Robust Model Predictions via Causal Ordering

Tineke Blom
Informatics Institute
University of Amsterdam
t.blom2@uva.nl

Joris M. Mooij
Korteweg-De Vries Institute for Mathematics
University of Amsterdam
j.m.mooij@uva.nl

Abstract

Often, mathematical models of the real world are simplified representations of complex systems. A caveat to using models for analysis is that predicted causal effects and conditional independences may not be robust under model extensions, and therefore the interpretation of parameters in the model may be uncertain. In this work, we consider conditions under which qualitative model predictions are preserved when two models are combined. We show how to use the technique of *causal ordering* to efficiently assess the robustness of qualitative model predictions and characterize a large class of model extensions that preserve these predictions. For certain dynamical systems at equilibrium, we demonstrate how novel insights help to select appropriate model extensions. We apply our ideas to a viral infection model with immune responses.

1 Introduction

Key aspects of the scientific method include generating a model or hypothesis that explains a phenomenon, deriving testable predictions from this model or hypothesis, and designing an experiment to test these predictions in the real world. There are quite some interesting statistical systems for which simple Structural Causal Models [3, 13] do not model all causal and Markov properties of the system [1, 2]. In those cases the causal ordering algorithm, first introduced by Simon [17], can be used to better understand these properties [2]. In this paper we consider what happens when two systems are combined and we give conditions under which the properties of the whole system can be understood in terms of properties of its parts. We discuss how a holistic approach towards causal modelling may result in novel insights when we derive and test the predictions of systems, for which new properties emerge from the combination of its parts.

We consider the practical issue of assessing whether qualitative model predictions are robust under model extensions. We revisit the observations of De Boer [5] concerning a viral infection model and demonstrate that the qualitative causal predictions of this model can change dramatically when the model is extended with extra equations describing simple immune responses. To assess the robustness of predicted causal relations or conditional independences it would be useful to gain a better understanding of the class of model extensions that lead to changes in these predictions. We propose the technique of causal ordering [17] as an efficient method to assess the robustness of qualitative causal predictions. This allows us to characterize a large class of model extensions under which these predictions are preserved. We also consider the class of models that are obtained from the equilibrium equations of dynamical models where each variable is *self-regulating*. For this class, we show that the predicted presence of causal relations and absence of conditional independences is robust when the model is extended with new equations.

The promise of causal discovery algorithms is that they are able to learn causal relations from a combination of background knowledge and data. The general idea of many constraint-based approaches (e.g. PC or FCI and variants thereof [4, 18, 20]) is to exploit information about conditional independences

Causal Discovery & Causality-Inspired Machine Learning Workshop at Neural Information Processing Systems, 2020.

in a probability distribution to construct an equivalence class of graphs that encode certain aspects of the probability distribution, and then draw conclusions about the causal relations from the graphs. There is a large amount of literature concerning particular algorithms for which the learned structure expresses causal relations under certain conditions (e.g. linearity, causal sufficiency, absence of feedback loops), see for example [4, 6, 8, 9, 10, 15, 18, 19, 20]. In this work, our main interest is in dynamical models with the property that graphs encoding the conditional independences of their equilibrium distribution should not be interpreted causally at all. Given a model for a subsystem, we present novel insights that enable us to reject model extensions based on conditional independences in equilibrium data of the subsystem. We hope that, in future work, existing algorithms that are designed for causal discovery could be useful for reasoning about appropriate model extensions from a combination of partial models and observational data of a subsystem.

1.1 Causal ordering graph and the effects of interventions

Here, we give a concise introduction to the technique of causal ordering, introduced by Simon [16].[1] In short, the causal ordering algorithm takes a set of equations as input and returns a *causal ordering graph* that encodes the effects of interventions and a *Markov ordering graph* that implies conditional independences between variables in the model. Compared with the popular framework of Structural Causal Models [13], the distinction between the causal ordering and Markov ordering graphs does not provide new insights for acyclic models but it results in non-trivial conclusions for models with feedback, as suggested in the discussion in Section 2.4 and thoroughly explained in Blom et al. [2].

We consider models consisting of equations F that contain endogenous variables V, independent exogenous random variables W, and constant parameters P. The structure of equations and the endogenous variables that appear in them can be represented by the *associated bipartite graph* $\mathcal{B} = \langle V, F, E\rangle$, where each endogenous variable is associated with a distinct vertex in V, and each equation is associated with a distinct vertex in F. There is an edge $(v - f) \in E$ if and only if variable $v \in V$ appears in equation $f \in F$. The causal ordering algorithm constructs a *directed cluster graph* $\langle \mathcal{V}, \mathcal{E}\rangle$, where $\mathcal{V}$ is a partition of vertices V into clusters and $\mathcal{E}$ is a set of edges from vertices in V to clusters in $\mathcal{V}$. Given a bipartite graph $\mathcal{B} = \langle V, F, E\rangle$ with a perfect matching M, the causal ordering algorithm proceeds with the following three steps [2, 12]:[2]

(i) For $v \in V$, $f \in F$ orient edges $(v - f)$ as $(v \leftarrow f)$ when $(v - f) \in M$ and as $(v \rightarrow f)$ otherwise.

(ii) Find all strongly connected components $S_1, S_2, \ldots, S_n$. Let $\mathcal{V}$ be the set of clusters $S_i \cup M(S_i)$ for $i \in \{1, \ldots, n\}$, where $M(S_i)$ denotes the set of vertices that are matched to vertices in S_i in matching M.

(iii) Let $\mathrm{cl}(f)$ denote the cluster in $\mathcal{V}$ containing f. For each $(v \rightarrow f)$ such that $v \notin \mathrm{cl}(f)$ add an edge $(v \rightarrow \mathrm{cl}(f))$ to $\mathcal{E}$.

Optionally, independent exogenous random variables and parameters can be added as singleton clusters with edges towards the clusters of the equations in which they appear. It was shown that the resulting directed cluster graph $\mathrm{CO}(\mathcal{B}) = \langle \mathcal{V}, \mathcal{E}\rangle$, which we refer to as the *causal ordering graph*, is independent of the choice of perfect matching [2]. Example 1 shows how the algorithm works and a graphical illustration of the algorithm for a more elaborate cyclic model can be found in the supplement.

Example 1. Let $V = \{v_1, v_2\}$, $W = \{w_1, w_2\}$, and $P = \{p_1, p_2\}$ be index sets. Consider model equations f_1 and f_2 with endogenous variables $(X_v)_{v\in V}$, exogenous random variables $(U_w)_{w\in W}$ and constant parameters C_p with $p \in P$ below.

$$f_1: \qquad C_{p_1} X_{v_1} - U_{w_1} = 0, \tag{1}$$

$$f_2: \qquad C_{p_2} X_{v_2} + X_{v_1} + U_{w_2} = 0. \tag{2}$$

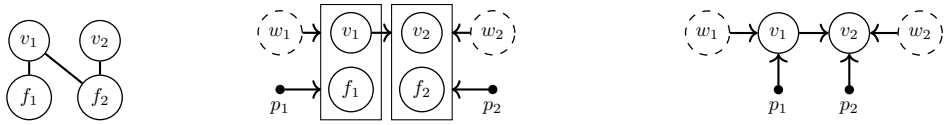
The bipartite graph $\mathcal{B} = \langle V, F, E\rangle$ in Figure 1a, with $E = \{(v_1 - f_1), (v_1 - f_2), (v_2 - f_2)\}$ is a compact representation of the model structure. This graph has a perfect matching $M = \{(v_1 - f_1), (v_2 - f_2)\}$. By orienting edges in $\mathcal{B}$ according to the rules in step (i) of the causal ordering algorithm we obtain the directed graph $\langle V \cup F, E_{\mathrm{dir}}\rangle$ with $E_{\mathrm{dir}} = \{(f_1 \rightarrow v_1), (f_2 \rightarrow v_2), (v_1 \rightarrow$

[1] Actually, we consider an equivalent algorithm for causal ordering that was shown to be more computationally efficient by [7, 12]. For more details, see [2].

[2] A perfect matching M is a subset of edges in a bipartite graph so that every vertex is adjacent to exactly one matched edge. Note that not every bipartite graph has a perfect matching.

$f_2)\}$. The clusters $C_1 = \{v_1, f_1\}$ and $C_2 = \{v_2, f_2\}$ are added to $\mathcal{V}$ in step (ii) of the algorithm, and the edge $(v_1 \to C_2)$ is added to $\mathcal{E}$ in step (iii). Finally, we may add the parameters P and independent exogenous random variables W as singleton clusters to $\mathcal{V}$, and the edges $(p_1 \to C_1)$, $(w_1 \to C_1)$, $(p_2 \to C_2)$, and $(w_2 \to C_2)$ to $\mathcal{E}$. The resulting causal ordering graph is given in Figure 1b. $\triangle$

Throughout this work, we will assume that models are uniquely solvable with respect to the causal ordering graph (see Blom et al. [2] for details). A perfect intervention on a cluster with endogenous variables represents a model change where the equations in the targeted cluster are replaced by equations that set the endogenous variables in that cluster equal to a constant. A soft intervention targets an equation, parameter, or exogenous variable, but does not affect which variables appear in the equations. We say that there is a directed path from a vertex x to a vertex y in a causal ordering graph $\langle \mathcal{V}, \mathcal{E} \rangle$ if either $\mathrm{cl}(x) = \mathrm{cl}(y)$ or there is a sequence of clusters $V_1 = \mathrm{cl}(x), V_2, \ldots, V_{k-1}, V_k = \mathrm{cl}(y)$ so that for all $i \in \{1, \ldots, k-1\}$ there is a vertex $z_i \in V_i$ such that $(z_i \to V_{i+1}) \in \mathcal{E}$. It can be shown that a) the presence of a directed path from a cluster, equation, parameter, or exogenous variable that is targeted by an intervention towards a certain variable in the causal ordering graph implies that the intervention has a generic effect on that variable and b) if no such path exists there is no causal effect of the intervention on that variable [2].



(a) Bipartite graph. (b) Causal ordering graph. (c) Markov ordering graph.

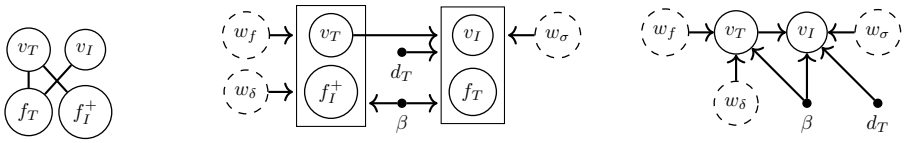
Figure 1: The bipartite graph in Figure 1a is a compact representation of the model in Example 1. The corresponding causal ordering graph and Markov ordering graph are given in Figures 1b and 1c respectively. Exogenous variables are denoted by dashed circles and parameters by black dots.

1.2 Markov ordering graph and causal discovery

The causal ordering graph $\mathrm{CO}(\mathcal{B}) = \langle \mathcal{V}, \mathcal{E} \rangle$ of model equations F with endogenous variables V, exogenous random variables W, constant parameters P, and bipartite graph $\mathcal{B}$ can be used to construct the *Markov ordering graph*, which is a DAG $\mathrm{MO}(\mathcal{B}) = \langle V \cup W, E \rangle$, with $(x \to y) \in E$ if and only if $(x \to \mathrm{cl}(y)) \in \mathcal{E}$. The Markov ordering graph for the model equations in Example 1 is given in Figure 1c. Throughout this work, we include constant parameters (e.g. p_1 and p_2) in the Markov ordering graph to show that these settable constants play a role in determining the value of endogenous variables. It has been shown that, under the assumption of unique solvability w.r.t. the causal ordering graph, d-separations in the Markov ordering graph imply conditional independences between the corresponding variables [2]. Note that constant parameters in the display of the Markov ordering graph play no role in this Markov property. Henceforth, we will assume that the probability distribution of the solution $(X_v)_{v \in V}$ to a set of model equations is faithful to the Markov ordering graph. Under the assumption that data is generated from such a model, some causal discovery algorithms, such as the PC algorithm [18], aim to construct the Markov equivalence class of the Markov ordering graph. In this work, we will specifically focus on feedback models for which the Markov ordering graph of the equilibrium distribution, and consequently the output of many causal discovery algorithms, does not have a straightforward causal interpretation.

2 Causal ordering for a viral infection model

This work was inspired by a viral infection model in De Boer [5], who showed through explicit calculations that the predictions of the model are not robust under addition of an immune response. This sheds doubt on the correct interpretation of variables and parameters in the model. For many systems it is intrinsically difficult to study their behaviour in detail. The use of simplified mathematical models that capture key characteristics aids in the analysis of a certain properties of the system. The hope is that the explanations inferred from model equations are legitimate accounts of the true underlying system [5]. In reality, a modeller must take into account that the outcome of these studies may be contingent on the specifics of the model design. Here, we demonstrate how causal



(a) Bipartite graph. (b) Causal ordering graph. (c) Markov ordering graph.

Figure 2: Graphical representations of the viral infection model in equations (5) and (6). Vertices v_i and w_j correspond to variables X_i and U_j, respectively. The causal ordering graph represents generic effects of interventions. The d-separations in Figure 2c imply conditional independences.

ordering can be used as a scalable tool to assess the robustness of model predictions without explicit calculations.

2.1 Viral infection without immune response

Let U_σ be a production term for target cells, d_T the death rate for target cells, U_f the fraction of successful infections, and U_δ the death rate of productively infected cells. Define $\beta = \frac{bp}{c}$, where b is the infection rate, p the amount of virus produced per infected cell, and c the clearance rate of viral particles. The following first-order differential equations describe how the amount of target cells $X_T(t)$ and the amount of infected cells $X_I(t)$ evolve over time [5]:

$$\dot{X}_T(t) = U_\sigma - d_T X_T(t) - \beta X_T(t) X_I(t), \tag{3}$$

$$\dot{X}_I(t) = (U_f \beta X_T(t) - U_\delta) X_I(t), \tag{4}$$

Suppose that we want to use this simple viral infection model to explain why the *set-point viral load* (i.e. the total amount of virus circulating in the bloodstream) of chronically infected HIV-patients differs by several orders of magnitude, as De Boer [5] does. To analyse this problem we look at the equilibrium equations that are implied by equations (3) and (4):[3]

$$f_T: \quad U_\sigma - d_T X_T - \beta X_T X_I = 0, \tag{5}$$

$$f_I^+: \quad U_f \beta X_T - U_\delta = 0. \tag{6}$$

Throughout the remainder of this work we will use this *natural labelling* of equilibrium equations, where the equation derived from the derivative $\dot{X}_i(t)$ is labelled f_i. For first-order differential equations that are written in canonical form, $\dot{X}_i(t) = g_i(X(t))$, the natural labelling always exists.

Suppose that U_σ, U_f and U_δ are independent exogenous random variables taking values in $\mathbb{R}_{>0}$ and d_T, β are strictly positive parameters. The associated bipartite graph, causal ordering graph, and Markov ordering graph are given in Figure 2. The causal ordering graph tells us that soft interventions targeting U_σ, U_f, U_δ, d_T, or β generically have an effect on the equilibrium distribution of the amount of infected cells X_I. From here on, we say that the causal ordering graph of a model predicts the *generic presence* or *absence* of causal effects. The Markov ordering graph shows that v_T and w_σ are d-separated. This implies that the amount of target cells X_T should be independent of the production rate U_σ when the system is at equilibrium. Henceforth, we will say that the Markov ordering graph predicts the *generic presence* or *absence* of conditional dependences.

2.2 Viral infection with a single immune response

The viral infection model in equations (3) and (4) can be extended with a simple immune response $X_E(t)$ by adding the following dynamic and static equations:

$$\dot{X}_E(t) = (U_a X_I(t) - d_E) X_E(t), \tag{7}$$

$$X_\delta(t) = d_I + U_k X_E(t), \tag{8}$$

where U_a is an activation rate, d_E and d_I are turnover rates and U_k is a mass-action killing rate [5]. Note that the exogenous random variable U_δ is now treated as an endogenous variable $X_\delta(t)$ instead.

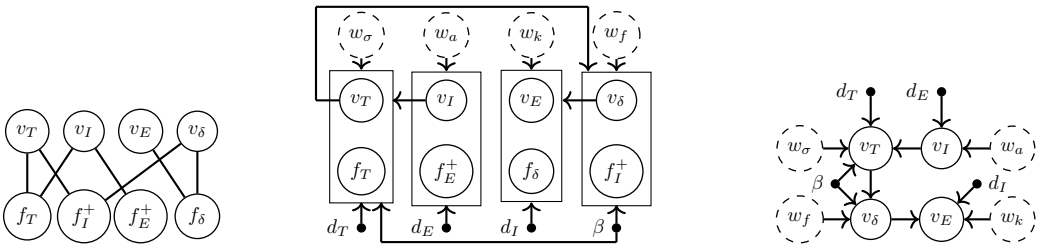
[3] Since we are only interested in strictly positive solutions we removed X_I from the equilibrium equation $f_I: \ (U_f \beta X_T - U_\delta) X_I = 0$ to obtain f_I^+.

We derive the following equilibrium equations using the natural labelling provided by equations (7) and (8):[4]

$$f_E^+ : \qquad U_a X_I - d_E = 0, \tag{9}$$

$$f_\delta : \qquad X_\delta - d_I - U_k X_E = 0, \tag{10}$$

Henceforth, we will call the addition of equations F_+ to F a *model extension*. Generally, equations F_+ may contain variables in V but they may also contain additional endogenous variables V_+. Parameters and exogenous variables in equations F can appear as endogenous variables in V_+ and in the extended model $F_{\text{ext}} = F \cup F_+$.



(a) Bipartite graph. (b) Causal ordering graph. (c) Markov ordering graph.

Figure 3: Graphical representations of the viral infection model with a single immune response. The presence or absence of causal relations and d-connections implied by the graphs in Figure 2 are not preserved if a single immune response is added.

Suppose that U_a and U_k are independent exogenous random variables taking values in $\mathbb{R}_{>0}$ and d_E, d_I are parameters taking value in $\mathbb{R}_{>0}$. The bipartite graph, causal ordering graph, and Markov ordering graph associated with equations (5), (6), (9), and (10) (with X_δ replacing U_δ) are given in Figure 3. The causal ordering graph predicts a causal effect of U_σ and d_T on X_T but not on X_I. By comparing with the predictions of the causal ordering graph in Figure 2b, we find that effects of interventions targeting U_σ and d_T are not robust under the model extension. The Markov ordering graph of the extended model shows that w_σ is d-connected to v_T, and hence U_σ and X_T are dependent. We conclude that the independence between U_σ and X_T that was implied by the Markov ordering graph of the viral infection model without immune response is not robust under the model extension.

The systematic graphical procedure followed here easily leads to the same causal conclusions as De Boer [5] obtained by explicitly solving the equilibrium equations. In addition, it leads to predictions regarding the conditional (in)dependences in the equilibrium distribution.

2.3 Viral infection with multiple immune responses

The following static and dynamical equations describe multiple immune responses:

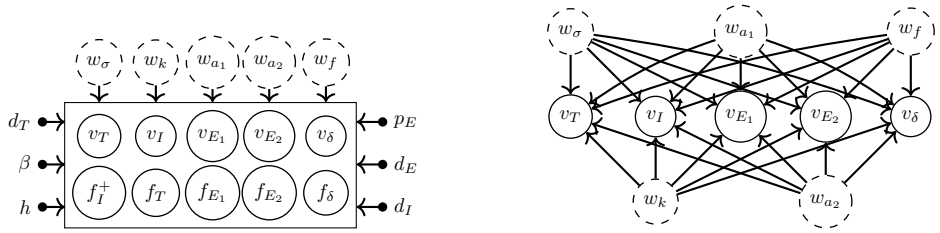
$$\dot{X}_{E_i}(t) = \frac{p_E X_{E_i}(t) U_{a_i} X_I(t)}{h + X_{E_i}(t) + U_{a_i} X_I(t)} - d_E X_{E_i}(t), \qquad i = 1, 2, \ldots, n \tag{11}$$

$$X_\delta(t) = d_I + U_k \sum_i^n U_{a_i} X_{E_i}(t), \tag{12}$$

where there are n immune responses, U_{a_i} is the avidity of immune response i, p_E is the maximum division rate, and h is a saturation constant [5]. For $n = 2$ we can derive equilibrium equations f_{E_1}, f_{E_2}, and f_δ using the natural labelling as we did for the equilibrium equations in the previous section. Together with the equilibrium equations (5) and (6) for the viral infection model this is another extended model. The bipartite graph of this extended model is given in Figure 5a, while the causal ordering graph can be found in Figure 4a. By comparing the directed paths in this causal ordering graph with that of the original viral infection model (i.e. the model without an immune response) in Figure 2b, it can be seen that the predicted presence of causal relations is preserved

[4] Analogous to changing f_I to f_I^+ for strictly positive solutions, we will look at f_E^+ instead of f_E.

under extension of the model with multiple immune responses, while the predicted absence of causal relations is not. Similarly, by comparing d-separations in the Markov ordering graphs in Figure 2c with those in Figure 4b, we find that predicted conditional dependencies are preserved under the extensions, while the predicted conditional independences are not.



(a) Causal ordering graph.

(b) Markov ordering graph.

Figure 4: Graphical representations of the viral infection model with multiple immune responses. The presence of causal relations and d-connections in Figure 2 is preserved.

2.4 Markov ordering graphs and causal interpretations

The Markov ordering graphs for the equilibrium equations of the viral infection models neither have a straightforward causal interpretation in terms of soft interventions targeting parameters, exogenous variables, or equations *nor* in terms of perfect interventions on variables in the dynamical model. To see this, consider the Markov ordering graph in Figure 3c for the viral infection with a single immune response. The edge $(v_I \to v_T)$ cannot correspond to the effect of a soft intervention targeting f_I^+, because the causal ordering graph in Figure 3b shows that there is no such effect. Clearly, directed paths in the Markov ordering graph do not necessarily represent the effects of soft interventions. The natural way to model a perfect intervention targeting a variable in the Markov ordering graph is to replace the (differential) equation of that variable with an equation setting that variable equal to a certain value in the underlying dynamical model [11]. By explicitly solving equilibrium equations it is easy to check that replacing f_δ with an equation setting X_δ equal to a constant generically changes the distribution of X_I. Since there is no directed path from v_δ to v_I in the Markov ordering graph, the effect of this perfect intervention would not have been predicted by the Markov ordering graph, if it would have been interpreted causally. Therefore, the Markov ordering graph does not have a causal interpretation in terms of soft or perfect interventions on the true underlying dynamical model.

3 Robust causal predictions under model extensions

One way to gauge the robustness of model predictions is to check to what extent they depend on the model design. The example of a viral infection with different immune responses in the previous section indicates that qualitative causal predictions entailed by the causal ordering graph of a mathematical model may strongly depend on the particulars of the model. Both the implied presence or absence of causal relations at equilibrium and the implied presence or absence of conditional independences at equilibrium may change under certain model extensions. Under what conditions are these qualitative model predictions preserved under extensions? In this section, we characterize a large class of model extensions under which qualitative equilibrium predictions are preserved.

Theorem 1 gives a sufficient condition on model extensions under which the predicted presence of causal relations and predicted presence of conditional dependences at equilibrium is preserved. The proof is given in the supplement.

Theorem 1. *Consider model equations F containing endogenous variables V with bipartite graph $\mathcal{B}$. Suppose F is extended with equations F_+ containing endogenous variables in $V \cup V_+$, where V_+ contains endogenous variables that are added by the model extension.*[5] *Let $\mathcal{B}_{\text{ext}}$ be the bipartite graph associated with $F_{\text{ext}} = F \cup F_+$ and $V_{\text{ext}} = V \cup V_+$, and $\mathcal{B}_+$ the bipartite graph associated with the extension F_+ and V_+, where variables in V appearing in F^+ are treated as exogenous variables (i.e. they are not added as vertices in $\mathcal{B}_+$). If $\mathcal{B}$ and $\mathcal{B}_+$ both have a perfect matching then:*

[5] V_+ may also contain parameters or exogenous variables that appear in F and become endogenous in the extended model.

(i) $\mathcal{B}_{\text{ext}}$ *has a perfect matching,*
(ii) ancestral relations in $\text{CO}(\mathcal{B})$ *are also present in* $\text{CO}(\mathcal{B}_{\text{ext}})$*,*
(iii) d-connections in $\text{MO}(\mathcal{B})$ *are also present in* $\text{MO}(\mathcal{B}_{\text{ext}})$*.*

This result characterizes a large set of extensions under which the implied causal effects and conditional dependences of a model are preserved. Consider again the equilibrium behaviour of the viral infection models in Section 2. We already showed explicitly that the extension of the viral infection model with multiple immune responses preserved the predicted presence of causal relations and conditional dependences, but with the help of Theorem 1 we only would have needed to check whether the bipartite graph in Figure 5c has a perfect matching to arrive at the same conclusion. The bipartite graph for the extension with a single immune response in Figure 5b does not have a perfect matching and hence the conditions of Theorem 1 do not hold. Recall that this model extension did not preserve the predicted presence of causal relations.

The theorem below gives a stronger condition under which (conditional) independence relations and the absence of causal relations that are implied by a model are also predicted by the extended model. The proof is provided in the supplement.

Theorem 2. *Let* F*,* F_+*,* F_{ext}*,* V*,* V_+*,* V_{ext}*,* $\mathcal{B}$*,* $\mathcal{B}_+$*, and* $\mathcal{B}_{\text{ext}}$ *be as in Theorem 1. If* $\mathcal{B}$ *and* $\mathcal{B}_+$ *both have perfect matchings and no vertex in* V_+ *is adjacent to a vertex in* F *in* $\mathcal{B}_{\text{ext}}$ *then:*[6]

(i) ancestral relations absent in $\text{CO}(\mathcal{B})$ *are also absent in* $\text{CO}(\mathcal{B}_{\text{ext}})$*,*
(ii) d-connections absent in $\text{MO}(\mathcal{B})$ *are also absent in* $\text{MO}(\mathcal{B}_{\text{ext}})$*.*

This result characterizes a large class of model extensions under which all qualitative model predictions are preserved. Consider again the equilibrium models for the viral infection in Section 2. The bipartite graph for the extension with a single immune response (i.e. by adding equations (9) and (10)) does not have a perfect matching. In the bipartite graph associated with the viral infection model with multiple immune responses the additional endogenous variable v_δ is adjacent to f_I. Neither of the model extensions satisfies the conditions of Theorem 2. We already demonstrated that neither of the model extensions preserves all qualitative model predictions. An example of a model extension that does satisfy the conditions in Theorem 1 and 2 is an acyclic Structural Causal Model that is extended with another acyclic Structural Causal Model such that the additional variables are non-ancestors of the original ones. Together, Theorem 1 and 2, can be used to understand when the properties of a system can be understood by studying the properties of its parts.

4 Selection of model extensions

As the example of the viral infection at equilibrium shows, different model extensions may imply different independences between variables in the original model. For the viral infection model and its extensions the Markov ordering graphs in Figures 2c, 3c, and 4b imply the following (in)dependences:

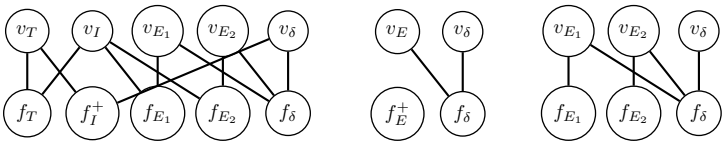
(i) Viral infection without immune response: $U_\sigma \perp\!\!\!\perp X_T, U_\sigma \not\perp\!\!\!\perp X_I$.
(ii) Viral infection with single immune response: $U_\sigma \not\perp\!\!\!\perp X_T, U_\sigma \perp\!\!\!\perp X_I$.
(iii) Viral infection with multiple immune responses: $U_\sigma \not\perp\!\!\!\perp X_T, U_\sigma \not\perp\!\!\!\perp X_I$.

Given a model for variables X_T, X_I only, we can reject model extensions based on the (conditional) independences for variables X_T, X_I, U_σ. Using this holistic modelling approach, we can already reason about an unknown model extension without observing the new mechanisms or variables. In the remainder of this section, we further discuss how this idea can be applied to the equilibrium data of dynamical systems.

4.1 Dynamical models with self-regulating variables

We say that a variable in a set of first-order differential equations in canonical form is *self-regulating* if it can be solved uniquely from the equilibrium equation that is constructed from its derivative. Interestingly, the Markov ordering graph for the equilibrium equations of a dynamical model with only self-regulating variables always has a causal interpretation. This can be seen as follows. For these models, there exists a perfect matching where each variable v_i is matched to the associated

[6]Note that V_+ is adjacent to F when one of the exogenous random variables or parameters in F becomes an endogenous variable in the model extension.



(a) Extended model. (b) Single. (c) Multiple.

Figure 5: The bipartite graphs associated with the viral infection model with multiple immune responses, the single immune response extension, and the multiple immune response extension are given in Figures 5a, 5b, and 5c, respectively.

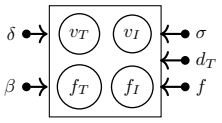


Figure 6: Causal ordering graph for positive and non-positive solutions of the viral infection model.

equilibrium equation f_i. By construction we know that a vertex v_i then always appears in a cluster with f_i in the causal ordering graph. The presence or absence of directed paths in the Markov ordering graph can then easily be associated with the presence or absence of directed paths in the causal ordering graph. Consequently, the Markov ordering graph can be interpreted in terms of both soft interventions targeting equations and perfect interventions that set variables equal to a constant by replacement of the associated dynamical and equilibrium equations. Note that dynamical systems with only self-regulating variables were also considered in Mooij et al. [11], where it was shown that their equilibria are Structural Causal Models without self-loops.

4.2 Robust conditional dependences under model extensions

It can easily be shown that when all variables in a dynamical model are self-regulating the corresponding equilibrium equations have a perfect matching provided by the natural labelling, see Lemma 1 in the supplement for more details. It then follows from Theorem 1 that the presence of ancestral relations and d-connections is robust under dynamical model extensions in which each variable is self-regulating, as is stated more formally in Corollary 1 below.

Corollary 1. *Consider a first-order dynamical model in canonical form for endogenous variables V and an extension consisting of canonical first-order differential equations for additional endogenous variables V_+. Let F and $F_{\mathrm{ext}} = F \cup F_+$ be the equilibrium equations of the original and extended model respectively. If all variables in $V \cup V_+$ are self-regulating then (ii) and (iii) of Theorem 1 hold.*

This characterizes a class of models under which qualitative predictions for the equilibrium distribution are robust, but the result can also be interpreted from a slightly different angle. Suppose that we have equilibrium data that is generated by an extended dynamical model with equilibrium equations F_{ext}, but we only have a *partial* model consisting of equations in F for a subset $V \subseteq V_{\mathrm{ext}} = V \cup V_+$ that appear in $F_{\mathrm{ext}} = F \cup F_+$. If we would find conditional independences between variables in V that do not correspond to d-separations in the Markov ordering graph of the partial model, this does not necessarily mean that the model equations are wrong. It could also be the case, for example, that we are wrong to assume that the system can be studied in a reductionist manner and that the model should be extended. Furthermore, under the assumption that data is generated from the equilibrium distribution of a dynamical model, Corollary 1 tells us that conditional independences in the data that are not predicted by the equations of a partial model imply the presence of variables that are not self-regulating, if we assume faithfulness. We consider the use of existing structure learning algorithms for the detection of feedback loops in models with variables that are not self-regulating from a combination of background knowledge and observational equilibrium data to be an interesting topic for future work.

5 Discussion

In this work we revisited several models of viral infections and immune responses. In our treatment of these models we closely followed the approach in De Boer [5] and therefore we only considered strictly positive solutions. If we would have modelled all solutions then, for example, we would have considered the equilibrium equation $f_I : \ (f\beta T - \delta)I = 0$ instead of f_I^+ in equation (6). In that case, we would have obtained the causal ordering graph in Figure 6 instead of that in Figure 2b. Clearly,

the model predictions of the causal ordering graph for the positive solutions in Figure 2b are more informative. The choice of only modelling strictly positive solutions depends on the application.

In many application domains mathematical models are used to predict the equilibrium behaviour of complex systems. An important issue is that (causal) predictions may strongly depend on the specifics of the model design. We revisited an example of a viral infection model [5], in which implied causal relations and conditional independences change dramatically when equations, describing immune reactions, are added. Analysis of this behaviour through explicit calculations is neither insightful nor scalable. We showed how the technique of causal ordering can be used to efficiently analyse the robustness of implied causal effects and conditional independences. Using key insights provided by this approach we characterized a large class of model extensions under which predicted causal relations and conditional independences are robust. We hope that the results presented in this paper are a step towards bringing the world of causal inference closer to practical applications.

Our results for the characterization of the robustness of model extensions can also be used to reason about the properties of models that are the combination of two submodels. This way, we can study systems whose causal and Markov properties can be understood in a reductionistic manner by considering the properties of its parts. When the properties of the whole model differ from those of its parts, a holistic modelling approach would be required. For models of the equilibrium distribution of dynamical systems, we proved that extensions of dynamical models where each variable is self-regulating preserve the predicted presence of causal effects an d-connections in the original model. Based on those insights, we proposed a novel approach to model selection, where information about conditional independences can be used in combination with model equations to reason about possible model extensions. For dynamical models with feedback, the output of structure learning algorithms does not always have a causal interpretation in terms of soft or perfect interventions for the equilibrium distribution [2]. We have shown that in dynamical systems where each variable is self-regulating the identifiable directed edges in the learned graph do express causal relations between variables. In future work we plan to further develop these ideas.

Broader Impact

In this work we presented novel ideas that can be used in the context of dynamical and mathematical modelling of real-world systems. Therefore there is no direct societal impact of our work.

Acknowledgments and Disclosure of Funding

We thank Johannes Textor for interesting discussions about the viral infection model. This work was supported by the ERC under the European Union's Horizon 2020 research and innovation programme (grant agreement 639466).

References

[1] T. Blom, S. Bongers, and J. M. Mooij. "Beyond Structural Causal Models: Causal Constraints Models". In: *Proceedings of the 35th Annual Conference on Uncertainty in Artificial Intelligence (UAI-19)*. 2019.

[2] T. Blom, M. M. van Diepen, and J. M. Mooij. "Conditional Independences and Causal Relations implied by Sets of Equations". In: *arXiv.org preprint* arXiv:2007.07183 [cs.AI] (2020).

[3] S. Bongers, P. Forré, J. Peters, B. Schölkopf, and J. M. Mooij. "Foundations of Structural Causal Models with Cycles and Latent Variables". In: *arXiv.org preprint* arXiv:1611.06221 [stat.ME] (2020).

[4] D. Colombo, M. Maathuis, M. Kalisch, and T. S. Richardson. "Learning high-dimensional directed acyclic graphs with latent and selection variables". In: *Annals of Statistics* 40 (2012), pp. 294–321.

[5] R. J. De Boer. "Which of Our Modeling Predictions are Robust?" In: *PLOS Computational Biology* 8.7 (2012).

[6] P. Forré and J. M. Mooij. "Constraint-based Causal Discovery for Non-Linear Structural Causal Models with Cycles and Latent Confounders". In: *Proceedings of the 34th Annual Conference on Uncertainty in Artificial Intelligence (UAI-18)*. 2018.

[7] B. Gonçalves and F. Porto. "A note on the complexity of the causal ordering problem". In: *Artificial Intelligence* 238 (2016), pp. 154–165.

[8] A. Hyttinen, F. Eberhard, and P. O. Hoyer. "Learning linear cyclic causal models with latent variables". In: *The Journal of Machine Learning Research* 13.1 (2012), pp. 3387–3439.

[9] G. Lacerda, P. Spirtes, J. Ramsey, and P. O. Hoyer. "Discovering Cyclic Causal Models by Independent Components Analysis". In: *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence* (2008), pp. 1159–1168.

[10] J. M. Mooij and T. Claassen. "Constraint-Based Causal Discovery using Partial Ancestral Graphs in the presence of Cycles". In: *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI-20)*. Ed. by J. Peters and D. Sontag. Vol. 124. PMLR, Aug. 2020, pp. 1159–1168.

[11] J. M. Mooij, D. Janzing, and B. Schölkopf. "From Ordinary Differential Equations to Structural Causal Models: the deterministic case". In: *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*. 2013, pp. 440–448.

[12] P. Nayak. *Automated Modeling of Physical Systems*. Springer-Verlag Berlin Heidelberg, 1995.

[13] J. Pearl. *Causality : models, reasoning, and inference*. Cambridge University Press, 2000.

[14] A. Pothen and C.-J. Fan. "Computing the block triangular form of a sparse matrix". In: *ACM Transactions on Mathematical Software (TOMS)* 16.4 (1990), pp. 303–324.

[15] T. Richardson and P. Spirtes. "Automated discovery of linear feedback models". In: *Computation, Causation, and Discovery* (1999), pp. 254–304.

[16] H. A. Simon. "On the Definition of the Causal Relation". In: *The Journal of Philosophy* 49.16 (1952), pp. 517–528.

[17] H. A. Simon. "Causal Ordering and Identifiability". In: *Studies in Econometric Methods*. John Wiley & Sons, 1953, pp. 49–74.

[18] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search*. MIT press, 2000.

[19] E. V. Strobl. "A constraint-based algorithm for causal discovery with cycles, latent variables and selection bias". In: *International Journal of Data Science and Analytics* 8 (2019), pp. 33–56.

[20] J. Zhang. "On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias". In: *Artificial Intelligence* 172.16–17 (2008), pp. 1873–1896.