Info Intervention

Gong Heyang
University of
Science and Technology of China
gonghy@mail.ustc.edu.cn

Zhu Ke
University of HongKong
mazhuke@hku.hk

Abstract

We highlight the information transfer and process understanding of causal models by proposing the info intervention, which intervening the information on the output edges of a node. We point out issues of existing definition of *do intervention* (also known as "surgical" or "atomatic" or "perfect" intervention) and other notations for causation. We show that *info intervention* are competitive in formalizing causal queries, identification of cause-effect, and it can solve (or alleviate) some issues in the existing causal modeling frameworks.

1 Introduction

The hard open problems of machinine learning are intrinsically related to causality [24]. Three fundamental obstacles are standing in our way to strong AI, including robustnesss(or adaptability), explainability and lacking of understanding cause-effect connections. Pearl asserts that all these obstacles can be overcome using causal modeling tools, in particular, causal diagrams and their associated logic [20]. How can machines represent causal knowledge in a way that would enable them to access the necessary information swiftly, answer questions correctly, and do it with ease, as a human can? This question, which is referred as "mini-Turning test" for AI, has been claimed as Pearl's life work [21].

Pearl propose the framework of structural causal models(SCM) which deploys three parts, including graphical models, structural equations, and counterfactual and interventional logic. Graphical models serve as a language for representing what we know about the world, counterfactuals help us to articulate what we want to know, while structural equations serve to tie the two together in a solid semantics. In the meanwhile, competitive causal notations and frameworks, such as Dawid's regime indicator framework and Rubin's potential outcome, are also used with prons and cons [14, 16, 17]. These different approaches to formalizing causal inquiries which, despite subtle differences, all build on a probabilistic graphical representation of the problem at hand.

Back to the very nature of modeling, a model is an idealized representation of reality that hightlights some apects and ignores others. Wheeler claims all things physical are information-theoretic in origin [1] and Bernhard highlights the information processing aspect of causal modeling in his work [24]. The main focus of current causal models are modeling the distangled form of the joint distritution, while there are certain degree of ignoring the information processing aspect.

Pearl's *do*-operator is one of the most important causal semantics or notations during the past three decades of the *causal revolution*. The haunting "confounding" problem has been demystified with *do*-calculus and many other achievements, but still criticisms of this notation exist, especially on the empirical interpretation when applied to non-manipulable variables such as race, obesity, or cholesterol level [19]. Physically, Einstein's theory of relativity tells us that time, length, and quality can change in different coordinate systems, but the causal relationship of events remain. This provides us an intuition of invariant causal mechanism and strong objection to "removing of causal mechanisms" or "a minimal change on mechanisms" which is exactly the meaning of Pearl's intervention. Moreover, intensive theoretical effects torward the development of SCM have been made in the recent years. One central topic is how to do causal inference beyond DAGs. Some of them focus on the theoretical aspects of cyclic SCM [4, 9, 10], and some of

[1] It from bit symbolizes the idea that every itme of the physical world has at bottom...an immaterial source and explanation...that all things physical are information-theoretic in origin and that this is a participatory universe. —- John A. Wheeler

them connect equilibrums to feedbacks/cycles [2, 3, 15]. However, many measure-theoretic and other complications arise in the presence of cyclesxu, and these difficulties may explain why in most of the SCM literature so far acyclicity has been assumed, even though many causal systems in nature involve feedbacks/cycles.

The main idea of regime indicator is to consider the 'seeing' and 'doing' as two types of regimes, a natural one and a set of interventional ones [5]. This notation has advantange in representing soft interventions which do not fix a variable at a value, but just 'nudge' it, adding a random error, or somehow shift its distribution [7] and dynamic intervention. The property of stability [6] or invariance [11] of conditional distributions across regimes is required to predict effects of interventions from observational data and thus making regime less popular. The proposing of this notation reveals some weakness of do intervention in representing variaous interventions used in empirical applications. Robins' potential outcome is defined as the hypothetical value of a variable if we force an observed variable to a given value which is essentially based on do intervention. An attractive feature of Robins' approach is that it largely avoids making counterfactual independence assumptions that are experimentally untestable and potential outcomes can be combined with graphs though may not be immediately obvious [16,22,23]. Much of the work in economics is closer in spirit to the potential outcome framework [13].

All these approaches to causality can be combined with graphs and focus on modeling the joint distribution associated with the graph. Each approach has some minor issues, but we are not going to solving issues within the original framework. In contrast, we will highlight the information-theoretic aspect of causal models by proposing a new concept of info intervention to go beyond traditional causal frameworks [2]. We will show that some issues can be circumvented and some can be alleviated.

In this articles, we first will review the frameworks of causal models and discuss issues of graphical or structural causal models. Then we propose the info intervention as a substitute for Pearl' s do intervention, emphasis the information aspects of causal models, and show how (or exhibit potentials) to solve (or alleviate) those issues. After that, we use the simplest case for causal mdoels — directed acyclic graphs to explain info intervention and assocated three-level question with example. The next part of this article presents the causal calculus for info intervention and its applications. We conclude with a discussion of how our work is related to the interpretation of caulity in a view of information transfer and process.

[2]We are NOT saying traditional causal framework must be replaced or not related to information processing, rather we mean they are not information-theoretic enough

2 Preliminaries

The notation of causality has been much examined, discussed and debated in science and philosophy over many centuries. The randomized controlled experiment(RCT) used to be a "golden" standard for causal inference, while causal inference for observational data usually relies on graphical assumptions of the underlying data generating process. Three main frameworks for causality have been developed with there unique concepts of do-calculus, potential outcomes, and regime indicator.

Definition 2.1 (Structural Causal Model)**.** *A* structural causal model (SCM) *by definition consists of:*

1. *A set of nodes $V^+ = U \dot\cup V$ [3], where elements of V correspond to endogenous variables and elements of U to exogenous(or Latent) variables,*
2. *An endogenous/exogenous space $\mathcal{X}_v$ for every $v \in V^+$, $\mathcal{X} := \prod_{v \in V^+} \mathcal{X}_v$,*
3. *A product probability measure $\mathbb{P} := \mathbb{P}_U = \otimes_{u \in U} \mathbb{P}_u$ on the latent space $\prod_{u \in U} \mathcal{X}_u$.*
4. *A directed graph structure $G^+ = (V^+, E^+)$, with a system of structural equations $f_V = (f_v)_{v \in V}$:*

$$f_v : \prod_{s \in \mathrm{Pa}^{G^+}(v)} \mathcal{X}_s \to \mathcal{X}_v,$$

where $\mathrm{Ch}^{G^+}(U) \subseteq V$ and all functions f_V are measurable.

The SCM can be summarized by the tuple $M = (G^+, \mathcal{X}, \mathbb{P}, f)$. G^+ is referred as the augmented functional graph *while the* functional graph *which includes only endogenous variables, denoted as G .*

To model an action $do(X = x)$ one performs a mininal change necessary for establishiing the antecedent $X = x$, while leaving the rest of the model intact. This calls for removing the mechanism equation that nominally assigns values to variable X, and replacing it with a new equation, $X = x$, that enforces the intent of the specified action. Formally, the do intervention is defined by:

Definition 2.2 (do intervention)**.** *Given an **SCM** $M = (G^+, \mathcal{X}, \mathbb{P}, f)$, $I \subseteq V$, the* do intervention [4] $do(X_I =$

[3]$U \dot\cup V$ means the disjoint union of sets U and V.

[4]A do intervention is usually imposed on endogenous variables.

x_I) *maps* M *to the intervened model* $M_{do(X_I=x_I)} = (G^+, \mathcal{X}, \mathbb{P}, \tilde{f})$ *where*

$$\tilde{f}_i(X_U, X_V) := \begin{cases} x_i & i \in I \\ f_i(X_U, X_V) & i \in V \setminus I\,. \end{cases}$$

For the special case of DAGs, causal semantics could be defined without any complications in the following way:

Definition 2.3 (Causal DAG)**.** *Consider a DAG* $G = (V, E)$ *and a random vector* $X = (X_1, ..., X_K)$ *with distribution* p. *Then* G *is called a causal DAG for* X *if* p *satisfies the following:*

(i) p *factorizes, and thus is Markov, according to* G, *and*

(ii) for any $A \subset V$ *and any* $\tilde{x}_A, x_B$ *in the domains of* X_A, X_B, *where* $B = V/A$,

$$p(x; do(\tilde{x}_A)) = \prod_{k \in B} p(x_k | x_{pa(k)}) \prod_{i \in A} \mathbb{I}(x_j = \tilde{x}_j)$$

The major obstacle to drawing causal inference from data known as "deconfounded" is demystified through a graphical criterion called "backdoor" for causal DAGs. And for models where the backdoor criterion does not hold, a symbolic engine called "do-calculus"(Pearl's 3-rules of causal calculus) is available.

There are other two notations for causal modeling. The key idea of *regime indicator* is to consider 'seeing' and 'doing' as two types of regimes, a natural one and a set of interventional ones. $p(x; \sigma = s) = p(x; s)$. Here, regimes refer to external circumstances under which we expect some aspects of the joint distribution of X to differ. The other notation in the context of causal inference uses *potential(counterfactual) outcomes*. We consider some causal effect of action A on an outcome Y, we define the potential outcome $Y(\tilde{a})$ to be the value of Y that we would observe if A were set (forced) to $\tilde{a}$. This approach is essentially based on perfect(atomatic) interventions. The condition for cause-effect estimation from observational data(counterpart of backdoor criteria) for potential outcome is *exchangeability*.

Definition 2.4 (Exchangeability or ignorability)**.** *Exchangeability means that the counterfactual outcome and the actual treatment(action) are independent, or* $Y^a \perp\!\!\!\perp A$, *for all values* a.

By the **common cause principle** that for two statistically dependent variables there will always be a third cause variable explains there dependence, we can assume the independence of exogenous variables. In other words,

Definition 2.5 (Causal Sufficient)**.** *If all dependencies among variables are captured by the causal model, then we call it causally sufficient.*

Pros and cons are among the above causal notations, we now present our semantic of intervention under the information view of causal models.

3 Info Intervention for Understanding Causality with Information View

Causal questions, such as what if I make something happen, can be formalized by do operator, but still, controversial on empirical understanding for do intervention exist. In many settings, a do intervention which forces some variable to a given value is somewhat idealized or hypothetical. How would one manipulate variables such as race, obesity, or cholesterol level and how would one, for instance, fix the dietary fat intake or BMI of a person exactly at a given value?

Moreover, the notation suggests that the manner in which a variable is manipulated is irrelevant to the intervened causal model. However, in practice it may matter whether a medical treatment is, for example, given orally or as an injection [5]. For greater generality, we may therefore want to consider a possibly larger and more detailed set S of different regimes describing different circumstances under which a system might be observed and manipulated. Each regime then induces a different probability measure for the joint distribution of X. In short, both how and what actions are taken have effectiong on the result [16].

In order to achieve a better represention of reality, our modified interventional semantic is formalized with the information view of causal models rather than effects of actions:

Principle 3.1 (Info view of causal models)**.** *Axioms of our causal model of extended word:*

i) Directed edges are used to represent cause-effect relationships.

ii) Cause events precede effect events.

iii) Every node/variable in the causal model represents a causal mechanism which accepts informations from its input edges and sends out information of its default state to the output edges.

Before we go any further, we now just present the definiton of *info intervention* for causal diagrams.

Definition 3.2 (Info intervention)**.** *Given an* ***SCM*** $M = (G^+, \mathcal{X}, \mathbb{P}, f)$, $I \subseteq V$, *the* ***info intervention*** $\sigma(X_I =$

[5] Some researchers refer it as different version of treatment [12].

x_I) (or in short $\sigma(x_I)$) maps M to the intervened model $M_{\sigma(X_I=x_I)} = (G^+, \mathcal{X}, \mathbb{P}, \tilde{f})$, $\tilde{f} := f(X_U, \tilde{X}_V)$ where

$$\tilde{X}_i := \begin{cases} x_i & i \in I \\ X_i & i \in V \setminus I\,. \end{cases}$$

the causal graph of info intervened SCM $M_{\sigma(X_I=x_I)}$ is the graph that removes all output edges from I in G^+.

To determine the causal content [25] of an evolving system we perform perturbations. In a network, for example, a perturbation can be deleting a node or deleting a link. The *do* operator only consider the situation of deleting a node but not a link. Our info intervention refers to deleting all output links or edges rather than mechanisms represented by a node. The critical difference between definitions of *do* intervention and info intervention is that the causal mechanisms do not change. Instead, the hypothesis minimal change by info intervention is the information seeding out to the output edges of a node. Particularly, structural equations f_V changed for *do* intervention, in contrast, it keeps the same for info intervention. Let's see an example.

Example 3.3. *For a SCM with treatment T, confounders X and outcome Y, see Figure 1. The structural equations are:*

$$\begin{cases} X = f_X(\epsilon_X) \\ T = f_T(\epsilon_T, X) \\ Y = f_Y(X, T) \end{cases}$$

The do intervened SCM is:

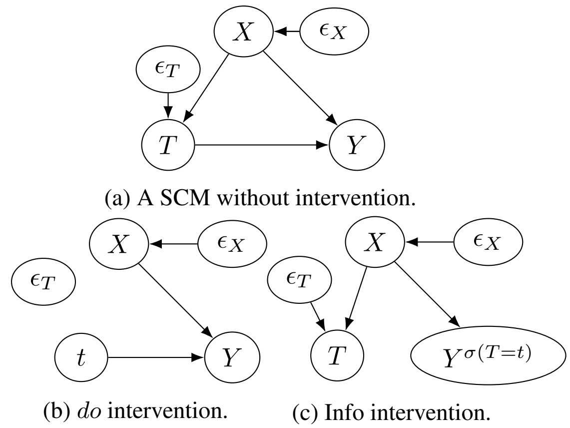
$$\begin{cases} X = f_X(\epsilon_X) \\ T = t \\ Y = f_Y(X, T) \end{cases}$$

The info intervened SCM is:

$$\begin{cases} X = f_X(\epsilon_X) \\ T = f_T(\epsilon_T, X) \\ Y = f_Y(X, t) \end{cases}$$

We use $Y^{\sigma(T=t)} \triangleq f_Y(X, t)$ to denote the variable of Y recieve the information $T = t$. Then the causal model is show in Figure 1a, and the do intervened SCM with Figure 1b, adn the info intervened SCM with Figure 1c. And the functional graph of the two intervened SCM can be seen in Figure 2.

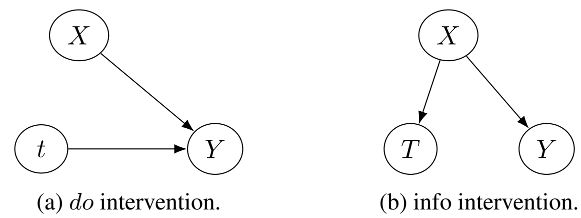
It is clear that those two intervened SCM are different from the intervened causal mechanisms f_T, f_X, f_Y. Particularly, info intervention keeps the causal mechanisms unchange while *do* intervention doesn't. Moreover, for empirical interpretation of info intervetion, we



(a) A SCM without intervention.

(b) *do* intervention. (c) Info intervention.

Figure 1: A simple SCM with treatment T, confounders X, outcome Y and two latent variables ϵ_X, ϵ_T.



(a) *do* intervention. (b) info intervention.

Figure 2: SCM with only endogenous variables.

don't have problem of treatment variables such as race, obesity or cholesterol level. For example, if T represents race and Y represent whether hired or not, then an *do* intervention questions would what if I manipulate my race to black? This causes an issue of non-manipulate variables problem. But for info intervention view, the question becomes "what if other nodes recieve an information that the race is black?". In our example, the intervened variable T becomes a constant for *do* intervention while it remains the same for info intervention.

4 Information Processing view of Causal Modeling

A model is an idealized representation of reality that highlights some aspects and ignores others. The info intervention semantic is to highlight the information processing aspect of causal model.

The phyical nature of information. In Principle 3.1, causal models are considered as systems which process information. The way of how to generate and convert forms of energy has lead to the first and second industrial revolution, while information plays the same role of energy in the recent AI revolution [24]. If we think it broadly, every physical system is about processing information even physical are information-theoretic in origin [6]. One of the most intriguing problems would be that the conservation of information might also be a consequence of symmetries [7].

Modeling physical phenomena with a set of coupled differential equations is a gold standard. It allows us to predict the future behavior of a system, to reason about the effect of interventions in the stystem, and predict statistical dependences [24]. But such models requires an intelligent human to come up with it thus restrict its usage only in certain applications. Statistical models can be viewd as a much more superficial one are widely used in almost every scientific area with data, but very limited explainability and unable to answer causal question. Causal modeling lies in between these two extremes. It's true that we human can live well without knowing causal mechansims of a liver, but what should we do when your liver goes wrong? Causal models should be convinient approximation of physical reality which could be help us fix problems when things go wrong, thus consistent with current scientific understanding of physical world matters, especially for building true intelligent systems.

Control VS Interfere. In the control system theory, say, a typical PID control systems, uses a feedback controller to make the system output to be a certain value. The widely used control system reminds us *do* intervention which forces some variables take given values should interpreted with a feedback control system, thus *do* intervention is more of controlling the state of a system with a sequence of control input signals. In contrast, our info intervention is just interfering the output information of the system at one particular timestap, which reflects the local property of causality. The issues of empirical understanding of *do*-intervention on non-manipulable variables do not exist for info intervention.

When it comes to causality, there will alway a temporal or order structure underlying the causal model which is the essential difference comparing to statistical models. For directed acyclic graphs, the temporal structure is implied by the second statement of Principle 3.1. But for feedback loops which shown up in extensive kinds of real-world problems and complex systems, the temporal sturcture is messed up. The information view of causal models can be help in this situation. In real-world situations, it spends some time for a node that accepts, processes, and sends information while the variables in feedback loops are usually measured at the same give timestap. This inconsistency between data and causal mechanisms suggest that state of a node could be affected by historical information of its cause nodes.

Info Markov property. Causal diagrams perform as parsimonious representation to encode causal knowledges transparently, access swifty and do it with ease. Conditional independence structure among variables plays a central role which induces different kind of Markov properties corresponding to a form of factorization.

Definition 4.1 (Info Markov property)**.** *For an info causal model G, every endogenous variable X is indenpent of its non-descents given informations from their parents, then we G satistfies info Markov property.*

For the causal DAGs, this info markov property is the same with causal Markov property, but when it comes to causal graphs with feedback loops where nodes may accept historical input informations, these two Markov property can be different.

Our information view of the causal model of extended world is shown in Figure 3 that data information consist of enviroment information and mechanism information, which is similar to SCM framework used in [1], and it can be dated back to the concept of signal-flow graph in the 1950s while many controversials on causal interpretation of it. There are many details about this information view of causal models, we start from the simplest case of DAGs.

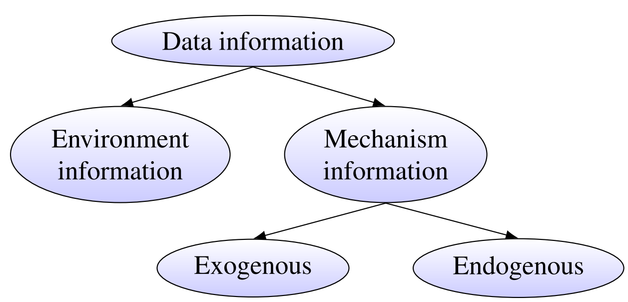


Figure 3: An event or record includes informations from the environment and causal mechanisms.

5 Info Causal DAGs

Similar to the definition of causal DAGs, we have:

Definition 5.1 (Info causal DAG)**.** *Consider a DAG $G = (V, E)$ and a random vector $X = (X_1, ..., X_K)$ with dis-*

[6] It from bit symbolizes the idea that every itme of the physical world has at bottom...an immaterial source and explanation...that all things physical are information-theoretic in origin and that this is a participatory universe. —- John A. Wheeler

[7] Emmy Noether claims that energy conservation is due to a symmetry of the fundamental laws of physics: they look the same no matter how we shift the time, in present, past, and future. Einsten was relying on covariance principles when he established the equivalence between energy and mass.

tribution p. Then G is called a info causal DAG *for X if p satisfies the following:*

(i) p factorizes, and thus is Markov, according to G,

(ii) for any $A \subseteq V$ and any $\tilde{x}_A$ in the domains of X_A,

$$p(x; \sigma(X_A = \tilde{x}_A)) = \prod_{k \in V} p(x_k | x^*_{pa(k)})$$

where $x^*_k = x_k$ if $k \notin A$ else $\tilde{x}_k$.

The causal graph of intervened SCM is defined as removing all input edges of those corresponding intervened variables. In contrast, intervened causal graph can be induced by removing all output edges for info intervention. Since causal DAGs are a special case of SCMs, the *info intervened causal graph* is just removing all output edges from info intervened varaibles, which is equivalent to SWIGs [22] removing constant splited nodes. .

Example 5.2 (A real world example)**.** *We consider domain variables* Exercise, Cholesterol, Occupation, Income, Diet*, and assumming a causal relationship among them in Figure 4.*

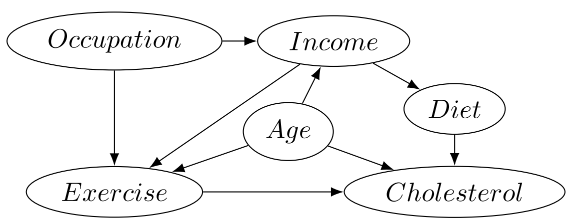


Figure 4: Causal effect of exercise on cholesterol level.

The three-level causal hierarchy. Causal questions can be classified into three-level hierarchy in the sense that questions at level $i (i = 1, 2, 3)$ can be answered only if information from level $j (j > i)$ is available. Here we have our into intervention version of causal questions with the above example.

I) Association $P(y|x)$

- Typical activity: Seeing
- Typical questions: What is? How would seeing X change my belief in Y?
- Examples: What does the habit of exercise tell me about cholesterol level?

II) Intervention $P(y|\sigma(x), z)$

- Typical activity: Intervening
- Typical questions: What if? What if I manipulate the information output from X ?
- Examples: What if my age is 18, will my cholesterol level higher?

III) Counterfactuals $P(y^{\sigma(x)}|x', y')$

- Typical activity: Imagining, Retrospection
- Typical questions: Why? Was it X that caused Y?
- Examples: Was it the exercise cause me low level cholestoral?

For do intervention on non-manipulable variable age, the empirical understanding of intervening question that what if I force or manipulate my age to 18, i.e. $P(\text{Cholestoral}|do(\text{age} = 18))$, will be a problem. Although Pearl suggests that we should interpret it in other dimensions [19], there is no problem like that for info intervention. In our causal model of extended world, every node recieves information from its input edges and sends out information to its output edges, so the value of each variable is determined by informations on its input edges. The information on a specific edge is default as the state of its start node, but an info intervention on the causal model may change informations to some events.

Science thrives on standards, because standards serve (at least) two purposes: communication and theoretical focus. On one hand, the info intervention is a nice standard for communicating about causal questions with less issues than do operators which also serve the purpose. On the other hand, many of the variants of causal queries can be reduced to "do," or to several applications of "do." Theoretical results established for "do" are then applicable to those variants. Then, what about info intervention?

6 Causal Calculus for Info Intervention

The major obstacle "Confounding" to drawing causal inference from data has been demystified through a graphical criterion called "backdoor criteria". First, an adjustment formula for info intervention is given for the most common treatment-outcome case in Example 3.3.

Theorem 6.1 (Adjustment formula)**.** *For an info causal DAG G, $G^{\sigma(X=x)}$ is the info intervened causal graph. If $T \perp\!\!\!\perp Y|X$ in $G^{\sigma(X=x)}$(Fig 2b), then*

$$P(Y|\sigma(T = t))) = \sum_x P(Y|T = t, X = x)P(X = x).$$

This theory is corresponding to the backdoor criteria and adjustment formula for causal DAGs. Moreover, Pearl's 3 rules for do intervention enables us to identify a causal query, it enables us turn a causal question into a statisitcal estimation problem. Pearl's 3 rules:

1) delete a observation
2) delete a action
3) observation/action exchange

For our info intervention on info causal DAGs, we have also developed our modified version of Pearl's three rules, Our version of causal calculus reveal a more straightforward representation.

Theorem 6.2 (Three rules of info intervention)**.** *For an info causal DAG $G = (V, E)$, X, Y, Z and W are arbitrary disjoint sets of visible variables, then for info intervention $\sigma(X = x)$ (or just $\sigma(x), \sigma(X)$ when no ambiguity exist for simplicity.),*

Rule 1*(Insertion/deletion of observations)*
$P(Y|\sigma(X = x), Z, W) = P(Y|\sigma(X = x), W)$ *in the case that* $Y \perp\!\!\!\perp_d Z|W$ *in* $G^{\sigma(X=x)}$

Rule 2*(Action/observation exchange)*
$P(Y|\sigma(Z = z), X) = P(Y|Z, X)$ *in the case that* $Y \perp\!\!\!\perp_d Z|X$ *in* $G^{\sigma(Z=z)}$

Rule 3*(Insertion/deletion of actions)*
$P(Y|\sigma(Z = z)) = P(Y)$ *in the case where no causal paths connect Z and Y in G.*

The single-world intervention graph (SWIG) which is quite similar to info intervened graph with the only difference of additional splitted constant intervened nodes [22, 23]. It is a simple graphical theory unifying causal directed acyclic graphs (DAGs) and potential (aka counterfactual) outcomes via a node-splitting transformation, thus the three rules of causal calculus hold. Correspondingly, the they will hold for the info intervened model. See the proof of this theory in the appendix. Let's see a hard inferfential problem for Pearl's causal calculus but easy with the three rules of info intervention.

Theorem 6.3 (Complete Version of Info Causal Calculus)**.** *For an info causal DAG $G = (V, E)$, X, Y, Z and W are arbitratry disjoint sets of visible variables, then for info intervention $\sigma(X = x)$ (or just $\sigma(x), \sigma(X)$ when no ambiguity exist for simplicity.),*

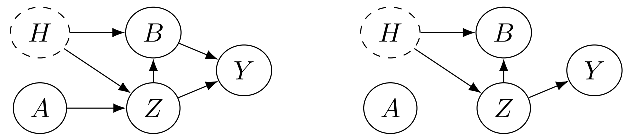
Rule 1*(Insertion/deletion of observations)*
$P(y|\sigma(x), z, w) = P(y|\sigma(x), w)$ *in the case that* $Y \perp\!\!\!\perp_d Z|W$ *in* $G^{\sigma(x)}$

Rule 2*(Action/observation exchange)*
$P(y|\sigma(x), \sigma(z), w) = P(y|\sigma(x), z, w)$ *in the case that* $Y \perp\!\!\!\perp_d Z|W$ *in* $G^{\sigma(x,z)}$

Rule 3*(Insertion/deletion of actions)*
$P(y|\sigma(x), \sigma(z), w) = P(y|\sigma(x), w)$ *in the case where no causal paths connect Z and Y in* $G^{\sigma(x)}$ [8].

[8] $Z = z$ is a cause of $Y = y$ if $P(y|\sigma(z)) \neq P(y)$ which mean information $Z = z$ have a effect on the event $Y = y$.

Example 6.4 (Independence of counterfactuals)**.** *If we have a causal DAG G, see Figure 5a, H is a hidden variable and $Y(a, b), B(a), Z(a)$ are corresponding counterfactual variables. then does the conditional independent relation $Y(a, b) \perp\!\!\!\perp B(a)|\{A, Z(a)\}$ hold?*



(a) Causal graph. (b) Info intervened graph.

Figure 5: Independence of Counterfactuals.

The answer is yes. Since the info intervened causal graph $G^{\sigma(A=a,B=b)}$ factorizes, we read conditional independence from the info intervened graph Figure 6b:

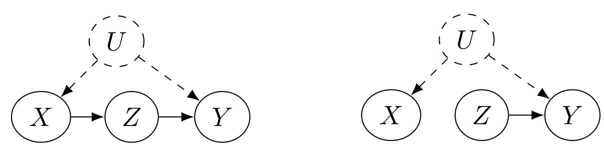
$$Y \perp\!\!\!\perp B|\{A, Z\}\, in\, G^{\sigma(A=a,B=b)}$$

By the definition of info casual DAG, $A(a, b) = A$ for that A doesn't recieve any information on eges $A \to Z, B \to Y$, and similarly for $B(a, b) = B(a), Z(a, b) = Z(a)$. Then we plug them into the above conditional relation we have $Y(a, b) \perp\!\!\!\perp B(a)|\{A, Z(a)\}$.

Similarly, we have the front-door criterion for causal inference.

Theorem 6.5 (Front-door Criteria)**.** *Z satisfies the front-door criterion when (i) Z intercepts all directed paths from X to Y , (ii) there are no unblocked back-door paths from X to Z, and (iii) X blocks all back-door paths from Z to Y . Then $P(Y = y|\sigma(X = \tilde{x}))$ can be wrtten as sum-product of conditional probabilities,*

$$P(y|\sigma(\tilde{x})) = \sum_z P(z|\tilde{x}) \sum_x P(y|z, x)P(x) \qquad (1)$$



(a) G. (b) $G^{\sigma(X=\tilde{x})}$.

Figure 6: Front-door Criteria.

Proof. The causal graph with hidden variables U of this front-door problem is given by Pearl [18] shown in Fig. 6a. By the factorization property of the info causal DAG in Fig. 6:

$$P(y, z, x, u|\sigma(\tilde{x})) = P(y|u, z)P(z|\tilde{x})P(x|u)P(u) \qquad (2)$$

Then we have

$$
\begin{aligned}
P(y|\sigma(\tilde{x})) &= \sum_x \sum_u \sum_z P(y, z, u, x|\sigma(\tilde{x})) \\
&= \sum_x \sum_u \sum_z P(y|u, z)P(z|\tilde{x})P(x|u)P(u) \\
&= \sum_z P(z|\tilde{x}) \sum_x \sum_u P(y|u, z)P(x|u)P(u) \\
&= \sum_z P(z|\tilde{x}) \sum_x \sum_u P(y|u, z, x)P(u|x)P(x) \\
&= \sum_z P(z|\tilde{x}) \sum_x P(x) \sum_u P(y|u, z, x)P(u|x) \\
&= \sum_z P(z|\tilde{x}) \sum_x P(x)P(y|z, x)P(x)
\end{aligned}
$$

□

Remark 6.6. *Our proof method is better than Pearl's method using do-calculus in Eq. 2 for that X with a non-trival distribution which faciltitates the following calculations.*

7 Discussion and Conclusion

This article proposes the info intervention which highlights the information apspect of causal model. We have not only addressed the logic behind this concept, but also developed associated causal semantics and presented our version of causal calculus.

It could be beneficial and applicable to the mind-brain debate to understand causality as information transfer for two reasons: (1) the brain is an information processing machine, and (2) the way we talk about the mental to physical relationship seems similar to the relationship between the abstract content of information and its physical realization. And two challenges are: (1) if the mental is similar to the meaning of a piece of information, how can that be shown, and (2) how would causation as information transfer be able to account for the causal efficacy of the meaning of said piece of information. Or, in other words, how we could develop an explanation that avoids the challenges of causal compatibilism and interventionism, while at the same time accounts for the causal efficacy of the mental [8]. We need mathematical notations of causation to formulize three-level causal questions in an information view of causal models, and info intervention plays the role that unifies interventionism and information transfer understanding to causality.

References

[1] Martin Arjovsky, Léon Bottou, Ishaan Gulrajani, and David Lopez-Paz. Invariant risk minimization. *arXiv preprint arXiv:1907.02893*, 2019.

[2] Tineke Blom, Stephan Bongers, and Joris M Mooij. Beyond Structural Causal Models: Causal Constraints Models. may 2018.

[3] Stephan Bongers and Joris M. Mooij. From Random Differential Equations to Structural Causal Models: the stochastic case. mar 2018.

[4] Stephan Bongers, Jonas Peters, Bernhard Schölkopf, and Joris M. Mooij. Theoretical Aspects of Cyclic Structural Causal Models. nov 2016.

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

[6] A Philip Dawid, Vanessa Didelez, et al. Identifying the consequences of dynamic treatment strategies: A decision-theoretic overview. *Statistics Surveys*, 4:184–231, 2010.

[7] Frederick Eberhardt and Richard Scheines. Interventions and causal inference. *Philosophy of Science*, 74(5):981–995, 2007.

[8] Bernard Feltz, Marcus Missal, and Andrew Sims. Free will, causality and neuroscience. 2019.

[9] Patrick Forré and Joris M Mooij. Constraint-based Causal Discovery for Non-Linear Structural Causal Models with Cycles and Latent Confounders. 2018.

[10] Patrick Forré and Joris M Mooij. Causal Calculus in the Presence of Cycles, Latent Confounders and Selection Bias. 2019.

[11] Daniel M Hausman and James Woodward. Independence, invariance and the causal markov condition. *The British journal for the philosophy of science*, 50(4):521–583, 1999.

[12] Miguel A. Hernán, John Hsu, and Brian Healy. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. *CHANCE*, 32(1):42–49, 2019.

[13] Guido W. Imbens. Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics *. Technical report, 2019.

[14] Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press, apr 2015.

[15] Steffen L. Laurizen and Thomas S. Richardson. Chain graph models and their causal interpretations. *Journal of the Royal Statistical Society Series G Statistical Methodology 64 Part 3*, 64(3):348–361, 2002.

[16] Steffen Lauritzen Marloes Maathuis, Mathias Drton and Martin Wainwright. *Handbook of Graphical Models*. CRC Press, Boca Raton, Florida : CRC Press, c2019., 1 ed edition, nov 2018.

[17] James M. Robins Migual A. Hernan. *Causal Inference: What If*. 2019.

[18] Judea Pearl. *Causality: Models, Reasoning and Inference*. 2009.

[19] Judea Pearl. On the Interpretation of do(x). Technical report, 2019.

[20] Judea Pearl. The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3):54–60, 2019.

[21] Judea Pearl and Mackenzie. *The Book of Why: The New Science of Cause and Effect*. 2018.

[22] Thomas S Richardson and James M Robins. Single world intervention graphs (swigs): A unification of the counterfactual and graphical approaches to causality. *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper*, 128(30):2013, 2013.

[23] Ts Richardson and Jm Robins. Single world intervention graphs: A primer. *Wiki.Math.Mcgill.Ca*, 2011.

[24] Bernhard Schölkopf. Causality for Machine Learning. nov 2019.

[25] Hector Zenil, Narsis A Kiani, Allan A Zea, and Jesper Tegnér. Causal deconvolution by algorithmic generative models. *Nature Machine Intelligence*, 1(1):58–66, 2019.