcock, C. [2001]: ‘The Intransitivity of Causation Revealed in Equations and Graphs’, *Journal of Philosophy*, **98**, pp. 273–99.

Ijäs, T. and Koskinen, R. [2021]: 'Exploring Biological Possibility through Synthetic Biology', *European Journal for Philosophy of Science*, **11**, pp. 1–17.

Kinney, D. [2018]: 'Imprecise Bayesian Networks as Causal Models', *Information*, **9**, pp. 1–18.

Knight, F. [1921]: *Risk, Uncertainty, and Profit*, Boston, PA: Houghton Mifflin.

Knuuttila, T. and Loettgers, A. [2022]: '(Un)Easily Possible Synthetic Biology', *Philosophy of Science*', pp. 1–14.

Lauritzen, S. [1996]: *Graphical Models*, Oxford: Oxford University Press.

Meek, C. and Glymour, C. [1994]: 'Conditioning and intervening', *British Journal for the Philosophy of Science*, **45**, pp. 1001–21.

Menzies, P. and Price, H. [1993]: 'Causation as a Secondary Quality', *British Journal for the Philosophy of Science*, **44**, pp. 187–203.

Norton, J. [2003]: 'Causation as Folk Science', *Philosophers' Imprint* **3**, pp. 1–22.

Norton, J. [2021a]: 'Eternal Inflation: When Probabilities Fail', *Synthese*, **198**, pp. 3853–75.

Norton, J. [2021b]: *The Material Theory of Induction*, Calgary: University of Calgary Press.

Pearl, J. [2000]: *Causality*, Cambridge: Cambridge University Press.

Pearl, J. and Paz, A. [1985]: 'Graphoids: A Graph-Based Logic for Reasoning about Relevance Relations', Technical Report R-53 CSD-850038, UCLA, available at <ftp.cs.ucla.edu/tech-report/198_-reports/850038.pdf>.

Runde J. [1998]: 'Clarifying Frank Knight's Discussion of the Meaning of Risk and Uncertainty', *Cambridge Journal of Economics*, **22**, pp. 539–46.

Spohn, W. [2012]: *The Laws of Belief: Ranking Theory and Its Philosophical Applications*, Oxford: Oxford University Press.

Spirtes, P., Glymour, C. and Scheines, R. [2001]: *Causation, Prediction, and Search*, Cambridge, MA: MIT Press.

Spence, M. [1973]: 'Job Market Signaling', *Quarterly Journal of Economics*, **87**, pp. 355–74.

Verma, T. and Pearl. J. [1990]: 'Causal Networks: Semantics and Expressiveness', *Machine Intelligence and Pattern Recognition*, **9**, pp. 69–76.

Weinberger, N. [2018]: 'Faithfulness, Coordination, and Causal Coincidences', *Erkenntnis* **83**, pp. 113–33.

Woodward, J. [2003]: *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.

Wysocki, T. [unpublished-a]: ‘Decision Theory for the Probabilistically Blinded’.

Wysocki, T. [unpublished-b]: ‘An Event Algebra for Causal Counterfactuals’.

Wysocki, T. [unpublished-c]: ‘Underdeterministic Causation: A Proof of Concept’.