

## Team semantics for interventionist counterfactuals and causal dependence

Barbero, Fausto

2017-12-22

---

Barbero , F & Sandu , G 2017 , ' Team semantics for interventionist counterfactuals and causal dependence ' arXiv.org .

---

<http://hdl.handle.net/10138/233361>

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# Team semantics for interventionist counterfactuals and causal dependence

Fausto Barbero\*    Gabriel Sandu\*  
University of Helsinki

December 22, 2017

## Abstract

We introduce a generalization of team semantics which provides a framework for manipulationist theories of causation based on structural equation models, such as Woodward’s and Pearl’s; our causal teams incorporate (partial or total) information about functional dependencies that are invariant under interventions. We give a unified treatment of observational and causal aspects of causal models by isolating two operators on causal teams which correspond, respectively, to conditioning and to interventionist counterfactual implication.

The evaluation of counterfactuals may involve the production of partially determined teams. We suggest a way of dealing with such cases by 1) the introduction of formal entries in causal teams, and 2) the introduction of weaker truth values (falsifiability and admissibility), for which we suggest some plausible semantical clauses.

We introduce formal languages for both deterministic and probabilistic causal discourse, and study in some detail their inferential aspects. Finally, we apply our framework to the analysis of direct and total causation, and other notions of dependence and invariance.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Theories of causation: background</b>	<b>3</b>
2.1	Conditional probabilities . . . . .	3
2.2	Counterfactual dependence . . . . .	5
2.3	Various notions of cause . . . . .	7

---

\*The present work has been developed within the Academy of Finland Project n.286991, “Dependence and Independence in Logic: Foundations and Philosophical Significance”.

<b>3</b>	<b>Causal team semantics</b>	<b>8</b>
3.1	Teams . . . . .	8
3.2	Causal teams . . . . .	10
3.3	Explicit causal teams . . . . .	12
3.4	Causal subteams . . . . .	13
3.5	A basic language and its semantics . . . . .	14
3.6	Selective implication . . . . .	14
3.7	Interventions on teams: some examples . . . . .	16
3.8	Formal definition of interventions and counterfactuals . . . . .	19
3.9	Logical languages . . . . .	23
<b>4</b>	<b>Some logical principles</b>	<b>25</b>
4.1	The law of excluded middle . . . . .	25
4.2	Conditional excluded middle . . . . .	26
4.3	Interactions between implications and dependencies . . . . .	27
4.4	Permutation and exportation/importation of antecedents . . . . .	31
4.5	Generalized import-export and permutation rules . . . . .	33
4.6	Composition and effectiveness rules . . . . .	34
<b>5</b>	<b>The nonparametric case: falsifiability and admissibility</b>	<b>39</b>
<b>6</b>	<b>Probabilistic semantics</b>	<b>41</b>
6.1	A probabilistic language . . . . .	41
6.2	Comparison with causal models . . . . .	45
6.3	Pearl’s example in causal teams . . . . .	46
6.4	Definability in the probabilistic language . . . . .	50
6.5	The Markov Condition . . . . .	52
<b>7</b>	<b>Causal modeling</b>	<b>55</b>
7.1	Direct and total cause . . . . .	55
7.2	Higher-order effects . . . . .	58
7.3	Invariance . . . . .	59
<b>8</b>	<b>Conclusions</b>	<b>61</b>

## 1 Introduction

Notions of dependence and independence entered the realm of logical investigation in the early days of mathematical logic, essentially with the introduction, by Frege, of nested quantification; this aspect of quantification was made explicit by the notion of Skolem function ([26]). It is only in the last decades, however, that a systematical analysis of (in)dependence notions within predicative, propositional, modal logical languages has been undertaken. One of the main unifying tools in this enterprise is the so-called *team semantics* ([14],[15],[32]), whose key idea is that formulas involving dependencies acquire meaning only when evaluated over *sets* of assignments. Variations of this methodology have allowed

a systematical study of logical systems enriched with dependencies that arise from database theory, probabilistic theory and quantum information theory. In many cases, distinct notions of (in)dependence can coexist in one and the same formal language, and this kind of interplay has been systematically investigated from the point of view of definability and complexity. However, to the best of our knowledge, the framework of team semantics has not yet been used to investigate notions of causal and counterfactual dependence. In the present paper, we provide a generalization of team semantics that is also adequate to capture causal, counterfactual and probabilistic notions of (in)dependence which arise from modern manipulationist theories of causation such as Pearl’s ([23]) and Woodward’s ([39]). The generalization is not trivial. It is usually acknowledged in the literature that causal relationships cannot be reduced to mere correlations of data (the latter can be represented by a set of assignments.) Instead, a richer structure, encoding counterfactual assumptions, is needed.

The plan of the paper is as follows. In section 2 we give a short motivation for the manipulationist (interventionist) approach to causation. In section 3 we present team semantics and show how to adequately enrich it so that it can handle interventionist counterfactuals. We will introduce several languages to express various (deterministic) notions of dependence. In section 4 we analyze the logical properties of these languages, up to some basic soundness and completeness proofs; we use some of these properties to compare our counterfactuals with those of Stalnaker ([28]), Lewis ([18]) and Galles&Pearl ([8]). Section 5 is dedicated to the logical issues that arise from nonparametric models. In section 6 we introduce probabilistic causal languages. In section 7 we will discuss various notions of causation (mainly taken from Woodward) and invariance, in the light of the logic developed in earlier sections.

## 2 Theories of causation: background

The reductive approach to causation has been philosophers’ favourite tool. It aims, roughly, at finding necessary and sufficient conditions for causal relationships like “ $X$  causes  $Y$ ”. Two such conditions have been prominent in the literature: those formulated in terms of conditional probabilities, and those based on counterfactuals (counterfactual dependence). We discuss them shortly in the next two sections. The main purpose of presenting them is to understand some of the reasons why the reductive approach has been found unsatisfactory and replaced by non-reductive approaches, such as the *manipulationist* or *interventionist* accounts of counterfactuals and causation (Pearl, Woodward, Halpern, Hitchcock, Briggs, among others).

### 2.1 Conditional probabilities

Two well known endeavours to connect causal relationships with conditional probabilities are due to P. Suppes ([30]) and N. Cartwright ([4]). For instance, Cartwright ([4], p. 26) requires that causes raise the probabilities of their effect:

(CC)  $C$  causes  $E$  iff  $Pr(E/C, K_j) > Pr(E/K_j)$  for all state descriptions  $K_j$  which satisfy certain conditions.

Woodward ([38]) finds (CC) defective in two ways. Firstly, the requirement to conditionalize on all  $K_j$  is too strong, and a weaker, existential condition would suffice. Secondly, (CC) holds, as initially intended, only for positive causes. When negative (inhibiting) changes are taken into account, it is natural to replace  $Pr(E/C, K_j) > Pr(E/K_j)$  by  $Pr(E/C, K_j) \neq Pr(E/K_j)$  which only requires  $C$  and  $E$  to be (probabilistically) dependent. With these two points in mind, Woodward ([38]) proposes to replace (CC) with something of the following sort:

(\*)  $X$  causes  $Z$  if and only if  $X$  and  $Z$  are dependent conditional on certain other factors  $F$ .

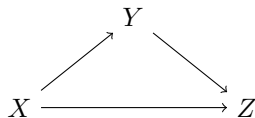
where  $X$  and  $Z$  are variables standing for properties.

The problem now becomes that of specifying the other factors  $F$ . Cartwright suggests that they include other causes of  $Z$  with the exception of those which are on a causal chain from  $X$  to  $Z$ . She recognizes, however ([4], p. 30; [5], p. 95 ff; [38], p. 58) that the claim that we should never condition on all such intermediate variables is too strong. Woodward ([38]) makes it clear that in order to understand these restrictions, and more broadly, in order for the project of specifying the other factors  $F$  to have any hope of success, we need to refer to other causes of  $Y$  and the way they are connected. For instance, to see why it is inappropriate to conditionalize on the variables which lie on a causal chain from  $X$  to  $Z$ , as Cartwright first suggestion goes, it is enough to consider the causal structure

$$X \rightarrow Y \rightarrow Z$$

Now if we were to conditionalize on  $Y$ , we would expect, intuitively,  $X$  and  $Z$  to be independent, which according to the definition above would result in  $X$  not being a cause of  $Z$ . But this is not what we want.

On the other side, to see that the requirement of never conditionalizing on the variables on causal chains from  $X$  to  $Z$  is too strong, it is enough to consider the causal structure in which both  $X$  and  $Y$  are “direct” causes of  $Z$  and are



on the causal paths between  $X$  and  $Z$ . If the causal connection between  $X$  and  $Z$  is to be reflected in the probabilistic dependence of  $Z$  on  $X$  conditional on some other properties  $F$ , then these other properties must include  $Y$ . In other words, to determine the causal influence of  $X$  on  $Z$  we must take into account the influence of  $Y$  on  $Z$ .

What all this shows, according to Woodward, is that the project to connect causal relationships between  $X$  and  $Y$  with conditional probabilities, which finds

its expression in (\*), goes via a mechanism which provides information about other causes (contributing or or total) of  $y$  besides  $X$  and how these causes are connected with one another and with  $Z$ . ([38], p. 58.)

One such mechanism is that of *causal Bayesian networks* (Pearl 2000/2009 [23], Spirtes, Glymour and Scheines 1993/2001 [27]). They are built on Directed Acyclic Graphs (DAGs) of the kind we have already encountered in our earlier examples. We start with a set of variables  $V = \{X_1, \dots, X_n\}$  whose causal relationships we want to investigate and a set of (directed) edges. A directed edge from the variable  $X$  (parent) to the variable  $Z$  (child) is intended to represent the fact that  $X$  is a *direct cause* of  $Z$ . The set of all parents of  $Z$  is denoted as  $PA_Z$ . The project is now to investigate the connection between causal relationships generated by the edges of the graph and conditional probabilities  $P(Z/PA_Z)$  determined by a joint probability distribution  $P$  over  $V$ . More exactly, let  $P$  be a joint probability distribution on the set  $V$ . We denote by  $x_1, x_2, \dots, x_n$  the values associated with the variables in  $V$ , that is,  $X_1 = x_1, \dots, X_n = x_n$ . A DAG  $G$  is said to *represent*  $P$  if the equation

$$P(x_1, \dots, x_n) = \prod_i P(x_i/pa_i) \tag{1}$$

holds, where the variables  $PA_i$  obey the parent-child relation of  $G$ :  $Y \in PA_i$  if and only if there is an arrow in  $G$  from  $Y$  to  $X_i$ .

A well known result states the following:

**Theorem 2.1** (Markov Condition, [35]). *Let  $G$  be a DAG with  $V$  its set of variables.  $G$  represents a probability distribution  $P$  if and only if every variable in  $V$  is independent of all its nondescendants (in  $G$ ) conditional on its parents.*

A DAG  $G$  which represents a probability distribution  $P$  is often referred as *causal Bayesian network*. The Markov Condition is thought to be important because it establishes a connection between causal relationships as represented by the arrows of a DAG and dependence relationships. For instance, if we take “ $Z$  is a nondescendant of  $X$ ” to stand for “ $X$  does not cause  $Z$ ”, then the Markov condition implies that if “ $X$  does not cause  $Z$ ” then conditional on its parents,  $X$  is independent of  $Z$ , which by contraposition gives us the right-to-left direction of (\*):

- If variables  $X$  and  $Z$  are dependent (that is:  $Pr(X/Parents(X), Z) \neq Pr(X/Parents(X))$ ), then  $X$  causes  $Z$ .

(Cf. e.g. [38])

As handy as causal Bayesian networks are to connect causes with conditional probabilities, they fail to represent caunterfactual reasoning ([23], p. 37). In other words, if we want a robust notion of cause which sustains counterfactuals, we need to supplement causal Bayesian networks with an additional, deterministic component.

## 2.2 Counterfactual dependence

Accounts of causal relationships based on counterfactual dependence have been available starting with the works of David Lewis and G. H. von Wright. Lewis ([17], [20]) reduces causal relationships to counterfactual dependence, which, in the end, is defined in terms of similarity between possible worlds. Von Wright ([36]) distinguishes a causal connection between  $p$  and  $q$  from an accidental generalization (the concomitance of  $p$  and  $q$ ) on the basis of the fact that the former, unlike the latter, sustains a counterfactual assumption of the form "on occasions where  $p$ , in fact, was not the case,  $q$  would have accompanied it, had  $p$  been the case"; that is,  $p$  is a state of affairs which we can *produce or suppress at will*. It is thus the *manipulativity* of the antecedent which is the individuating aspect of the cause factor ([36], p. 70). Lewis's account has been criticized for relying too much on "dubious metaphysics" and von Wright's account for being too "anthropomorphic".

Both Lewis's and von Wright's accounts contain ingredients which have been incorporated later on into interventionist accounts of counterfactuals and causation. Roughly, one needs a mechanism to represent the exogenous process which is the manipulation of the (variables of the) antecedent of a counterfactual. This mechanism has become known as *intervention*<sup>1</sup>. In addition, we need another mechanism to measure the *effects* the changes of the intervened variables have on the (variables of the) consequent. This mechanism is encoded into the so-called *structural (functional) equations* ([23], [38]).

In more details, we divide the set of variables whose causal relationships we want to investigate into two disjoint sets: a set  $V$  of endogenous variables and a set  $U$  of exogenous variables. With each endogenous variable  $X \in V$ , an equation of the form

$$X = f_X(PA_X)$$

is associated, where the variables  $PA_X \subseteq U \cup V \setminus \{X\}$  are called the *parents of  $X$* . The standard interpretation of a functional equation  $X = f_X(PA_X)$  is that of a law which specifies the value of  $X$  given every possible combination of the values of  $PA_X$ . If we draw an arrow from each variable in  $PA_X$  to the variable  $X$  we obtain directed graphs as in the case of the causal Bayesian frameworks mentioned in the previous section. The crucial difference between the two frameworks is that in the present case, instead of characterizing the child-parent relationships stochastically in terms of conditional probabilities  $P(X/PA_X)$ , we characterize them deterministically using the equations.

Various notions of intervention have been proposed, both in the context of causal Bayesian networks and in that of structural equations (e.g., [27], [37], [12], and [23]). A detailed discussion of this variety is outside the scope of this paper. Suffice it to say that an intervention  $do(X = x)$  on a variable  $X$  is an action which disconnects the variable  $X$  from all the incoming arrows into  $X$  while preserving all the other arrows of the graph including those directed

---

<sup>1</sup>Lewis does not speak of interventions, but their role is mirrored, in the work of Lewis, by the notions of "local miracle" and "non-backtracking counterfactual".

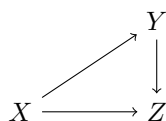
out of  $X$ . This is known as the *arrow-breaking* conception of interventions. In the structural equations framework where the causal graph is induced by the appropriate set of equations, the intervention  $do(X = x)$  results also in the alteration of that set: the equation  $X = f_X(PA_X)$  associated with the variable  $X$  is replaced with a new equation  $X = x$ , while keeping intact the other equations in the set.

It may be useful to illustrate these notions by way of an example ([38]).

Consider the following two equations:

$$(6) \quad Y = aX \qquad (7) \quad Z = bX + cY$$

This set of equations induces the DAG



If we intervene on  $Y$  and set its value to 1 (i.e.,  $do(Y = 1)$ ), the result will be the altered system of equations:

$$(6') \quad Y = 1 \qquad (7) \quad Z = bX + cY$$

corresponding to the new DAG:



### 2.3 Various notions of cause

Woodward ([37],[38]) uses interventions in the context of structural equations to *define* “ $X$  is a cause of  $Y$ ”. It turns out, however, that there are several distinct notions of cause, each satisfying the central commitment of the manipulability theory of causation: there is a causal relationship between  $X$  and  $Y$  whenever there is a possible intervention that changes the value of  $X$  such that carrying it out changes the value of  $Y$  (or its probability). [38], p. 54). Here are several notions of cause:

- (DC) (Direct cause) A necessary and sufficient condition for  $X$  to be a direct cause of  $Y$  with respect to some variable set  $Z$  is that there be a possible intervention on  $X$  that will change  $Y$  (or the probability distribution of  $Y$ ) when all the other variables in  $Z$  besides  $X$  and  $Y$  are held fixed at some values by interventions. ([38], p. 52)



For Woodward, direct causes correspond to the parent-child relationships in the underlying DAG. The other notions cannot be always recovered from the arrows of the DAG:

**(TC)** (Total cause)  $X$  is a total cause of  $Y$  if and only if it has a non-null total effect on  $Y$ - that is, if and only if there is some intervention on  $X$  alone such that for some values of the other variables, this intervention on  $X$  will change  $Y$ . The total effect of a change  $dx$  in  $X$  on  $Y$  is the change in the value of  $Y$  that would result from an intervention on  $X$  alone that changes it by amount  $dx$  (given the values of other variables that are not descendants of  $X$ ). ([38], p. 54)

**(CC)** (Contributing cause)  $X$  is a contributing cause of  $Y$  if and only if it makes non-null contribution to  $Y$  along some directed path in the sense that there is some set of values of variables that are not on this path such that if these variables were fixed at those values, there is some intervention on  $X$  that will change the value of  $Y$ . The contribution to a change in the value of  $Y$  due to a change  $dx$  in the value of  $X$  along some directed path is the change in the value of  $Y$  that would result from this change in  $X$ , given that the values of off path variables are fixed by independent interventions. ([38], pp. 54-55)

For instance in our earlier example consisting of the set of equations (6) and (7),  $X$  is a direct cause of  $Z$ . To make this more transparent we shall assume that the coefficients  $a, b, c$  are all equal to 1.

We first intervene on  $Y$  and set its value to 1 as we did above (i.e.,  $do(Y = 1)$ ). The result will be the altered system of equations:

$$(6') \quad Y = 1 \qquad (7) \quad Z = X + Y.$$

Next we perform two (independent) interventions on the system (6')-(7):  $do(X = 1)$  yields  $Z = 2$ ; and  $do(X = 2)$  yields  $Z = 3$ . We conclude that  $X$  is a direct cause of  $Z$  in the sense of (DC).

Consider now the set of equations

$$(8) \quad Y = aX \qquad (9) \quad Z = dY$$

which corresponds to the DAG

$$X \longrightarrow Y \longrightarrow Z$$

Any intervention  $do(Y = e)$  leads to the system of equations

$$(8') \quad Y = e \qquad (9) \quad Z = dY$$

Now it is obvious that no change in the value of  $X$  will have any influence on the value of  $Z$ , hence  $X$  is not a direct cause of  $Z$  as expected. On the other side, it is easy to see that  $X$  is a total cause of  $Z$  in the sense of (TC) – except in the special case that  $d = \frac{1}{a}$ .

In section 7 we shall represent some of these causal notions in the framework of causal team semantics, to which we now turn.

### 3 Causal team semantics

#### 3.1 Teams

Team semantics was introduced by W. Hodges ([14],[15]) in order to provide a compositional presentation of the (game-theoretically defined) semantics of Independence-Friendly logic ([13],[21]). In the following years, team semantics has been used to extend first-order logic with database dependencies (e.g. Dependence logic [32], Independence logic [10], Inclusion logic [9]); similar approaches have been applied to propositional logics ([41],[42]) and modal logics ([33], [31], [1]). Appropriate generalizations of teams have been used as descriptive languages for probabilistic dependencies ([7]), for quantum phenomena ([16]), for Bayes networks ([6]).

The basic idea of team semantics is that notions such as dependence and independence, which express properties of relations (instead of individuals), cannot be captured by Tarskian semantics, which evaluates formulas on single assignments<sup>2</sup>; the appropriate unit for semantical evaluation is instead the *team*, i.e., a *set* of assignments (all sharing a common variable domain). In the standard approach, all the values of the variables come from a unique domain of individuals associated with an underlying model. However, in order to model causal and counterfactual dependence, we shall need to relax the assumption that variables may take as values only individuals coming from a common domain. Instead we shall take variables to represent properties which may have all kinds of values (as specified by their range). That is, once a set  $Dom$  of variables is fixed, each assignment will be a mapping  $s : Dom \rightarrow \bigcup_{X \in Dom} Ran(X)$  such that  $s(X) \in Ran(X)$  for each  $X \in Dom$ . A **team**  $T$  of domain  $dom(T) = Dom$  will be any set of such assignments.

As an example, recall Woodward’s DAG corresponding to the structural equations (6) and (7) (subsection 2.2). We may take  $X$  to express the property “(whether it is) winter”,  $Y$  the property “(whether it is) cloudy” and  $Z$  the property “(whether it is) snowing” and take them to be represented in the team  $T = \{s_1, s_2\}$ :

	$X$	$Y$	$Z$
$s_1$	0	0	0
$s_2$	1	1	1

The basic semantic relation is now  $T \models \psi$ : the team  $T$  satisfies the formula  $\psi$ .

One can define a team semantics already for classical propositional languages. However such a semantics does not really add anything new, in the sense that

$T \models \psi$  if and only if for all  $s \in T$ ,  $s \models \psi$  (in the Tarskian sense).

That is, the meaning of a first-order formula is always reducible to a property of single assignments. However, once their semantics is expressed in terms

---

<sup>2</sup>This can be formally proved, see [3].

of teams, propositional languages can be extended in ways that would be unavailable within Tarskian semantics; for instance, one can add (functional) dependence atoms  $\models(X_1, \dots, X_n; Y)$  whose semantics is defined by

(\*)  $T \models(X_1, \dots, X_n; Y)$  if and only if for all  $s, s' \in X$ , if  $s(X_1) = s'(X_1), \dots, s(X_n) = s'(X_n)$ , then  $s(Y) = s'(Y)$

expressing that  $Y$  is functionally determined by  $\{X_1, \dots, X_n\}$ ; and this is a global property of the team, not reducible to properties of the single assignments<sup>3</sup>.

Thus, in our example, it holds that whether it is snowing depends completely on whether it is winter,  $T \models(X; Z)$ , and whether it is cloudy,  $T \models(Y; Z)$ ; and whether it is cloudy depends entirely on whether it is winter,  $T \models(X; Y)$ .

A team therefore might be used, for example, to represent a set of individual records coming from a statistical or experimental investigation; or, to represent all possible configurations that are compatible with the ranges of each variable; or yet, a subset of all possible configurations. This last case is particularly important in our context: it may well happen that some configurations are forbidden, even though they respect all variable ranges; this in particular happens if there are functional dependencies between the variables. For what regards the first possibility (analysis of statistical data) for many purposes it may be more suitable to use *multiteams*, which allow multiple copies of assignments. Team-theoretical logical languages could thus be used to express global properties of distributions of values. On a second, different interpretation, teams and multiteams may be used to represent epistemic uncertainty about the current state of affairs; if one thinks of each assignment as a possible world, then the team may be thought as representing a set of equally plausible worlds. An intervention on such an object should, therefore, produce a new set of equally plausible candidates for the actual world.

### 3.2 Causal teams

Despite some claims to the contrary in the literature, teams are insufficient to represent counterfactuals and causal notions based on them. As explained in subsection 2.2, a proper treatment of causal notions requires an account of counterfactual information, as encoded e.g. in invariant structural equations. It is true that a team sustains a number of functional dependencies among variables; but such dependencies may well be contingent, and disappear if the system is intervened upon. For this reason, we must extend teams so that they incorporate invariant dependencies or functions; and we must explain what may count as an intervention *on a team*.

Besides the ideal case in which a causal model contains a complete description of the functions involved in the structural equations (*parametric* case), we develop our semantics with enough generality as to accommodate the more realistic case in which we possess only partial information about the functions

<sup>3</sup>Actually, it is a property of *pairs* of assignments of  $X$ , but more complex formulas arise whose meaning cannot be analogously reduced.

(*nonparametric* case); as an extreme case, the model might only incorporate information as to which functional *dependencies* are invariant. In the nonparametric case, not all counterfactual statements can be evaluated; the additional logical complications related to this case will be examined in section 5. Most of the paper will focus on the parametric case.

Before proceeding, we want to fix some notational conventions. As is often done in the literature on causal models, we use the symbol  $PA_Y$  ambiguously, so that it may mean either the set of parents of  $Y$ , or a sequence of the same variables in some fixed alphabetical ordering. For other sets/sequences of variables, we will adhere to the following conventions:

**Notation 3.1.** • We use boldface letters such as  $\mathbf{X}$  to denote either a set  $\{X_1, \dots, X_n\}$  of variables or a sequence of the same variables (in the fixed alphabetical order)

- We use  $\mathbf{x}$  to denote a set or sequence of values, each of which is a value for exactly one of the variables in  $\mathbf{X}$ . We leave the details of these correspondences between variables and values as non-formalized.
- Writing  $s(\mathbf{X})$  we mean the set/sequence of values  $s(X_1), \dots, s(X_n)$  that the assignment  $s$  assigns to each of the variables in  $\mathbf{X}$
- $Ran(\mathbf{X})$  is an abbreviation for  $\prod_{X \in \mathbf{X}} Ran(X)$
- By  $\mathbf{X} \setminus \mathbf{Y}$  we denote the set/the sequence (in alphabetical order) of variables occurring in  $\mathbf{X}$  but not in  $\mathbf{Y}$ , and by  $\mathbf{x} \setminus \mathbf{y}$  a corresponding set/sequence of values
- By  $\mathbf{X} \cap \mathbf{Y}$  we denote the set/sequence of variables occurring in both  $\mathbf{X}$  and  $\mathbf{Y}$ , and by  $\mathbf{x} \cap \mathbf{y}$  the corresponding set/sequence of values and so on.

Given a team  $T^-$  and a variable  $X \in dom(T^-)$ , we write  $T^-(X)$  for the set of values that are obtained for  $X$  in the team  $T^-$ ; that is,  $T^-(X) = \{s(X) | s \in T^-\}$ . As before, we say that a team  $T^-$  satisfies a functional dependence  $= (X_1, \dots, X_n; Y)$ , and we write  $T^- \models (X_1, \dots, X_n; Y)$ , if:

for all  $s, s' \in T^-$ , if  $s(X_i) = s'(X_i)$  for all  $i = 1..n$ , then  $s(Y) = s'(Y)$ .

**Definition 3.2.** A *causal team*  $T$  over variable domain  $dom(T)$  with endogenous variables  $\mathbf{V} \subseteq dom(T)$  is a quadruple  $T = (T^-, G(T), \mathcal{R}_T, \mathcal{F}_T)$ , where:

1.  $T^-$  is a team.
2.  $G(T) = (dom(T), E)$  is a graph over the set of variables. For any  $X \in dom(T)$ , we denote as  $PA_X$  the set of all variables  $Y \in dom(T)$  such that the arrow  $(Y, X)$  is in  $E$ .
3.  $\mathcal{R}_T = \{(X, Ran(X)) | X \in dom(T)\}$  (where the  $Ran(X)$  may be arbitrary sets) is a function which assigns a range to each variable

4.  $\mathcal{F}_T$  is a function  $\{(V_i, f_{V_i}) | V_i \in \mathbf{V}\}$  that assigns to each endogenous variable a  $|PA_{V_i}|$ -ary function  $f_{V_i} : \text{dom}(f_{V_i}) \rightarrow \text{ran}(V_i)$  (for some  $\text{dom}(f_{V_i}) \subseteq \prod \text{Ran}(X_i)$ , where  $X_i \in PA_{V_i}$ )

which satisfies the further restrictions:

- a)  $T^-(X) \subseteq \text{Ran}(X)$  for each  $X \in \text{dom}(T)$   
b) If  $PA_Y = \{X_1, \dots, X_n\}$ , then  $T^- \models (X_1, \dots, X_n; Y)$   
c) if  $s \in T^-$  is such that  $s(PA_Y) \in \text{dom}(f_Y)$ , then  $s(Y) = f_Y(s(PA_Y))$ .

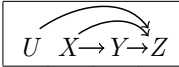
Clause b) is there to ensure that whenever the graph contains an arrow  $X_i \rightarrow Y$ , and  $\{X_1, \dots, X_n\}$  is the maximal set of variables whence arrows come to  $Y$ , then the team satisfies the corresponding functional dependency  $\models (X_1, \dots, X_n; Y)$ . Clause c) further ensures that such functional dependency is in accordance with the (partial description of the) function  $f_Y \in \mathcal{F}_T$ .

The functional component  $\mathcal{F}_T$  induces an associated system of structural equations, say

$$Y := \mathcal{F}_T(Y)(PA_Y)$$

for each variable  $Y \in \text{dom}(T)$ .

**Example 3.3.** Consider a causal team  $T$  which has underlying team  $T^- = \{\{(U, 2), (X, 1), (Y, 2), (Z, 4)\}, \{(U, 3), (X, 1), (Y, 2), (Z, 4)\}, \{(U, 1), (X, 3), (Y, 3), (Z, 1)\}, \{(U, 1), (X, 4), (Y, 1), (Z, 1)\}, \{(U, 4), (X, 4), (Y, 1), (Z, 1)\}\}$ , graph  $G(T) = (\{U, X, Y, Z\}, \{(U, Z), (X, Y), (Y, Z), (X, Z)\})$ , ranges  $\text{Ran}(U) = \text{Ran}(X) = \text{Ran}(Y) = \text{Ran}(Z) = \{1, 2, 3, 4\}$ , and partial description of (one value of) the invariant function for  $Z$ :  $\mathcal{F}(Z)(4, 1, 2) := 3$ . We represent the  $T^-$  and  $G(T)$  components of  $T$  by means of a decorated table:

			
$U$	$X$	$Y$	$Z$
2	1	2	4
3	1	2	4
1	3	3	1
1	4	1	1
4	4	1	1

### 3.3 Explicit causal teams

For many purposes – first of all, to keep a smoother account of the operations of taking subteams, and of applying iterated interventions – it will be convenient to restrict attention to causal teams of a special form. This restriction causes no loss of generality within the developments pursued in the present paper.

**Definition 3.4.** A causal team  $T = (T^-, G(T), \mathcal{R}_T, \mathcal{F}_T)$  with endogenous variables  $\mathbf{V}$  is **explicit** if, for every  $V \in \mathbf{V}$ , the following additional condition holds:

- d) Let  $(X_1, \dots, X_n)$  be the list of the parents of  $V$  in the fixed alphabetical order. For every  $s \in T^-$ ,  $(s(X_1), \dots, s(X_n)) \in \text{dom}(f_Y)$ .

(Here, as in the definition of causal team,  $f_Y$  is a shorthand for  $\mathcal{F}_T(Y)$ ).

In words, the  $\mathcal{F}_T$  component of an explicit causal team encodes *all* the information of the team that concerns invariant functions; no further values of the functions can be reconstructed from the team component  $T^-$ .

Given any causal team, it is always possible to construct in a canonical way an explicit causal team that corresponds to it, in the sense that it encodes exactly the same information (but it is more stable under the operations of taking subteams or interventions – to be defined in the following subsections). For this purpose, given a causal team  $T = (T^-, G(T), \mathcal{R}_T, \mathcal{F}_T)$  with endogenous variables  $\mathbf{V}$ , to any variable  $Z \in \mathbf{V}$  we may associate an **explicit function**  $h_Z^T : \text{Ran}(W) \rightarrow \text{Ran}(Z)$  as follows; given  $pa_Z \in \text{Ran}(W)$ , define

$$h_Z^T(pa_Z) = \begin{cases} f_Z(pa_Z) & \text{if this is defined in } T \\ s(Z) & \text{if there is some row } s \text{ in } T \text{ with } s(PA_Z) = pa_Z \end{cases}$$

(Conditions b) and c) in the definition of causal team ensure that  $h_Z^T$  is well-defined.)

We collect these explicit functions in a single function  $\mathcal{H}_T$  such that, for each variable  $Z \in \mathbf{V}$ ,  $\mathcal{H}_T(Z) := h_Z^T$ . Then:

The explicit causal team associated to  $T$  is  $T^E := (T^-, G(T), \mathcal{R}_T, \mathcal{H}_T)$ .

### 3.4 Causal subteams

It will be important, in order to define a semantics for our languages, to talk about causal subteams. A causal subteam  $S$  of a causal team  $T$  is meant to express a condition of lesser uncertainty; this will be encoded by the fact that the assignments in  $S^-$  form a subset of the assignments of  $T^-$ , which may be interpreted as the fact that less configurations are considered possible. At the same time, the transition to a subteam should not erase information concerning the graph, the ranges of variables, and the invariant functions. The definitions in the previous subsection should make it clear that this proviso is easily guaranteed if  $T$  is an *explicit* causal team. In this case, we can define:

**Definition 3.5.** *Given an explicit causal team  $T$ , a **causal subteam**  $S$  of  $T$  is a causal team with the same domain and the same set of endogenous variables, which satisfies the following conditions:*

1.  $S^- \subseteq T^-$
2.  $G(S) = G(T)$
3.  $\mathcal{R}_S = \mathcal{R}_T$
4.  $\mathcal{F}_S = \mathcal{F}_T$ .

In case the team  $T$  is not explicit, what can go wrong is that, for some endogenous variable  $Y$ , there may be some assignment  $s \in T^- \setminus S^-$  such that

$(s(X_1), \dots, s(X_n)) \notin \text{dom}(f_Y)$  (where  $X_1, \dots, X_n$  list  $PA_Y$  in alphabetical order); in such case, the team  $T$  encodes the fact that the invariant function which produces  $Y$  assigns, to the list of arguments  $(s(X_1), \dots, s(X_n))$ , the value  $s(Y)$ ; but this information is lost in the subteam  $S$ . To avoid this problem, we can define more generally:

**Definition 3.6.** *Given a causal team  $T$ , a **causal subteam**  $S$  of  $T$  is a causal subteam of the associated explicit team  $T^E$  (as defined in subsection 3.3)*

This second definition obviously coincides with the previous one over explicit causal teams.

### 3.5 A basic language and its semantics

Before discussing interventions and counterfactuals, we need to specify what it means for a causal team to satisfy a propositional formula. The kind of propositional language we consider, for now, contains atomic dependence statements, as defined above; atomic formulas of the form  $Y = y$  and  $Y \neq y$ , where  $Y \in \text{dom}(T^-)$  and  $y \in \text{Ran}(Y)$ ; connectives  $\wedge$  and  $\vee$ .

By analogy with the other kinds of team semantics that have been proposed in the literature, we can define satisfaction of a propositional formula by a causal team by the clauses:

- $T \models (\mathbf{X}; Y)$  if for all  $s, s' \in T$ ,  $s(\mathbf{X}) = s'(\mathbf{X})$  implies  $s(Y) = s'(Y)$ .
- $T \models Y = y$  if, for all  $s \in T^-$ ,  $s(Y) = y$ .
- $T \models Y \neq y$  if, for all  $s \in T^-$ ,  $s(Y) \neq y$ .
- $T \models \psi \wedge \chi$  if  $T \models \psi$  and  $T \models \chi$ .
- $T \models \psi \vee \chi$  if there are two causal subteams  $T_1, T_2$  of  $T$  such that  $T_1^- \cup T_2^- = T^-$ ,  $T_1 \models \psi$  and  $T_2 \models \chi$ .<sup>4</sup>

### 3.6 Selective implication

Our main goal is to give an exact semantics to counterfactual statements of the form “If  $\psi$  had been the case, then  $\chi$  would have been the case”. Very often, however, one find examples in the literature where these statements are embedded into a larger context. We have seen that von Wright ([36]) considers examples of the form “on occasions where  $p$ , in fact was not the case,  $q$  would have accompanied it, had  $p$  been the case”. Pearl ([23]) analyzes the following query: “what is the probability  $Q$  that a subject who died under treatment ( $X = 1, Y = 1$ ) would have recovered ( $Y = 0$ ) had he or she not been treated ( $X = 0$ )?”

---

<sup>4</sup>Notice that defining the union of any pair of *causal* teams of the same domain is problematic, as the information on invariant functions given by each team might be incompatible with the information encoded in the other team.

The appropriate representation of the last statement seems to be:

$$(X = 1 \wedge Y = 1) \supset (X = 0 \Box\rightarrow Y = 0).$$

where the symbol  $\Box\rightarrow$  stands for counterfactual implication, while the implication  $\supset$  is a form of restriction of the range of application of the counterfactual to the available evidence. The symbol  $\supset$ , that we may call *selective implication*, may be thought as an abbreviation for  $(\neg\dots)\vee\dots$  as in classical logic; however, the resulting operation is very different from classical material implication. What it does is to generate a subteam by selecting those assignments which satisfy the antecedent; and then it is checked whether the consequent holds in this subteam.

Given a causal team  $T$ , and a classical formula  $\psi$  (that is, a formula as in subsection 3.5, but without dependence atoms) define the subteam  $T^\psi$  by the condition:

- $(T^\psi)^- = \{s \in T^- \mid \{s\} \models \psi\}$ .

Then, we define selective implication by the clause:

- $T \models \psi \supset \chi$  iff  $T^\psi \models \chi$ .

Here the consequent  $\chi$  can be any formula of our current logical language; therefore, it might happen not to be a property of single assignments. Instead, we require for now the antecedent to be of classical form. The general idea is that selective implication is a reasonable operator only for antecedents which are *flat* formulas, in the sense with which this word is used in the literature on logics of dependence (which will be reviewed in the following subsections). Actually, typical applications involve at most conjunctions of atomic formulas of the type  $Z = z$ .

**Example 3.7.** *We observe that the selective implication*

$$T \models Z = 3 \supset Y = 2$$

*holds on any causal team that is based on the team  $T$  which is depicted in the figure:*

$Z$	$Y$	$X$
1	2	3
2	1	1
3	2	1
3	2	2

*To see that the formula holds on it, we have to construct the reduced subteam  $T^{Z=3}$ :*

$Z$	$Y$	$X$
3	2	1
3	2	2



which is obtained by selecting the third and fourth row of the previous table (the rows that satisfy  $Z = 3$ ). We can see, then, that  $Y = 2$  is satisfied by each row of this smaller table. Therefore, by the semantical clause for this kind of atomic formulas,  $T^{Z=3} \models Y = 2$ . Then, the semantical clause for selective implication allows us to conclude that  $T \models Z = 3 \supset Y = 2$ .

Notice that we only needed the team structure in order to evaluate a selective implication; all the information required to evaluate it is already encoded in the team, and no information about the structural equations or the underlying DAG is needed.

### 3.7 Interventions on teams: some examples

We want to define (interventionist) *counterfactual implication*. The task is more complicated than in the case of selective implication; we first illustrate the idea with some examples; formal definitions will be provided after that. Informally, the idea is that a counterfactual  $X = x \square \rightarrow \psi$  is true in the causal team  $T$  if  $\psi$  is true in the causal team which results from an intervention  $do(X = x)$  on the team  $T$ .

**Example 3.8.** Consider any causal team  $T$  with  $T^- = \{(X, 1), (Y, 2), (Z, 3)\}, \{(X, 2), (Y, 1), (Z, 4)\}, \{(X, 4), (Y, 1), (Z, 4)\}, \{(X, 3), (Y, 3), (Z, 4)\}$ , and underlying graph  $G(T) = (\{X, Y, Z\}, \{(X, Y), (Y, Z)\})$ . We can represent it as an annotated table as below.

$X \rightarrow Y \rightarrow Z$		
1	2	3
2	1	4
4	1	4
3	3	4

We want to establish whether the conditional  $Y = 2 \square \rightarrow Z = 3$  holds in  $T$ . The idea is to intervene in  $T$  by setting the value of  $Y$  to 2; updating all other variables that invariantly depend on  $Y$ ; and removing all the arrows that enter into  $Y$ . The causal team thus produced will be denoted by  $T_{Y=2}$ . So, first we intervene on  $Y$ :

$X$	$Y \rightarrow Z$
1	<b>2</b> ...
2	<b>2</b> ...
4	<b>2</b> ...
3	<b>2</b> ...

$Z$  is the only variable which has an invariant dependence on  $Y$ ; therefore, we have to update its column. Since the function that determines  $Z$  has  $Y$  as its only parameter, we just have to consult  $T$  and see that, in rows where  $Y$  has value 2,  $Z$  takes value 3. Therefore,  $T_{Y=2}$  looks like this:

$X$	$Y \rightarrow Z$
1	<b>2</b> 3
2	<b>2</b> 3
4	<b>2</b> 3
3	<b>2</b> 3

Now  $T_{Y=2} \models Z = 3$ , therefore we conclude that  $T \models Y = 2 \square \rightarrow Z = 3$ . A remark: the team  $T_{Y=2}$  describes a counterfactual situation in which we have certainty about the values of  $Y$  and  $Z$ , but not about the value of  $X$ .

Notice that this example uses both the team and the graph structure, but the specific invariant functions are not needed in the evaluation of this specific sentence. However, the observations above do not cover, for example, evaluation of counterfactuals of the form  $Y = 4 \square \rightarrow \dots$ , because team and graph structure do not tell us anything about what value should  $Z$  take in circumstances in which  $Y = 4$ .

**Example 3.9.** We consider a slightly less trivial example. Here we evaluate the sentence  $Y = 2 \square \rightarrow (Z = 2 \vee Z = 3)$  in the causal team  $T$  shown in the picture below

$X$	$Y \rightarrow Z$
1	1 1
1	2 2
2	2 3

So we must produce again  $T_{Y=2}$ . First we intervene on  $Y$

$X$	$Y \rightarrow Z$
1	<b>2</b> ...
1	<b>2</b> ...
2	<b>2</b> ...

and next we update the value of  $Z$ , taking into account both the values of  $X$  and  $Y$  in each row (since  $Z$  is reached both by arrows coming from  $X$  and from  $Y$ ). Notice that now two of the modified rows become equal, so that only one appears in the table:

$X$	$Y \rightarrow Z$
1	<b>2</b> 2
2	<b>2</b> 3

This team can be partitioned into two causal subteams:

<table border="1"> <thead> <tr> <th><math>X</math></th> <th><math>Y \rightarrow Z</math></th> </tr> </thead> <tbody> <tr> <td>1</td> <td>2 2</td> </tr> </tbody> </table>	$X$	$Y \rightarrow Z$	1	2 2	<table border="1"> <thead> <tr> <th><math>X</math></th> <th><math>Y \rightarrow Z</math></th> </tr> </thead> <tbody> <tr> <td>2</td> <td>2 3</td> </tr> </tbody> </table>	$X$	$Y \rightarrow Z$	2	2 3
$X$	$Y \rightarrow Z$								
1	2 2								
$X$	$Y \rightarrow Z$								
2	2 3								

the first of which satisfies  $Z = 2$ , while the latter satisfies  $Z = 3$ . Therefore, by the semantical clause for disjunction,  $T_{Y=2}$  satisfies  $Z = 2 \vee Z = 3$ . We may conclude that  $T \models Y = 2 \square \rightarrow (Z = 2 \vee Z = 3)$ .

Again, the team and graph structure sufficed for the evaluation of the formula.

**Example 3.10.** Here is a still more complicated example. We want to evaluate  $X = 1 \square \rightarrow Y = 2$  in the nonparametric team  $T$  shown in the picture:

U $X \rightarrow Y \rightarrow Z$			
2	1	2	4
3	1	2	4
1	3	3	1
1	4	1	1

In order to evaluate  $X = 1 \square \rightarrow Y = 2$  we need to generate the causal team  $T_{X=1}$ . First we intervene on  $X$ ; this will affect all descendants of  $X$ , which in this case are the (children)  $Y$  and  $Z$ .

U $X \rightarrow Y \rightarrow Z$			
2	<b>1</b>	...	...
3	<b>1</b>	...	...
1	<b>1</b>	...	...
1	<b>1</b>	...	...

Notice, however, that we cannot yet evaluate  $Z$  (which is a function of  $U, X$  and  $Y$ ) unless we first update  $Y$ :

U $X \rightarrow Y \rightarrow Z$			
2	1	<b>2</b>	...
3	1	<b>2</b>	...
1	1	<b>2</b>	...

Finally we update  $Z$ , but we have a surprise:

U $X \rightarrow Y \rightarrow Z$			
2	1	2	<b>4</b>
3	1	2	<b>4</b>
1	1	2	$\hat{f}_Z(1, 1, 2)$

The information contained in the team  $T^-$  is insufficient for the evaluation of the  $Z$ -value of the last row of  $T_{X=1}$ . Since the causal team is nonparametric, we do not have further information about the function which generates  $Z$  in terms

of  $U, X, Y$  beyond what is encoded in  $T^-$ . Therefore, the best we can do is to fill that part of the table with a formal term. Here we wrote  $\hat{f}_Z$  as a formal symbol distinguished from the function  $f_Z$ . In general, in case of repeated interventions, we can expect also complex terms, with incapsulated function symbols, to be produced. Notice however that we have no uncertainties about the  $Y$  column; so, it is natural to state that  $T_{X=1} \models Y = 2$ , and that, therefore,  $T \models X = 1 \square \rightarrow Y = 2$ . (Notice also that, in this specific case, it was not really necessary to generate the  $Z$  column in order to evaluate the counterfactual; this suggests a possible computational simplification in the evaluation of counterfactuals, but also a possible generalization; we might be able to answer some counterfactual questions also in cases where the process which generates the counterfactual team does not halt). What if we were instead trying to evaluate some statement about  $Z$  under these same counterfactual circumstances? We might renounce completely to assign truth values to such statements; or we might perhaps add a further truth value, which holds of such statement if the statement is admissible given the form of the terms involved. A precise semantical definition is not trivial, and we address it later.

Finally notice that, if our initial causal team  $T$  had had  $\mathcal{F}_T(Z) = f_Z$  with  $(1, 1, 2) \in \text{dom}(f_Z)$ , we could have used this function to explicitly evaluate the last entry of  $T_{X=1}$  with the number  $f_z(1, 1, 2)$ .

### 3.8 Formal definition of interventions and counterfactuals

All the examples in the previous subsection have one aspect in common: the graphs underlying causal teams are acyclic. In this case the corresponding causal team is called *recursive*, in analogy with the literature on structural equation models; the examples show that, for these kinds of causal teams, the notion of intervention is naturally conceived in algorithmic terms. We now move towards a precise definition.

We have learned a number of lessons from the previous examples:

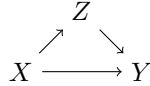
- A. An intervention  $do(X = x)$  amounts to 1) setting the whole  $X$ -column to  $x$ ; 2) eliminating all arrows that enter into  $X$ ; 3) updating the columns that correspond to descendants of  $X$ .
- B. It might not be possible to update all the descendants of  $X$  in a single step (actually, the order of updating might not be trivial to decide).
- C. The information encoded in the causal team might be insufficient for generating, under intervention, a proper causal team; we must then admit teams which assign formal terms to some variables.

We begin by addressing this last problem. Given a graph  $G$  whose set of vertices  $V$  is a set of variables, we call  $L_G$  the set of function symbols  $\hat{f}_X$  (of arity  $\text{card}(PA_X)$ ), for each  $X \in V$  (actually, we only need one symbol for each endogenous variable, that is, for variables which have indegree at least one). We call  $G$ -terms the terms generated from variables in  $V$  and from symbols

in  $L_G$  by the obvious inductive rules; the set of  $G$ -terms will be denoted as  $Term_G$ . When speaking of a causal team  $T$ , we will implicitly assume, from now on, that the ranges of all variables contain the set of terms  $Term_{G(T)}$ . Actually, it is not difficult to prove that (iterated) interventions (as defined below) on a recursive causal team with finite variable ranges can generate only a finite number of formal terms. Therefore, in principle a finite causal team could always be extended to a *finite* causal team with formal terms, by using an appropriate finite subset of  $Term_{G(T)}$  to extend the ranges of variables.

We may combine these ideas with those of subsection 3.3. Given a causal team  $T$ , let  $T'$  be the causal team obtained by extending the ranges of variables with formal terms, as described above. Form then the corresponding explicit team  $T^{FE} := (T')^E$  following the construction given in subsection 3.3. We call  $T^{FE}$  the **fully explicit** causal team corresponding to  $T$ . The invariant functions of  $T^{FE}$  now encode sufficient information for performing any kind of intervention that we will consider in this paper. The fact that the ranges of variables contain formal terms solves issue C. The fact that the team is explicit ensures that no information is lost after an intervention.

Now we address problem *B*. How do we decide the order of updating for the columns? To understand the problem, observe the graph in the figure:



$Y$  is connected to  $X$  by an arrow, so one might think that, in an intervention on  $X$ , it is possible to update  $Y$  immediately after updating  $X$ . However, this is impossible, because also the updated values of  $Z$  are needed in the evaluation of the new values for  $Y$ . The point is that, from  $X$  to  $Y$ , there is a longer path than the direct one. This suggests defining a distance from  $X$ ,  $d(X, \cdot)$ , given by the maximum length of paths from  $X$  to  $\cdot$ . Nodes at distance 1 can be immediately evaluated after updating  $X$ ; at this point it is safe to update nodes at distance 2; and so on. Nodes that are not accessible by directed paths from  $X$  will be assigned a negative distance and excluded from the updating procedure. Of course, this strategy will work provided there are no loops in the graph, or at least no loops in the part of the graph which is accessible by directed paths from  $X$  (the set of descendants of  $X$ ). Since in the applications it is very common to have conjunctive interventions, say  $do(X_1 = x_1 \wedge \dots \wedge X_n = x_n)$ , we will define, more generally, the notion of distance from a *set* of variables  $\mathbf{X}$ . In this more complex case, one must consider a reduced graph in which all arrows entering in  $\mathbf{X}$  have been removed, and examine the directed paths of this reduced graph. The reader may think of the  $card(\mathbf{X}) = 1$  case for ease of visualization.

**Definition 3.11.** *Given a graph  $G = (\mathbf{V}, E)$  and  $\mathbf{X} \subseteq \mathbf{V}$ ,*

- *We denote as  $G_{-\mathbf{X}} = (\mathbf{V}, E_{-\mathbf{X}})$  the graph obtained by removing all arrows going into some vertex of  $\mathbf{X}$  (i.e., an edge  $(V_1, V_2)$  is in  $E_{-\mathbf{X}}$  iff it is in*

$E$  and  $V_2 \notin \mathbf{X}$ ). Notice that, in the special case that  $\mathbf{X} = \{X\}$ , the set of directed paths of  $G_{-\mathbf{X}}$  starting from  $X$  coincides with the set of directed paths of  $G$  starting from  $X$ .

- Let  $Y \in \mathbf{V}$ . We call (**evaluation**) **distance** between  $\mathbf{X}$  and  $Y$  the value  $d_G(\mathbf{X}, Y) = \sup\{\text{lenght}(P) \mid P \text{ dir. path of } G_{-\mathbf{X}} \text{ going from some } X \in \mathbf{X} \text{ to } Y\}$ . In case no such path exists, we set  $d_G(\mathbf{X}, Y) = -1$ . Clearly, if the graph is finite and acyclic,  $d_G(\mathbf{X}, Y) \in \mathbb{N} \cup \{-1\}$  for any pair  $\mathbf{X}, Y$ . When the graph is clear from the context, we simply write  $d(\mathbf{X}, Y)$ .

Let  $T$  be a *fully explicit, recursive* causal team of endogenous variables  $\mathbf{V}$ . Let  $\mathbf{X} = \{X_1, \dots, X_n\} \subseteq \text{dom}(T)$  and  $x_1 \in \text{Ran}(X_1), \dots, x_n \in \text{Ran}(X_n)$  corresponding values with the additional property that, if  $X_i$  and  $X_j$  denote the same variable, then  $x_i = x_j$ . We define the **intervention**  $do(\bigwedge_{i=1..n} X_i = x_i)$  (in short,  $do(\mathbf{X} = \mathbf{x})$ ) as an algorithm<sup>5</sup>:

Stage 0. Delete all arrows coming into  $\mathbf{X}$ , and replace each assignment  $s \in T$  with  $s(\mathbf{x}/\mathbf{X})$ . Replace  $\mathcal{F}_T$  with its restriction to  $\mathbf{V} \setminus \mathbf{X}$ . Denote the resulting causal team (of endogenous variables  $\mathbf{V} \setminus \mathbf{X}$ ) as  $T_0$ .

Stage  $n + 1$ . If  $\{Z_1, \dots, Z_{k_{n+1}}\}$  is the set of all the variables  $Z_j$  such that  $d_G(\mathbf{X}, Z_j) = n + 1$ , define a new team  $T_{n+1}$  by replacing each  $s \in T_n$  with the assignment  $s(f_{Z_1}(s(PA_{Z_1}))/Z_1, \dots, f_{Z_{k_{n+1}}}(s(PA_{Z_{k_{n+1}}})))/Z_{k_{n+1}}$ .

End the procedure after step  $\hat{n} = \sup\{d_{G(T)}(\mathbf{X}, Z) \mid Z \in \text{dom}(T)\}$ .

Notice that  $\mathcal{R}_T$  is not modified by the algorithm; and that, except for the modifications to  $G(T)$ , it would be the same to apply the algorithm to each assignment separately. In case the causal team  $T$  is recursive but not fully explicit, we should begin the algorithm with an additional step:

Step  $-1$ . Replace  $T$  with the corresponding fully explicit causal team  $T^{FE}$ .

In case the intervention  $do(\mathbf{X} = \mathbf{x})$  is a terminating algorithm on  $T$ , we define the team  $T_{\mathbf{X}=\mathbf{x}}$  as the team which is returned when  $do(\mathbf{X} = \mathbf{x})$  is applied to  $T$ . Actually, relaxing a bit the notion of algorithm, the *do* algorithm applies as well to causal teams with infinite variable ranges:

**Theorem 3.12.** *If  $G(T)$  is a finite acyclic graph, then  $T_{\mathbf{X}=\mathbf{x}}$  is well-defined.*

*Proof.* We can assume without loss of generality that  $T$  is fully explicit.

Assume also, at first, that  $T^-$  is finite.

---

<sup>5</sup>Notice that our definition allows the presence of many occurrences of the same variable in different conjuncts, provided that all such conjuncts assign the same value to the given variable. We do *not* define interventions associated to contradictory conjuncts. E.g., the intervention  $do(X = x \wedge X = x \wedge Y = y)$  is defined, and coincides with the intervention  $do(X = x \wedge Y = y)$ ; while the intervention  $do(X = 1 \wedge X = 2)$  is not defined.

Suppose that, for  $\mathbf{X} \subseteq \text{dom}(T), Y \in \text{dom}(T), d(\mathbf{X}, Y) > \text{card}(G(T))$ . Then there is a path  $P$  from some  $X \in \mathbf{X}$  to  $Y$  such that  $\text{length}(P) > \text{card}(G(T))$ . Then there is a node which is crossed at least twice by  $P$ ; so  $G(T)$  contains a cycle: contradiction.

Therefore,  $\sup\{d(\mathbf{X}, Z) | Z \in \text{dom}(T)\} \leq \text{card}(G(T))$ ; this means that the “for” cycle in the algorithm goes through a finite number of iterations over the variable  $n$ .

Finally, notice that, for each  $n$ , there are a finite number of variables  $Z$  such that  $d(\mathbf{X}, Z) = n$  (due to finiteness of  $G(T)$ ) and a finite number of assignments  $t$  in the team that is undergoing modification (due to the finiteness of  $T^-$ ). Therefore, the algorithm terminates after a finite number of steps.

If instead  $T^-$  is infinite, we can replace  $\text{do}(\mathbf{X} = \mathbf{x})$  with an infinitary algorithm which, in each iteration of the “for” cycle, performs simultaneously the substitution  $t \mapsto t(f_{Z_1}(s(PA_{Z_1}))/Z_1, \dots, f_{Z_{k_{n+1}}}(s(PA_{Z_{k_{n+1}}})))/Z_{k_{n+1}}$  for all the assignments  $t$  in the current team. By the arguments above, such an “algorithm” terminates and yields a well-defined causal team.  $\square$

In case a causal team is not recursive, the algorithm above may well fail to terminate. In this case, a definition of the intervention  $\text{do}(\mathbf{X} = \mathbf{x})$  could still be given in terms of the set of solutions of the structural equations, provided the team is parametric. Each assignment  $s \in T$  is a solution of the system of structural equations encoded in  $\mathcal{F}_T$  (this is ensured by parametricity and conditions a) and c) in the definition of causal team). Galles and Pearl ([8]) considered the case of systems with unique solutions (meaning that: 1) for fixed values of the exogenous variables, the system has a unique solution, and 2) each “intervened” system of equations obtained from the initial one replacing some equations of the form  $X := f_X(PA_X)$  with constant equations  $X := x$  still has a unique solution for each choice of values for the exogenous variables). In case the  $\mathcal{F}_T$  component of the team encodes a system with unique solutions, then the natural way to define an intervention  $\text{do}(X = x)$  on the team is: replace each assignment  $s \in T$  with the (unique) assignment  $t$  which encodes the solution of the intervened system for the assignment  $s(\mathbf{U})$  of values to the exogenous variables<sup>6</sup>.

It is not difficult to extend the previous ideas to systems that may have no solution at all. In case the intervened system obtained replacing the appropriate equations with  $\mathbf{X} = \mathbf{x}$  has no solution corresponding to the choice  $s(\mathbf{U})$  for the exogenous variables, then no assignment should correspond to  $s$  in the intervened team  $T_{\mathbf{X}=\mathbf{x}}$ . Although this extension is straightforward, we can expect significant changes in the underlying logic.

The case of systems with multiple solutions is more problematic. If many solutions correspond to an assignment of the initial team, should we include them all in the intervened team? And, for what regards the probabilistic developments of the following sections, should we consider all the assignments thus produced as equiprobable? These matters should probably be settled according

<sup>6</sup>In case the intervention acts also on some of the exogenous variables, this idea should be modified in an obvious way.

to the kind of interpretation we want to give to nonrecursive causal teams in a given application, and one is lead to the classical problems of interpretation of nonrecursive causal models (see e.g [29]). It might also be reasonable to model such an intervention as producing not one, but multiple teams, corresponding to possible different outcomes of the intervention. This set of “accessible teams” would then induce a nontrivial modality, and it would then be reasonable to treat counterfactuals as necessity operators in a dynamic logic setting (in the spirit of [11]). This would lead us far away from the approach of the present paper, in which we focus on the nonproblematic recursive case. We now return to it.

Having defined the intervened team  $T_{\mathbf{X}=\mathbf{x}}$ , we are immediately led to a semantical clause for counterfactuals of the form  $\mathbf{X} = \mathbf{x} \square \rightarrow \psi$ :

$$T \models \mathbf{X} = \mathbf{x} \square \rightarrow \psi \iff T_{\mathbf{X}=\mathbf{x}} \models \psi.$$

In case the antecedent is inconsistent (i.e., it contains two conjuncts  $X_i = x_i, X_i = x'_i$  with  $x_i \neq x'_i$ ), the corresponding intervention is not defined; in this case, we postulate the counterfactual to be (trivially) true.

### 3.9 Logical languages

We call the (basic) *language of causal dependence*,  $\mathcal{CD}$ , the language formed by the following rules:

$$Y = y \mid Y \neq y \mid =(\mathbf{X}; Y) \mid \psi \wedge \chi \mid \psi \wedge \chi \mid \theta \supset \chi \mid \mathbf{X} = \mathbf{x} \square \rightarrow \chi$$

for  $Y, \mathbf{X}$  variables,  $y, \mathbf{x}$  values,  $\psi, \chi$  formulae of  $\mathcal{CD}$ ,  $\theta$  classical formula.

If we interpret this language by the semantical clauses introduced so far, it can be shown that:

**Theorem 3.13.** *The logic  $\mathcal{CD}$  is downwards closed, that is: if  $\varphi \in \mathcal{CD}$ ,  $T$  is a causal team,  $T'$  is a causal subteam of  $T$ , and  $T \models \varphi$ , then also  $T' \models \varphi$ .*

*Proof.* We prove it by induction on the syntactical structure of  $\varphi$ . The atomic and propositional cases are routine.

Suppose  $T \models \psi \supset \chi$ . Then  $T^\psi \models \chi$ . Now  $T'^\psi \subseteq T^\psi$ ; therefore, by inductive hypothesis,  $T'^\psi \models \chi$ . Thus, by the semantic clause for selective implication,  $T' \models \psi \supset \chi$ .

Suppose  $T \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$ . Then  $T_{\mathbf{X}=\mathbf{x}} \models \chi$ . But since the algorithm  $do(\mathbf{X} = \mathbf{x})$  acts on each assignment separately, we can conclude that  $(T'_{\mathbf{X}=\mathbf{x}})^- \subseteq (T_{\mathbf{X}=\mathbf{x}})^-$ . So, by the inductive hypothesis,  $T'_{\mathbf{X}=\mathbf{x}} \models \chi$ . Applying again the semantic clause,  $T' \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$ .  $\square$

**Theorem 3.14.** *The logic  $\mathcal{CD}$  has the empty set property, that is: for every  $\varphi \in \mathcal{CD}$ ,  $\emptyset \models \varphi$ .*

*Proof.* By induction on the syntax of the formula. The new cases are those for  $\supset$  and  $\square \rightarrow$ .



Let  $\varphi$  be  $\psi \supset \chi$ . Notice that  $\emptyset^\psi = \emptyset$ . Trivially, any assignment in it satisfies  $\chi$ . So  $\emptyset \models \psi \supset \chi$ .

Let  $\varphi$  be of the form  $\mathbf{X} = \mathbf{x} \square \rightarrow \chi$ . Any *do* algorithm, applied to  $\emptyset$ , produces  $\emptyset$  again; therefore, by the inductive hypothesis,  $\emptyset_{\mathbf{X}=\mathbf{x}} = \emptyset \models \chi$ . Thus, by the clause for counterfactuals we obtain  $\emptyset \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$ .  $\square$

We also give names to some relevant fragments of  $\mathcal{CD}$ :

- $\mathcal{C}$ : the *counterfactual fragment*, where symbols  $=(\cdot; \cdot)$  and  $\supset$  are not allowed
- $\mathcal{CO}$ : the *causal-observational fragment* where  $=(\cdot; \cdot)$  is not allowed.

We may also define variants  $\mathcal{CO}^{neg}, \mathcal{C}^{neg}$  that are closed (except for antecedents of counterfactuals) under a dual negation  $\neg$ , which satisfies the additional semantical clause:

- $T \models \neg\psi \iff$  for all  $s \in T^-$ ,  $\{s\} \not\models \psi$ .

We can show that the language  $\mathcal{CO}^{neg}$  (and thus also  $\mathcal{CO}, \mathcal{C}^{neg}, \mathcal{C}$ ) satisfies a much more restrictive condition: flatness.

**Theorem 3.15.** *The logic  $\mathcal{CO}^{neg}$  over recursive causal teams is flat, that is: for every formula  $\varphi$  of  $\mathcal{CO}^{neg}$  and every recursive causal team  $T$ ,  $T \models \varphi$  iff  $\{s\} \models \varphi$  for every assignment  $s \in T^-$ .*

*Proof.* It can be shown by induction on the complexity of the involved formulas. The cases related to propositional connectives are routine. We only have to check the cases of  $\varphi = \mathbf{X} = \mathbf{x} \square \rightarrow \chi$  and  $\varphi = \psi \supset \chi$ .

Case  $\square \rightarrow$ ) Suppose first that  $T \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$ , with  $\chi$  flat. Then by the semantical clause for counterfactuals,  $T_{\mathbf{X}=\mathbf{x}} \models \chi$ . By flatness of  $\chi$ ,  $\{t\} \models \chi$  for each  $t \in (T_{\mathbf{X}=\mathbf{x}})^-$ . Now, notice that, for every  $s \in T^-$ ,  $\{s\}_{\mathbf{X}=\mathbf{x}}$  is again a singleton causal team (see the details of the *do* algorithm). Therefore, by the remarks above,  $\{s\}_{\mathbf{X}=\mathbf{x}} \models \chi$  for each  $s \in T$ ; so, by the semantical clauses,  $\{s\} \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$  for every  $s \in T$ .

In the other direction, suppose  $\{s\} \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$  for every  $s \in T^-$ . Then for every  $s \in T^-$ ,  $\{s\}_{\mathbf{X}=\mathbf{x}} \models \chi$ . Again, any such causal team is a singleton set, say  $\{s\}_{\mathbf{X}=\mathbf{x}} =: \{t_s\}$ . But notice that, if  $t$  is any assignment in  $(T_{\mathbf{X}=\mathbf{x}})^-$ , then  $t = t_s$  for some  $s \in T^-$ . So we can apply the inductive hypothesis and conclude that  $T_{\mathbf{X}=\mathbf{x}} \models \chi$ . Therefore  $T \models \mathbf{X} = \mathbf{x} \square \rightarrow \chi$ .

Case  $\supset$ ) Suppose  $T \models \psi \supset \chi$ , with  $\psi, \chi$  flat. Then, by the semantical clause for  $\supset$ ,  $T^\psi \models \chi$ . By induction hypothesis,  $\{s\}^\psi = \{s\} \models \chi$  for each  $s \in (T^\psi)^-$ . So  $\{s\} \models \psi \supset \chi$ . Suppose instead  $s' \in T$  is such that  $\{s'\} \not\models \psi$ . Then  $\{s'\}^\psi = \emptyset$ . Now  $\{s'\}^\psi = \emptyset \models \chi$ ; so, by the clause for  $\supset$ ,  $\{s'\} \models \psi \supset \chi$ . In conclusion, for any  $s \in T^-$ ,  $\{s\} \models \psi \supset \chi$ .

Viceversa, suppose  $\{s\} \models \psi \supset \chi$  for all  $s \in T^-$ . Let  $s \in T^-$  be such that  $\{s\} \models \psi$ . Then  $\{s\}^\psi = \{s\}$ ; our assumption then yields  $\{s\} \models \chi$ . Since  $\chi$  is flat, we can conclude that  $T^\psi \models \chi$ . So  $T \models \psi \supset \chi$ .  $\square$

This flatness theorem uses in an essential way the recursiveness of the causal team. The previous downward closure and empty set property theorems may work as well in the nonrecursive case, provided a definition of intervention is given also in that case, along the lines sketched at the end of subsection 3.8.

Notice also that, unlike in similar results for team semantics (see e.g. [32]), the truth of a flat formula in a causal team based on a singleton team uses, through structural equations, information that goes beyond that which is encoded in the single assignment of the team.

## 4 Some logical principles

In this section, we analyze some validities and inference rules for the languages  $\mathcal{CD}$ ,  $\mathcal{CO}$  and  $\mathcal{C}$  when their semantics ranges over *parametric, recursive* causal teams (we will assume this restriction throughout the section). Most observations extend also to unique solution causal teams.

We shall write  $\psi \models \chi$  to mean that  $\chi$  holds on any parametric, recursive causal team on which  $\psi$  holds. By  $\models \chi$  we mean that  $\chi$  holds on any parametric, recursive causal team.

### 4.1 The law of excluded middle

Does the fact that the logic  $\mathcal{CO}$  is flat entail that its propositional logic is classical? The issue is somewhat problematic.

It is well known that in predicate logics of dependence and independence the law of excluded middle fails; in the propositional case this point is more subtle. We propose two formulations of the law of excluded middle, a weak one and a stronger one that comes closer to the classical version:

$$\begin{aligned} (WEM) : \quad & \models \psi \vee \neg\psi \\ (SEM) : \quad & \text{For any team } T, T \models \psi \text{ or } T \models \neg\psi \end{aligned}$$

Neither formulation covers the full classical scheme, because for many formulas we lack negation. Within  $\mathcal{CO}^{neg}$ , instead, SEM fundamentally corresponds to the classical law of excluded middle.

It is very easy to find a counterexample to SEM, even at the atomic level. Notice indeed that the following team

X
1
2

does not satisfy  $X = 1$  nor its negation  $X \neq 1$ .

With regard to WEM, we get:

**Theorem 4.1.** *All formulas in  $\mathcal{CO}^{neg}$  satisfy WEM.*

*Proof.* Let  $\varphi$  be a formula of  $\mathcal{CO}^{neg}$ . Let  $T$  be a team. Let  $T_1 = \{s \in T^- \mid \{s\} \models \varphi\}$ , and  $T_2 = \{s \in T^- \mid \{s\} \not\models \varphi\} = \{s \in T^- \mid \{s\} \models \neg\varphi\}$ . Then, by flatness,  $T_1 \models \varphi$  and  $T_2 \models \neg\varphi$ . By the clause for disjunction,  $T \models \varphi \vee \neg\varphi$ .  $\square$

For trivial reasons (since it lacks negation of all but atomic formulas),  $\mathcal{CD}$  itself satisfies WEM. It is not obvious how to extend  $\mathcal{CD}$  itself with a negation. One way followed in the literature is to treat the negation of a dependence atom as satisfied only by the empty team. In this case, one immediately finds counterexamples to WEM.

Another possibility would be to extend  $\mathcal{CD}$  with a contradictory negation of the dependence atom:

- $T \models \neq(\mathbf{X}; Y) \iff$  there are  $s, s' \in T$  such that  $s(\mathbf{X}) = s'(\mathbf{X})$  but  $s(Y) \neq s'(Y)$ .

This option does not save WEM, either: just let  $T$  be any team that satisfies the atom  $=(\mathbf{X}; Y)$ . Then, in any partition  $T_1, T_2$  of  $T$ , both parts satisfy the atom, and not its negation (also in case one of the two subteams is the empty subteam). Notice also the failure of the empty set property.

## 4.2 Conditional excluded middle

In the present section we consider the law that is commonly called *conditional excluded middle*:

$$(CEM) : \quad \models \psi \Box \rightarrow (\chi \vee \neg\chi).$$

This is often considered in pair with the formally similar law called *distribution*:

$$(D) : \quad \psi \Box \rightarrow (\chi \vee \chi') \models [(\psi \Box \rightarrow \chi) \vee (\psi \Box \rightarrow \chi')]$$

These principles are known to hold for Stalnaker counterfactuals and fail for Lewis counterfactuals. We will show in a later subsection (4.6) that distribution is valid for our languages. Here we focus mainly on the CEM law.

What about systems of interventionist counterfactuals, such as those given by Galles and Pearl ([8]), Halpern ([11]) and Briggs ([2])? In the system considered by Galles and Pearl, only atomic consequents are allowed; therefore CEM and D have no instances. For what regards the remaining systems, these fundamentally evaluate formulas on single assignments, that we might think as singleton causal teams. Each intervention produces again an assignment/singleton team. Therefore, CEM trivially holds for atomic consequents. Simple induction arguments allow to extend this observation to all the formulas of the Halpern language. Briggs also allows disjunctions and negations in the antecedents of counterfactuals; we do not consider this possibility in the paper. The languages of these papers also have an operator corresponding to classical implication; therefore, D can be internalized and shown to hold in there, as is done explicitly in [11].

As far as causal teams are concerned, by the flatness of  $\mathcal{CO}^{neg}$ , CEM holds in  $\mathcal{CO}^{neg}$ , and therefore also in  $\mathcal{CD}$ . However, it is natural to wonder about a corresponding semantical property:

(CEM') : For every causal team  $T$ ,  $T \models \theta \Box \rightarrow \chi$  or  $T \models \theta \Box \rightarrow \neg\chi$

We show that CEM' can fail in very simple examples on causal teams; therefore, our logic does *not* coincide with that of usual interventionist counterfactuals. The simplest possible counterexample is of the form:

$$T: \begin{array}{|c|c|} \hline X & Y \\ \hline 1 & 1 \\ \hline 1 & 2 \\ \hline \end{array}$$

The intervention  $do(X = 1)$  does not modify this team, that is:  $T_{X=1} = T$ . On the other hand,  $T \not\models Y = 1$ , and  $T \not\models Y \neq 1$ . So,  $T \not\models X = 1 \Box \rightarrow Y = 1$ , and  $T \not\models X = 1 \Box \rightarrow Y \neq 1$ . Thus, CEM' is falsified.

The peculiarity of this example might suggest that the failure of these laws could be attributed to the fact that there are no arrows between the variables of concern. This is not the case, as we can show with a slightly more complex example. Consider the team

$$S: \begin{array}{|c|c|c|} \hline X & Z \rightarrow Y & \\ \hline 1 & 1 & 2 \\ \hline 2 & 2 & 4 \\ \hline \end{array}$$

where the behavior of  $Y$  is determined by the function  $\mathcal{F}_S(Y)(X, Z) := X + Z$ . The intervention  $do(X = 1)$  generates the team

$$S_{X=1}: \begin{array}{|c|c|c|} \hline X & Z \rightarrow Y & \\ \hline 1 & 1 & 2 \\ \hline 1 & 2 & 3 \\ \hline \end{array}$$

It is then immediate to verify that  $S$  neither satisfies  $X = 1 \Box \rightarrow Y = 2$  nor  $X = 1 \Box \rightarrow Y \neq 2$ , thereby falsifying CEM'.

### 4.3 Interactions between implications and dependencies

As causal team semantics defines both dependencies based on correlations (functional dependence) and causal and counterfactual dependencies, it allows us to investigate the logical principles which govern their interaction. Here we list some of the inference rules that connect the various kinds of dependence that are referred to within the language  $\mathcal{CD}$ ; the simple soundness proofs are left to the reader.

Given that the selective implication is strongly analogous to classical implication, the following inference rules hold (assuming the classicality of all antecedents):

$$\frac{\theta \supset \chi}{(\theta \wedge \psi) \supset \chi} \quad \frac{\theta \supset \chi}{\theta \supset (\chi \vee \psi)} \quad \frac{\theta \supset (\chi \supset \psi)}{(\theta \wedge \chi) \supset \psi}$$

the third of which is also invertible. The second rule can be expected to fail in logics that lack the empty set property.

It is time to make the relationship between selective and classical implication more explicit. First of all, notice that a restricted version of a semantic Deduction Theorem holds for selective implication: if any team that satisfies a *classical* formula  $\theta$  also satisfies  $\chi$ , then this in particular holds for any team of the form  $T^\theta$ ; therefore  $T^\theta \models \chi$ , from which we obtain  $T \models \theta \supset \chi$  (for any causal team  $T$ ). Therefore, all the valid inference rules that are stated in this section can be internalized in the formal language as selective implications, *provided the antecedents are classical formulas*<sup>7</sup>.

Secondly, we wish to establish whether our approach, of allowing selective implications, is equivalent or not to the approach of adding to our languages the dual negations of all classical formulas; could then a selective implication  $\theta \supset \chi$  be replaced by an expression of the form  $\neg\theta \vee \chi$ ? The answer is *yes for downward closed logics*, as those that we have considered thus far; but the approach via selective implication is more general. The point is this: applying the clause for disjunction,  $T \models \neg\theta \vee \chi$  holds if and only if there are  $S_1, S_2 \subseteq T^-$  such that  $S_2 \models \chi$  and, for all  $s \in S_1$ ,  $\{s\} \not\models \theta$ . Then  $(T^\theta)^- \subseteq S_2$ ; if the logic under consideration is downward closed, then, we can conclude that  $T^\theta \models \chi$ , that is,  $T \models \theta \supset \chi$ . But it is easy to find counterexamples in logics that are not downward closed. Consider for example the (marginal) *independence atoms*  $X \perp Y$  defined (along the lines of [10]) by the semantical clause:

- $T \models X \perp Y \iff$  for every  $s, s' \in T^-$  there is an  $s'' \in T^-$  s.t.  $s''(X) = s(X)$  and  $s''(Y) = s'(Y)$ .

Then the following team

$$U: \begin{array}{|c|c|} \hline X & Y \\ \hline 1 & 1 \\ \hline 1 & 2 \\ \hline 2 & 1 \\ \hline 2 & 2 \\ \hline \end{array}$$

satisfies  $X \neq 1 \vee X \perp Y$  (just consider the partition of  $U$  into  $\emptyset$  and  $U$ ;  $\emptyset \models X \neq 1$  and  $U \models X \perp Y$ ). But  $U \not\models X = 1 \supset X \perp Y$ .

<sup>7</sup>We might consider extending our languages by allowing *flat* instead of classical antecedents for selective implications. This would extend the range of application of the Deduction Theorem, but would have the drawback that the syntax of the resulting language might not be decidable anymore.

We now turn to the relationship between selective and counterfactual implication. We start by showing that their commutativity, that is, the equivalence between  $\theta \square \rightarrow (\chi \supset \psi)$  and  $\chi \supset (\theta \square \rightarrow \psi)$  fails in causal teams in both directions. We show this by a couple of counter-examples. For all the teams involved, the graph will be  $\{\{X, Y, Z\}, \{(X, Y), (Z, Y)\}\}$  (so, we do not draw it every time) and the only structural equation will be  $Y := X + Z$ ; it is not really important to specify the ranges of variables. Consider the following teams:

T:	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>1</td><td>2</td><td>3</td></tr><tr><td>2</td><td>3</td><td>5</td></tr></table>	X	Z	Y	1	1	2	1	2	3	2	3	5
X	Z	Y											
1	1	2											
1	2	3											
2	3	5											

$T_{X=1}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>1</td><td>2</td><td>3</td></tr><tr><td>1</td><td>3</td><td>4</td></tr></table>	X	Z	Y	1	1	2	1	2	3	1	3	4
X	Z	Y											
1	1	2											
1	2	3											
1	3	4											

$(T_{X=1})^{X=1}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>1</td><td>2</td><td>3</td></tr><tr><td>1</td><td>3</td><td>4</td></tr></table>	X	Z	Y	1	1	2	1	2	3	1	3	4
X	Z	Y											
1	1	2											
1	2	3											
1	3	4											

$T^{X=1}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>1</td><td>2</td><td>3</td></tr></table>	X	Z	Y	1	1	2	1	2	3
X	Z	Y								
1	1	2								
1	2	3								

$(T^{X=1})_{X=1}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>1</td><td>2</td><td>3</td></tr></table>	X	Z	Y	1	1	2	1	2	3
X	Z	Y								
1	1	2								
1	2	3								

From the tables, it is immediate to see that  $T \models X = 1 \supset (X = 1 \square \rightarrow (Y = 2 \vee Y = 3))$ , but  $T \not\models X = 1 \square \rightarrow (X = 1 \supset (Y = 2 \vee Y = 3))$ .

A counterexample for the opposite direction is given by the following teams, with graph and equation as before:

S:	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>1</td><td>2</td><td>3</td></tr><tr><td>2</td><td>1</td><td>3</td></tr></table>	X	Z	Y	1	1	2	1	2	3	2	1	3
X	Z	Y											
1	1	2											
1	2	3											
2	1	3											

$S_{Z=1}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>2</td><td>1</td><td>3</td></tr></table>	X	Z	Y	1	1	2	2	1	3
X	Z	Y								
1	1	2								
2	1	3								

$(S_{Z=1})^{Y=3}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>2</td><td>1</td><td>3</td></tr></table>	X	Z	Y	2	1	3
X	Z	Y					
2	1	3					

$S^{Y=3}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>2</td><td>3</td></tr><tr><td>2</td><td>1</td><td>3</td></tr></table>	X	Z	Y	1	2	3	2	1	3
X	Z	Y								
1	2	3								
2	1	3								

$(S^{Y=3})_{Z=1}$ :	<table border="1" style="display: inline-table;"><tr><th>X</th><th>Z</th><th>Y</th></tr><tr><td>1</td><td>1</td><td>2</td></tr><tr><td>2</td><td>1</td><td>3</td></tr></table>	X	Z	Y	1	1	2	2	1	3
X	Z	Y								
1	1	2								
2	1	3								

which show that  $S \models Z = 1 \square \rightarrow (Y = 3 \supset Y = 3)$  but  $S \not\models Y = 3 \supset (Z = 1 \square \rightarrow Y = 3)$ .

The noncommutativity of  $\square \rightarrow$  with  $\supset$  is a phenomenon that, to the best of our knowledge, has no analogous in the usual notations of the causal inference literature; we return to this point in Section 6.4.

In our search for further rules linking  $\square \rightarrow$  and  $\supset$ , we can now have a look at the literature on Stalnaker-Lewis counterfactuals, and check whether in some case  $\supset$  may take the role that in Stalnaker-Lewis rules is taken by material implication. For example, it is known that 1) a Stalnaker-Lewis counterfactual  $\theta \square \rightarrow \chi$  implies the corresponding material conditional  $\theta \rightarrow \chi$ , and 2) the converse holds, *provided  $\theta$  holds*. It might be interesting to inquire whether similar principles hold for our counterfactual and selective implication, respectively. The literature on causation seems to suggest a negative answer: looking for example at rule 2 of Pearl's *do calculus* ([23], sec. 3.4) we see that, in some languages that also allow the discussion of probabilities, replacing an intervention with an observation, or viceversa, can be done only under some graph-theoretical

(and not logical) assumptions. Explicit counterexamples to 1) and 2) can be constructed in our languages, following the lines of [2] (sect. 3.1 and 4), for some formulas in which the consequent  $\chi$  is itself a counterfactual. The counterexample in [2] is based on the famous “execution line” example ([23]), but we can construct a much simpler one. Consider the following causal teams, with boolean ranges and invariant functions  $f_Y(X) := X$  and  $f_Z(X, Y) := X \wedge Y$ :

$$\begin{array}{cc}
T = T^{Y=0}: & \begin{array}{|c|c|c|} \hline \text{X} \curvearrowright \text{Y} \rightarrow \text{Z} \\ \hline 0 & 0 & 0 \\ \hline \end{array} & (T^{Y=0})_{X=1}: & \begin{array}{|c|c|c|} \hline \text{X} \curvearrowright \text{Y} \rightarrow \text{Z} \\ \hline 1 & 1 & 1 \\ \hline \end{array} \\
\\
T_{Y=0}: & \begin{array}{|c|c|c|} \hline \text{X} \quad \text{Y} \rightarrow \text{Z} \\ \hline 0 & 0 & 0 \\ \hline \end{array} & (T_{Y=0})_{X=1}: & \begin{array}{|c|c|c|} \hline \text{X} \quad \text{Y} \rightarrow \text{Z} \\ \hline 1 & 0 & 0 \\ \hline \end{array}
\end{array}$$

It is then immediate to see that  $T \models Y = 0 \square \rightarrow (X = 1 \square \rightarrow Z \neq 1)$  but  $T \not\models Y = 0 \supset (X = 1 \square \rightarrow Z \neq 1)$ . The important element in this example (and similar ones) is that we have an intervention  $do(Y = 0)$  which does not modify the team component of the causal team, but modifies the graph – which encodes counterfactual relations.

Other simple examples could be constructed for consequents which contain probabilistic statements – we will define a semantics for them in later sections. We could not find, instead, a counterexample with consequents that involve only selective implications, dependence atoms and connectives; but here is a counterexample using selective implications and *contradictory negations* of dependence atoms, as defined in subsection 4.1. Consider the following causal teams, with invariant function  $f_Z(X, Y) := Y$ :

$$\begin{array}{c}
S: \begin{array}{|c|c|c|} \hline \text{X} \quad \text{Y} \rightarrow \text{Z} \\ \hline 0 & 0 & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \\
\\
(S_{X=0}) = (S_{X=0})^{X=0}: \begin{array}{|c|c|c|} \hline \text{X} \quad \text{Y} \rightarrow \text{Z} \\ \hline 0 & 0 & 0 \\ \hline 0 & 1 & 1 \\ \hline \end{array} & S^{X=0} = (S^{X=0})^{X=0}: \begin{array}{|c|c|c|} \hline \text{X} \quad \text{Y} \rightarrow \text{Z} \\ \hline 0 & 0 & 0 \\ \hline \end{array}
\end{array}$$

Then  $S \models X = 0 \square \rightarrow (X = 0 \supset \neq (X; Z))$  but  $S \not\models X = 0 \supset (X = 0 \supset \neq (X; Z))$ . Of course, it is questionable whether we should allow this kind of arrow  $X \rightarrow Z$  in our definition of causal teams.

Let us now see how dependence atoms interact with the two kinds of implication. For what regards selective implication, the following rules are valid:

$$\frac{=(\mathbf{X}; Y)}{\mathbf{X} = \mathbf{x} \supset =(Y)} \quad \frac{\bigvee_{y \in \text{Ran}(Y)} (\mathbf{X} = \mathbf{x} \supset Y = y)}{=(\mathbf{X}; Y)}.$$

As a simple special case of the second rule, we have that  $Y = y$  implies  $=(Y)$ . An analogous rule holds for counterfactual implication:

$$\frac{\bigwedge_{\mathbf{x} \in \text{Ran}(\mathbf{X})} \bigvee_{y \in \text{Ran}(Y)} (\mathbf{X} = \mathbf{x} \square \rightarrow Y = y)}{=(\mathbf{X}; Y)}.$$

but the implication from  $=(\mathbf{X}; Y)$  to  $\mathbf{X} = \mathbf{x} \square \rightarrow =(Y)$  only holds on teams where the tuple  $\mathbf{x}$  occurs.

#### 4.4 Permutation and exportation/importation of antecedents

Consider the following laws of *permutation*, *exportation* and *importation* of antecedents:

$$\begin{aligned} (P) : \psi \square \rightarrow (\psi' \square \rightarrow \chi) &\equiv \psi' \square \rightarrow (\psi \square \rightarrow \chi) \\ (E) : (\psi \wedge \psi') \square \rightarrow \chi &\models \psi \square \rightarrow (\psi' \square \rightarrow \chi) \\ (I) : \psi \square \rightarrow (\psi' \square \rightarrow \chi) &\models (\psi \wedge \psi') \square \rightarrow \chi \end{aligned}$$

These laws famously fail for Stalnaker's ([28]) and Lewis' ([18]) counterfactuals; see [25], chap. 8; they are also claimed to fail, in general, for natural language counterfactuals. The purpose of this section is to show that under some restriction (that the variables mentioned in  $\psi$  and, resp.  $\psi'$  be distinct) these laws are valid for our counterfactuals. Under this restriction (which will be somewhat relaxed in the next section), exportation/importation amounts to the statement that an intervention over a set of variables can be split into successive interventions over disjoint subsets of variables; permutation corresponds to the assertion the two interventions of this kind can be performed in any order.

**Remark 4.2.** *We assume in this subsection that causal teams are explicit. However, the results presented here hold for non-explicit causal teams as well, although the proofs in that case are more involved.*

**Lemma 4.3.** *Let  $G$  be a finite acyclic graph,  $\mathbf{X}, Y, W$  vertices of  $G$  with  $W \in PA_Y$ ,  $Y \notin \mathbf{X}$  and  $Y$  reachable from  $\mathbf{X}$  (that is,  $d(\mathbf{X}, Y) \geq 0$ ). Then  $d(\mathbf{X}, W) < d(\mathbf{X}, Y)$ .*

*Proof.*  $d(\mathbf{X}, Y) = \sup\{\text{lenght}(P) \mid P \text{ directed path of } G_{-\mathbf{X}} \text{ from some } X \in \mathbf{X} \text{ to } Y\}$ . But any directed path from  $\mathbf{X}$  to  $W$  is included in some path from this set. So  $d(\mathbf{X}, W) \leq d(\mathbf{X}, Y)$ .

Assume first that at least one path from  $\mathbf{X}$  to  $W$  exists in  $G_{-\mathbf{X}}$ . Since  $G$  is finite and acyclic,  $d(\mathbf{X}, W) = n \in \mathbb{N}$ . Let  $P$  be a path of  $G_{-\mathbf{X}}$  from some  $X \in \mathbf{X}$  to  $W$  of lenght  $n$ . Then  $P \cup \{(W, Y)\}$  is a path of  $G_{-\mathbf{X}}$  from  $X$  to  $Y$  of lenght  $n + 1$ . This holds for any  $X \in \mathbf{X}$ . Since  $\mathbf{X}$  is finite,  $d(\mathbf{X}, Y) = \max_{X \in \mathbf{X}} d(X, Y) > \max_{X \in \mathbf{X}} d(X, W) = d(\mathbf{X}, W)$ .

If instead no path from  $\mathbf{X}$  to  $W$  exists, then  $d(\mathbf{X}, W) = -1$ , and the claim of the Lemma is satisfied.  $\square$

In the next theorem exportation and importation of antecedents are expressed as rules. A more general version will be proved later on (Theorem 4.8).



**Theorem 4.4.** *Let  $T$  be a causal team with  $G(T)$  finite acyclic,  $\mathbf{X}, \mathbf{Y} \in \text{dom}(T)$  such that  $\mathbf{X} \cap \mathbf{Y} = \emptyset$ ,  $\mathbf{x} \in \text{Ran}(\mathbf{X})$ , and  $\mathbf{y} \in \text{Ran}(\mathbf{Y})$ . Then  $T_{\mathbf{X}=\mathbf{x}\mathbf{Y}=\mathbf{y}} = (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}}$ . Therefore, the following rules*

$$(IMP) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow (\mathbf{Y} = \mathbf{y} \square \rightarrow \chi)}{(\mathbf{X} = \mathbf{x} \wedge \mathbf{Y} = \mathbf{y}) \square \rightarrow \chi} \quad (EXP) : \frac{(\mathbf{X} = \mathbf{x} \wedge \mathbf{Y} = \mathbf{y}) \square \rightarrow \chi}{\mathbf{X} = \mathbf{x} \square \rightarrow (\mathbf{Y} = \mathbf{y} \square \rightarrow \chi)}$$

are valid, under the above restrictions, over recursive teams.

*Proof.* Notice first that for any  $s' \in (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}}$  and any  $s'' \in T_{\mathbf{X}=\mathbf{x}\mathbf{Y}=\mathbf{y}}$ ,  $s'(\mathbf{X}) = s''(\mathbf{X}) = \mathbf{x}$  and  $s'(\mathbf{Y}) = s''(\mathbf{Y}) = \mathbf{y}$ .

Secondly, the restriction to explicit causal teams implies that  $\mathcal{F}_T \supseteq \mathcal{F}_{T_{\mathbf{X}=\mathbf{x}}} \supseteq \mathcal{F}_{(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}}} = \mathcal{F}_{T_{\mathbf{X}=\mathbf{x}\mathbf{Y}=\mathbf{y}}}$ .

Fix  $s \in T$ . Let  $s' \in (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}}$  be the assignment obtained by applying to  $\{s\}$  the intervention  $do(\mathbf{X} = \mathbf{x})$  followed by  $do(\mathbf{Y} = \mathbf{y})$ , and  $s''$  the assignment obtained by applying  $do(\mathbf{X} = \mathbf{x} \wedge \mathbf{Y} = \mathbf{y})$  to  $s$ . We shall prove, by a simultaneous induction on  $m$  and  $n$ , that  $s'(Z) = s''(Z)$  for any variable  $Z$  such that  $d(\mathbf{X}, Z) \leq m$  and  $d(\mathbf{Y}, Z) \leq n$ . In the base case  $Z \in \mathbf{X} \cup \mathbf{Y}$ , this result is obvious.

Suppose that the statement holds for  $m$  and  $n$ , and let  $Z \in \text{dom}(T) \setminus (\mathbf{X} \cup \mathbf{Y})$  be a variable such that  $d(\mathbf{X}, Z) \leq m$  and  $d(\mathbf{Y}, Z) = n + 1$  (the proof of the symmetric case is analogous).

From the algorithm  $do(\mathbf{Y} = \mathbf{y})$ , it follows that  $s'(Z) = \mathcal{F}_{T_{\mathbf{X}=\mathbf{x}}}(Z)(s'(PA_Z))$  (notice indeed that, thanks to acyclicity, the algorithm does not modify the  $PA_Z$ -columns after modifying the  $Z$ -column), which, by the observation above, is equal to  $\mathcal{F}_T(Z)(s'(PA_Z))$ ; and from the algorithm  $do(\mathbf{X} = \mathbf{x} \wedge \mathbf{Y} = \mathbf{y})$ , it follows that  $s''(Z) = \mathcal{F}_T(Z)(s''(PA_Z))$ . By Lemma 4.3 (which we can apply because  $d(\mathbf{Y}, Z) = n + 1$  implies  $Z \notin \mathbf{Y}$ ), the variables in  $PA_Z$  have strictly smaller distances from  $\mathbf{Y}$  than  $Z$  (and, obviously, at most the same distance from  $\mathbf{X}$ ). Then, by the inductive hypothesis,  $s'(PA_Z) = s''(PA_Z)$ . So, applying  $\mathcal{F}_T(Z)$ , we obtain  $s'(Z) = s''(Z)$ .  $\square$

As a corollary, we obtain the following variant of the permutation rule:

**Theorem 4.5.** *Let  $T$  be a causal team with  $G(T)$  finite acyclic,  $\mathbf{X}, \mathbf{Y} \in \text{dom}(T)$  such that  $\mathbf{X} \cap \mathbf{Y} = \emptyset$ ,  $\mathbf{x} \in \text{ran}(\mathbf{X})$ , and  $\mathbf{y} \in \text{ran}(\mathbf{Y})$ . Then  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (T_{\mathbf{Y}=\mathbf{y}})_{\mathbf{X}=\mathbf{x}}$ . Therefore, the following rule*

$$(PERM) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow (\mathbf{Y} = \mathbf{y} \square \rightarrow \chi)}{\mathbf{Y} = \mathbf{y} \square \rightarrow (\mathbf{X} = \mathbf{x} \square \rightarrow \chi)}$$

is valid, under the above restrictions, over recursive causal teams.

*Proof.* From Theorem 4.4 we obtain the equalities  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = T_{\mathbf{X}=\mathbf{x}\mathbf{Y}=\mathbf{y}}$  and  $(T_{\mathbf{Y}=\mathbf{y}})_{\mathbf{X}=\mathbf{x}} = T_{\mathbf{Y}=\mathbf{y}\mathbf{X}=\mathbf{x}}$ . Since the order of variables is irrelevant in the definition of the  $do$  algorithm, we also have  $T_{\mathbf{X}=\mathbf{x}\mathbf{Y}=\mathbf{y}} = T_{\mathbf{Y}=\mathbf{y}\mathbf{X}=\mathbf{x}}$ . Transitivity yields the desired result.  $\square$

## 4.5 Generalized import-export and permutation rules

The purpose of this section is to show that the assumptions of the permutation and the import-export rules can be relaxed in the following way: if the interventions involved are over two sets of variables  $\mathbf{X}$  and  $\mathbf{Y}$  with nonempty intersection, we have only to require that the two interventions act “in the same way” over the common set of variables  $\mathbf{X} \cap \mathbf{Y}$ .

**Lemma 4.6.** *For any  $\mathbf{X} \subseteq \text{dom}(T)$ ,  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{X}=\mathbf{x}} = T_{\mathbf{X}=\mathbf{x}}$ .*

*Proof.* By a simple induction on the steps of the “for” cycle in the *do* algorithm.  $\square$

**Lemma 4.7.** *Let  $T$  be a causal team with  $G(T)$  finite acyclic,  $\mathbf{X}, \mathbf{Y} \in \text{dom}(T)$ ,  $\mathbf{x} \in \text{Ran}(\mathbf{X})$ , and  $\mathbf{y} \in \text{Ran}(\mathbf{Y})$ . Suppose also that, if  $X_i = Y_j \in \mathbf{X} \cap \mathbf{Y}$ , then  $x_i = y_j$ . Let  $\mathbf{Y}' = \mathbf{Y} \setminus \mathbf{X}$ , and  $\mathbf{y}' = \{y_j | Y_j \in \mathbf{Y}'\}$ . Then  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}'=\mathbf{y}'}$ .*

*Proof.* Let  $\mathbf{X}' = \mathbf{X} \setminus \mathbf{Y}$ , and  $\mathbf{x}' = \{x_j | X_j \in \mathbf{X}'\}$ . Let  $\mathbf{Z} = \mathbf{X} \cap \mathbf{Y}$  and  $\mathbf{z} = \mathbf{x} \cap \mathbf{y}$ . Applying twice Theorem 4.4 (import-export), we get  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (((T_{\mathbf{X}'=\mathbf{x}'}))_{\mathbf{Z}=\mathbf{z}})_{\mathbf{Y}'=\mathbf{y}'}$ . But this last term is equal to  $((T_{\mathbf{X}'=\mathbf{x}'})_{\mathbf{Z}=\mathbf{z}})_{\mathbf{Y}'=\mathbf{y}'}$ , by Lemma 4.6. Given that  $\mathbf{Z}$  and  $\mathbf{X}'$  are disjoint, we can apply again Theorem 4.4 to obtain  $((T_{\mathbf{X}'=\mathbf{x}'}))_{\mathbf{Z}=\mathbf{z}})_{\mathbf{Y}'=\mathbf{y}'} = (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}'=\mathbf{y}'}$ . The chain of equalities yields the desired result.  $\square$

Next we show that the import-export rule works under the weaker hypothesis that the two interventions act in the same way on the shared set of variables.

**Theorem 4.8.** *Let  $T$  be a causal team with  $G(T)$  finite acyclic,  $\mathbf{X}, \mathbf{Y} \in \text{dom}(T)$ ,  $\mathbf{x} \in \text{ran}(\mathbf{X})$ , and  $\mathbf{y} \in \text{ran}(\mathbf{Y})$ . Suppose also that, if  $X_i = Y_j \in \mathbf{X} \cap \mathbf{Y}$ , then  $x_i = y_j$ . Then  $T_{\mathbf{X}=\mathbf{x} \wedge \mathbf{Y}=\mathbf{y}} = (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}}$ .*

*Proof.* Apply Lemma 4.7 to see that  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}'=\mathbf{y}'}$ , where  $\mathbf{Y}' = \mathbf{Y} \setminus \mathbf{X}$ . Given that  $\mathbf{X}$  and  $\mathbf{Y}'$  are disjoint sets of variables, we can use the weak import-export rule (Theorem 4.4) to obtain  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}'=\mathbf{y}'} = T_{\mathbf{X}=\mathbf{x} \wedge \mathbf{Y}'=\mathbf{y}'}$ . But now the formula  $\mathbf{X} = \mathbf{x} \wedge \mathbf{Y} = \mathbf{y}$  differs from  $\mathbf{X} = \mathbf{x} \wedge \mathbf{Y}' = \mathbf{y}'$  only in that it contains some repetitions of conjuncts that are already present in  $\mathbf{X} = \mathbf{x} \wedge \mathbf{Y}' = \mathbf{y}'$ . Therefore, the two formulas define the same interventions (as can be seen by checking the *do* algorithm); this immediately implies that  $T_{\mathbf{X}=\mathbf{x} \wedge \mathbf{Y}=\mathbf{y}} = T_{\mathbf{X}=\mathbf{x} \wedge \mathbf{Y}'=\mathbf{y}'}$ . The whole chain of equalities yields the result.  $\square$

Again, under these hypotheses the order of the interventions is irrelevant.

**Theorem 4.9.** *Let  $T$  be a causal team with  $G(T)$  finite acyclic,  $\mathbf{X}, \mathbf{Y} \in \text{dom}(T)$ ,  $\mathbf{x} \in \text{ran}(\mathbf{X})$ , and  $\mathbf{y} \in \text{ran}(\mathbf{Y})$ . Suppose also that, if  $X_i = Y_j \in \mathbf{X} \cap \mathbf{Y}$ , then  $x_i = y_j$ . Then  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (T_{\mathbf{Y}=\mathbf{y}})_{\mathbf{X}=\mathbf{x}}$ .*

*Proof.* This works like the proof of Theorem 4.5, using the relaxed import-export rule (Theorem 4.8) instead of Theorem 4.4.  $\square$

## 4.6 Composition and effectiveness rules

Galles and Pearl ([8]) present a complete and sound system for interventionist counterfactuals over recursive causal models (to be more precise, the result holds only for causal models that are parametric and under the restriction that the range of each variable is *finite*). This system is based on two main rules called *composition* and *effectiveness*. In informal presentations ([23]), Pearl claims that these two axioms exhaust the system; a closer look at the original paper (and at a subsequent clarifying paper by Halpern, [11]) shows that two further axioms (*definiteness* and *uniqueness*), plus “the rules of logic” are part of the system. Let us first point out that Pearl uses a very different syntax from ours which originates in a tradition of representation of counterfactuals common in the statistical literature (the “potential outcome” approach, [22],[24]). He uses

$$Y_x(u) = y$$

to mean that, in an actual state in which the (set of) exogenous variables  $U$  take values  $u$ , if we fixed the value of  $X$  to  $x$  by an intervention, then  $Y$  would take value  $y$ . Given that in a recursive structural equation model the endogenous variables are functionally determined by the exogenous ones, we can represent such an assignment  $u$  of values to the exogenous variables as a unique assignment  $s_u$ . Therefore, we might represent Pearl’s counterfactual in a notation closer to the logical tradition as:

$$s_u \models X = x \square \rightarrow Y = y.$$

Assuming that, unless otherwise specified, variables and values are implicitly universally quantified, here is now the list of properties that Galles and Pearl ascribe to causal models with unique solutions (which include recursive models):

$$(Composition) : W_x(u) = w \Rightarrow Y_{xw}(u) = Y_x(u)$$

$$(Effectiveness) : X_{xw}(u) = x$$

$$(Definiteness) : \text{there is an } x \text{ such that } X_{\mathbf{y}}(u) = x$$

$$(Uniqueness) : X_{\mathbf{y}}(u) = x \wedge X_{\mathbf{y}}(u) = x' \Rightarrow x = x'$$

to which, according to Halpern ([11]), we should add all classical propositional tautologies plus Modus Ponens. A further axiom scheme (call it REC) forces the system to be recursive (i.e., the graph to be acyclic). Translating the REC axiom into our language is somewhat problematic; we consider this issue towards the end of the present subsection.

Do these axioms hold for our causal teams? First of all, we should decide how these axioms should be interpreted in our framework. One possible way to go is to conceive these counterfactuals as encoding properties of a team instead of an actual state; that is, to have a team  $T$  play the role that was previously played by  $u$  (or by  $s_u$ ). It is then straightforward to translate effectiveness as the statement that each causal team  $T$  satisfies:

$$(EFF) : \models (X = x \wedge \mathbf{W} = \mathbf{w}) \square \rightarrow X = x.$$

Translating composition is a more complex task. First of all, Galles&Pearl have an operator  $\Rightarrow$  which is most likely intended as a material implication (which, by the deduction theorem, reflects logical implication). We do not have in  $\mathcal{CD}$  any such operator (we know that our selective implication  $\supset$  obeys a deduction theorem, but only for  $\mathcal{CO}$  antecedents). For this reason, we will focus here on the languages  $\mathcal{C}^{neg}$  and  $\mathcal{CO}^{neg}$ , which are flat and closed under dual negation. We can therefore add modus ponens in the following form:

$$MP_{\vee} : \frac{\theta \quad \neg\theta \vee \chi}{\chi}.$$

However, in order to facilitate the transition to more general languages like  $\mathcal{CD}$ , which are not closed under negation, we may also adopt the strategy of rewriting all the implications corresponding to Pearl’s axioms as inference rules. For example, it will be reasonable to move from the formal language to the semantic metalanguage, and think of composition as a rule of inference rather than an axiom. The composition rule poses a further problem: the formula  $Y_{xw}(u) = Y_x(u)$  is not a counterfactual. We may think instead of the equivalent expression (for all  $y \in \mathcal{R}(Y)$ ):  $Y_{xw}(u) = y \iff Y_x(u) = y$ . It then becomes natural to decompose “composition” into *two* rules of inference:

$$(CE) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow W = w \quad (\mathbf{X} = \mathbf{x} \wedge W = w) \square \rightarrow Y = y}{\mathbf{X} = \mathbf{x} \square \rightarrow Y = y}$$

$$(CI) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow W = w \quad \mathbf{X} = \mathbf{x} \square \rightarrow Y = y}{(\mathbf{X} = \mathbf{x} \wedge W = w) \square \rightarrow Y = y}$$

Definiteness (“Decidability”) contains an existential quantification. *If we assume the ranges of variables to be finite*, then we can formulate this axiom in a way that works relativized to teams that have a specific finite range for  $X$ :

$$(DEC) : \forall_{x \in \text{Ran}(X)} (\mathbf{Y} = \mathbf{y} \square \rightarrow X = x)$$

What this axiom expresses about a team is that the team can be decomposed into subteams, each of which satisfies, after the intervention  $do(\mathbf{Y} = \mathbf{y})$ , one of the  $X = x$  atoms; it is true of any parametric causal team (in the nonparametric case, the disjunction should not be taken over  $\text{Ran}(X)$ , but over  $\text{Ran}(X) \cup \text{Term}_{G(T)}$ ).

The uniqueness axiom mentions formulas of the form  $x = x'$ , which are absent from our language. But we can think of a reformulation of these axioms due to Halpern (which he simply calls the “equality” axiom scheme), which rephrases it as a scheme of rules of inference (one rule for each variable  $X$ , set of variables  $\mathbf{Y}$ , and pair of values  $x, x' \in \text{Ran}(X)$  with  $x \neq x'$ ):

$$(UNI) : \frac{\mathbf{Y} = \mathbf{y} \square \rightarrow X = x}{\mathbf{Y} = \mathbf{y} \square \rightarrow X \neq x'}$$

The import of this axiom seems much weaker than in the Galles-Pearl case, given that  $X = x$  and  $X \neq x'$  are global statements over many assignments.

Our translations of the axioms and rules all fall within the language  $\mathcal{C}^{neg}$ . By its flatness, it immediately follows that the axioms EFF and the appropriate instances of DEC also hold for our causal teams.

**Theorem 4.10.** *The rules CE, CI and UNI are sound for acyclic causal teams.*

*Proof.* CE) Assume  $T$  satisfies the two assumptions of the rule. Since the consequent  $Y = y$  of the counterfactuals involved is a formula that does not contain counterfactuals, in order to prove that  $T$  satisfies the conclusion of the rule, it is sufficient to show that  $(T_{\mathbf{X}=\mathbf{x}})^- = (T_{\mathbf{X}=\mathbf{x}\wedge W=w})^-$ . In case  $W \in \mathbf{X}$ , the rule holds for trivial reasons. Otherwise, a previous theorem (4.4) tells us that  $T_{\mathbf{X}=\mathbf{x}\wedge W=w} = (T_{\mathbf{X}=\mathbf{x}})_{W=w}$ . The first assumption of CE tells us that  $T_{\mathbf{X}=\mathbf{x}} \models W = w$ . But then, obviously, applying  $do(W = w)$  to  $T_{\mathbf{X}=\mathbf{x}}$  leaves its underlying team unvaried.

CI) An analogous argument.

UNI) Immediate.  $\square$

Finally, notice that, by the flatness of  $\mathcal{C}^{neg}$ , all classical validities that can be formulated in  $\mathcal{C}^{neg}$  are sound for acyclic teams. (But remember that some classical semantic principles, like SEM, fail). Halpern ([11]) considers a language  $L_{uniq}$  that allows boolean formulas as consequents, and boolean combinations of counterfactuals, but does not allow embeded counterfactuals. This language can be identified with an appropriate fragment of  $\mathcal{C}^{neg}$ , call it  $\mathcal{C}_u$ . What Halpern shows is that the axioms described above form a sound and complete system for  $L_{uniq}$  over recursive (i.e. acyclic) causal models. In particular, they axiomatize the set of valid formulas of the Galles-Pearl language (which is restricted to conjunctions of counterfactuals of the form  $\mathbf{X} = \mathbf{x} \square \rightarrow Y = y$ ), although derivations may use formulas that are not part of the Galles-Pearl language.

As we have seen, these axioms and rules are sound also for  $\mathcal{C}^{neg}$ , over recursive causal teams. But unlike  $\mathcal{C}_u$ , the language  $\mathcal{C}^{neg}$  also allows counterfactual conditionals to occur in the consequents. Thus it may be that some more axioms are needed in order to obtain a complete system for  $\mathcal{C}^{neg}$  over recursive causal teams.

Our strategy for the rest of this section is to show that any counterfactual in  $\mathcal{C}^{neg}$  is equivalent to a counterfactual with a "simple" consequent (i.e., a consequent which does not contain embedded counterfactuals, or even atomic formulas of the form  $Y \neq y$ ), and to isolate the rules that allow this transformation.

Our first observation is that Halpern [11] proves that recursive causal models (or, more generally, causal models whose equations have unique solutions) satisfy some rules that allow removing connectives from consequents. In our language, these rules may be stated as follows:

$$(OR - OUT) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow \psi \vee \psi'}{(\mathbf{X} = \mathbf{x} \square \rightarrow \psi) \vee (\mathbf{X} = \mathbf{x} \square \rightarrow \psi')}$$

$$(AND - OUT) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow \psi \wedge \psi'}{(\mathbf{X} = \mathbf{x} \square \rightarrow \psi) \wedge (\mathbf{X} = \mathbf{x} \square \rightarrow \psi')}$$

$$(NEG - OUT) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow \neg \psi}{\neg(\mathbf{X} = \mathbf{x} \square \rightarrow \psi)}$$

Halpern also shows the soundness of their inverses (call them OR-IN, AND-IN, NEG-IN). We check now that the same holds in our causal teams.

**Lemma 4.11.** *Suppose a causal team of the form  $T_{\mathbf{X}=\mathbf{x}}$  has causal subteams  $U', V'$  such that  $(U')^- \cup (V')^- = (T_{\mathbf{X}=\mathbf{x}})^-$ . Then  $T$  has causal subteams  $U, V$  such that  $U_{\mathbf{X}=\mathbf{x}} = U'$  and  $V_{\mathbf{X}=\mathbf{x}} = V'$ .*

*Proof.* Let  $T, U', V'$  as above. For each  $t \in (T_{\mathbf{X}=\mathbf{x}})^-$ , define the set of assignments  $S_t = \{s \in T^- \mid \{s\}_{\mathbf{X}=\mathbf{x}} = \{t\}\}$ . Define then  $U^- := \bigcup_{t \in U'} S_t$  and  $V^- := \bigcup_{t \in V'} S_t$ . Define  $G(U), \mathcal{F}_U$ , etc. in the obvious way.

Let us show that  $U, V$  behave as wanted. Let  $t \in (U')^-$ . Pick an  $s \in S_t \subseteq U^-$ .  $\{s\}_{\mathbf{X}=\mathbf{x}} = \{t\}$ ; so,  $t \in (U_{\mathbf{X}=\mathbf{x}})^-$ . This proves  $(U')^- \subseteq (U_{\mathbf{X}=\mathbf{x}})^-$ .

Let  $t \notin (U')^-$ . Suppose for the sake of contradiction that there is  $s \in U$  such that  $\{s\}_{\mathbf{X}=\mathbf{x}} = \{t\}$ . But then there is a  $t' \in U'$  such that  $s \in S_{t'} \subseteq U^-$ . But this implies that  $\{t\} = \{s\}_{\mathbf{X}=\mathbf{x}} = \{t'\}$ . Therefore  $t \in (U')^-$ : contradiction.  $\square$

**Theorem 4.12.** *The rules OR-OUT, AND-OUT, NEG-OUT, OR-IN, AND-IN, NEG-IN are sound for  $\mathcal{C}^{neg}$  on causal teams with unique solutions; except for NEG-IN/OUT, this also holds in  $\mathcal{CD}$ .*

*Proof.* The AND and NOT case are straightforward. The soundness of OR-OUT/OR-IN follows immediately from Lemma 4.11.  $\square$

How should we treat counterfactuals occurring in the consequent? The importation rule presented in previous sections (Theorem 4.8) can eliminate them only under some syntactical restrictions. We need a different inference rule that may work without restrictions. We prove here that, if two consecutive interventions affect some common set of variables, then the action performed by the first intervention on these common variables is completely overwritten by the second intervention.

**Lemma 4.13.** *Let  $T$  be a causal team,  $\mathbf{X} \subseteq \text{dom}(T)$ , and  $\mathbf{x}, \mathbf{x}' \in \text{Ran}(\mathbf{X})$ . Then  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{X}=\mathbf{x}'} = T_{\mathbf{X}=\mathbf{x}'}$ .*

*Proof.* Since all the interventions act on the same set of variables  $\mathbf{X}$ , they give rise to the same graph. A straightforward induction on  $d(\mathbf{X}, Y)$  can be used to show that, for any  $s \in T$ , if  $\{t\} = \{s\}_{\mathbf{X}=\mathbf{x}}_{\mathbf{X}=\mathbf{x}'}$  and  $\{t'\} = \{s\}_{\mathbf{X}=\mathbf{x}'}$ , then  $t(Y) = t'(Y)$  for any  $Y \in \text{dom}(T)$ .  $\square$

**Theorem 4.14.** *Let  $\mathbf{X}, \mathbf{Y} \subseteq \text{dom}(T)$ ,  $\mathbf{X}' = \mathbf{X} \setminus \mathbf{Y}$  and  $\mathbf{x}' = \mathbf{x} \setminus \mathbf{y}$ . Then  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (T_{\mathbf{X}'=\mathbf{x}'})_{\mathbf{Y}=\mathbf{y}}$ .*

*Proof.* Let  $\mathbf{Y}' = \mathbf{Y} \setminus \mathbf{X}$  and  $\mathbf{y}' = \mathbf{y} \setminus \mathbf{x}$ . Let  $\mathbf{Z} = \mathbf{X} \cap \mathbf{Y}$ . Suppose  $\mathbf{Z} = (X_{i_1}, \dots, X_{i_k}) = (Y_{j_1}, \dots, Y_{j_k})$ . Write  $\mathbf{z}_1$  for  $(x_{i_1}, \dots, x_{i_k})$  (the restriction of  $\mathbf{x}$  to values for  $\mathbf{Z}$ ), and  $\mathbf{z}_2$  for  $(y_{j_1}, \dots, y_{j_k})$  (the restriction of  $\mathbf{y}$  to values for  $\mathbf{Z}$ ). Then, by Theorem 4.4,  $(T_{\mathbf{X}=\mathbf{x}})_{\mathbf{Y}=\mathbf{y}} = (((T_{\mathbf{X}'=\mathbf{x}'})_{\mathbf{Z}=\mathbf{z}_1})_{\mathbf{Z}=\mathbf{z}_2})_{\mathbf{Y}'=\mathbf{y}'}$ . By Lemma 4.13, this is equal to  $((T_{\mathbf{X}'=\mathbf{x}'}))_{\mathbf{Z}=\mathbf{z}_2})_{\mathbf{Y}'=\mathbf{y}'}$ . Applying Theorem 4.4 again, we obtain  $(T_{\mathbf{X}'=\mathbf{x}'})_{\mathbf{Y}=\mathbf{y}}$ .  $\square$

This result, combined with Theorem 4.4, can be immediately turned into a pair of inference rules:

$$(CF - OUT) : \frac{\mathbf{X} = \mathbf{x} \square \rightarrow (\mathbf{Y} = \mathbf{y} \square \rightarrow \psi)}{(\mathbf{X}' = \mathbf{x}' \wedge \mathbf{Y} = \mathbf{y}) \square \rightarrow \psi}$$

where  $\mathbf{X}' = \mathbf{X} \setminus \mathbf{Y}$  and  $\mathbf{x}' = \mathbf{x} \setminus \mathbf{y}$ . Its inverse, CF-IN, works for  $\mathbf{X}' \cap \mathbf{Y} = \emptyset$  and  $\mathbf{X} \supseteq \mathbf{X}'$ . This pair of rules corresponds essentially to the axiom CM11 proposed by Briggs ([2], sect. 4).

Finally, we return to the axiom REC in Halpern ([11]), which characterizes recursive systems. Halpern shows how to represent, in the potential outcome formalism, the notion “ $X$  affects  $Y$  under some intervention” (we believe this notion is the same as Woodward’s notion of *contributing cause*). We were not able to express this notion in  $C^{neg}$ . Instead we can express it in an extended language  $C^{neg}(\sqcup)$  where we allow positive occurrences of the so-called *intuitionistic disjunction*  $\sqcup$ , which follows the semantical clause:

$$T \models \psi \sqcup \psi' \iff T \models \psi \text{ or } T \models \psi'$$

With this additional connective, “ $X$  affects  $Y$ ” can be written as follows; write  $ND_X$  for the set of nondescendants of  $X$ , and  $\mathcal{W} := \{\mathbf{Z} \subset \text{dom}(T) \setminus \{X, Y\} \mid ND_X \subseteq \mathbf{Z}\}$ :

$$CC(X, Y) : \bigsqcup \{ \mathbf{Z} = \mathbf{z} \square \rightarrow [(X = x \square \rightarrow Y = y) \wedge (X = x' \square \rightarrow Y = y')] \}.$$

where the disjunction is taken over all  $\mathbf{Z} \in \mathcal{W}$ ,  $\mathbf{z} \in \text{Ran}(\mathbf{Z})$ ,  $x \neq x' \in \text{Ran}(X)$ ,  $y \neq y' \in \text{Ran}(Y)$ . Then in  $C^{neg}(\sqcup)$  we can write:

$$(REC_k) : \frac{CC(X_1, X_2) \quad \dots \quad CC(X_{n-1}, X_n)}{\neg CC(X_n, X_1)}$$

and take REC to be the set of all the  $REC_k$  rules, for  $k \in \mathbb{N}$ ,  $k \geq 2$ .

Given a causal team  $T$ , let  $AX_C[T]$ <sup>8</sup> be the axiom system constituted of axiom schemes EFF; DEC; UNI; and the rules CE, CI, OR/AND/NEG/CF-OUT, OR/AND/NEG/CF-IN,  $MP_\vee$  and REC. Now  $AX_C[T]$  is a set of valid  $C^{neg}(\sqcup)$  formulas and sound rules for  $C^{neg}(\sqcup)$ , but we can show it to be complete for the restricted set of formulas  $C^{neg}$ .<sup>9</sup>

**Theorem 4.15.**  *$AX_C[T]$  is a sound and complete axiom system for the language  $C^{neg}$  on recursive causal teams of domain  $\text{dom}(T)$  and range function  $\mathcal{R}_T$ .*

*Proof.* Soundness in  $C^{neg}$  was shown above; it is not difficult to show that soundness extends to  $C^{neg}(\sqcup)$ . Let  $\Gamma \subseteq C^{neg}$ ,  $\varphi \in C^{neg}$  be such that  $\varphi$  holds

<sup>8</sup>The extra parameter  $T$  is there because we need, as was done in [11], to restrict the axioms to teams which share the same “signature” as  $T$ , that is, they have the same domain of variables and the same ranges for each of the variables.

<sup>9</sup>The approach which consists of defining a formal system over a language more general than the set of target formulas is not unheard of; it is for example the approach taken in [8].

in any recursive causal team on which  $\Gamma$  holds. Notice that any  $\psi \in \Gamma$  can be transformed into a  $\psi' \in \mathcal{C}_u$  by repeated applications of OR/AND/NEG/CF-OUT; call  $\Gamma'$  the set of the  $\psi'$ . By the soundness and invertibility of the rules,  $\Gamma$  and  $\Gamma'$  hold in the same causal teams. Similarly,  $\varphi$  can be transformed into a  $\varphi' \in \mathcal{C}_u$  such that  $\varphi$  and  $\varphi'$  hold in the same causal teams. Therefore, in order to prove completeness, we only have to show that, whenever  $\Gamma' \vdash \varphi'$  according to the (translation of the) axiom system of Halpern, we can construct a derivation  $\Gamma \vdash_{AX_C[T]} \varphi$ .

For suppose we have a derivation  $\Gamma' \vdash \varphi'$  in Halpern's axiom system. Then we can compose it with a derivation of  $\Gamma'$  from  $\Gamma$  (using the OUT rules) and a derivation of  $\varphi$  from  $\varphi'$  (using the IN rules) to obtain a derivation of  $\varphi$  from  $\Gamma$  within  $AX_C[T]$  (after Halpern's rules and axioms are replaced with our translations).  $\square$

The logic  $\mathcal{CO}^{neg}$  is, again, flat and closed under dual negation. The obvious way to obtain a complete axiomatization for  $\mathcal{CO}^{neg}$  is then to add to  $AX_C[T]$  the conversion rule

$$SEL_E : \frac{\theta \supset \chi}{\neg\theta \vee \chi}$$

and its inverse  $SEL_I$ , which allow to reduce the problem to that of completeness of  $\mathcal{C}^{neg}$ .

Notice that  $SEL_E$  is valid also for non-flat logics, of course under the restriction that  $\theta$  be flat (and its negation be present in the language); the validity of  $SEL_I$ , instead, requires downward closure of the logic.

## 5 The nonparametric case: falsifiability and admissibility

We have allowed, in our causal teams, the possibility of having incomplete information concerning the invariant functions that relate the variables, up to the extreme case where no information at all is present (except for the set of arguments of the function, and the range of its allowed values). As a consequence, it is in general impossible to fill the counterfactual causal teams generated by interventions with proper values; if we do not have sufficient information to fill an entry with a value, we fill it instead with a formal term which describes the functional dependence of this entry on the values of other variables.

But how should we evaluate counterfactual statements which involve variables whose columns cannot be filled with proper values? We think it should not be possible to ascribe truth values to such statements by the usual clauses. For example, we are not entitled to assert that  $Y = 3$  in a team whose non-formal entries for  $Y$  are all equal to 3.

Yet in some cases we might be able to observe the falsity of similar statements; therefore, to state their contradictory negation. Let us write  $\downarrow s(X)$  to signify that  $s(X)$  is a proper value and not a formal term. Let  $T$  be a causal team, possibly having some formal entries. We read  $T \models^f \psi$  as “ $\psi$  is falsifiable



in  $T^m$ , but notice that this is not falsifiability in the dual sense; rather, it is a partial notion of contradictory negation. We propose the following semantical clauses:

- $T \models^f X = x$  if there is  $s \in T^-$  such that  $\downarrow s(X)$  and  $s(X) \neq x$
- $T \models^f X \neq x$  if there is  $s \in T^-$  such that  $\downarrow s(X)$  and  $s(X) = x$
- $T \models^f =(\mathbf{X}; Y)$  if there are  $s, s' \in T^-$  such that  $\downarrow s(Y), \downarrow s'(Y), s(\mathbf{X}) = s'(\mathbf{X})$  and  $s(Y) \neq s'(Y)$  (notice that the  $s(\mathbf{X})$  need not all be proper values)
- $T \models^f \psi \wedge \chi$  if  $T \models^f \psi$  or  $T \models^f \chi$
- $T \models^f \psi \vee \chi$  if for all subteams  $T_1, T_2$  of  $T$  with  $T_1^- \cup T_2^- = T^-$ , we have  $T_1 \models^f \psi$  and  $T_2 \models^f \chi$ .

(Notice that the clause for disjunction involves universal quantification over teams).

Can we extend these clauses also to selective and counterfactual implication? Let us consider as an example the team

$X \rightarrow Y$	
2	1
1	$\hat{f}_Y(1)$

It seems unreasonable to assert that this team falsifies the formula  $Y = 1 \supset X = 2$ , because, as long as we do not know the function  $f_Y$ , we cannot know whether  $\hat{f}_Y(1)$  is meant to denote 1 or some other number; therefore, we do not know whether the second assignment is compatible or not with our selection (if it were, then the formula would be falsified, otherwise it would not be). We now give a clause that takes into account this kind of problem.

First, let  $\psi$  be a classical formula; let  $\mathbf{V}$  be the set of variables occurring in  $\psi$ ; define  $T_*^\psi := T^\psi \cup \{s \in T^- \mid \not\downarrow s(V) \text{ for some } V \in \mathbf{V}\}$ . Then, for  $\chi$  *downward closed* formula (as we know formulas of  $\mathcal{CD}$  to be), a somewhat reasonable clause for selective implication seems to be:

- $T \models^f \psi \supset \chi$  if  $T_*^\psi \models^f \chi$

This clause is not completely satisfying, because there might be variables in  $\mathbf{V}$  that are irrelevant for determining the truth value of  $\psi$ .

For counterfactuals, the natural clause is more immediate:

- $T \models^f \mathbf{X} = \mathbf{x} \square \rightarrow \chi$  if  $T_{\mathbf{X}=\mathbf{x}} \models^f \chi$ .

Thirdly, notice that, even though we cannot most often decide of the truth of a statement when columns are incomplete, we might however be interested in asserting that some proposition is *admissible* in the given team, that is, consistent with the data we possess. The following seem to be reasonable clauses for the atomic formulas:

- $T \models^a X = x$  if for all  $s \in T^-$  such that  $\downarrow s(X), s(X) = x$
- $T \models^a X \neq x$  if for all  $s \in T^-$  such that  $\downarrow s(X), s(X) \neq x$
- $T \models^a = (\mathbf{X}; Y)$  if for all  $s, s' \in T^-$  such that  $\downarrow s(Y) \downarrow s'(Y), s(\mathbf{X}) = s'(\mathbf{X})$ , we have  $s(Y) = s'(Y)$ .

Finding general clauses for composite formulas appears to be a difficult task; for simplicity, we define this notion only for classical formulas in disjunctive normal form; we also require, in each conjunctive clause, that atomic formulas  $P_j^i$  contain distinct variables.

- $T \models^a \bigvee_{i=1..m} \bigwedge_{j=1..n(i)} P_j^i$  ( $P_j^i$  being of the form  $X_j^i = x_j^i$  or  $X_j^i \neq x_j^i$ ) if there are subteams  $T_i$  of  $T$ , for  $i = 1..m$ , such that
  1.  $T_i \models^a P_j^i$ , for all  $j$ .
  2. for each  $j, j' = 1..n(i)$ , if  $j \neq j'$ ,  $P_j^i$  is  $X_j^i = a$  and  $P_{j'}^i$  is  $X_{j'}^i = b$  (with  $a \neq b$ ), then for all  $s \in T_i$ ,  $s(X_j^i) \neq s(X_{j'}^i)$ .
  3. for each  $j, j' = 1..n(i)$ , if  $j \neq j'$ ,  $P_j^i$  is  $X_j^i = a$  and  $P_{j'}^i$  is  $X_{j'}^i \neq a$ , then for all  $s \in T_i$ ,  $s(X_j^i) \neq s(X_{j'}^i)$ .

The clauses 2. and 3. above refer to formal inequality between terms. To have an idea of the intuition behind clause 2., the reader may think, for example, of the problem of checking the admissibility of  $X = 1 \wedge Y = 2$ ; imagine that there is a row in which both the  $X$ -column and the  $Y$ -column contain the formal term  $f(3, g(2))$ ; then, surely, the formula is not admissible (for  $X = 1 \wedge Y = 2$  to hold in the team, it is necessary that the  $X$  and  $Y$ -column differ on each row).

An intuition for clause 3. can be similarly obtained by thinking of checking the admissibility of  $X = 1 \wedge Y \neq 1$ ; again, for admissibility, we need the term  $f(3, g(2))$  not to occur both in the  $X$  and  $Y$  column.

## 6 Probabilistic semantics

### 6.1 A probabilistic language

In Woodward's approach, probabilistic notions of causality are considered next to the deterministic ones; for example, a variable  $X$  is said to be a *total cause* (in the probabilistic sense) of a different variable  $Y$  if there is an intervention  $do(X = x)$  which changes the probability of  $Y = y$  for some value of  $y$ .

In Pearl ([23]), and Spirtes, Glymour and Scheines ([27]), one finds a representation of causal and counterfactual reasoning in a mixed framework which combines a deterministic component (functional equations) with a probabilistic one (probability distributions over exogenous variables). Such structures, which have been widely studied in the literature, induce causal Bayesian networks, as shortly mentioned in the introduction.

In the context of team semantics, probabilities have been recently introduced via the notion of multiteam. A multiteam differs from a team in that it may feature multiple copies of the same assignment; it is therefore closer to a collection of experimental data than teams are. There have been at least two different approaches to the formalization of multiteams in the literature ([34], [7]). However, the approach we take in the following is that of simulating multiteams by means of teams; this can be easily accomplished by assuming that each team has an extra variable *Key* (which is never mentioned in our formal languages) which takes distinct values on distinct assignments of the same causal team. In this way, we can have two assignments that agree on all the significant variables and just differ on *Key*. With this assumption in mind, the definition of *causal multiteam* can follow word by word the definition of causal team.

We will now work towards the definition of languages that can be appropriate for the discussion and study of probabilistic notions as interpreted over causal multiteams. If we wish to talk about probabilities, it seems natural to add appropriate atomic formulas to our formal language. In theory, it would make sense to assign probabilities to any *flat* formula of the resulting language. However, for technical simplicity, we only allow our formal language to talk of probabilities of  $\mathcal{CO}$  formulas (which were proved to be flat in earlier sections).

**Definition 6.1.** *The set of probabilistic literals is given by:*

$$\neg\alpha | Pr(\chi) \leq \epsilon | Pr(\chi) \geq \epsilon | Pr(\chi) \leq Pr(\theta) | Pr(\chi) \geq Pr(\theta)$$

where  $\alpha$  is a probabilistic literal,  $\chi, \theta$  are formulas of  $\mathcal{CO}$  and  $\epsilon \in \mathbb{R} \cap [0, 1]$ . Literals and probabilistic literals without negation will be called atomic formulas.

The (basic) probabilistic causal language ( $\mathcal{PCD}$ ) is given by the following clauses:

$$\alpha | \psi \wedge \chi | \psi \vee \chi | \psi \sqcup \chi | \theta \supset \psi | \mathbf{X} = \mathbf{x} \square \rightarrow \psi$$

where  $\alpha$  is a literal or probabilistic literal,  $\psi, \chi$  are  $\mathcal{PCD}$  formulas, and  $\theta$  a boolean combination of (non probabilistic) atoms.

As in the non-probabilistic case, we also define restricted languages:

- $\mathcal{PC}$  if selective implication is not allowed
- $\mathcal{PO}$  if counterfactuals are not allowed
- $\mathcal{P}$  if neither selective implications nor counterfactuals are allowed.

Some words are needed to motivate our choice of languages and their restrictions. First of all, notice that, besides the usual (tensor) disjunction  $\vee$  that most naturally arises in the context of logics of dependence, we also add to our probabilistic languages what is called in the literature *intuitionistic disjunction*; that is, a binary connective  $\sqcup$  which satisfies the following semantical clause (already introduced in an earlier section):

- $T \models \psi \sqcup \chi \iff T \models \psi \text{ or } T \models \chi$ .

If one wants to express that the disjunctive event “ $X = x$  or  $Y = y$ ” has probability  $\epsilon$ , it can be done using the formula  $Pr(X = x \vee Y = y) = \epsilon$ . The usage of intuitionistic disjunction is unavoidable if one wishes to make disjunctive statements about probabilities. For example, the statement “either  $X = x$  has probability less than one third, or probability greater than two thirds” may be expressed as  $Pr(X = x) < 1/3 \sqcup Pr(X = x) > 2/3$ . Intuitionistic disjunction turns out to be particularly useful for defining other operators. For example, dependence atoms and conditional independence atoms turn out to be (model-theoretically) *definable* already in  $\mathcal{PO}$  (this will be shown later on in the section). This is the reason why we have omitted such atoms from the syntax. We could not come up with any reasonable semantics for selective implications with probabilistic antecedents. This is the reason why we excluded such constructs from the syntax.

We will use some obvious abbreviations, such as  $Pr(\chi) = \epsilon$  for  $Pr(\chi) \leq \epsilon \wedge Pr(\chi) \geq \epsilon$ , or  $Pr(\chi) < \epsilon$  for  $Pr(\chi) \leq \epsilon \wedge \neg Pr(\chi) \geq \epsilon$ .

The semantic clauses are the same as for the language  $\mathcal{CD}$ , enriched with clauses for the probabilistic atoms. Given any  $\mathcal{CO}$  formula<sup>10</sup>  $\chi$  and any team  $T$  with *finite, nonempty* support  $T^-$ , we can define the **probability of  $\chi$  in  $T$**  as:

$$Pr_T(\chi) := \frac{\text{card}(\{s \in T^- \mid \{s\} \models \chi\})}{\text{card}(T^-)}.$$

For  $T^- = \emptyset$ , we conventionally assume  $Pr_T(\chi)$  to be undefined; and consequently, expressions like  $Pr_T(\chi) \leq \epsilon$ , etc. to be false.

Then the clauses for probabilistic atoms follow quite smoothly:

- $T \models Pr(\chi) \leq \epsilon$  iff  $Pr_T(\chi) \leq \epsilon$
- $T \models Pr(\chi) \geq \epsilon$  iff  $Pr_T(\chi) \geq \epsilon$
- $T \models Pr(\chi) \leq Pr(\theta)$  iff  $Pr_T(\chi) \leq Pr_T(\theta)$
- $T \models Pr(\chi) \geq Pr(\theta)$  iff  $Pr_T(\chi) \geq Pr_T(\theta)$

It is easy to see that the resulting logic is *not* downward closed. For example, a team which is such that less than half of its assignments satisfy  $\chi$  will satisfy  $Pr(\chi) \leq \frac{1}{2}$ . But the subteam  $T^\chi$  consisting *only* of the assignments that satisfy  $\chi$  will not satisfy  $Pr(\chi) \leq \frac{1}{2}$ . Notice also that the atoms of the form  $Pr(\chi) \leq Pr(\theta)$  (and  $Pr(\chi) \geq Pr(\theta)$ ) could be eliminated if we allowed quantification over real values: we could then replace them with  $\exists \epsilon (Pr(\chi) \leq \epsilon \wedge Pr(\theta) \geq \epsilon)$ .

A crucial question which arises right away is: does our definition really define a probability distribution on a team  $T$  of finite support? Let  $\mathcal{E}_T$  be the set of all subsets  $S$  of  $T$  which are definable by a  $\mathcal{CO}$  formula  $\chi$  (that is,  $S = \{s \in T^- \mid \{s\} \models \chi\}$ ). This will be our set of *events*. For any event  $S \in \mathcal{E}_T$ , define  $P_T(S) := Pr_T(\psi)$ , where  $\psi$  is any formula which defines  $S$ .

<sup>10</sup>This definition applies as well to any flat formula.

Now, a nice thing about  $\mathcal{CO}$  formulas is that one can explicitly find (within  $\mathcal{CO}$  itself) their complementary negations (a formula is a complementary negation of  $\psi$  if it is satisfied exactly by those singleton teams that do not satisfy  $\psi$ ). We define a canonical complementary negation  $\psi^c$  for each formula  $\psi$  of  $\mathcal{CO}$ :

- $(X = x)^c$  is  $X \neq x$  (for any variable  $X$  and value  $x$ )
- $(X \neq x)^c$  is  $X = x$  (for any variable  $X$  and value  $x$ )
- $(\theta \wedge \chi)^c$  is  $(\theta^c \vee \chi^c)$
- $(\theta \vee \chi)^c$  is  $(\theta^c \wedge \chi^c)$
- $(\theta \supset \chi)^c$  is  $\theta \wedge \chi^c$
- $(\theta \Box \rightarrow \chi)^c$  is  $\theta \Box \rightarrow \chi^c$

Complementary negations of  $\mathcal{CO}$  formulas may fail to be boolean formulas, but they turn out to be useful because of the following:

**Lemma 6.2.** *Let  $T$  be a causal multiteam,  $\psi \in \mathcal{CO}$ , and  $S = \{s \in T^- \mid \{s\} \models \psi\}$ . Then  $T^- \setminus S = \{s \in T^- \mid \{s\} \models \psi^c\}$ .*

*Proof.* By induction on the syntactical complexity of  $\psi \in \mathcal{CO}$ . We consider the less obvious cases.

Let  $\psi = \theta \supset \chi$ . Then  $\{s\} \models \psi^c$  iff  $\{s\} \models \theta \wedge \chi^c$  iff (by inductive hypothesis)  $\{s\} \models \theta$  and  $\{s\} \not\models \chi$  iff  $\{s\} \not\models \theta \supset \chi$ .

Let  $\psi = \theta \Box \rightarrow \chi$ . Then  $\{s\} \models \psi^c$  iff  $\{s\} \models \theta \Box \rightarrow \chi^c$  iff  $\{s\}_\theta \models \chi^c$  iff (by inductive hypothesis)  $\{s\}_\theta \not\models \chi$  iff  $\{s\} \not\models \theta \Box \rightarrow \chi$ .  $\square$

**Theorem 6.3.** *Let  $T$  be a finite causal multiteam. Then  $(T, \mathcal{E}_T, P_T)$  is a probability space.*

*Proof.* 1)  $T^-$  is definable by the formula  $X = x \vee X \neq x$ . Therefore  $T^- \in \mathcal{E}_T$ .

2)  $\mathcal{E}$  is obviously closed under unions (due to finiteness).

3) We want to show that  $\mathcal{E}$  is closed under complementation. Let  $S \in \mathcal{E}$ . Then  $S$  is defined by some  $\mathcal{CO}$  formula  $\psi$ . But then  $T^- \setminus S$  is defined by the complementary formula  $\psi^c$ , therefore  $T^- \setminus S \in \mathcal{E}$ .

4)  $Pr_T(T^-) = 1$ .

5) Let  $S_1, S_2 \in \mathcal{E}_T$ , with  $S_1 \cap S_2 = \emptyset$ .  $S_1$  and  $S_2$  are defined by two  $\mathcal{CO}$  formulas  $\psi_1, \psi_2$ , respectively. Then  $S_1 \cup S_2$  is defined by  $\psi_1 \vee \psi_2$ . So

$$\begin{aligned}
Pr_T(S_1 \cup S_2) &= Pr_T(\psi_1 \vee \psi_2) = \frac{card(\{s \in T^- \mid \{s\} \models \psi_1 \vee \psi_2\})}{card(T^-)} = \\
&= \frac{card(\{s \in T^- \mid \{s\} \models \psi_1\}) + card(\{s \in T^- \mid \{s\} \models \psi_2\})}{card(T^-)} = \\
&= \frac{card(\{s \in T^- \mid \{s\} \models \psi_1\})}{card(T^-)} + \frac{card(\{s \in T^- \mid \{s\} \models \psi_2\})}{card(T^-)} = \\
&= Pr_T(\psi_1) + Pr_T(\psi_2) = P_T(S_1) + P_T(S_2);
\end{aligned}$$

the third equality holds due to the assumption  $S_1 \cap S_2 = \emptyset$ .  $\square$

This result allows us to use the usual rules of probability calculus.

## 6.2 Comparison with causal models

At this point we are in the position to notice a difference between our approach and the usual causal models. Recall that in the semi-deterministic approach of [23] and [27], by a *causal model* one usually understands (the terminology is by no means used consistently) a structural equation model enriched with a probabilistic distribution over the exogenous variables  $\mathbf{U}$ . If the causal model is recursive (i.e. the graph is acyclic), then each endogenous variable can be represented as a function of the exogenous variables (just consider the structural equation defining an endogenous variable,  $V := f_V(PA_V)$ ; iteratively replace each of the endogenous variables occurring in the right term of this expression with its structural equation until you obtain a function  $f'_V(\mathbf{U})$ ). Given recursivity, the probability distribution over exogenous variables induces a joint probability distribution over *all* the variables of the system given by the following equation:

$$P(X_1 = x_1 \wedge \cdots \wedge X_n = x_n) := \sum_{\{(u_1, \dots, u_k) \mid \text{for all } i=1..n : f'_{X_i}(u_1, \dots, u_k) = x_i\}} P(U_1 = u_1 \wedge \cdots \wedge U_k = u_k)$$

From this joint probability, one can obtain by marginalization the distributions of each set of variables.

This approach fails, however, for non-recursive causal models, in case their system of structural equations has multiple solutions or no solutions at all for some configuration(s) of values of the exogenous variables. In this case some endogenous variables may fail to be functions of the exogenous variables (consider e.g. a system with variables  $U, X, Y$  all having range  $\{0, 1\}$ , and consisting of the equations  $X := Y$  and  $Y := X$ , which induce the arrows  $X \rightarrow Y$  and  $Y \rightarrow X$ ; here  $U$  is the only exogenous variable; given a fixed value  $u$  for  $U$ , both triples  $(u, 0, 0)$  and  $(u, 1, 1)$  are solutions to the system; therefore, neither  $X$  nor  $Y$  can be determined as functions of  $U$ ).

Unlike the semi-deterministic approach of causal models, our definition of probabilities for causal teams applies to both recursive and non-recursive systems. This is due to the fact that we do not conflate in our framework two different accounts of exogeneity, as it is often the case in the literature. In the first account, exogenous variables are treated as quantities whose causal mechanisms are not further examined in the model. In the second account, exogenous variables are treated as *unobserved* variables. The treatment of exogenous variables in our causal teams supports the first account: they have no structural equations, but the assignments of the causal teams encode quantitative information about them. To the extent to which causal teams stick to the first account, they are more general than causal models. (Of course, if we further extend causal teams to allow for unobserved variables, probabilities in non-recursive systems might fail again to be determined by the exogenous distribution; such a generalisation is beyond the scope of the present paper).

We might ask at this point whether our probabilities extend correctly (“conservatively”) the probabilities typical of recursive causal models. That is, if we take the team-defined probability over the exogenous variables  $\mathbf{U}$ , and we define the joint probability by the formula above, do we obtain the team-defined joint probability? The answer is affirmative:

**Theorem 6.4.** *Let  $T$  be a causal multiteam with exogenous variables  $\mathbf{U} = U_1, \dots, U_k$ . For each  $(u_1, \dots, u_k) \in \prod_{i=1..k} \text{Ran}(U_i)$ , let  $P(U_1 = u_1 \wedge \dots \wedge U_k = u_k) := \text{Pr}_T(U_1 = u_1 \wedge \dots \wedge U_k = u_k)$ . Let  $X_1, \dots, X_n$  be a list of all variables of  $T$ , and  $P(X_1, \dots, X_n)$  be defined as above. Then  $P(X_1 = x_1, \dots, X_n = x_n) = \text{Pr}_T(X_1 = x_1 \wedge \dots \wedge X_n = x_n)$  for any tuple  $(x_1, \dots, x_n) \in \prod_{X_i \in \text{dom}(T)} \text{Ran}(X_i)$ .*

*Proof.* By definition,

$$\text{Pr}_T(X_1 = x_1 \wedge \dots \wedge X_n = x_n) = \frac{\text{card}(\{s \in T \mid s(X_1) = x_1 \wedge \dots \wedge s(X_n) = x_n\})}{\text{card}(T)}.$$

Clause c) in the definition of causal teams entails that this last term is equal to

$$\frac{\text{card}(\{s \in T \mid f'_{X_i}(s(U_1), \dots, s(U_k)) = x_i \text{ for } i = 1..n\})}{\text{card}(T)}$$

where the  $f'_{X_i}(u_1, \dots, u_k)$  are defined as above. We can further decompose the numerator according to the values taken by  $U_1, \dots, U_k$ , which ultimately leads to

$$\sum_{(u_1, \dots, u_k) \mid \text{for all } i, f'_{X_i}(u_1, \dots, u_k) = x_i} \frac{\text{card}(\{s \in T \mid s(U_1) = u_1, \dots, s(U_k) = u_k\})}{\text{card}(T)}$$

that is

$$\sum_{(u_1, \dots, u_k) \mid \text{for all } i, f'_{X_i}(u_1, \dots, u_k) = x_i} P(U_1 = u_1 \wedge \dots \wedge U_k = u_k)$$

which is the definition of  $P(X_1 = x_1, \dots, X_n = x_n)$ .  $\square$

### 6.3 Pearl’s example in causal teams

Before proceeding with the analysis of the probabilistic semantics for causal teams, we illustrate its workings by means of an example taken from Pearl ([23], 1.4.4).

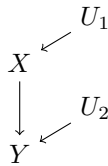
For Pearl the purpose of the example is to show that causal Bayesian networks are insufficient to answer counterfactual questions, which require structural equations. Our concern here would only be with that aspect of the example (Pearl’s “model 2”) which illustrates the three stages in his treatment of counterfactuals in the structural equation framework. The following quote gives the spirit of the enterprise:

The preceding considerations further imply that the three tasks listed in the beginning of this section - prediction, intervention, and counterfactuals - form a natural hierarchy of causal reasoning tasks, with increasing levels of refinement and increasing demands on the knowledge required for accomplishing the tasks. Prediction is the simplest of the three, requiring only a specification of a joint distribution function. The analysis of interventions requires a causal structure in addition to a joint distribution. Finally, processing counterfactuals is the hardest task because it requires some information about the functional relationships and/or the distribution of the omitted factors. (p. 38)

The example is about a clinical test. There are two exogenous variables  $X$  (being treated:  $X = 1$ ; not being treated:  $X = 0$ ) and  $Y$  (which measures the subject's response: recovery  $Y = 0$ , death,  $Y = 1$ ). There are two exogenous (hidden), independent variables  $U_1$  and  $U_2$ . The equations of the model are:

$$\begin{cases} x = u_1 \\ y = xu_2 + (1-x)(1-u_2) \end{cases}$$

They induce the following *DAG*:



The two independent binary variables are governed by the probability distribution  $P(u_1 = 1) = P(u_2 = 1) = \frac{1}{2}$ , which generates joint probability distributions  $P(x, y, u_2)$  and  $P(x, y)$  that may be calculated using the formula from subsection 6.2. For instance,  $P(X = x \wedge Y = y) =$

$$= \sum_{\{(u_1, u_2) | u_1 = x \wedge u_1 u_2 + (1-u_1)(1-u_2) = y\}} P(U_1 = u_1 \wedge U_2 = u_2) = 0,25$$

for all  $x$  and  $y$ . Here is a table which lists  $P(x, y, u_2)$  and  $P(x, y)$  :

	$u_2 = 1$		$u_2 = 0$		<i>Marginal</i>	
	$x = 1$	$x = 0$	$x = 1$	$x = 0$	$x = 1$	$x = 0$
$y = 1$	0	0.25	0.25	0	0.25	0.25
$y = 0$	0.25	0	0	0.25	0.25	0.25

Pearl considers the following counterfactual query:

- (\*) What is the probability  $Q$  that a subject who died under treatment ( $x = 1, y = 1$ ) would have recovered ( $y = 0$ ) had he or she not been treated ( $x = 0$ )?



He calls the fact that the subject died under treatment the evidence

$$e = \{(x = 1, y = 1)\}$$

and points out that answering the above query involves three steps: (1) given the evidence  $e$ , (2) we compute the probability  $Y = y$  (2) under the hypothetical condition  $X = x$ .

In his example this amounts to:

1. We apply the evidence  $e$  to the set of equations, and conclude that it is compatible with only one realization:  $u_1 = 1$  and  $u_2 = 1$ . This gives us a value for  $u_2$ .
2. We make an intervention  $do(X = 0)$  on the variable  $X$ . It results in the first equation being replaced with  $x = 0$ .
3. We solve the second equation for  $x = 0$ ,  $u_2 = 1$ . We get  $y = 0$  from which we conclude that the probability  $Q$  is 1.

We are told that in the general case

- step (1) corresponds to an abduction: update the probability  $P(\mathbf{u})$  to obtain  $P(\mathbf{u}/e)$
- step (2) corresponds to an action: replace the equations corresponding to each intervened variable  $X$  by  $X = x$
- step (3) corresponds to a prediction: compute  $P(y)$  in the new model.

(Pearl 2009, p.61)

Pearl is explicit on the point that if for each value  $\mathbf{u}$  of the exogenous variables  $\mathbf{U}$  there is a unique solution for  $Y$ , then step (3) always gives a unique solution for the needed probability: we simply sum up the probabilities  $P(\mathbf{u}/e)$  assigned to all  $\mathbf{u}$  which yield  $Y = y$  as a solution. More technically:

$$P(Y = y) = \sum_{\{\mathbf{u}: f'_Y(\mathbf{u})=y\}} P(\mathbf{u}/e)$$

Pearl's model 2 can be represented in our framework by using the causal team  $T$

$U_1 \overset{\curvearrowright}{\curvearrowleft} U_2 \overset{\curvearrowright}{\curvearrowleft} X \rightarrow Y$			
0	0	0	1
0	1	0	0
1	0	1	0
1	1	1	1

where the range of each variable is  $\{0, 1\}$  and the arrows of the DAG are induced by the two structural equations given earlier, which are part of the

description of the team. Notice also that this team, interpreted as a multiteam, bears the same probabilities that are carried by the system described by Pearl.

It can be checked that  $T$  is an explicit causal team. We formulate Pearl’s query in our language as

$$X = 1 \wedge Y = 1 \supset (X = 0 \square \rightarrow Y = 0) \quad (2)$$

or, in its probabilistic version, as

$$X = 1 \wedge Y = 1 \supset (X = 0 \square \rightarrow Pr(Y = 0) = 1). \quad (3)$$

By the clause for selective implication, the formula (2) is true in the causal team  $T$  if and only if the formula  $(X = 0 \square \rightarrow Y = 0)$  is true in the causal team represented by the annotated table

$U_1$	$U_2$	$X \rightarrow Y$	
1	1	1	1

and the two structural equations. Now the formula  $(X = 0 \square \rightarrow Y = 0)$  is true, given that the intervention  $do(X = 0)$  generates the causal team represented by the annotated team

$U_1$	$U_2$	$X \rightarrow Y$	
1	1	0	0

and the same system of equations, except for the first one, replaced by  $X = 0$ <sup>11</sup>. Obviously  $Y = 0$  is true in this causal team, which shows that the formula (3) is true.

Later on in section 7, where we will introduce various notions of cause within causal team semantics, we will see that in this example it is also true that  $X$  is a *direct cause* of  $Y$  in  $T$ .

In the next subsection we will see that conditional probabilities of the form  $Pr(\chi_2 \mid \chi_1) \leq \epsilon$ ,  $Pr(\chi_2 \mid \chi_1) = \epsilon$ , etc. may be defined in our probabilistic language. There it will be shown that a formula like

$$X = 1 \wedge Y = 1 \supset (X = 0 \square \rightarrow Pr(Y = 0) = 1)$$

is true in a causal team  $T$  if and only if

$$Pr_T(X = 0 \square \rightarrow Y = 0 \mid X = 1 \wedge Y = 1) = 1.$$

We notice that the formula inside the brackets shows nicely Pearl’s three “mental steps”.

---

<sup>11</sup>Actually, in the causal team formalism this is represented by the removal of the invariant function  $\mathcal{F}_T^{X=1 \wedge Y=1}(X)$ , and the change of role of  $X$ , from endogenous to exogenous variable.

## 6.4 Definability in the probabilistic language

Can we define conditional probabilities in our framework? Semantically, the obvious definition of a conditional probability over a team is ( assuming that  $\chi_1, \chi_2$  are  $\mathcal{CO}$  formulas, and  $\chi_1$  is satisfied by at least one assignment of the team):

$$Pr_T(\chi_2|\chi_1) := \frac{Pr_T(\chi_1 \wedge \chi_2)}{Pr_T(\chi_1)}.$$

It is more complicated to assess whether our current formal languages can talk about conditional probabilities. We can see that, in a number of circumstances, the selective implication can take the role of conditioning. For example, under the same assumptions as above, we can write  $Pr(\chi_2|\chi_1) \leq \epsilon$  as an abbreviation for

$$\chi_1 \supset Pr(\chi_2) \leq \epsilon$$

Here is the proof that the abbreviation has the intended meaning:

$$\begin{aligned} T \models \chi_1 \supset Pr(\chi_2) \leq \epsilon &\iff \\ &\iff T^{x_1} \models Pr(\chi_2) \leq \epsilon \\ &\iff \frac{\text{card}(\{s \in (T^{x_1})^- \mid \{s\} \models \chi_2\})}{\text{card}((T^{x_1})^-)} \leq \epsilon \\ &\iff \frac{\text{card}(\{s \in (T^{x_1})^- \mid \{s\} \models \chi_2\})}{\text{card}(T^-)} \frac{\text{card}(T^-)}{\text{card}((T^{x_1})^-)} \leq \epsilon \\ &\iff \frac{Pr_T(\chi_1 \wedge \chi_2)}{Pr_T(\chi_1)} \leq \epsilon \iff Pr_T(\chi_2|\chi_1) \leq \epsilon \end{aligned}$$

provided  $Pr_T(\chi_1) > 0$ . In case  $Pr_T(\chi_1) = 0$ , then the causal team is empty, so that our earlier conventions impose that  $T^{x_1} \not\models Pr(\chi_2|\chi_1) \leq \epsilon$ . Therefore, in this case we treat conventionally the expression  $Pr(\chi_2|\chi_1) \leq \epsilon$  as not being satisfied by  $T$ . Things work analogously for inequalities in the opposite direction, and one can similarly prove that  $T \models \chi \supset Pr(\psi_1) \leq Pr(\psi_2)$  iff  $Pr_T(\psi_1|\chi) \leq Pr_T(\psi_2|\chi)$ . A mistake to be avoided, at this point, is to think that all semantical statements about  $Pr_T(\chi \supset \theta)$  can *always* be directly translated into formulas of the formal language by replacing  $Pr_T(\chi \supset \theta)$  with  $Pr(\theta|\chi)$ . Such an identification is impossible for any reasonable notion of conditional, as shown by Lewis ([19]).

We would like, however, to point out that our logical language is in some respect more general than other languages typically used in the causation literature (see e.g. [23], chap. 3). To give an example, Pearl would write  $P(y|do(x), u) = \epsilon$ , or  $P(y|\hat{x}, u) = \epsilon$ , to denote the probability, given the observation that  $U$  is equal to  $u$ , of  $Y$  being equal to  $\epsilon$ , in the model where  $X$  has been fixed to  $x$  by an intervention (see e.g. [23] sec. 3.4). This is expressed in our notation by  $U = u \supset (X = x \square \rightarrow P(Y = y) = \epsilon)$ . The equivalence can be proved as follows:

$$T \models U = u \supset (X = x \square \rightarrow P(Y = y) = \epsilon) \iff$$

$$\begin{aligned}
&\iff T^{U=u} \models X = x \Box \rightarrow P(Y = y) = \epsilon \\
&\iff (T^{U=u})_{X=x} \models P(Y = y) = \epsilon \\
&\iff \frac{\text{card}(\{s \in ((T^{U=u})_{X=x})^- \mid \{s\} \models Y = y\})}{\text{card}(((T^{U=u})_{X=x})^-)} = \epsilon \\
&\iff \frac{\text{card}(\{t \in (T^{U=u})^- \mid \{t\}_{X=x} \models Y = y\})}{\text{card}((T^{U=u})^-)} = \epsilon \\
&\iff \frac{\text{card}(\{t \in (T^{U=u})^- \mid \{t\}_{X=x} \models Y = y\})}{\text{card}(T^-)} \cdot \frac{\text{card}(T^-)}{\text{card}((T^{U=u})^-)} = \epsilon \\
&\iff \frac{\text{Pr}_T(U = u \wedge (X = x \Box \rightarrow Y = y))}{\text{Pr}_T(U = u)} = \epsilon \\
&\iff \text{Pr}_T(X = x \Box \rightarrow Y = y \mid U = u) = \epsilon
\end{aligned}$$

(the fourth equivalence makes an essential use of the fact that we are using multiteams, not teams). But in our language we also have available formulas like  $X = x \Box \rightarrow (U = u \supset \text{Pr}(Y = y) = \epsilon)$ . A straightforward computation shows that this expression is equivalent to the statement that  $\text{Pr}_{T_{X=x}}(Y = y \mid U = u) = \epsilon$ , that is, that the probability of  $Y = y$  conditional on  $U = u$  *in the intervened team* is  $\epsilon$ . This formula is not necessarily equivalent to the previous one. The nonequivalence of these two formulas can be seen (using different names for variables) by looking back at an example considered in subsection 4.3 to show the non-commutativity of selective and counterfactual implications. The team  $S$  in that example satisfies  $Z = 1 \Box \rightarrow (Y = 3 \supset \text{Pr}(Y = 3) = 1)$  but does not satisfy  $Y = 3 \supset (Z = 1 \Box \rightarrow \text{Pr}(Y = 3) = 1)$ . The difference between the two formulas is that in the second we consider the observational evidence that *in the intervened model*,  $U$  has value  $u$ . It seems to us that in Pearl's notation, there is always the implicit convention that what occurs as a conditioning variable (in this case  $u$ ) has to be conditioned on in the initial model. This is also the reason why  $u$  and  $do(x)$  can occur in the same position in the notation, without any problems of non-commutativity arising. But this limitation seems problematic in the context of causal reasoning. We might want, for example, to discuss what might happen under certain observations after an hypothetical future experiment, and Pearl's notation does not seem to be of any help in this case. The same problem seems to afflict other notational representations of counterfactuals, such as the one coming from the potential outcome tradition ([22], [24], [8]). In that notational system we can write, say,  $Y_x(u) = y$  to express that, under the observation that  $U$  is  $u$ , fixing  $X$  to  $x$  forces  $Y$  to take value  $y$  ( $U = u \supset (X = x \Box \rightarrow Y = y)$ ), but there seems to be no way to express the formula  $X = x \Box \rightarrow (U = u \supset Y = y)$ .

Let us now see what kind of (probabilistic) dependencies and independencies are definable in our framework. First of all, we can give a model-dependent definition of *two CO formulas being probabilistically independent*:  $\chi_1 \perp\!\!\!\perp \chi_2$ . To this purpose, notice that all probabilities expressible over a team  $T$  are multiple of the minimum unit  $u = 1/\text{card}(T^-)$ . Thanks to this, probabilistic atoms of

the form  $Pr(\chi_2|\chi_1) \leq Pr(\theta)$  can be given a model-dependent definition within  $\mathcal{PO}$  (under the assumption  $Pr_T(\chi_1) > 0$ ):

$$T \models Pr(\chi_2|\chi_1) \leq Pr(\theta) \iff \\ \iff T \models \bigsqcup_{m=0..card(T^-)} [(\chi_1 \supset Pr(\chi_2) \leq mu) \wedge Pr(\theta) \geq mu]$$

Assuming by default that, if either  $\chi_1$  or  $\chi_2$  has probability 0, then the two formulas are probabilistically independent, we can take  $\chi_1 \perp\!\!\!\perp \chi_2$  as an abbreviation for  $Pr(\chi_1) = 0 \sqcup Pr(\chi_2) = 0 \sqcup Pr(\chi_2|\chi_1) = Pr(\chi_2)$ . More generally, we can talk of **conditional independence atoms**  $\chi_1 \perp\!\!\!\perp_{\chi_3} \chi_2$  as abbreviations for  $Pr(\chi_1|\chi_3) = 0 \sqcup Pr(\chi_2|\chi_3) = 0 \sqcup Pr(\chi_2|\chi_1 \wedge \chi_3) = Pr(\chi_2|\chi_3)$ .

It would be important to understand whether independence atoms  $\chi_1 \perp\!\!\!\perp_{\chi_3} \chi_2$  ([10]) are definable or not in this framework, or whether it would be worth to add them to the language. This would lead to a simple uniform definition of dependence atoms; but we do not know the answer to this question. However, we can see that it is possible to give a model-dependent definition (in  $\mathcal{PO}$ ) of dependence atoms in terms of conditional independence atoms. Given a sequence  $(x_1, \dots, x_n)$  of values (implicitly associated to variables  $X_1, \dots, X_n \in dom(T)$ ), we say that  $(x_1, \dots, x_n)$  *occurs* in  $T$  if there is an  $s \in T$  such that  $s(X_1) = x_1, \dots, s(X_n) = x_n$ .

**Theorem 6.5.** *For any finite causal multiteam  $T$ ,  $T \models (\mathbf{X}; Y)$  if and only if*

$$T \models \bigwedge_{(\mathbf{x}, y) \in Ran(\mathbf{X}) \times Ran(Y)} Y = y \perp\!\!\!\perp_{\mathbf{X}=\mathbf{x}} Y = y.$$

*Proof.* The conjunction says that, for every tuple  $\mathbf{x}, y$  of appropriate values for  $\mathbf{X}, Y$ , either  $Pr_T(Y = y|\mathbf{X} = \mathbf{x}) = 0$ , or  $Pr_T(Y = y|Y = y \wedge \mathbf{X} = \mathbf{x}) = Pr_T(Y = y|\mathbf{X} = \mathbf{x})$ .

That is, for each  $\mathbf{x}, y \in Ran(\mathbf{X}) \times Ran(Y)$ , either this tuple does not occur in  $T$  or  $Pr_T(Y = y|\mathbf{X} = \mathbf{x}) = 1$ , which means, more explicitly:

$$\frac{card(\{s \in T^- | s(Y) = y \text{ and } s(\mathbf{X}) = \mathbf{x}\})}{card(\{s \in T^- | s(\mathbf{X}) = \mathbf{x}\})} = 1$$

which is equivalent to

$$card(\{s \in T^- | s(Y) = y \text{ and } s(\mathbf{X}) = \mathbf{x}\}) = card(\{s \in T^- | s(\mathbf{X}) = \mathbf{x}\})$$

Given the finiteness of the team, this last equality is equivalent to the statement that any assignment  $s$  which satisfies  $s(\mathbf{X}) = \mathbf{x}$  also satisfies  $s(Y) = y$ .  $\square$

## 6.5 The Markov Condition

An assumption underlying a vast literature on causal studies is the so-called Markov Condition. As a condition on a causal team  $T$ , it seems natural to paraphrase it as follows:

$$Pr_T(X = x|PA_X = pa_X) = Pr_T(X = x|PA_X = pa_X, \mathbf{N} = \mathbf{n})$$

for any set  $\mathbf{N}$  of nondescendants of  $X$  disjoint from  $PA_X$   
and any tuple  $(x, pa_X, \mathbf{n})$  occurring in  $T$ .

This is a scheme of conditions which state that each variable is independent of its nondescendants, conditional on its parents. This condition is considered important in causal inference, for example, because it allows to write the joint probability of the system in a compact and usually easy to compute form (product formula):

$$Pr_T(X_1 = x_1 \wedge \dots \wedge X_n = x_n) = \prod_{i=1..n} Pr_T(X_i = x_i|PA_i = pa_i)$$

where  $pa_i$  is  $x_1, \dots, x_n$  restricted to the values for  $PA_i$ . This formula, in turn, gives a modular method for calculating the joint probability in intervened teams:

$$Pr_{T_{\mathbf{z}=\mathbf{z}}}(X_1 = x_1 \wedge \dots \wedge X_n = x_n) = \prod_{X_i \notin \mathbf{Z}} Pr_T(X_i = x_i|PA_i = pa_i).$$

However, in causal (multi)teams the Markov Condition is not automatically satisfied. As a simple counterexample, consider the team

X	Y
1	1
2	1
1	2

Since the parent set of  $X$  is the empty set, and the tuple  $X = 1, Y = 2$  occurs in the team, the Markov Condition would imply:  $P(X = 1) = P(X = 1|Y = 2)$ . However, here we have  $Pr(X = 1) = 2/3$ , but  $P(X = 1|Y = 2) = 1$ .

The Markov Condition is a rather complex condition to require on a team; for example, it might be noticed that its verification can require an exponential time in the number of variables of the domain (since we have to verify such a condition for every appropriate *set*  $\mathbf{Z}$  of variables). It is probably desirable to find a simpler axiom that implies the Markov Condition. We suggest the following *Markov Axiom Scheme* (MA):

$$Pr_T(X = x|PA_X = pa_X) = Pr_T(X = x|PA_X = pa_X, \mathbf{Y} = \mathbf{y})$$

where  $\mathbf{Y}$  is the set of *all* nondescendants of  $X$  not occurring in  $PA_X$   
and for any tuple  $(x, \mathbf{y})$  occurring in  $T$ .

Notice that in this case we have one axiom for each variable in the domain; therefore, the complexity of verifying the Markov Axiom Scheme is linear in the number of variables. We have to prove that this condition is sufficient to enforce the Markov condition. This is obtained by a merely probabilistic reasoning, which we are allowed to use thanks to our earlier proof that teams endowed with our probability function form a  $\sigma$ -algebra. The following lemma is an immediate consequence of basic probability calculus:

**Lemma 6.6.** For any  $\mathcal{CO}$  formulas  $\psi, \chi, \theta$  and multiteam  $T$ , and assuming that  $Pr_T(\chi \wedge \theta) > 0$ , it follows that  $Pr_T(\psi \wedge \chi | \theta) = Pr_T(\psi | \chi \wedge \theta) \cdot Pr_T(\chi | \theta)$ .

**Theorem 6.7.** Suppose a multiteam  $T$  satisfies the Markov Axiom Scheme. Then it satisfies the Markov Condition.

*Proof.* Let  $X_i \in \text{dom}(T)$ ,  $PA_i$  its set of parents,  $\mathbf{Z}$  a set of variables such that  $PA_i \subseteq \mathbf{Z} \subseteq \text{dom}(T) \setminus \{X_i\}$ , and such that any element of  $\mathbf{Z} \setminus PA_i$  is a nondescendant of  $X_i$ . We want to show that  $Pr_T(X_i = x_i | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) = Pr_T(X_i = x_i | \bigwedge_{X_j \in PA_i} X_j = x_j)$  for any tuple  $(x_1, \dots, x_n)$  occurring in  $T$ .

Let  $\mathbf{W} = \{X_j \in \text{dom}(T) | X_j \text{ nondescendant of } X_i \text{ and } X_j \notin \mathbf{Z}\}$ . In case  $\mathbf{W}$  is empty, this means that  $\mathbf{Z} \setminus PA_i$  is the set of all nondescendants of  $X_i$  that are not in  $PA_i$ ; therefore the equality we are trying to prove is simply an instance of the Markov Axiom.

Suppose then that  $\mathbf{W}$  is nonempty. We now write  $\mathbf{z}$  for the subsequence of values of the specific sequence  $(x_1, \dots, x_n)$  that correspond to the variables in  $\mathbf{Z}$ ; and  $\mathbf{wz}$  for the concatenation of sequence  $\mathbf{z}$  with some sequence  $\mathbf{w}$  of values for  $\mathbf{W}$ . Define  $T_{\mathbf{Z}}(\mathbf{W}) = \{\mathbf{w} | \text{there is } s \in T^- \text{ such that } s(\mathbf{W}) = \mathbf{w} \text{ and } s(\mathbf{Z}) = \mathbf{z}\} = \{\mathbf{w} | Pr_T(\mathbf{W} = \mathbf{w} \wedge \mathbf{Z} = \mathbf{z}) > 0\}$ . Notice then that, if  $Pr_T(\mathbf{W} = \mathbf{w} \wedge \mathbf{Z} = \mathbf{z}) = 0$ , it follows that  $Pr_T(X_i = x_i \wedge \mathbf{W} = \mathbf{w} | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) = 0$ . Therefore we can write:

$$\begin{aligned} Pr_T(X_i = x_i | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) &= \\ &= \sum_{\mathbf{w} \in T_{\mathbf{Z}}(\mathbf{W})} Pr_T(X_i = x_i \wedge \mathbf{W} = \mathbf{w} | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) \end{aligned}$$

because the events  $X_i = x_i \wedge \mathbf{W} = \mathbf{w}$ , as  $\mathbf{w}$  varies among tuples occurring in  $T_{\mathbf{Z}}(\mathbf{W})$ , are disjoint; and for other values  $\mathbf{w} \notin T_{\mathbf{Z}}(\mathbf{W})$ ,  $Pr_T(X_i = x_i \wedge \mathbf{W} = \mathbf{w} | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) = 0$ .

Since, furthermore, for any  $\mathbf{w} \in T_{\mathbf{Z}}(\mathbf{W})$  it holds that  $Pr_T(\mathbf{W} = \mathbf{w} \wedge \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) > 0$ , we are entitled to use Lemma 6.6 on each summand, which yields the expression

$$\sum_{\mathbf{w} \in T_{\mathbf{Z}}(\mathbf{W})} [Pr_T(X_i = x_i | \mathbf{W} = \mathbf{w}, \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j) \cdot Pr_T(\mathbf{W} = \mathbf{w} | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j)].$$

Now, applying the appropriate instances of the Markov Axiom, one obtains

$$\sum_{\mathbf{w} \in T_{\mathbf{Z}}(\mathbf{W})} [Pr_T(X_i = x_i | \bigwedge_{X_j \in PA_i} X_j = x_j) \cdot Pr_T(\mathbf{W} = \mathbf{w} | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j)].$$

Since the left factor does not depend anymore on  $\mathbf{w}$ , we can extract it:

$$Pr_T(X_i = x_i | \bigwedge_{X_j \in PA_i} X_j = x_j) \cdot \sum_{\mathbf{w} \in T_{\mathbf{Z}}(\mathbf{W})} Pr_T(\mathbf{W} = \mathbf{w} | \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j).$$

Since the events form a  $\sigma$ -algebra, and  $T_{\mathbf{Z}}(\mathbf{W})$  contains all the values  $\mathbf{w}$  such that  $\mathbf{W} = \mathbf{w} \wedge \bigwedge_{X_j \in \mathbf{Z}} X_j = x_j$  has positive probability, the right factor sums up to 1, and we obtain the desired equality.  $\square$

## 7 Causal modeling

### 7.1 Direct and total cause

What might be interesting to express in our languages? We consider the two basic type-causal notions from Woodward ([39]), *direct* and *total* cause.

Let us consider the case of **direct cause**. Recall Woodward's definition of a direct cause in section 2.3:

A necessary and sufficient condition for  $X$  to be a direct cause of  $Y$  with respect to some variable set  $V$  is that there be a possible intervention on  $X$  that will change  $Y$  (or the probability distribution of  $Z$ ) when all other variables in  $V$  besides  $X$  and  $Z$  are held fixed at some value by interventions.

Woodward's definition is slightly ambiguous in that it talks about a change in  $Y$ , but does not say with respect to what the change is made; to  $Y$ 's actual value? To some possible value of  $Y$ , i.e., some  $y \in \text{Ran}(Y)$ ? To a  $y$  occurring in some allowed configuration (assignment)? We resolve the ambiguity by stipulating that the values of  $Y$  to be compared be generated by *two* distinct interventions.

The kind of intervention needed for the evaluation of direct causation of  $X$  on  $Y$  is a simultaneous intervention on all variables in the system except for  $X$  and  $Y$ . As an example, let  $T$  be the team

$X \rightarrow Z \rightarrow Y$		
1	1	2
2	2	4
3	3	6

with  $\mathcal{F}_Z(x) := x$  and  $\mathcal{F}_Y(x, z) = x + z$ . It is an essential ingredient of this example that there is an arrow from  $X$  to  $Y$ . We show that  $X$  is a direct cause of  $Y$  in  $T$ . First of all we must fix all other variables (in this case, just  $Z$ ) to an appropriate value (we choose 1) by an intervention, which also updates  $Y$ :

$X \quad Z \rightarrow Y$		
1	1	2
2	1	3
3	1	4

Then we intervene in two different ways on  $X$ :

$X \quad Z \rightarrow Y$			$X \quad Z \rightarrow Y$		
1	1	2	2	1	3

The fact that the two interventions select distinct values for  $Y$  proves that  $X$  is a direct cause of  $Y$ .



The peculiar form of these kinds of interventions makes so that, *if there is an arrow from  $X$  to  $Y$* , after the intervention all the columns in the team are constant, that is, a singleton team is produced. This suggests that direct cause can also be defined as an operation on teams, instead of as an operation on assignments. Let  $Fix(\mathbf{z})$  be an abbreviation for  $\bigwedge_{Z \in Dom(T) \setminus \{X, Y\}} Z = z$ :

$T \models DC(X; Y)$  iff there are  $\mathbf{z}, x \neq x', y \neq y'$  such that

$$T \models (Fix(\mathbf{z}) \wedge X = x) \Box \rightarrow Y = y$$

$$\text{and } T \models (Fix(\mathbf{z}) \wedge X = x') \Box \rightarrow Y = y'.$$

For future ease of notation we give names to formulas of the following form:

$$\Phi_{x, x'; y, y'; \mathbf{z}} : ((Fix(\mathbf{z}) \wedge X = x) \Box \rightarrow Y = y) \wedge ((Fix(\mathbf{z}) \wedge X = x') \Box \rightarrow Y = y')$$

Direct causation might be explicitly defined in a first-order extension of our current logics:

$$\exists x \exists x' \exists y \exists y' \exists \mathbf{z} [x \neq x' \wedge y \neq y' \wedge \Phi_{x, x'; y, y'; \mathbf{z}}]$$

This definition is model-dependent in that it makes implicit reference to a set  $\mathbf{Z} = dom(T) \setminus \{X, Y\}$  of variables which is determined by the domain of the causal team. When the ranges of variables are finite, it is immediate to turn this into a non-uniform definition in  $\mathcal{CD}$  enriched with intuitionistic disjunction  $\sqcup$ :<sup>12</sup>

$$\bigsqcup_{x \neq x', y \neq y', \mathbf{z}} \Phi_{x, x'; y, y'; \mathbf{z}}$$

Indeed, if there is an arrow from  $X$  to  $Y$ , any intervention of this kind will collapse the team to a singleton team, so that the condition for direct cause can be checked (it can fail; see the example below). But if instead some intervention of this form does *not* collapse the team into a singleton team, then necessarily there is no arrow from  $X$  to  $Y$ . Therefore, the graph of direct causes is a subgraph of the graph of the team. It can be a *proper* subgraph, that is, there might be an arrow between two variables  $X$  and  $Y$ , and yet those fail to be in the direct cause relationship; this is what happens in the following causal team (with, say,  $Ran(X) = Ran(Y) = \{0, 1\}$ ):

$X \rightarrow Y$	
0	0
1	0

<sup>12</sup>The results of [41] entail that, in the case of propositional dependence logic, the intuitionistic disjunction is eliminable in terms of the other propositional connectives (although not uniformly definable, [40]). It is unclear, at the current state of investigation, whether these results translate into our framework.

The definition of probabilistic direct cause in  $\mathcal{PCD}$  presents a little surprise. As in the deterministic case, we use counterfactual antecedents of the form  $X = x \wedge Fix(\mathbf{z})$  to try to reduce to one the number of configurations occurring in the multiteam; however, now, even in presence of an arrow from  $X$  to  $Y$ , the resulting team is not necessarily a singleton team, but it will be in general a collection of identical copies of one and the same assignment; the formula  $Y = y$  will assume probability 0 or 1. This leads to simplifications in the definition of probabilistic direct cause. Write  $\Psi_{x,x';y;\mathbf{z}}$  for

$$((Fix(\mathbf{z}) \wedge X = x) \Box \rightarrow Pr(Y = y) = 0) \wedge ((Fix(\mathbf{z}) \wedge X = x') \Box \rightarrow Pr(Y = y) = 1).$$

Then, probabilistic direct cause can be formulated as:

$$T \models PDC(X; Y) \text{ iff there are } \mathbf{z}, x, x', y \text{ such that } T \models \Psi_{x,x';y;\mathbf{z}}$$

A non-uniform definition in  $\mathcal{PCD}$  is as follows:

$$\bigsqcup_{x \neq x', y, \mathbf{z}} \Psi_{x,x';y;\mathbf{z}}$$

We may conclude that the probabilistic notion of direct cause collapses to the deterministic one.

The case of **total cause** ([39]) is more complicated. Recall Woodward's definition from section 2.3:

$X$  is a total cause of  $Y$  if and only if there is a possible intervention on  $X$  that will change  $Y$  or the probability distribution of  $Y$ .

This formulation of Woodward is again very ambiguous in some respects; as before, we can think this “change” in terms of two interventions; but, are the “changes” mentioned relative to some fixed actual configuration of the system (i.e. assignment), or should one check all possible configurations? We opt here for the second interpretation. This gives us a reasonable condition for obtaining again the collapse of the team to a singleton. The idea to obtain this is to fix (to some value) all the variables that are not descendants of  $X$ . Let us call this set  $ND_X$ ; we use letter  $\mathbf{w}$  to indicate values for these variables. Write  $Fix'(\mathbf{w})$  for  $\bigwedge_{W \in ND_X} W = w$ , and  $\Xi_{x,x';y,y';\mathbf{w}}$  for

$$Fix'(\mathbf{w}) \Box \rightarrow [(X = x \Box \rightarrow Y = y) \wedge (X = x' \Box \rightarrow Y = y')].$$

We can then define total cause as:

$$T \models TC(X; Y) \text{ iff there are } x \neq x', \mathbf{w} \in T(ND_X), y \neq y' \text{ such that } T \models \Xi_{x,x';y,y';\mathbf{w}}.$$

This, once more, can be written as a single formula:

$$T \models \bigsqcup_{x \neq x', \mathbf{w} \in T(ND_X), y \neq y'} \Xi_{x,x';y,y';\mathbf{w}}.$$

Things work differently in the probabilistic framework. Notice that in the language  $\mathcal{PCD}$  we have (abbreviated) formulas of the form

$$X = x \square \rightarrow Pr(Y = y) = \epsilon$$

saying that an intervention fixing the value of  $X$  at  $x$  will affect in a certain way the probability distribution of  $Y$ . If a conjunction of the following form

$$\Theta_{x,x';y;\epsilon,\epsilon'} : (X = x \square \rightarrow Pr(Y = y) = \epsilon) \wedge (X = x' \square \rightarrow Pr(Y = y) = \epsilon')$$

holds for two distinct values  $\epsilon, \epsilon'$ , we can conclude that  $X$  is a total cause in the (probabilistic) sense given by Woodward. More precisely, we can treat probabilistic total causation as follows.  $X$  is a **(probabilistic) total cause** of  $Y$ ,  $T \models PTC(X; Y)$ , iff

$$\text{there are } x \neq x', y, \epsilon \neq \epsilon' \text{ such that } T \models \Theta_{x,x';y;\epsilon,\epsilon'}.$$

Compare this with the definition of deterministic total cause; the clauses have been considerably simplified. This follows from the fact that, differently from deterministic atoms of the form  $Y = y$ , the probabilistic atoms  $Pr(Y = y) = \epsilon$  are not flat; they instead express a global property of the team. Therefore, there is no need to use the trick that was applied in the deterministic case (fixing counterfactually the nondescendants of  $X$ , so that the team collapses to a single assignment). In this probabilistic context, total causation is genuinely a property of teams, not of single assignments.

Observing that the probability of any formula in the language is an integer multiple of the smallest unit  $u = \frac{1}{\text{card}(T^-)}$ , and that probabilities range within  $[0, 1] \cap \mathbb{R}$ , we can also explicitly define (in a non-uniform way)  $T \models PTC(X; Y)$  as:

$$\bigsqcup_{x \neq x', y, m \neq n, m, n = 0.. \text{card}(T^-)} \Theta_{x,x';y;mu,nu}.$$

## 7.2 Higher-order effects

In general, a variable  $X$  might be thought to have some kind of “causal effect” on another variable  $Y$  if intervening in some way on  $X$  can change the state of variable  $Y$ . In the kind of language that is usually applied to describe a causal model, the “state” of a variable  $Y$  is completely summarized by a formula of the form  $Y = y$ , asserting that  $Y$  is taking a specific value  $y$ . But, in a team, it may happen that a variable  $Y$  satisfies *none* of such formulas. Yet there are properties that can be ascribed or not to a variable  $Y$  in  $T$ , and that might cease or begin to hold after an intervention on  $X$ . One example are statements on the probability distribution of  $Y$ . Or, there might be some deterministic property, say  $\psi(Y)$ , which does not hold in  $T$ , but holds after an intervention that reduces the range of  $Y$ -values occurring in the team. This suggests that “ $X$  is a total cause of  $Y$ ” might be redefined, in contexts more general than the usual causal modeling, as the statement that “there is an intervention on  $X$

that changes some property of  $Y$ ”; and similarly for other notions of causation. Therefore, the arguments above rather show that the notion of direct cause is special, in that its evaluation over a team can be reduced to evaluation over causal models.

Our formalism allows many more counterfactuals than those that typically used in the definition of causal notions; it allows us to ponder, for example, what effects an intervention on a variable has on a *set* of variables. Let us consider on an example which is quite natural in the team context, but is unlikely to emerge in the usual discussions on interventionist causality. Consider a causal team  $T$  of the form

X→Y		Z
1	2	1
2	3	1

Clearly  $T \not\models (Z; Y)$ . However, the intervened team  $T_{X=1}$

X→Y		Z
1	2	1

is such that  $T_{X=1} \models (Z; Y)$ . Said otherwise,  $T \models X = 1 \square \rightarrow (Z; Y)$ . So, intervening on  $X$  has changed a property of variable  $Y$  (that of being functionally dependent, or not, on  $Z$ ). Clearly, such a team property is completely different from those that are usually considered in discussions about causation; and they may be thought as properties of the *set* of variables  $\{Y, Z\}$ . Said otherwise, intervening on  $X$  may affect the set  $\{Y, Z\}$ .

We can construct similar examples also in the language of probabilities, without need for counterfactuals nor selective implications. Consider the following multiteam with variables  $X, Y$  and no arrows, and the multiteam obtained from it by applying  $do(X = 0)$ :

S:	<table> <thead> <tr> <th>X</th> <th>Y</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> </tr> <tr> <td>0</td> <td>1</td> </tr> <tr> <td>1</td> <td>2</td> </tr> </tbody> </table>	X	Y	0	0	0	1	1	2
X	Y								
0	0								
0	1								
1	2								

$S_{X=0}$ :	<table> <thead> <tr> <th>X</th> <th>Y</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> </tr> <tr> <td>0</td> <td>1</td> </tr> <tr> <td>0</td> <td>2</td> </tr> </tbody> </table>	X	Y	0	0	0	1	0	2
X	Y								
0	0								
0	1								
0	2								

Notice that the intervened team contains a new assignment  $\{(X, 0), (Y, 2)\}$ . Therefore, the probability of formula  $X = 0 \wedge Y = 2$  raised from 0 (before the intervention) to 1/3 (after the intervention). In this case, intervening on  $X$  affected the set  $\{X, Y\}$ , in the sense that it changed a property of the relationship between  $X$  and  $Y$ .

### 7.3 Invariance

We want to make some remarks about one methodological aspect of Pearl’s approach to causation. In his causal models, arrows represent the statement, or guess, that some functional dependencies are invariant under *most* of the interventions that can be carried out on the system (a structural equation  $X :=$

$f_X(PA_X)$  is invariant under all interventions that do not act on  $X$ ). Graphs therefore draw a line between some of the dependencies occurring in the data, which might be considered either 1) contingent or 2) due to a third common cause, and those that 3) have to be considered causal dependencies. But it must be immediately observed that there may be dependencies that are invariant in the sense specified above, and yet are not represented by arrows in the causal graph.

Similarly, in our causal teams, the graph  $G(T)$  encodes a set of invariant functional dependencies of the form  $\text{=}(PA_Y; Y)$ , but the causal team could still sustain invariant dependencies which are not explicitated by  $G(T)$  (in this case, we say that a functional dependency  $\text{=}(X_1, \dots, X_n; Y)$  is *invariant in  $T$*  if it holds after any intervention on  $T$ <sup>13</sup>); call  $G_{inv}(T)$  the graph which represents all the invariant dependencies. Notice that while the structural equation for  $Y$  is not invariant under interventions that act on  $Y$ , the functional dependency  $\text{=}(X_1, \dots, X_n; Y)$  is invariant under such interventions, although its arrows are removed from the graph. What happens is this (we consider for simplicity an intervention over only one variable  $Y$ ): if  $T \models \text{=}(X_1, \dots, X_n; Y)$ , then  $(X_i, Y) \in G(T)$  but  $(X_i, Y) \notin G(T_{Y=y})$ . But since  $Y$  is constant in  $T_{Y=y}$ , no further intervention can disrupt the dependency  $\text{=}(X_1, \dots, X_n; Y)$ ; therefore  $(X_i, Y) \in G_{inv}(T_{Y=y})$ . This shows that there may be invariant dependencies in a causal team that are not represented by arrows of the graph.

In general, we can isolate three graphs of interest:  $G(T) \subseteq G_{inv}(T) \subseteq G_{cont}(T)$ , where  $G_{cont}(T)$  represents *all* the functional dependencies in  $T^-$ , including those that are not invariant. (Furthermore, as shown in a previous subsection, it might be of interest to consider the subgraphs of  $G(T)$  that represent direct causes, total causes and probabilistic total causes). The point is that the choice of  $G(T)$  determines what an intervention is/does; in turn, the set of interventions thus defined may determine some further invariant dependencies (set  $G_{inv}(T)$ ), those that hold both in the initial team and in all the intervened ones. The arrows of this new graph might happen to be inappropriate to capture a notion of “direct” or also of “indirect” causation.

It might be interesting to ask whether this kind of invariance is definable in our languages; the question makes sense not only for functional dependencies, but for any property definable by a formula. The answer is positive; we can give a model-dependent definition. A formula  $\psi$  is invariant for  $T$  if:

$$T \models INV(\psi) := \bigwedge \{ \mathbf{Z} = \mathbf{z} \square \rightarrow \psi \}$$

where the conjunction ranges over all consistent conjunctions of the form  $\mathbf{Z} = \mathbf{z}$ . Notice that this is not a single formula, but a scheme of formulas of  $\mathcal{CD}$ ; which formula is actually picked depends on the domain of the causal team  $T$  and

<sup>13</sup>One might raise the protest that it should be required that  $\text{=}(X_1, \dots, X_n; Y)$  be invariant under any *sequence* of interventions. However, the importation law we have proved in earlier sections shows that for any sequence of interventions, there is a single intervention generating the same modified causal team. Notice, furthermore, that in this context the actual team  $T$  can be thought of as the result of a null intervention.

the ranges of its variables. We do not know whether a uniform definition can be given, or if instead adding invariance statements (for sentences, or just for dependencies) might increase the expressivity of our languages.

## 8 Conclusions

We have shown how to generalize team semantics in a way that it may incorporate the kind of counterfactual and causal reasoning which is typical of recursive, possibly nonparametric structural equation models. We briefly explained how the semantics can be extended to cover also nonrecursive, parametric models with unique or no solutions. We introduced languages for causal discourse (both deterministic and probabilistic) over causal teams; we showed that these languages are in some ways more general than notations typically used in the literature (in that, for example, they may distinguish pre- and post-intervention observations).

We explored somewhat systematically the logic underlying deterministic and probabilistic languages for interventionist counterfactual reasoning when these are applied to recursive, parametric causal teams; we also shortly sketched what challenges arise in the treatment of the nonparametric case, and proposed tentative clauses for nonclassical semantical notions that naturally arise in this case. Finally, we analyzed how some of the notions of cause and invariance that arise from manipulationist theories of causation can be expressed and generalized in our framework.

These initial explorations of causal team semantics leave space for further investigation. On the side of causal modeling, parametric nonrecursive and nonparametric causal teams certainly deserve a more extensive study than what was sketched here (we find it unlikely, instead, that the study of the nonrecursive nonparametric teams might be fruitful). Another important case in causal modeling, which we only briefly grazed in this paper, are models with unobserved variables; defining a causal team semantics for this case would surely require some deep rethinking of the framework. Similar challenges might be posed by the study of forms of interventions beyond the “surgical” ones considered here.

On the logical side, it would be important to develop tools for the comparison of expressive power of distinct languages, and to investigate systematically whether the introduction of counterfactuals leads to significant new definability results. The inferential aspects of causal languages should also be further investigated; in particular, complete deduction systems for deterministic languages with (in)dependence atoms are missing; and the case of probabilistic languages is still unexplored. Further directions for investigation are comparisons with previous notion of counterfactuals, and the study of the modal structure induced by interventions on teams.

The present work has been developed within the Academy of Finland Project n.286991, “Dependence and Independence in Logic: Foundations and Philosophical Significance”.

## References

- [1] J. C. Bradfield, S. B. Fröschle, Independence-friendly modal logic and true concurrency, *Nordic J. of Computing* 9 (2) (2002) 102–117.
- [2] R. Briggs, Interventionist counterfactuals, *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 160 (1) (2012) 139–166.
- [3] P. Cameron, W. Hodges, Some combinatorics of imperfect information, *Journal of Symbolic Logic* 66 (2001) 673–684.
- [4] N. Cartwright, *How the Laws of Physics Lie*, Oxford University Press, 1983.
- [5] N. Cartwright, *Nature’s Capacities and Their Measurement*, Oxford University Press, 1989.
- [6] J. Corander, A. Hyttinen, J. Kontinen, J. Pensar, J. Väänänen, A logical approach to context-specific independence, in: *Logic, Language, Information, and Computation - 23rd International Workshop, WoLLIC 2016, Puebla, Mexico, August 16-19th, 2016. Proceedings, 2016.*
- [7] A. Durand, M. Hannula, J. Kontinen, A. Meier, J. Virtema, Approximation and dependence via multiteam semantics, in: *Proceedings of the 9th International Symposium on Foundations of Information and Knowledge Systems*, vol. LNCS 9616, Springer, 2016.
- [8] D. Galles, J. Pearl, An axiomatic characterization of causal counterfactuals, *Foundations of Science* 3 (1) (1998) 151–182.
- [9] P. Galliani, Inclusion and exclusion dependencies in team semantics - on some logics of imperfect information, *Annals of Pure and Applied Logic* 163 (1) (2012) 68–84.
- [10] E. Grädel, J. Väänänen, Dependence and independence, *Studia Logica* 101 (2013) 399–410.
- [11] J. Y. Halpern, Axiomatizing causal reasoning, *J. Artif. Int. Res.* 12 (1) (2000) 317–337.
- [12] D. M. Hausman, *Causal asymmetries*, Cambridge University Press, 1998.
- [13] J. Hintikka, G. Sandu, Informational independence as a semantical phenomenon, in: J. E. Fenstad et al (ed.), *Logic, Methodology and Philosophy of Science VIII*, Elsevier Science Publishers B.V., 1989, pp. 571–589.
- [14] W. Hodges, Compositional semantics for a language of imperfect information, *Logic Journal of the IGPL* 5 (1997) 539–563.
- [15] W. Hodges, Some strange quantifiers, in: J. Mycielski, G. Rozenberg, A. Salomaa (eds.), *Structures in Logic and Computer Science. Lecture Notes in Computer Sci.* vol.1261, Springer-Verlag London, UK, 1997, pp. 51–65.

- [16] T. Hyttinen, G. Paolini, J. Väänänen, Quantum team logic and Bell's inequalities, *Review of Symbolic Logic* 08(04) (2015) 722–742.
- [17] D. Lewis, Causation, *Journal of Philosophy* 70 (17) (1973) 556–567.
- [18] D. Lewis, *Counterfactuals*, Oxford: Blackwell Publishers, 1973.
- [19] D. Lewis, Probabilities of conditionals and conditional probabilities, in: *Ifs*, Springer, 1976, pp. 129–147.
- [20] D. Lewis, Counterfactual dependence and time's arrow, *Noûs* (1979) 455–476.
- [21] A. L. Mann, G. Sandu, M. Sevenster, Independence-Friendly Logic - a Game-Theoretic Approach, vol. 386 of London Mathematical Society lecture note series, Cambridge University Press, 2011.
- [22] J. Neyman, Sur les applications de la théorie des probabilités aux expériences agricoles: Essai des principes, *Roczniki Nauk Rolniczych* 10 (1923) 1–51.
- [23] J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, New York, NY, USA, 2000.
- [24] D. B. Rubin, Estimating causal effects of treatments in randomized and nonrandomized studies., *Journal of educational Psychology* 66 (5) (1974) 688.
- [25] T. Sider, *Logic for philosophy*, vol. 304, Oxford University Press, Oxford, 2010.
- [26] T. Skolem, Logisch-kombinatorische untersuchungen ber die erfüllbarkeit oder bewiesbarkeit mathematischer sätze nebst einem theorem ber dichte mengen, in: *Selected Works in Logic*, Universitetsforlaget, 1920, p. 136.
- [27] P. Spirtes, C. Glymour, R. N. Scheines, *Causation, Prediction, and Search*, vol. 81 of *Lecture Notes in Statistics*, Springer New York, 1993.
- [28] R. C. Stalnaker, A theory of conditionals, *Americal Philosophical Quarterly* (1968) 98–112.
- [29] R. H. Strotz, H. O. Wold, Recursive vs. nonrecursive systems: An attempt at synthesis (part i of a triptych on causal chain systems), *Econometrica: Journal of the Econometric Society* (1960) 417–427.
- [30] P. Suppes, *A probabilistic theory of causality*, North-Holland Publishing Company Amsterdam, 1970.
- [31] T. Tulenheimo, On IF modal logic and its expressive power, in: P. Balbiani, N.-Y. Suzuki, F. Wolter, M. Zacharyashev (eds.), *Advances in Modal Logic* 4, King's College Publications, 2003.



- [32] J. Väänänen, *Dependence Logic: A New Approach to Independence Friendly Logic*, vol. 70 of London Mathematical Society Student Texts, Cambridge University Press, 2007.
- [33] J. Väänänen, Modal dependence logic, in: K. Apt, R. van Rooij (eds.), *New Perspectives on Games and Interaction*, vol. 5 of Texts in Logic and Games, Amsterdam University press, 2008, pp. 237–254.
- [34] J. Väänänen, *The Logic of Approximate Dependence*, Springer International Publishing, Cham, 2017, pp. 227–234.
- [35] T. Verma, J. Pearl, Causal networks: Semantics and expressiveness, in: R. D. Shachter, T. S. Levitt, L. N. Kanal, J. F. Lemmer (eds.), *Uncertainty in Artificial Intelligence*, vol. 9 of Machine Intelligence and Pattern Recognition, North-Holland, 1990, pp. 69 – 76.
- [36] G. H. von Wright, *Explanation and Understanding*, Cornell University Press, Ithaca, NY, 1971.
- [37] J. Woodward, Explanation, invariance, and intervention, *Philosophy of Science* 64 (1997) S26–S41.
- [38] J. Woodward, Probabilistic causality, direct causes and counterfactual dependence, *Stochastic causality* (2001) 39–63.
- [39] J. Woodward, *Making Things Happen*, vol. 114 of Oxford Studies in the Philosophy of Science, Oxford University Press, 2003.
- [40] F. Yang, Uniform definability in propositional dependence logic, *The Review of Symbolic Logic* 10 (1) (2017) 65–79.
- [41] F. Yang, J. Väänänen, Propositional logics of dependence, *Annals of Pure and Applied Logic* 167 (7) (2016) 557 – 589.
- [42] F. Yang, J. Väänänen, Propositional team logics, *Annals of Pure and Applied Logic* 168 (7) (2017) 1406 – 1441.