Representing Independence Models with Elementary Triplets

Jose M. Peña
ADIT, IDA, Linköping University, Sweden
jose.m.pena@liu.se

Abstract

In an independence model, the triplets that represent conditional independences between singletons are called elementary. It is known that the elementary triplets represent the independence model unambiguously under some conditions. In this paper, we show how this representation helps performing some operations with independence models, such as finding the dominant triplets or a minimal independence map of an independence model, or computing the union or intersection of a pair of independence models, or performing causal reasoning. For the latter, we rephrase in terms of conditional independences some of Pearl's results for computing causal effects.

1 Introduction

In this paper, we explore a non-graphical approach to representing and reasoning with independence models. Specifically, in Section 2, we study under which conditions an independence model can unambiguously be represented by its elementary triplets. In Section 3, we show how this representation helps performing some operations with independence models, such as finding the dominant triplets or a minimal independence map of an independence model, or computing the union or intersection of a pair of independence models, or performing causal reasoning. We close the paper with some discussion in Section 4.

2 Representation

Let V denote a finite set of elements. Subsets of V are denoted by upper-case letters, whereas the elements of V are denoted by lower-case letters. Given three disjoint sets $I, J, K \subseteq V$, the triplet $I \perp J|K$ denotes that I is conditionally independent of J given K. Given a set of triplets G, also known as an independence model, $I \perp_G J|K$ denotes that $I \perp J|K$ is in G. A triplet $I \perp J|K$ is called elementary if $|I| = |J| = 1$. We shall not distinguish between elements of

V and singletons. We use IJ to denote $I \cup J$. Union has higher priority than set difference in expressions. Consider the following properties:

(CI0) $I \perp J|K \Leftrightarrow J \perp I|K$.

(CI1) $I \perp J|KL, I \perp K|L \Leftrightarrow I \perp JK|L$.

(CI2) $I \perp J|KL, I \perp K|JL \Rightarrow I \perp J|L, I \perp K|L$.

(CI3) $I \perp J|KL, I \perp K|JL \Leftarrow I \perp J|L, I \perp K|L$.

A set of triplets with the properties CI0-1/CI0-2/CI0-3 is also called a semigraphoid/graphoid/ compositional graphoid.[1] The CI0 property is also called symmetry property. The $\Rightarrow$ part of the CI1 property is also called contraction property, and the $\Leftarrow$ part corresponds to the so-called weak union and decomposition properties. The CI2 and CI3 properties are also called intersection and composition properties.[2] In addition, consider the following properties:

(ci0) $i \perp j|K \Leftrightarrow j \perp i|K$.

(ci1) $i \perp j|kL, i \perp k|L \Leftrightarrow i \perp k|jL, i \perp j|L$.

(ci2) $i \perp j|kL, i \perp k|jL \Rightarrow i \perp j|L, i \perp k|L$.

(ci3) $i \perp j|kL, i \perp k|jL \Leftarrow i \perp j|L, i \perp k|L$.

Note that CI2 and CI3 only differ in the direction of the implication. The same holds for ci2 and ci3.

Given a set of triplets $G = \{I \perp J|K\}$, let $\mathbb{P} = p(G) = \{i \perp j|M : I \perp_G J|K$ with $i \in I$, $j \in J$ and $K \subseteq M \subseteq (I \setminus i)(J \setminus j)K\}$. Given a set of elementary triplets $P = \{i \perp j|K\}$, let $\mathbb{G} = g(P) = \{I \perp J|K : i \perp_P j|M$ for all $i \in I$, $j \in J$ and $K \subseteq M \subseteq (I \setminus i)(J \setminus j)K\}$. The following two lemmas prove that there is a bijection between certain sets of triplets and certain sets of elementary triplets. The lemmas have been proven before when G and P satisfy CI0-1 and ci0-1 [11, Proposition 1]. We extend them to the cases where G and P satisfy CI0-2/CI0-3 and ci0-2/ci0-3.

Lemma 1. *If G satisfies CI0-1/CI0-2/CI0-3 then (a) $\mathbb{P}$ satisfies ci0-1/ci0-2/ci0-3, (b) $G = g(\mathbb{P})$, and (c) $\mathbb{P} = \{i \perp j|K : i \perp_G j|K\}$.*

[1]For instance, the conditional independences in a probability distribution form a semigraphoid, while the independences in a strictly positive probability distribution form a graphoid, and the independences in a regular Gaussian distribution form a compositional graphoid.

[2]Intersection is typically defined as $I \perp J|KL, I \perp K|JL \Rightarrow I \perp JK|L$. Note however that this and our definition are equivalent if CI1 holds. First, $I \perp JK|L$ implies $I \perp J|L$ and $I \perp K|L$ by CI1. Second, $I \perp J|L$ together with $I \perp K|JL$ imply $I \perp JK|L$ by CI1. Likewise, composition is typically defined as $I \perp JK|L \Leftarrow I \perp J|L, I \perp K|L$. Again, this and our definition are equivalent if CI1 holds. First, $I \perp JK|L$ implies $I \perp J|KL$ and $I \perp K|JL$ by CI1. Second, $I \perp K|JL$ together with $I \perp J|L$ imply $I \perp JK|L$ by CI1. In this paper, we will study sets of triplets that satisfy CI0-1, CI0-2 or CI0-3. So, the standard and our definitions are equivalent.

Proof. The proof of (c) is trivial. We now prove (a). That G satisfies CI0 implies that $\mathbb{P}$ satisfies ci0 by definition of $\mathbb{P}$.

Proof of CI1 $\Rightarrow$ ci1

Since ci1 is symmetric, it suffices to prove the $\Rightarrow$ implication of ci1.

1. Assume that $i \perp_{\mathbb{P}} j|kL$.

2. Assume that $i \perp_{\mathbb{P}} k|L$.

3. Then, it follows from (1) and the definition of $\mathbb{P}$ that $i \perp_G j|kL$ or $I \perp_G J|M$ with $i \in I$, $j \in J$ and $M \subseteq kL \subseteq (I \setminus i)(J \setminus j)M$. Note that the latter case implies that $i \perp_G j|kL$ by CI1.

4. Then, $i \perp_G k|L$ by the same reasoning as in (3).

5. Then, $i \perp_G jk|L$ by CI1 on (3) and (4), which implies $i \perp_G k|jL$ and $i \perp_G j|L$ by CI1. Then, $i \perp_{\mathbb{P}} k|jL$ and $i \perp_{\mathbb{P}} j|L$ by definition of $\mathbb{P}$.

Proof of CI1-2 $\Rightarrow$ ci1-2

Assume that $i \perp_{\mathbb{P}} j|kL$ and $i \perp_{\mathbb{P}} k|jL$. Then, $i \perp_G j|kL$ and $i \perp_G k|jL$ by the same reasoning as in (3), which imply $i \perp_G j|L$ and $i \perp_G k|L$ by CI2. Then, $i \perp_{\mathbb{P}} j|L$ and $i \perp_{\mathbb{P}} k|L$ by definition of $\mathbb{P}$.

Proof of CI1-3 $\Rightarrow$ ci1-3

Assume that $i \perp_{\mathbb{P}} j|L$ and $i \perp_{\mathbb{P}} k|L$. Then, $i \perp_G j|L$ and $i \perp_G k|L$ by the same reasoning as in (3), which imply $i \perp_G j|kL$ and $i \perp_G k|jL$ by CI3. Then, $i \perp_{\mathbb{P}} j|kL$ and $i \perp_{\mathbb{P}} k|jL$ by definition of $\mathbb{P}$.

Finally, we prove (b). Clearly, $G \subseteq g(\mathbb{P})$ by definition of $\mathbb{P}$. To see that $g(\mathbb{P}) \subseteq G$, note that $I \perp_{g(\mathbb{P})} J|K \Rightarrow I \perp_G J|K$ holds when $|I| = |J| = 1$. Assume as induction hypothesis that the result also holds when $2 < |IJ| < s$. Assume without loss of generality that $1 < |J|$. Let $J = J_1 J_2$ such that $J_1, J_2 \neq \varnothing$ and $J_1 \cap J_2 = \varnothing$. Then, $I \perp_{g(\mathbb{P})} J_1|K$ and $I \perp_{g(\mathbb{P})} J_2|J_1 K$ by definition of $g(\mathbb{P})$ and, thus, $I \perp_G J_1|K$ and $I \perp_G J_2|J_1 K$ by the induction hypothesis, which imply $I \perp_G J|K$ by CI1. □

Lemma 2. *If P satisfies ci0-1/ci0-2/ci0-3 then (a) $\mathbb{G}$ satisfies CI0-1/CI0-2/CI0-3, (b) $P = p(\mathbb{G})$, and (c) $P = \{i \perp j|K : i \perp_{\mathbb{G}} j|K\}$.*

Proof. The proofs of (b) and (c) are trivial. We prove (a) below. That P satisfies ci0 implies that $\mathbb{G}$ satisfies CI0 by definition of $\mathbb{G}$.

Proof of ci1 $\Rightarrow$ CI1

The $\Leftarrow$ implication of CI1 is trivial. We prove below the $\Rightarrow$ implication.

1. Assume that $I \perp_{\mathbb{G}} j|KL$.

2. Assume that $I \perp_{\mathbb{G}} K|L$.

3. Let $i \in I$. Note that if $i \not\perp_P j|M$ with $L \subseteq M \subseteq (I \setminus i)KL$ then (i) $i \not\perp_P j|kM$ with $k \in K \setminus M$, and (ii) $i \not\perp_P j|KM$. To see (i), assume to the contrary that $i \perp_P j|kM$. This together with $i \perp_P k|M$ (which follows from (2)

by definition of $\mathbb{G}$) imply that $i \perp_P j|M$ by ci1, which contradicts the assumption of $i \not\perp_P j|M$. To see (ii), note that $i \not\perp_P j|M$ implies $i \not\perp_P j|kM$ with $k \in K \setminus M$ by (i), which implies $i \not\perp_P j|kk'M$ with $k' \in K \setminus kM$ by (i) again, and so on until the desired result is obtained.

4. Then, $i \perp_P j|M$ for all $i \in I$ and $L \subseteq M \subseteq (I \setminus i)KL$. To see it, note that $i \perp_P j|KM$ follows from (1) by definition of $\mathbb{G}$, which implies the desired result by (ii) in (3).

5. $i \perp_P k|M$ for all $i \in I$, $k \in K$ and $L \subseteq M \subseteq (I \setminus i)(K \setminus k)L$ follows from (2) by definition of $\mathbb{G}$.

6. $i \perp_P k|jM$ for all $i \in I$, $k \in K$ and $L \subseteq M \subseteq (I \setminus i)(K \setminus k)L$ follows from ci1 on (4) and (5).

7. $I \perp_{\mathbb{G}} jK|L$ follows from (4)-(6) by definition of $\mathbb{G}$.

Therefore, we have proven above the $\Rightarrow$ implication of CI1 when $|J| = 1$. Assume as induction hypothesis that the result also holds when $1 < |J| < s$. Let $J = J_1 J_2$ such that $J_1, J_2 \neq \varnothing$ and $J_1 \cap J_2 = \varnothing$.

8. $I \perp_{\mathbb{G}} J_1|KL$ follows from $I \perp_{\mathbb{G}} J|KL$ by definition of $\mathbb{G}$.

9. $I \perp_{\mathbb{G}} J_2|J_1 KL$ follows from $I \perp_{\mathbb{G}} J|KL$ by definition of $\mathbb{G}$.

10. $I \perp_{\mathbb{G}} J_1 K|L$ by the induction hypothesis on (8) and $I \perp_{\mathbb{G}} K|L$.

11. $I \perp_{\mathbb{G}} JK|L$ by the induction hypothesis on (9) and (10).

<u>Proof of ci1-2 $\Rightarrow$ CI1-2</u>

12. Assume that $I \perp_{\mathbb{G}} j|kL$ and $I \perp_{\mathbb{G}} k|jL$.

13. $i \perp_P j|kM$ and $i \perp_P k|jM$ for all $i \in I$ and $L \subseteq M \subseteq (I \setminus i)L$ follows from (12) by definition of $\mathbb{G}$.

14. $i \perp_P j|M$ and $i \perp_P k|M$ for all $i \in I$ and $L \subseteq M \subseteq (I \setminus i)L$ by ci2 on (13).

15. $I \perp_{\mathbb{G}} j|L$ and $I \perp_{\mathbb{G}} k|L$ follows from (14) by definition of $\mathbb{G}$.

Therefore, we have proven the result when $|J| = |K| = 1$. Assume as induction hypothesis that the result also holds when $2 < |JK| < s$. Assume without loss of generality that $1 < |J|$. Let $J = J_1 J_2$ such that $J_1, J_2 \neq \varnothing$ and $J_1 \cap J_2 = \varnothing$.

16. $I \perp_{\mathbb{G}} J_1|J_2 KL$ and $I \perp_{\mathbb{G}} J_2|J_1 KL$ by CI1 on $I \perp_{\mathbb{G}} J|KL$.

17. $I \perp_{\mathbb{G}} J_1|J_2 L$ and $I \perp_{\mathbb{G}} J_2|J_1 L$ by the induction hypothesis on (16) and $I \perp_{\mathbb{G}} K|JL$.

18. $I \perp_{\mathbb{G}} J_1|L$ by the induction hypothesis on (17).

19. $I \perp_{\mathbb{G}} J|L$ by CI1 on (17) and (18).

20. $I \perp_{\mathbb{G}} K|L$ by CI1 on (19) and $I \perp_{\mathbb{G}} K|JL$.

Proof of ci1-3 ⇒ CI1-3

21. Assume that $I \perp_{\mathbb{G}} j|L$ and $I \perp_{\mathbb{G}} k|L$.

22. $i \perp_P j|M$ and $i \perp_P k|M$ for all $i \in I$ and $L \subseteq M \subseteq (I \setminus i)L$ follows from (21) by definition of $\mathbb{G}$.

23. $i \perp_P j|kM$ and $i \perp_P k|jM$ for all $i \in I$ and $L \subseteq M \subseteq (I \setminus i)L$ by ci3 on (22).

24. $I \perp_{\mathbb{G}} j|kL$ and $I \perp_{\mathbb{G}} k|jL$ follows from (23) by definition of $\mathbb{G}$.

Therefore, we have proven the result when $|J| = |K| = 1$. Assume as induction hypothesis that the result also holds when $2 < |JK| < s$. Assume without loss of generality that $1 < |J|$. Let $J = J_1 J_2$ such that $J_1, J_2 \neq \varnothing$ and $J_1 \cap J_2 = \varnothing$.

25. $I \perp_{\mathbb{G}} J_1|L$ by CI1 on $I \perp_{\mathbb{G}} J|L$.

26. $I \perp_{\mathbb{G}} J_2|J_1 L$ by CI1 on $I \perp_{\mathbb{G}} J|L$.

27. $I \perp_{\mathbb{G}} K|J_1 L$ by the induction hypothesis on (25) and $I \perp_{\mathbb{G}} K|L$.

28. $I \perp_{\mathbb{G}} K|JL$ by the induction hypothesis on (26) and (27).

29. $I \perp_{\mathbb{G}} JK|L$ by CI1 on (28) and $I \perp_{\mathbb{G}} J|L$.

30. $I \perp_{\mathbb{G}} J|KL$ and $I \perp_{\mathbb{G}} K|JL$ by CI1 on (29).

□

The following two lemmas generalize Lemmas 1 and 2 by removing the assumptions about G and P.

Lemma 3. *Let G^* denote the CI0-1/CI0-2/CI0-3 closure of G, and let $\mathbb{P}^*$ denote the ci0-1/ci0-2/ci0-3 closure of $\mathbb{P}$. Then, $\mathbb{P}^* = p(G^*)$, $G^* = g(\mathbb{P}^*)$ and $\mathbb{P}^* = \{i \perp j|K : i \perp_{G^*} j|K\}$.*

Proof. Clearly, $G \subseteq g(\mathbb{P}^*)$ and, thus, $G^* \subseteq g(\mathbb{P}^*)$ because $g(\mathbb{P}^*)$ satisfies CI0-1/CI0-2/CI0-3 by Lemma 2. Clearly, $\mathbb{P} \subseteq p(G^*)$ and, thus, $\mathbb{P}^* \subseteq p(G^*)$ because $p(G^*)$ satisfies ci0-1/ci0-2/ci0-3 by Lemma 1. Then, $G^* \subseteq g(\mathbb{P}^*) \subseteq g(p(G^*))$ and $\mathbb{P}^* \subseteq p(G^*) \subseteq p(g(\mathbb{P}^*))$. Then, $G^* = g(\mathbb{P}^*)$ and $\mathbb{P}^* = p(G^*)$, because $G^* = g(p(G^*))$ and $\mathbb{P}^* = p(g(\mathbb{P}^*))$ by Lemmas 1 and 2. Finally, that $\mathbb{P}^* = \{i \perp j|K : i \perp_{G^*} j|K\}$ is now trivial. □

Lemma 4. *Let P^* denote the ci0-1/ci0-2/ci0-3 closure of P, and let $\mathbb{G}^*$ denote the CI0-1/CI0-2/CI0-3 closure of $\mathbb{G}$. Then, $\mathbb{G}^* = g(P^*)$, $P^* = p(\mathbb{G}^*)$ and $P^* = \{i \perp j|K : i \perp_{\mathbb{G}^*} j|K\}$.*

Proof. Clearly, $P \subseteq p(\mathbb{G}^*)$ and, thus, $P^* \subseteq p(\mathbb{G}^*)$ because $p(\mathbb{G}^*)$ satisfies ci0-1/ci0-2/ci0-3 by Lemma 1. Clearly, $\mathbb{G} \subseteq g(P^*)$ and, thus, $\mathbb{G}^* \subseteq g(P^*)$ because $g(P^*)$ satisfies CI0-1/CI0-2/CI0-3 by Lemma 2. Then, $P^* \subseteq p(\mathbb{G}^*) \subseteq p(g(P^*))$ and $\mathbb{G}^* \subseteq g(P^*) \subseteq g(p(\mathbb{G}^*))$. Then, $P^* = p(\mathbb{G}^*)$ and $\mathbb{G}^* = g(P^*)$, because $P^* = p(g(P^*))$ and $\mathbb{G}^* = g(p(\mathbb{G}^*))$ by Lemmas 1 and 2. Finally, that $P^* = \{i \perp j|K : i \perp_{\mathbb{G}^*} j|K\}$ is now trivial. □

The parts (a) of Lemmas 1 and 2 imply that every set of triplets G satisfying CI0-1/CI0-2/CI0-3 can be paired to a set of elementary triplets P satisfying ci0-1/ci0-2/ci0-3, and vice versa. The pairing is actually a bijection, due to the parts (b) of the lemmas. Thanks to this bijection, we can use $\mathbb{P}$ to represent G. This is in general a much more economical representation: If $|V| = n$, then there are up to 4^n triplets,[3] whereas there are $n^2 \cdot 2^{n-2}$ elementary triplets at most. We can reduce further the size of the representation by iteratively removing from $\mathbb{P}$ an elementary triplet that follows from two others by ci0-1/ci0-2/ci0-3. Note that $\mathbb{P}$ is an unique representation of G but the result of the removal process is not in general, as ties may occur during the process.

Likewise, Lemmas 3 and 4 imply that there is a bijection between the CI0-1/CI0-2/CI0-3 closures of sets of triplets and the ci0-1/ci0-2/ci0-3 closures of sets of elementary triplets. Thanks to this bijection, we can use $\mathbb{P}^*$ to represent G^*. Note that $\mathbb{P}^*$ is obtained by ci0-1/ci0-2/ci0-3 closing $\mathbb{P}$, which is obtained from G. So, there is no need to CI0-1/CI0-2/CI0-3 close G and so produce G^*. Whether closing $\mathbb{P}$ can be done faster than closing G on average is an open question. In the worst-case scenario, both imply applying the corresponding properties a number of times exponential in $|V|$ [12]. We can avoid this problem by simply using $\mathbb{P}$ to represent G^*, because $\mathbb{P}$ is the result of running the removal process outlined above on $\mathbb{P}^*$. All the results in the sequel assume that G and P satisfy CI0-1/CI0-2/CI0-3 and ci0-1/ci0-2/ci0-3. Thanks to Lemmas 3 and 4, these assumptions can be dropped by replacing G, P, $\mathbb{G}$ and $\mathbb{P}$ in the results below with G^*, P^*, $\mathbb{G}^*$ and $\mathbb{P}^*$.

Let $I = i_1 \dots i_m$ and $J = j_1 \dots j_n$. In order to decide whether $I \perp_{\mathbb{G}} J|K$, the definition of $\mathbb{G}$ implies checking whether $m \cdot n \cdot 2^{(m+n-2)}$ elementary triplets are in P. The following lemma simplifies this for when P satisfies ci0-1, as it implies checking $m \cdot n$ elementary triplets. For when P satisfies ci0-2 or ci0-3, the lemma simplifies the decision even further as the conditioning sets of the elementary triplets checked have all the same size or form.

Lemma 5. *Let* $\mathbb{H}_1 = \{I \perp J|K : i_s \perp_P j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ *for all* $1 \le s \le m$ *and* $1 \le t \le n\}$, $\mathbb{H}_2 = \{I \perp J|K : i \perp_P j|(I \setminus i)(J \setminus j)K$ *for all* $i \in I$ *and* $j \in J\}$, *and* $\mathbb{H}_3 = \{I \perp J|K : i \perp_P j|K$ *for all* $i \in I$ *and* $j \in J\}$. *If* P *satisfies ci0-1, then* $\mathbb{G} = \mathbb{H}_1$. *If* P *satisfies ci0-2, then* $\mathbb{G} = \mathbb{H}_2$. *If* P *satisfies ci0-3, then* $\mathbb{G} = \mathbb{H}_3$.

Proof. Proof for ci0-1

It suffices to prove that $\mathbb{H}_1 \subseteq \mathbb{G}$, because it is clear that $\mathbb{G} \subseteq \mathbb{H}_1$. Assume that $I \perp_{\mathbb{H}_1} J|K$. Then, $i_s \perp_P j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ and $i_s \perp_P j_{t+1} | i_1 \dots i_{s-1} j_1 \dots j_t K$

[3] A triplet can be represented as a n-tuple whose entries state if the corresponding node is in the first, second, third or none set of the triplet.

by definition of $\mathbb{H}_1$. Then, $i_s \perp_P j_{t+1} | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ and $i_s \perp_P j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} j_{t+1} K$ by ci1. Then, $i_s \perp_{\mathbb{G}} j_{t+1} | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ and $i_s \perp_{\mathbb{G}} j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} j_{t+1} K$ by definition of $\mathbb{G}$. By repeating this reasoning, we can then conclude that $i_s \perp_{\mathbb{G}} j_{\sigma(t)} | i_1 \dots i_{s-1} j_{\sigma(1)} \dots j_{\sigma(t-1)} K$ for any permutation σ of the set $\{1 \dots n\}$. By following an analogous reasoning for i_s instead of j_t, we can then conclude that $i_{\varsigma(s)} \perp_{\mathbb{G}} j_{\sigma(t)} | i_{\varsigma(1)} \dots i_{\varsigma(s-1)} j_{\sigma(1)} \dots j_{\sigma(t-1)} K$ for any permutations σ and ς of the sets $\{1 \dots n\}$ and $\{1 \dots m\}$. This implies the desired result by definition of $\mathbb{G}$.

Proof for ci0-2

It suffices to prove that $\mathbb{H}_2 \subseteq \mathbb{G}$, because it is clear that $\mathbb{G} \subseteq \mathbb{H}_2$. Note that $\mathbb{G}$ satisfies CI0-2 by Lemma 2. Assume that $I \perp_{\mathbb{H}_2} J | K$.

1. $i_1 \perp_{\mathbb{G}} j_1 | (I \setminus i_1)(J \setminus j_1) K$ and $i_1 \perp_{\mathbb{G}} j_2 | (I \setminus i_1)(J \setminus j_2) K$ follow from $i_1 \perp_P j_1 | (I \setminus i_1)(J \setminus j_1) K$ and $i_1 \perp_P j_2 | (I \setminus i_1)(J \setminus j_2) K$ by definition of $\mathbb{G}$.

2. $i_1 \perp_{\mathbb{G}} j_1 | (I \setminus i_1)(J \setminus j_1 j_2) K$ by CI2 on (1), which together with (1) imply $i_1 \perp_{\mathbb{G}} j_1 j_2 | (I \setminus i_1)(J \setminus j_1 j_2) K$ by CI1.

3. $i_1 \perp_{\mathbb{G}} j_3 | (I \setminus i_1)(J \setminus j_3) K$ follows from $i_1 \perp_P j_3 | (I \setminus i_1)(J \setminus j_3) K$ by definition of $\mathbb{G}$.

4. $i_1 \perp_{\mathbb{G}} j_1 j_2 | (I \setminus i_1)(J \setminus j_1 j_2 j_3) K$ by CI2 on (2) and (3), which together with (3) imply $i_1 \perp_{\mathbb{G}} j_1 j_2 j_3 | (I \setminus i_1)(J \setminus j_1 j_2 j_3) K$ by CI1.

By continuing with the reasoning above, we can conclude that $i_1 \perp_{\mathbb{G}} J | (I \setminus i_1) K$. Moreover, $i_2 \perp_{\mathbb{G}} J | (I \setminus i_2) K$ by a reasoning similar to (1-4) and, thus, $i_1 i_2 \perp_{\mathbb{G}} J | (I \setminus i_1 i_2) K$ by an argument similar to (2). Moreover, $i_3 \perp_{\mathbb{G}} J | (I \setminus i_3) K$ by a reasoning similar to (1-4) and, thus, $i_1 i_2 i_3 \perp_{\mathbb{G}} J | (I \setminus i_1 i_2 i_3) K$ by an argument similar to (4). Continuing with this process gives the desired result.

Proof for ci0-3

It suffices to prove that $\mathbb{H}_3 \subseteq \mathbb{G}$, because it is clear that $\mathbb{G} \subseteq \mathbb{H}_3$. Note that $\mathbb{G}$ satisfies CI0-3 by Lemma 2. Assume that $I \perp_{\mathbb{H}_3} J | K$.

1. $i_1 \perp_{\mathbb{G}} j_1 | K$ and $i_1 \perp_{\mathbb{G}} j_2 | K$ follow from $i_1 \perp_P j_1 | K$ and $i_1 \perp_P j_2 | K$ by definition of $\mathbb{G}$.

2. $i_1 \perp_{\mathbb{G}} j_1 | j_2 K$ by CI3 on (1), which together with (1) imply $i_1 \perp_{\mathbb{G}} j_1 j_2 | K$ by CI1.

3. $i_1 \perp_{\mathbb{G}} j_3 | K$ follows from $i_1 \perp_P j_3 | K$ by definition of $\mathbb{G}$.

4. $i_1 \perp_{\mathbb{G}} j_1 j_2 | j_3 K$ by CI3 on (2) and (3), which together with (3) imply $i_1 \perp_{\mathbb{G}} j_1 j_2 j_3 | K$ by CI1.

By continuing with the reasoning above, we can conclude that $i_1 \perp_{\mathbb{G}} J | K$. Moreover, $i_2 \perp_{\mathbb{G}} J | K$ by a reasoning similar to (1-4) and, thus, $i_1 i_2 \perp_{\mathbb{G}} J | K$ by an argument similar to (2). Moreover, $i_3 \perp_{\mathbb{G}} J | K$ by a reasoning similar to (1-4) and, thus, $i_1 i_2 i_3 \perp_{\mathbb{G}} J | K$ by an argument similar to (4). Continuing with this process gives the desired result. □

We are not the first to use some distinguished triplets of G to represent it. However, most other works use dominant triplets for this purpose [1, 8, 10, 17]. The following lemma shows how to find these triplets with the help of $\mathbb{P}$. A triplet $I \perp J|K$ dominates another triplet $I' \perp J'|K'$ if $I' \subseteq I$, $J' \subseteq J$ and $K \subseteq K' \subseteq (I \smallsetminus I')(J \smallsetminus J')K$. Given a set of triplets, a triplet in the set is called dominant if no other triplet in the set dominates it.

Lemma 6. *If G satisfies CI0-1, then $I \perp J|K$ is a dominant triplet in G iff $I = i_1 \dots i_m$ and $J = j_1 \dots j_n$ are two maximal sets such that $i_s \perp_{\mathbb{P}} j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ for all $1 \leq s \leq m$ and $1 \leq t \leq n$ and, for all $k \in K$, $i_s \not\perp_{\mathbb{P}} k | i_1 \dots i_{s-1} J(K \smallsetminus k)$ and $k \not\perp_{\mathbb{P}} j_t | I j_1 \dots j_{t-1}(K \smallsetminus k)$ for some $1 \leq s \leq m$ and $1 \leq t \leq n$. If G satisfies CI0-2, then $I \perp J|K$ is a dominant triplet in G iff I and J are two maximal sets such that $i \perp_{\mathbb{P}} j | (I \smallsetminus i)(J \smallsetminus j)K$ for all $i \in I$ and $j \in J$ and, for all $k \in K$, $i \not\perp_{\mathbb{P}} k | (I \smallsetminus i)J(K \smallsetminus k)$ and $k \not\perp_{\mathbb{P}} j | I(J \smallsetminus j)(K \smallsetminus k)$ for some $i \in I$ and $j \in J$. If G satisfies CI0-3, then $I \perp J|K$ is a dominant triplet in G iff I and J are two maximal sets such that $i \perp_{\mathbb{P}} j|K$ for all $i \in I$ and $j \in J$ and, for all $k \in K$, $i \not\perp_{\mathbb{P}} k | K \smallsetminus k$ and $k \not\perp_{\mathbb{P}} j | K \smallsetminus k$ for some $i \in I$ and $j \in J$.*

Proof. We prove the lemma for when G satisfies CI0-1. The other two cases can be proven in much the same way. To see the if part, note that $I \perp_G J|K$ by Lemmas 1 and 5. Moreover, assume to the contrary that there is a triplet $I' \perp_G J'|K'$ that dominates $I \perp_G J|K$. Consider the following two cases: $K' = K$ and $K' \subset K$. In the first case, CI0-1 on $I' \perp_G J'|K'$ implies that $I i_{m+1} \perp_G J|K$ or $I \perp_G J j_{n+1}|K$ with $i_{m+1} \in I' \smallsetminus I$ and $j_{n+1} \in J' \smallsetminus J$. Assume the latter without loss of generality. Then, CI0-1 implies that $i_s \perp_{\mathbb{P}} j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ for all $1 \leq s \leq m$ and $1 \leq t \leq n+1$. This contradicts the maximality of J. In the second case, CI0-1 on $I' \perp_G J'|K'$ implies that $Ik \perp_G J|K \smallsetminus k$ or $I \perp_G Jk|K \smallsetminus k$ with $k \in K$. Assume the latter without loss of generality. Then, CI0-1 implies that $i_s \perp_{\mathbb{P}} k | i_1 \dots i_{s-1} J(K \smallsetminus k)$ for all $1 \leq s \leq m$, which contradicts the assumptions of the lemma.

To see the only if part, note that CI0-1 implies that $i_s \perp_{\mathbb{P}} j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ for all $1 \leq s \leq m$ and $1 \leq t \leq n$. Moreover, assume to the contrary that for some $k \in K$, $i_s \perp_{\mathbb{P}} k | i_1 \dots i_{s-1} J(K \smallsetminus k)$ for all $1 \leq s \leq m$ or $k \perp_{\mathbb{P}} j_t | I j_1 \dots j_{t-1}(K \smallsetminus k)$ for all $1 \leq t \leq n$. Assume the latter without loss of generality. Then, $Ik \perp_G J|K \smallsetminus k$ by Lemmas 1 and 5, which implies that $I \perp_G J|K$ is not a dominant triplet in G, which is a contradiction. Finally, note that I and J must be maximal sets satisfying the properties proven in this paragraph because, otherwise, the previous paragraph implies that there is a triplet in G that dominates $I \perp_G J|K$. □

Inspired by [12], if G satisfies CI0-1 then we represent $\mathbb{P}$ as a DAG. The nodes of the DAG are the elementary triplets in $\mathbb{P}$ and the edges of the DAG are $\{i \perp_{\mathbb{P}} k|L \rightarrow i \perp_{\mathbb{P}} j|kL\} \cup \{k \perp_{\mathbb{P}} j|L \dashrightarrow i \perp_{\mathbb{P}} j|kL\}$. See Figure 1 for an example. For the sake of readability, the DAG in the figure does not include symmetric elementary triplets. That is, the complete DAG can be obtained by adding a second copy of the DAG in the figure, replacing every node $i \perp_{\mathbb{P}} j|K$ in the copy with $j \perp_{\mathbb{P}} i|K$, and replacing every edge $\rightarrow$ (respectively $\dashrightarrow$) in the copy with $\dashrightarrow$ (respectively $\rightarrow$). We say that a subgraph over $m \cdot n$ nodes of the DAG

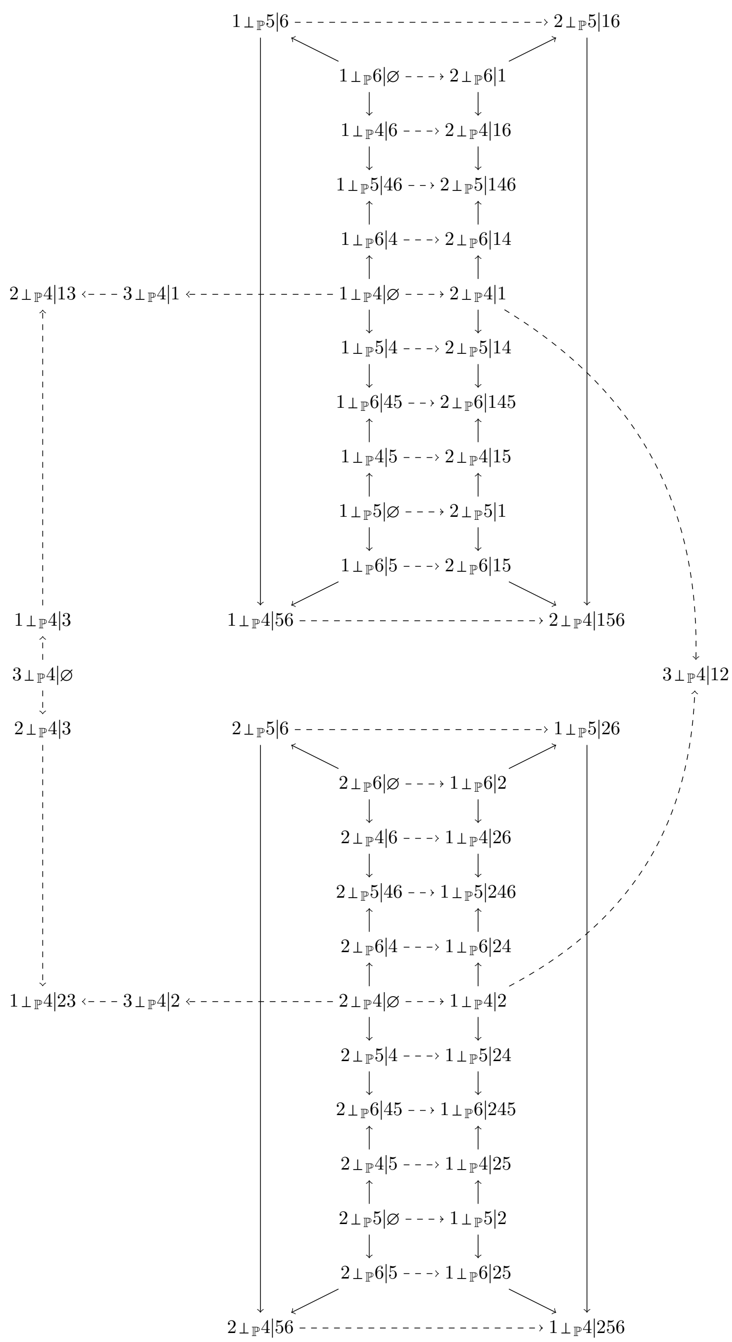

Figure 1: DAG representation of $\mathbb{P}$ (up to symmetry).

is a grid if there is a bijection between the nodes of the subgraph and the labels $\{v_{s,t} : 1 \leq s \leq m, 1 \leq t \leq n\}$ such that the edges of the subgraph are $\{v_{s,t} \to v_{s,t+1} : 1 \leq s \leq m, 1 \leq t < n\} \cup \{v_{s,t} \dashrightarrow v_{s+1,t} : 1 \leq s < m, 1 \leq t \leq n\}$. For instance, the following subgraph of the DAG in Figure 1 is a grid:

$$\begin{array}{ccc} 2 \perp_{\mathbb{P}} 5|4 & \dashrightarrow & 1 \perp_{\mathbb{P}} 5|24 \\ \downarrow & & \downarrow \\ 2 \perp_{\mathbb{P}} 6|45 & \dashrightarrow & 1 \perp_{\mathbb{P}} 6|245 \end{array}$$

The following lemma is an immediate consequence of Lemmas 1 and 5.

Lemma 7. *Let G satisfy CI0-1, and let $I = i_1 \dots i_m$ and $J = j_1 \dots j_n$. If the subgraph of the DAG representation of $\mathbb{P}$ induced by the set of nodes $\{i_s \perp_{\mathbb{P}} j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K : 1 \leq s \leq m, 1 \leq t \leq n\}$ is a grid, then $I \perp_G J|K$.*

Thanks to Lemmas 6 and 7, finding dominant triplets can now be reformulated as finding maximal grids in the DAG. Note that this is a purely graphical characterization. For instance, the DAG in Figure 1 has 18 maximal grids: The subgraphs induced by the set of nodes $\{\sigma(s) \perp_{\mathbb{P}} \varsigma(t) | \sigma(1) \dots \sigma(s-1) \varsigma(1) \dots \varsigma(t-1) : 1 \leq s \leq 2, 1 \leq t \leq 3\}$ where σ and ς are permutations of $\{1,2\}$ and $\{4,5,6\}$, and the set of nodes $\{\pi(s) \perp_{\mathbb{P}} 4 | \pi(1) \dots \pi(s-1) : 1 \leq s \leq 3\}$ where π is a permutation of $\{1,2,3\}$. These grids correspond to the dominant triplets $12 \perp_G 456|\varnothing$ and $123 \perp_G 4|\varnothing$.

3 Operations

In this section, we discuss some operations with independence models that can be performed with the help of $\mathbb{P}$. See [2, 3] for how to perform some of these operations when independence models are represented by their dominant triplets.

3.1 Membership

We want to check whether $I \perp_G J|K$, where G denotes a set of triplets satisfying CI0-1/CI0-2/CI0-3. Recall that G can be obtained from $\mathbb{P}$ by Lemma 1. Recall also that $\mathbb{P}$ satisfies ci0-1/ci0-2/ci0-3 by Lemma 1 and, thus, Lemma 5 applies to $\mathbb{P}$, which simplifies producing G from $\mathbb{P}$. Specifically if G satisfies CI0-1, then we can check whether $I \perp_G J|K$ with $I = i_1 \dots i_m$ and $J = j_1 \dots j_n$ by checking whether $i_s \perp_{\mathbb{P}} j_t | i_1 \dots i_{s-1} j_1 \dots j_{t-1} K$ for all $1 \leq s \leq m$ and $1 \leq t \leq n$. Thanks to Lemma 7, this solution can also be reformulated as checking whether the DAG representation of $\mathbb{P}$ contains a suitable grid. Likewise, if G satisfies CI0-2, then we can check whether $I \perp_G J|K$ by checking whether $i \perp_{\mathbb{P}} j | (I \setminus i)(J \setminus j) K$ for all $i \in I$ and $j \in J$. Finally, if G satisfies CI0-3, then we can check whether $I \perp_G J|K$ by checking whether $i \perp_{\mathbb{P}} j|K$ for all $i \in I$ and $j \in J$. Note that in the last two cases, we only need to check elementary triplets with conditioning sets of a specific length or form.

3.2 Minimal Independence Map

We say that a DAG D is a minimal independence map (MIM) of a set of triplets G relative to an ordering σ of the elements in V if (i) $I \perp_D J|K \Rightarrow I \perp_G J|K$,[4] (ii) removing any edge from D makes it cease to satisfy condition (i), and (iii) the edges of D are of the form $\sigma(s) \rightarrow \sigma(t)$ with $s < t$. If G satisfies CI0-1, then D can be built by setting $Pa_D(\sigma(s))$[5] for all $1 \leq s \leq |V|$ to a minimal subset of $\sigma(1) \ldots \sigma(s-1)$ such that $\sigma(s) \perp_G \sigma(1) \ldots \sigma(s-1) \setminus Pa_D(\sigma(s)) | Pa_D(\sigma(s))$ [13, Theorem 9]. Thanks to Lemma 7, building a MIM of G relative to σ can now be reformulated as finding, for all $1 \leq s \leq |V|$, a longest grid in the DAG representation of $\mathbb{P}$ that is of the form $\sigma(s) \perp_{\mathbb{P}} j_1 | \sigma(1) \ldots \sigma(s-1) \setminus j_1 \ldots j_n \rightarrow \sigma(s) \perp_{\mathbb{P}} j_2 | \sigma(1) \ldots \sigma(s-1) \setminus j_2 \ldots j_n \rightarrow \ldots \rightarrow \sigma(s) \perp_{\mathbb{P}} j_n | \sigma(1) \ldots \sigma(s-1) \setminus j_n$, or $j_1 \perp_{\mathbb{P}} \sigma(s) | \sigma(1) \ldots \sigma(s-1) \setminus j_1 \ldots j_n \dashrightarrow j_2 \perp_{\mathbb{P}} \sigma(s) | \sigma(1) \ldots \sigma(s-1) \setminus j_2 \ldots j_n \dashrightarrow \ldots \dashrightarrow j_n \perp_{\mathbb{P}} \sigma(s) | \sigma(1) \ldots \sigma(s-1) \setminus j_n$ with $j_1 \ldots j_n \subseteq \sigma(1) \ldots \sigma(s-1)$. Then, we set $Pa_D(\sigma(s))$ to $\sigma(1) \ldots \sigma(s-1) \setminus j_1 \ldots j_n$. Moreover, if G represents the conditional independences in a probability distribution $p(V)$, then D implies the following factorization: $p(V) = \prod_{s=1}^{|V|} p(\sigma(s) | Pa_D(\sigma(s)))$ [13, Corollary 4].

We say that a MIM D relative to an ordering σ is a perfect map (PM) of a set of triplets G if $I \perp_D J|K \Leftarrow I \perp_G J|K$. If G satisfies CI0-1, then we can check whether D is a PM of G by checking whether G coincides with the CI0-1 closure of $\{\sigma(s) \perp \sigma(1) \ldots \sigma(s-1) \setminus Pa_D(\sigma(s)) | Pa_D(\sigma(s)) : 1 \leq s \leq |V|\}$ [13, Corollary 7]. This together with Lemma 7 lead to the following method for checking whether G has a PM: G has a PM iff $PM(\varnothing, \varnothing)$ returns true, where

$\underline{PM(Visited, Marked)}$

1 if $Visited = V$ then
2   if $\mathbb{P}$ coincides with the ci0-1 closure of $Marked$
3   then return true and stop
4 else
5   for each node $i \in V \setminus Visited$ do
6   if the DAG representation of $\mathbb{P}$ has no grid of the form
      $i \perp_{\mathbb{P}} j_1 | Visited \setminus j_1 \ldots j_n \rightarrow i \perp_{\mathbb{P}} j_2 | Visited \setminus j_2 \ldots j_n \rightarrow \ldots$
      $\ldots \rightarrow i \perp_{\mathbb{P}} j_n | Visited \setminus j_n$ or
      $j_1 \perp_{\mathbb{P}} i | Visited \setminus j_1 \ldots j_n \dashrightarrow j_2 \perp_{\mathbb{P}} i | Visited \setminus j_2 \ldots j_n \dashrightarrow \ldots$
      $\ldots \dashrightarrow j_n \perp_{\mathbb{P}} i | Visited \setminus j_n$ with $j_1 \ldots j_n \subseteq Visited$
7   then $PM(Visited \cup \{i\}, Marked)$
8   else
9   for each longest such grid do
10         $PM(Visited \cup \{i\}, Marked \cup p(\{i \perp_G j_1 \ldots j_n | Visited \setminus j_1 \ldots j_n\})$
        $\cup p(\{j_1 \ldots j_n \perp_G i | Visited \setminus j_1 \ldots j_n\}))$

Lines 5, 7 and 10 make the algorithm consider every ordering of the nodes in V. For a particular ordering, line 6 searches for the parents of the node i

[4] $I \perp_D J|K$ stands for I and J are d-separated in D given K.
[5] $Pa_D(\sigma(s))$ denotes the parents of $\sigma(s)$ in D.

in much the same way as we described above for when searching for a MIM. Specifically, $Visited \setminus j_1 \dots j_n$ correspond to these parents. Note that line 9 makes the algorithm consider every set of parents. Lines 7 and 10 mark i as processed, mark the elementary triplets used in the search for the parents of i and, then, launch the search for the parents of the next node in the ordering. Note that the parameters are passed by value in lines 7 and 10. Finally, note the need of computing the ci0-1 closure of $Marked$ in line 2. The elementary triplets in $Marked$ represent the triplets corresponding to the grids identified in lines 6 and 9. However, it is the ci0-1 closure of the elementary triplets in $Marked$ that represents the CI0-1 closure of the triplets corresponding to the grids identified in lines 6 and 9, by Lemma 3.

3.3 Inclusion

Let G and G' denote two sets of triplets satisfying CI0-1/CI0-2/CI0-3. We can check whether $G \subseteq G'$ by checking whether $\mathbb{P} \subseteq \mathbb{P}'$. If the DAG representations of $\mathbb{P}$ and $\mathbb{P}'$ are available, then we can answer the inclusion question by checking whether the former is a subgraph of the latter.

3.4 Intersection

Let G and G' denote two sets of triplets satisfying CI0-1/CI0-2/CI0-3. Note that $G \cap G'$ satisfies CI0-1/CI0-2/CI0-3. Likewise, $\mathbb{P} \cap \mathbb{P}'$ satisfies ci0-1/ci0-2/ci0-3. We can represent $G \cap G$ by $\mathbb{P} \cap \mathbb{P}'$. To see it, note that $I \perp_{G \cap G'} J|K$ iff $i \perp_{\mathbb{P}} j|M$ and $i \perp_{\mathbb{P}'} j|M$ for all $i \in I$, $j \in J$, and $K \subseteq M \subseteq (I \setminus i)(J \setminus j)K$. If the DAG representations of $\mathbb{P}$ and $\mathbb{P}'$ are available, then we can represent $G \cap G$ by the subgraph of either of them induced by the nodes that are in both of them.

Typically, a single expert (or learning algorithm) is consulted to provide an independence model of the domain at hand. Hence the risk that the independence model may not be accurate, e.g. if the expert has some bias or overlooks some details. One way to minimize this risk consists in obtaining multiple independence models of the domain from multiple experts and, then, combining them into a single consensus independence model. In particular, we define the consensus independence model as the model that contains all and only the conditional independences on which all the given models agree, i.e. the intersection of the given models. Therefore, the paragraph above provides us with an efficient way to obtain the consensus independence model. To our knowledge, it is an open problem how to obtain the consensus independence model when the given models are represented by their dominant triplets. And the problem does not get simpler if we just consider Bayesian network models, i.e. independence models represented by Bayesian networks: There may be several non-equivalent consensus Bayesian network models, and finding one of them is NP-hard [16, Theorems 1 and 2]. So, one has to resort to heuristics.

3.5 Context-specific Independences

Let I, J, K and L denote four disjoint subsets of V. Let l denote a subset of the domain of L. We say that I is conditionally independent of J given K in the context l if $p(I|JK, L = l) = p(I|K, L = l)$ whenever $p(JK, L = l) > 0$ [4, Definition 2.2]. We represent this by the triplet $I \perp_G J|K, L = l$. Since the context always appears in the conditioning set of the triplet, the results presented so far in this paper hold for independence models containing context-specific independences. We just need to rephrase the properties CI0-3 and ci0-3 to accommodate context-specific independences. We elaborate more on this in Section 3.8. Finally, note that a (non-context-specific) conditional independence implies several context-specific ones by definition, i.e. $I \perp_G J|KL$ implies $I \perp_G J|K, L = l$ for all subsets l of the domain of L. In such a case, we do not need to represent the context-specific conditional independences explicitly.

3.6 Union

Let G and G' denote two sets of triplets satisfying CI0-1/CI0-2/CI0-3. Note that $G \cup G'$ may not satisfy CI0-1/CI0-2/CI0-3. For instance, let $G = \{x \perp y|z, y \perp x|z\}$ and $G' = \{x \perp z|\emptyset, z \perp x|\emptyset\}$. We can solve this problem by simply introducing an auxiliary random variable e with domain $\{G, G'\}$, and adding the context $e = G$ (respectively $e = G'$) to the conditioning set of every triplet in G (respectively G'). In the previous example, $G = \{x \perp y|z, e = G, y \perp x|z, e = G\}$ and $G' = \{x \perp z|e = G', z \perp x|e = G'\}$. Now, we can represent $G \cup G'$ by first adding the context $e = G$ (respectively $e = G'$) to the conditioning set of every elementary triplet in $\mathbb{P}$ (respectively $\mathbb{P}'$) and, then, taking $\mathbb{P} \cup \mathbb{P}'$. This solution has advantages and disadvantages. The main advantage is that we represent $G \cup G'$ exactly. One of the disadvantages is that the same elementary triplet may appear twice in the representation, i.e. with different contexts in the conditioning set. Another disadvantage is that we need to modify slightly the procedures described above for building MIMs, and checking membership and inclusion. We believe that the advantage outweighs the disadvantages.

If the solution above is not satisfactory, then we have two options: Representing a minimal superset or a maximal subset of $G \cup G'$ satisfying CI0-1/CI0-2/CI0-3. Note that the minimal superset of $G \cup G'$ satisfying CI0-1/CI0-2/CI0-3 is unique because, otherwise, the intersection of any two such supersets is a superset of $G \cup G'$ that satisfies CI0-1/CI0-2/CI0-3, which contradicts the minimality of the original supersets. On the other hand, the maximal subset of $G \cup G'$ satisfying CI0-1/CI0-2/CI0-3 is not unique. For instance, let $G = \{x \perp y|z, y \perp x|z\}$ and $G' = \{x \perp z|\emptyset, z \perp x|\emptyset\}$. We can represent the minimal superset of $G \cup G'$ satisfying CI0-1/CI0-2/CI0-3 by the ci0-1/ci0-2/ci0-3 closure of $\mathbb{P} \cup \mathbb{P}'$. Clearly, this representation represents a superset of $G \cup G'$. Moreover, the superset satisfies CI0-1/CI0-2/CI0-3 by Lemma 2. Minimality follows from the fact that removing any elementary triplet from the closure of $\mathbb{P} \cup \mathbb{P}'$ so that the result is still closed under ci0-1/ci0-2/ci0-3 implies removing some elementary triplet in $\mathbb{P} \cup \mathbb{P}'$, which implies not representing some triplet in $G \cup G'$ by Lemma 1. Note that the DAG representation of $G \cup G'$ is not the union of the DAG representations of $\mathbb{P}$ and $\mathbb{P}'$, because we first have to close $\mathbb{P} \cup \mathbb{P}'$ under ci0-1/ci0-2/ci0-3. We can represent a maximal subset of $G \cup G'$ satisfying CI0-1/CI0-2/CI0-3 by a maximal subset U of $\mathbb{P} \cup \mathbb{P}'$ that is closed under ci0-1/ci0-2/ci0-3 and such that every triplet represented by U is in $G \cup G'$. Recall that we can check the latter as shown above. In fact, we do not need to check it for every triplet but only for the dominant triplets. Recall that these can be obtained from U as shown in the previous section.

3.7 Causal Reasoning

Inspired by [14, Section 3.2.2], we start by adding an exogenous random variable F_a for each $a \in V$, such that F_a takes values in $\{interventional, observational\}$. These values represent whether an intervention has been performed on a or not. We use I_a and O_a to denote that $F_a = interventional$ and $F_a = observational$. We assume to have access to an independence model G over $F_V V$, in the vein of the decision theoretic approach to causality in [7]. We assume that G represents the conditional independences in a probability distribution $p(F_V V)$. We aim to compute expressions of the form $p(Y|I_X O_{V \setminus X} XW)$ with X, Y and W disjoint subsets of V. However, we typically only have access to $p(V|O_V)$. So, we aim to identify cases where G enables us to compute $p(Y|I_X O_{V \setminus X} XW)$ from $p(V|O_V)$. For instance, if $Y \perp_G F_X|O_{V \setminus X} XW$ then $p(Y|I_X O_{V \setminus X} XW) = p(Y|O_V XW)$. Note that the conditional independence in this example is context-specific. This will be the case for most of the conditional independences in this section. Moreover, we assume that $p(V|O_V)$ is strictly positive. This prevents an intervention from setting a random variable to a value with zero probability under the observational regime, which would make our quest impossible. For the sake of readability, we assume that the random variables in V are in their observational regimes unless otherwise stated. Thus, hereinafter $\tilde{p}(Y|I_X XW)$ is a shortcut for $p(Y|I_X O_{V \setminus X} XW)$, $Y \tilde{\perp}_G F_X|XW$ is a shortcut for $Y \perp_G F_X|O_{V \setminus X} XW$, and so on.

The rest of this section shows how to perform causal reasoning with independence models by rephrasing some of the main results in [14, Chapter 4] in terms of conditional independences alone, i.e. no causal graphs are involved. We start by rephrasing Pearl's *do*-calculus [14, Theorem 3.4.1].

Theorem 1. *Let X, Y, W and Z denote four disjoint subsets of V. Then*

- *Rule 1 (insertion/deletion of observations).*
  If $Y \tilde{\perp}_G X|I_Z WZ$ then $\tilde{p}(Y|I_Z XWZ) = \tilde{p}(Y|I_Z WZ)$.
- *Rule 2 (intervention/observation exchange).*
  If $Y \tilde{\perp}_G F_X|I_Z XWZ$ then $\tilde{p}(Y|I_X I_Z XWZ) = \tilde{p}(Y|I_Z XWZ)$.
- *Rule 3 (insertion/deletion of interventions).*
  If $Y \tilde{\perp}_G X|I_X I_Z WZ$ and $Y \tilde{\perp}_G F_X|I_Z WZ$, then $\tilde{p}(Y|I_X I_Z XWZ) = \tilde{p}(Y|I_Z WZ)$.

Proof. Rules 1 and 2 are immediate. To prove rule 3, note that

$$\tilde{p}(Y|I_X I_Z XWZ) = \tilde{p}(Y|I_X I_Z WZ) = \tilde{p}(Y|I_Z WZ)$$

by deploying the conditional independences given. □

Recall that checking whether the antecedents of the rules above hold can be done as shown in Section 3.1. The antecedent of rule 1 should be read as, given that Z operates under its interventional regime and $V \setminus Z$ operates under its observational regime, X is conditionally independent of Y given W. The antecedent of rule 2 should be read as, given that Z operates under its interventional regime and $V \setminus Z$ operates under its observational regime, the conditional probability distribution of Y given XWZ is the same in the observational and interventional regimes of X and, thus, it can be transferred across regimes. The antecedent of rule 3 should be read similarly.

We say that the causal effect $\tilde{p}(Y|I_X XW)$ is identifiable if it can be computed from $p(V|O_V)$. Clearly, if repeated application of rules 1-3 reduces the causal effect expression to an expression involving only observed quantities, then it is identifiable. The following theorem shows that finding the sequence of rules 1-3 to apply can be systematized in some cases. The theorem likens [14, Theorems 3.3.2, 3.3.4 and 4.3.1, and Section 4.3.3].[6]

Theorem 2. *Let X, Y and W denote three disjoint subsets of V. Then, $\tilde{p}(Y|I_X XW)$ is identifiable if one of the following cases applies:*

- *Case 1 (back-door criterion). If there exists a set $Z \subseteq V \setminus XYW$ such that*
  - *Condition 1.1. $Y \perp_G F_X | XWZ$*
  - *Condition 1.2. $Z \perp_G X | I_X W$ and $Z \perp_G F_X | W$*

  then $\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|XWZ)\tilde{p}(Z|W)$.

- *Case 2 (front-door criterion). If there exists a set $Z \subseteq V \setminus XYW$ such that*
  - *Condition 2.1. $Z \perp_G F_X | XW$*
  - *Condition 2.2. $Y \perp_G F_Z | XWZ$*
  - *Condition 2.3. $X \perp_G Z | I_Z W$ and $X \perp_G F_Z | W$*
  - *Condition 2.4. $Y \perp_G F_Z | I_X XWZ$*
  - *Condition 2.5. $Y \perp_G X | I_X I_Z WZ$ and $Y \perp_G F_X | I_Z WZ$*

  then $\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Z|XW) \sum_X \tilde{p}(Y|XWZ)\tilde{p}(X|W)$.

[6] The best way to appreciate the likeness between our and Pearl's theorems is by first adding the edge $F_a \to a$ to the causal graphs in Pearl's theorems for all $a \in V$ and, then, using d-separation to compare the conditions in our theorem and the conditional independences used in the proofs of Pearl's theorems. We omit the details because our results do not build on Pearl's, i.e. they are self-contained.

- *Case 3. If there exists a set* $Z \subseteq V \setminus XYW$ *such that*
  - *Condition 3.1.* $\tilde{p}(Z|I_X XW)$ *is identifiable*
  - *Condition 3.2.* $Y \perp_G F_X | XWZ$

  then $\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|XWZ)\tilde{p}(Z|I_X XW)$.
- *Case 4. If there exists a set* $Z \subseteq V \setminus XYW$ *such that*
  - *Condition 4.1.* $\tilde{p}(Y|I_X XWZ)$ *is identifiable*
  - *Condition 4.2.* $Z \perp_G X | I_X W$ *and* $Z \perp_G F_X | W$

  then $\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|I_X XWZ)\tilde{p}(Z|W)$.

Proof. To prove case 1, note that

$$\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|I_X XWZ)\tilde{p}(Z|I_X XW) = \sum_Z \tilde{p}(Y|XWZ)\tilde{p}(Z|I_X XW)$$
$$= \sum_Z \tilde{p}(Y|XWZ)\tilde{p}(Z|W)$$

where the second equality is due to rule 2 and condition 1.1, and the third due to rule 3 and condition 1.2.

To prove case 2, note that condition 2.1 enables us to apply case 1 replacing X, Y, W and Z with X, Z, W and $\varnothing$. Then, $\tilde{p}(Z|I_X XW) = \tilde{p}(Z|XW)$. Likewise, conditions 2.2 and 2.3 enable us to apply case 1 replacing X, Y, W and Z with Z, Y, W and X. Then, $\tilde{p}(Y|I_Z WZ) = \sum_X \tilde{p}(Y|XWZ)\tilde{p}(X|W)$. Finally, note that

$$\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|I_X XWZ)\tilde{p}(Z|I_X XW) = \sum_Z \tilde{p}(Y|I_X I_Z XWZ)\tilde{p}(Z|I_X XW)$$
$$= \sum_Z \tilde{p}(Y|I_Z WZ)\tilde{p}(Z|I_X XW)$$

where the second equality is due to rule 2 and condition 2.4, and the third due to rule 3 and condition 2.5. Plugging the intermediary results proven before into the last equation gives the desired result.

To prove case 3, note that

$$\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|I_X XWZ)\tilde{p}(Z|I_X XW) = \sum_Z \tilde{p}(Y|XWZ)\tilde{p}(Z|I_X XW)$$

where the second equality is due to rule 2 and condition 3.2.

To prove case 4, note that

$$\tilde{p}(Y|I_X XW) = \sum_Z \tilde{p}(Y|I_X XWZ)\tilde{p}(Z|I_X XW) = \sum_Z \tilde{p}(Y|I_X XWZ)\tilde{p}(Z|W)$$

where the second equality is due to rule 3 and condition 4.2. □

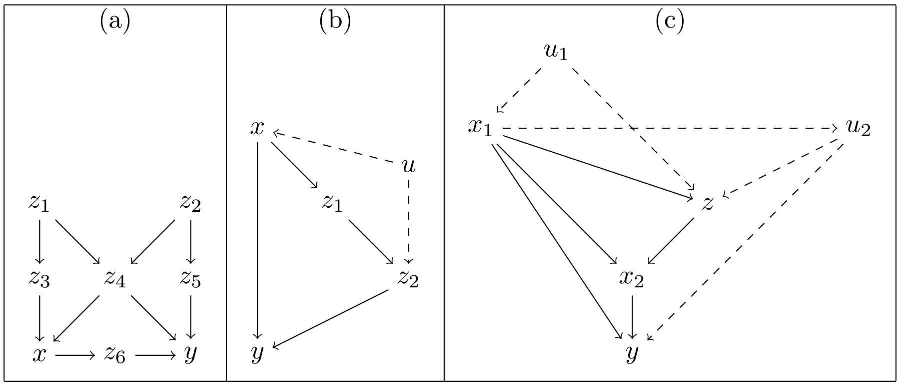

Figure 2: Causal graphs in the examples. All the nodes are observed except u, u_1 and u_2.

For instance, consider the causal graph (a) in Figure 2 [14, Figure 3.4]. Then, $\tilde{p}(y|I_x x z_3)$ can be identified by case 1 with $X = x$, $Y = y$, $W = z_3$ and $Z = z_4$ and, thus, $\tilde{p}(y|I_x x)$ can be identified by case 4 with $X = x$, $Y = y$, $W = \varnothing$ and $Z = z_3$. To see that each triplet in the conditions in cases 1 and 4 holds, we can add the edge $F_a \to a$ to the graph for all $a \in V$ and, then, apply d-separation in the causal graph after having performed the interventions in the conditioning set of the triplet, i.e. after having removed any edge with an arrowhead into any node in the conditioning set. See [14, 3.2.3] for further details. Given the causal graph (b) in Figure 2 [14, Figure 4.1 (b)], $\tilde{p}(z_2|I_x x)$ can be identified by case 2 with $X = x$, $Y = z_2$, $W = \varnothing$ and $Z = z_1$ and, thus, $\tilde{p}(y|I_x x)$ can be identified by case 3 with $X = x$, $Y = y$, $W = \varnothing$ and $Z = z_2$. Note that we do not need to know the causal graphs nor their existence to identify the causal effects. It suffices to know the conditional independences in the conditions of the cases in the theorem above. Recall again that checking these can be done as shown in Section 3.1. The theorem above can be seen as a recursive procedure for causal effect identification: Cases 1 and 2 are the base cases, and cases 3 and 4 are the recursive ones. In applying this procedure, efficiency may be an issue, though: Finding Z seems to require an exhaustive search.

The following theorem covers an additional case where causal effect identification is possible. It likens [14, Theorem 4.4.1]. See also [14, Section 11.3.7]. Specifically, it addresses the evaluation of a plan, where a plan is a sequence of interventions. For instance, we may want to evaluate the effect on the patient's health of some treatments administered at different time points. More formally, let $X_1, \ldots, X_n$ denote the random variables on which we intervene. Let Y denote the set of target random variables. Assume that we intervene on X_k only after having intervened on $X_1, \ldots, X_{k-1}$ for all $1 \leq k \leq n$, and that Y is observed only after having intervened on $X_1, \ldots, X_n$. The goal is to identify $\tilde{p}(Y|I_{X_1} \ldots I_{X_n} X_1 \ldots X_n)$. Let $N_1, \ldots, N_n$ denote some observed random variables besides $X_1, \ldots, X_n$ and Y. Assume that N_k is observed before intervening on X_k for all $1 \leq k \leq n$. Then, it seems natural to assume for all $1 \leq k \leq n$ and all $Z_k \subseteq N_k$ that Z_k does not get affected by future interventions, i.e.

$$Z_k \tilde{\perp}_G X_k \dots X_n | I_{X_k} \dots I_{X_n} X_1 \dots X_{k-1} Z_1 \dots Z_{k-1} \tag{1}$$

and

$$Z_k \tilde{\perp}_G F_{X_k} \dots F_{X_n} | X_1 \dots X_{k-1} Z_1 \dots Z_{k-1}. \tag{2}$$

Theorem 3. *If there exist disjoint sets $Z_k \subseteq N_k$ for all $1 \leq k \leq n$ such that*

$$Y \tilde{\perp}_G F_{X_k} | I_{X_{k+1}} \dots I_{X_n} X_1 \dots X_n Z_1 \dots Z_k \tag{3}$$

then $\tilde{p}(Y | I_{X_1} \dots I_{X_n} X_1 \dots X_n) =$

$$\sum_{Z_1 \dots Z_n} \tilde{p}(Y | X_1 \dots X_n Z_1 \dots Z_n) \prod_{k=1}^{n} \tilde{p}(Z_k | X_1 \dots X_{k-1} Z_1 \dots Z_{k-1}).$$

Proof. Note that

$$\begin{aligned} &\tilde{p}(Y | I_{X_1} \dots I_{X_n} X_1 \dots X_n) \\ &= \sum_{Z_1} \tilde{p}(Y | I_{X_1} \dots I_{X_n} X_1 \dots X_n Z_1) \tilde{p}(Z_1 | I_{X_1} \dots I_{X_n} X_1 \dots X_n) \\ &= \sum_{Z_1} \tilde{p}(Y | I_{X_2} \dots I_{X_n} X_1 \dots X_n Z_1) \tilde{p}(Z_1 | I_{X_1} \dots I_{X_n} X_1 \dots X_n) \\ &= \sum_{Z_1} \tilde{p}(Y | I_{X_2} \dots I_{X_n} X_1 \dots X_n Z_1) \tilde{p}(Z_1) \end{aligned}$$

where the second equality is due to rule 2 and Equation (3), and the third due to rule 3 and Equations (1) and (2). For the same reasons, we have that

$$\begin{aligned} &\tilde{p}(Y | I_{X_1} \dots I_{X_n} X_1 \dots X_n) \\ &= \sum_{Z_1 Z_2} \tilde{p}(Y | I_{X_2} \dots I_{X_n} X_1 \dots X_n Z_1 Z_2) \tilde{p}(Z_1) \tilde{p}(Z_2 | I_{X_2} \dots I_{X_n} X_1 \dots X_n Z_1) \\ &= \sum_{Z_1 Z_2} \tilde{p}(Y | I_{X_3} \dots I_{X_n} X_1 \dots X_n Z_1 Z_2) \tilde{p}(Z_1) \tilde{p}(Z_2 | X_1 Z_1). \end{aligned}$$

Continuing with this process for $Z_3, \dots, Z_n$ yields the desired result. □

For instance, consider the causal graph (c) in Figure 2 [14, Figure 4.4]. We do not need to know the graph nor its existence to identify the effect on y of the plan consisting of $I_{x_1} x_1$ followed by $I_{x_2} x_2$. It suffices to know that $N_1 = \emptyset$, $N_2 = z$, $y \tilde{\perp}_G F_{x_1} | I_{x_2} x_1 x_2$, and $y \tilde{\perp}_G F_{x_2} | x_1 x_2 z$. Recall also that $z \tilde{\perp}_G x_2 | I_{x_2} x_1$ and $z \tilde{\perp}_G F_{x_2} | x_1$ are known by Equations (1) and (2). Then, the desired effect can be identified thanks to the theorem above by setting $Z_1 = \emptyset$ and $Z_2 = z$.

In applying the theorem above, efficiency may be an issue again: Finding $Z_1, \dots, Z_n$ seems to require an exhaustive search. An effective way to carry out this search is as follows: Select Z_k only after having selected $Z_1, \dots, Z_{k-1}$, and such that Z_k is a minimal subset of N_k that satisfies Equation (3). If no such subset exists or all the subsets have been tried, then backtrack and set

Z_{k-1} to a different minimal subset of N_{k-1}. We now show that this procedure finds the desired subsets whenever they exist. Assume that there exist some sets $Z_1^*, \ldots, Z_n^*$ that satisfy Equation (3). For $k = 1$ to n, set Z_k to a minimal subset of Z_k^* that satisfies Equation (3). If no such subset exists, then set Z_k to a minimal subset of $(\bigcup_{i=1}^k Z_i^*) \smallsetminus \bigcup_{i=1}^{k-1} Z_i$ that satisfies Equation (3). Such a subset exists because setting Z_k to $(\bigcup_{i=1}^k Z_i^*) \smallsetminus \bigcup_{i=1}^{k-1} Z_i$ satisfies Equation (3), since this makes $Z_1 \ldots Z_k = Z_1^* \ldots Z_k^*$. In either case, note that $Z_k \subseteq N_k$. Then, the procedure outlined will find the desired subsets.

We can extend the previous theorem to evaluate the effect of a plan on the target random variables Y and on some observed non-control random variables $W \subseteq N_n$. For instance, we may want to evaluate the effect that the treatment has on the patient's health at intermediate time points, in addition to at the end of the treatment. This scenario is addressed by the following theorem, whose proof is similar to that of the previous theorem. The theorem likens [15, Theorem 4].

Theorem 4. *If there exist disjoint sets $Z_k \subseteq N_k \smallsetminus W$ for all $1 \leq k \leq n$ such that*

$$WY \tilde{\perp}_G F_{X_k} | I_{X_{k+1}} \ldots I_{X_n} X_1 \ldots X_n Z_1 \ldots Z_k$$

then $\tilde{p}(WY | I_{X_1} \ldots I_{X_n} X_1 \ldots X_n) =$

$$\sum_{Z_1 \ldots Z_n} \tilde{p}(WY | X_1 \ldots X_n Z_1 \ldots Z_n) \prod_{k=1}^{n} \tilde{p}(Z_k | X_1 \ldots X_{k-1} Z_1 \ldots Z_{k-1}).$$

Finally, note that in the previous theorem X_k may be a function of $X_1 \ldots X_{k-1}$ $W_1 \ldots W_{k-1} Z_1 \ldots Z_{k-1}$, where $W_k = (W \smallsetminus \bigcup_{i=1}^{k-1} W_i) \cap N_k$ for all $1 \leq k \leq n$. For instance, the treatment prescribed at any point in time may depend on the treatments prescribed previously and on the patient's response to them. In such a case, the plan is called conditional, otherwise is called unconditional. We can evaluate alternative conditional plans by applying the theorem above for each of them. See also [14, Section 11.4.1].

3.8 Context-specific Independences Revisited

As mentioned in Section 3.5, we can extend the results in this paper to independence models containing context-specific independences of the form $I \perp J | K, L = l$ by just rephrasing the properties CI0-3 and ci0-3 to accommodate them. In the causal setup described above, for instance, we may want to represent triplets with interventions in their third element as long as they do not affect the first two elements of the triplets, i.e. $I \tilde{\perp} J | K M I_M I_N$ with I, J, K, M and N disjoint subsets of V, which should be read as follows: Given that MN operates under its interventional regime and $V \smallsetminus MN$ operates under its observational regime, I is conditionally independent of J given KM. Note that an intervention is made on N but the resulting value is not considered in the triplet, e.g. we know that a treatment has been prescribed but we ignore which. The properties CI0-3 can be extended to these triplets by simply adding $M I_M I_N$ to the third member of the triplets. That is, let $C = M I_M I_N$. Then:

(CI0) $I \tilde{\perp} J | KC \Leftrightarrow J \tilde{\perp} I | KC$.

(CI1) $I \tilde{\perp} J | KLC, I \tilde{\perp} K | LC \Leftrightarrow I \tilde{\perp} JK | LC$.

(CI2) $I \tilde{\perp} J | KLC, I \tilde{\perp} K | JLC \Rightarrow I \tilde{\perp} J | LC, I \tilde{\perp} K | LC$.

(CI3) $I \tilde{\perp} J | KLC, I \tilde{\perp} K | JLC \Leftarrow I \tilde{\perp} J | LC, I \tilde{\perp} K | LC$.

Similarly for ci0-3.

Another case that we may want to consider is when a triplet includes interventions in its third element that affect its second element, i.e. $I \tilde{\perp} J | KMI_J I_M I_N$ with I, J, K, M and N disjoint subsets of V, which should be read as follows: Given that JMN operates under its interventional regime and $V \setminus JMN$ operates under its observational regime, the causal effect on I is independent of J given KM. These triplets liken the probabilistic causal irrelevances in [9, Definition 7]. The properties CI1-3 can be extended to these triplets by simply adding $MI_J I_K I_M I_N$ to the third member of the triplets. Note that CI0 does not make sense now, i.e. I is observed whereas J is intervened on. Let $C = MI_J I_K I_M I_N$. Then:

(CI1) $I \tilde{\perp} J | KLC, I \tilde{\perp} K | LC \Leftrightarrow I \tilde{\perp} JK | LC$.

(CI2) $I \tilde{\perp} J | KLC, I \tilde{\perp} K | JLC \Rightarrow I \tilde{\perp} J | LC, I \tilde{\perp} K | LC$.

(CI3) $I \tilde{\perp} J | KLC, I \tilde{\perp} K | JLC \Leftarrow I \tilde{\perp} J | LC, I \tilde{\perp} K | LC$.

(CI1') $I \tilde{\perp} J | I'LC, I' \tilde{\perp} J | LC \Leftrightarrow II' \tilde{\perp} J | LC$.

(CI2') $I \tilde{\perp} J | I'LC, I' \tilde{\perp} J | ILC \Rightarrow I \tilde{\perp} J | LC, I' \tilde{\perp} J | LC$.

(CI3') $I \tilde{\perp} J | I'LC, I' \tilde{\perp} J | ILC \Leftarrow I \tilde{\perp} J | LC, I' \tilde{\perp} J | LC$.

Similarly for ci1-3.

4 Discussion

In this work, we have proposed to represent semigraphoids, graphoids and compositional graphoids by their elementary triplets. We have also shown how this representation helps performing some operations with independence models, including causal reasoning. For this purpose, we have rephrased in terms of conditional independences some of Pearl's results for causal effect identification. We find interesting to explore non-graphical approaches to causal reasoning in the vein of [7], because of the risks of relying on causal graphs for causal reasoning, e.g. a causal graph of the domain at hand may not exist. See [5, 6] for a detailed account of these risks. Pearl also acknowledges the need to develop non-graphical approaches to causal reasoning [9, p. 10]. As future work, we consider seeking for necessary conditions for non-graphical causal effect identification (recall that the ones described in this paper are just sufficient). We also consider implementing and experimentally evaluating the efficiency of some of the operations discussed in this work.

References

[1] Marco Baioletti, Giuseppe Busanello, and Barbara Vantaggi. Conditional independence structure and its closure: Inferential rules and algorithms. *International Journal of Approximate Reasoning*, 50:1097–1114, 2009.

[2] Marco Baioletti, Giuseppe Busanello, and Barbara Vantaggi. Acyclic directed graphs representing independence models. *International Journal of Approximate Reasoning*, 52:2 – 18, 2011.

[3] Marco Baioletti, Davide Petturiti, and Barbara Vantaggi. Qualitative combination of independence models. In *Proceedings of the 12th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 37–48, 2013.

[4] Craig Boutilier, Nir Friedman, Moises Goldszmidt, and Daphne Koller. Context-specific independence in Bayesian networks. In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence*, pages 115–123, 1996.

[5] A. Philip Dawid. Beware of the DAG! *Journal of Machine Learning Research Workshop and Conference Proceedings*, 6:59–86, 2010.

[6] A. Philip Dawid. Seeing and doing: The Pearlian synthesis. In *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 309–325, 2010.

[7] A. Philip Dawid. Statistical causality from a decision-theoretic perspective. *Annual Review of Statistics and Its Applications*, 2:273–303, 2015.

[8] Peter de Waal and Linda C. van der Gaag. Stable independence and complexity of representation. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pages 112–119, 2004.

[9] David Galles and Judea Pearl. Axioms of causal relevance. *Artificial Intelligence*, 97:9–43, 1997.

[10] Stavros Lopatatzidis and Linda C. van der Gaag. Computing concise representations of semi-graphoid independency models. In *Proceedings of the 13th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 290–300, 2015.

[11] František Matúš. Ascending and descending conditional independence relations. In *Proceedings of the 11th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 189–200, 1992.

[12] František Matúš. Lengths of semigraphoid inferences. *Annals of Mathematics and Artificial Intelligence*, 35:287–294, 2002.

[13] Judea Pearl. *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann Publishers Inc., 1988.

[14] Judea Pearl. *Causality: Models, reasoning, and inference.* Cambridge University Press, 2009.

[15] Judea Pearl and James M. Robins. Probabilistic evaluation of sequential plans from causal models with hidden variables. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, pages 444–453, 1995.

[16] Jose M. Peña. Finding consensus Bayesian network structures. *Journal of Artificial Intelligence Research*, 42:661–687, 2011.

[17] Milan Studený. Complexity of structural models. In *Proceedings of the Joint Session of the 6th Prague Conference on Asymptotic Statistics and the 13th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pages 521–528, 1998.