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 view of causality as information transfer proposing the info intervention, which intervening the sending out information. We point out issues of existing definition of Pearl's *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 notations.

1 Introduction

The hard open problems of machinine learning are intrinsically related to causality [23]. 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 [19]. 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 [20].

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 [13, 15, 16]. 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 [23]. 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 [18]. 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

[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

mechanisms" or "a minimal change on mechanisms" which is exactly the meaning of Pearl's intervention. Moreover, intensive theoretical effects toward 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 [3, 8, 9], and some of them connect equilibrums to feedbacks/cycles [1, 2, 14]. However, many measure-theoretic and other complications arise in the presence of cycles, 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 [4]. 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 [6] and dynamic intervention. The property of stability [5] or invariance [10] 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 [15, 21, 22]. Much of the work in economics is closer in spirit to the potential outcome framework [12].

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 its own issues, here we are not working on those issue within the original framework. In contrast, we will interpret causality as information transfer to highlight the information-theoretic aspect of causal models by proposing a new concept of info intervention to go beyond traditional causal notations. We then will see that some existing issues can be circumvented or 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 semantic , 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 using the $\sigma(\cdot)$ operator corresponding to $do(\cdot)$ operator, and examples of application with this information causal calculus formulas. 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 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 $V^+ = U \dot\cup V$ [2], 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.

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

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 [3] $do(X_I = x_I)$ *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.*

We have reviewed the main causal notations currently used by researchers. Now it's time for us to present our new notations for causal models.

3 Info Intervention

Causality as Information Transfer. From the philosophical point of view, the study of the concept of causality started with Aristotle (384 bc322 bc) who proposed four different types of causes. David Hume initiated the modern approach of causality. He recognized the importance of causal beliefs for human understanding. The modern concept of causality has been deeply influenced by physics and psychology during the 20th century and has a deep impact on

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

causality in neurosciences. "law-like" causality and interventionalist causality are among the most popular interpretations. There are two main reasons to think that understanding causality as information transfer could be beneficial and applicable to the mind-brain debate: (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 [7].

What is causality as information transfer. Collier himself defines it as the transfer of a particular quantity of information from one state of a system to another. Physical causation is a special case in which physical information instances are transferred from one state of physical system to another. (Collier, 1999, p. 215). For the Causal Exclusion question: was it the flashing light or the belief that it signifies the low salt level that made us go and add the salt to our dishwasher? we think the belief that it signifies the low salt level is the cause of our action.

Problems of do-intervention. 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 [4]. 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 [15]. In order to avoid those problems and achieve a better represention of reality, we could use causal semantics and notations based on understanding causality as information transfer.

Now we presents our main ideas:

Principle 3.1 (Causality as Information Transfer)**.** *For causal model in view of causality as information transfer:*

i) If the information of a node X has an effect on another node Y while fix other nodes, then X is a direct cause of Y.

ii) Directed edges are information channels which can accept information from its input node.

iii) Nodes are causal mechanisms to process the information accepts from its input edges.

We give the definiton of *info intervention* for SCMs and Causal DAGs.

For SCMs, we could define an intervention output information of a node. The formal definiton is given below:

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

For interventiaonal causality, to determine the causal content [24] of an evolving system we perform perturbations. In a network, for example, a perturbation can be deleting a node or deleting a link. The $do(X = x)$ intevention, which removes the causal mechanism represented by the node from the model, can be consider as deleting of a node but not a link. Our info intervention $\sigma(X = x)$, which force the output edges of node X accept the fixed information $X = x$, can be interpreted as cutting off information transfer form node X to its children but keeping the mechanisms represented by X. 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 accepted by the output edges of a node. In other words, structural equations f_V changed for do intervention, in contrast, it keeps the same for info intervention. Let's see an example.

[4] Some researchers refer it as different version of treatment [11].

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

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

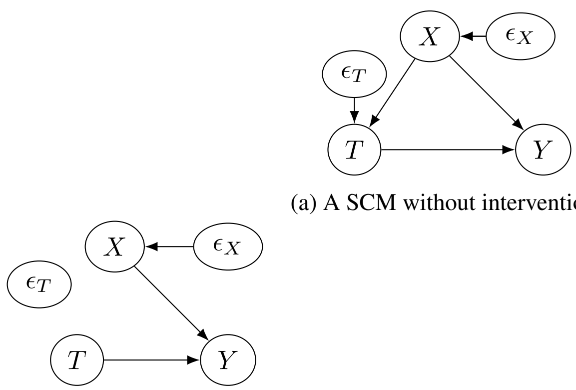
The do *intervened SCM is:*

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

The info intervened SCM is:

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

The causal model is show in Figure 1a, and the do *intervened SCM with Figure 1b, adn the info intervened SCM with Figure 1c.*



(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.

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 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 Info Causal DAGs

In the previous section, we have illustrated the idea of understanding as information transfer, and given the defintion of information intervention $\sigma(x)$ for SCMs. Now we are going to focus on the directed acyclic graphs(DAGs) for illustration of causal semantics and notations based on $\sigma(\cdot)$-operator.

First, we give the definition of info causal DAGs based on info intervention, corresponding to causal DAGs based on do intervention.

Definition 4.1 (Info 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* 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,

$$p(x|\sigma(X_A = \tilde{x}_A)) = \prod_{k \in V} p(x_k|x^*_{pa(k)}) \tag{1}$$

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

Lemma 4.2. *For arbitrary disjoint sets S and T of an info causal DAG $G = (V, E)$, we have*

$$P(X_S = x_s|X_T = x_T, \sigma(X_T = x_T)) = P(X_S = x_s|X_T = x_T)$$

Proof. According to Equation 1,

$$P(X = x|\sigma(X_T = \tilde{x}_T)) = \prod_{i \in V} P(X_i = x_i|X_{pa(i)} = x^*_{pa(i)})$$

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

Let $\tilde{x}_k = x_k$ for $k \in T$, then

$$\begin{aligned} P(X = x|\sigma(X_T = \tilde{x}_T)) &= P(X_T = \tilde{x}_T, X_{V/T} = x_{V/T}|\sigma(X_T = \tilde{x}_T)) \\ &= \prod_{i \in V} P(X_i = x_i|X_{pa(i)} = x^*_{pa(i)}) \\ &= \prod_{i \in V} P(X_i = x_i|X_{pa(i)} = x_{pa(i)}) \\ &= P(X = x) \end{aligned}$$

Then with marginalization on $S \cup T$ and S we have

$$P(X_S = x_S, X_T = x_T|\sigma(X_T = x_T)) = P(X_S = x_S, X_T = x_T) P(X_T = x_T|\sigma(X_T = x_T)) = P(X_T = x_T)$$

With the above two equations lead to:

$$P(X_S = x_s|X_T = x_T, \sigma(X_T = x_T)) = P(X_S = x_s|X_T = x_T)$$

□

It have been proved that causal DAGs are a special case of SCMs, then we can analogy to consider the *info causal DAG* as an info intervened SCM. The intervened causal graph of do intervention $G^{do(x)}$ is given by the intervened joint probability distribution, similary we can definte the intervened causal graph of info intervention.

Definition 4.3 (Intervened causal graph of info intervention)**.** *Consider an info causal DAG $G = (V, E)$ and the info intervention $\sigma(X_A = \tilde{x}_A)$, the* info intervened causal graph *is defined as the graph of X with respect to the joint distribution $p(x; \sigma(X_A = \tilde{x}_A))$.*

Comparing to intervened graph for intervention $do(X = x)$ which deletes all input edges, the intervened graph for intervention $\sigma(X = x)$ is obtained by removing all output edges of X.

Example 4.4 (A real world example)**.** *We consider domain variables* Exercise, Cholesterol, Occupation, Income, Diet*, and assumming a causal relationship among them in Figure 2.*

The three-level causal hierarchy. We can not talk about causal-effect without specifying a particular sub population. In the above example, the causal graph serves as a graph of information transfering for a particular person. 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.

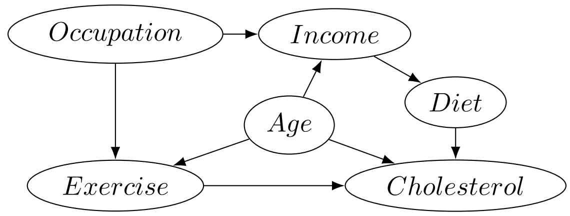


Figure 2: Causal effect of exercise on cholesterol level.

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 information 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 sending out from X?
- Examples: What will my income be if I the company accepts the information that my age is 28?

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 age cause me low income? what if the company accept the information that my age is 32?

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. $P(\text{Cholestoral}|do(\text{age} = 38)$, will be a problem. Although Pearl suggests that we should interpret it in other dimensions [18], there is no problem like that for info intervention. In our information transfer causality, 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. Interventional question $P(\text{Income}|\sigma(\text{Age} = 38)$ can be interpreted as what will happen to my income/salary if the company recieves the message of my age is 38, counterfactual or retrospective question $P(\text{Income}^{\sigma(\text{Age}=32)}|\text{Income} = low, \text{Age} = 22)$ can be interpreted as what if the company accept the information that I am 32, given that I am 22 and my income is low.

Science thrives on standards, because standards serve (at least) two purposes: communication and theoretical focus. We have already show that info intervention is a nice standard for communicating about causal questions as convenient pool as do intervention. The do operator set a theoretical focus on causal inference, many of the variants of causal queries can be reduced to expressions with do operator, or to several applications of do. Theoretical results established for "do" are then applicable to those variants. Then, what about info intervention?

5 Causal Calculus for Info Intervention

Pearl created the do-operator for do intervention, and the operator becomes essentially important for formulating causal queries and theories. Likewisely, we have proposed σ-operator for info intervention, and now we are going to using it to formulate corresponding causal propositions. First, an adjustment formula for info intervention is given for the most common treatment-outcome case.

Theorem 5.1 (Adjustment formula)**.** *For an info causal DAG G, T is the treatment, X is the covariates and Y is the outcome, then*

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

Proof. According to Equation 1,

$$p(Y=y,T=t,X=x|\sigma(T=\tilde{t}))=p(Y=y|T=\tilde{t},X=x)p(T=t|X=x)p(X=x)$$

Then we have

$$\begin{aligned}
p(Y=y|\sigma(T=\tilde{t})) &= \sum_x\sum_t p(Y=y,T=t,X=x|\sigma(T=\tilde{t})) \\
&= \sum_x\sum_t p(Y=y|T=\tilde{t},X=x)p(T=t|X=x)p(X=x) \\
&= \sum_x p(Y=y|T=\tilde{t},X=x)p(X=x)\sum_t p(T=t|X=x) \\
&= \sum_x P(Y=y|T=\tilde{t},X=x)P(X=x)
\end{aligned}$$

□

The major obstacle "Confounding" to drawing causal inference from data has been demystified through a graphical criterion called "backdoor criteria". For a DAG with a joint distribution p factorizes according to the graph structure, d-seperation implies conditional independence. Moreover, Pearl's 3 rules for do intervention enables us to identify a causal query, turn a causal question into a statisitcal estimation problem. Pearl's 3 rules describe graphical rules for

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

For the info intervention on info causal DAGs, we propose three rules for $\sigma(\cdot)$ operator, which reveal a more straightforward representation as follows:

Theorem 5.2 (Three rules for σ operator)**.** *For an info causal DAG $G=(V,E)$, X,Y,Z and W are arbitratry disjoint sets of 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)*
$P(Y=y|\sigma(X=x),Z=z,W=w)=P(Y=y|\sigma(X=x),W=w)$ *in the case that* $Y \perp\!\!\!\perp_d Z|W$ *in* $G^{\sigma(X=x)}$

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

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

Proof. For Rule 1, since $G^{\sigma(X=x)}$ is a DAG w.r.t. a factorization $P(X=x,Y=y,Z=x,W=w|\sigma(X=x))$, and d-separation implies conditional independence, then

$$P(Y=y|\sigma(X=x),Z=z,W=w)=P(Y=y|\sigma(X=x),W=w)$$

For Rule 2, since $Y \perp\!\!\!\perp_d Z|X$ in $G^{\sigma(Z=z)}$, then we have the conditional independence:

$$\begin{aligned}
P(Y=y|X=x,\sigma(Z=z)) &= P(Y=y|X=x,Z=z,\sigma(X=x)) \\
&= P(Y=y,Z=z|X=x,\sigma(X=x))/P(Z=z|X=x,\sigma(X=x))
\end{aligned}$$

and by Lemma 4.2,

$$P(Y=y,Z=z|X=x,\sigma(X=x))=P(Y=y,Z=z|X=x)$$
$$P(Z=z|X=x,\sigma(X=x))=P(Z=z|X=x)$$

Then we have

$$P(Y=y|X=x,\sigma(Z=z))=P(Y=y,Z=z|X=x)/P(Z=z|X=x)=P(Y=y|X=x,Z=z)$$

For Rule 3, we rephrase the notations for simplicity and need to prove

$$P(X_B=x_B|\sigma(X_A=\tilde{x}_A))=P(X_B=x_B)$$

where there is no causal paths form A to B.

Note that $A\notin Anc(B)$ ($Anc(B)$ represent the set of ancestors of B and itself), and for any $j\in Anc(B)$, $pa(j)\subset Anc(B)$, then $x^*_{pa(j)}=x_{pa(j)}$ by definition, which leads to

$$\prod_{j\in Anc(B)}P(X_j=x_j|X_{pa(j)}=x^*_{pa(j)})=\prod_{j\in Anc(B)}P(X_j=x_j|X_{pa(j)}=x_{pa(j)})=P(X_{Anc(B)}=x_{Anc(B)})$$

According to Equation 1,

$$\begin{aligned}P(X=x|\sigma(X_A=\tilde{x}_A))&=\prod_{k\in V}p(x_k|x^*_{pa(k)})\\&=\prod_{k\in Anc(B)}p(x_k|x^*_{pa(k)})\cdot\prod_{k\notin Anc(B)}p(x_k|x^*_{pa(k)})\\&=\prod_{k\in Anc(B)}p(x_k|x_{pa(k)})\cdot\prod_{k\notin Anc(B)}p(x_k|x^*_{pa(k)})\end{aligned}$$

Marginlize on $Anc(B)$, we have:

$$\begin{aligned}P(X_{Anc(B)}=x_{Anc(B)}|\sigma(X_A=\tilde{x}_A))&=\prod_{k\in Anc(B)}p(x_k|x_{pa(k)})\\&=P(X_{Anc(B)}=x_{Anc(B)})\end{aligned}$$

Since $B\in Anc(B)$, then $P(X_B=x_B|\sigma(X_A=\tilde{x}_A))=P(X_B=x_B)$ hold.

□

The above three rules for $\sigma(\cdot)$ operater have the following equivalent form, and in some case it brings some convenience for application.

Theorem 5.3. *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* $G^{\sigma(x)}$

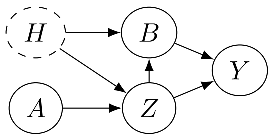
Rule 2*(Action/observation exchange)*
$P(y|\sigma(x),\sigma(z),w)=P(y|\sigma(x),z,w)$ *in the case that* $Y\perp\!\!\!\perp_d Z|W$ *in* $G^{\sigma(x,z)}$

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* $G^{\sigma(x)}$ [5].

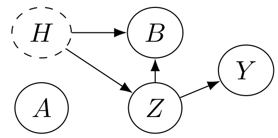
[5] $Z=z$ is a cause of $Y=y$ if $P(y|\sigma(z))\neq P(y)$ which mean information $Z=z$ have a effect on the event $Y=y$.

The single-world intervention graph (SWIG) which is quite similar to info intervened graph with the only difference of additional splitted constant intervened nodes [21, 22]. 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.

Example 5.4 (Independence of counterfactuals)**.** *If we have a causal DAG G, see Figure 3a, 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 3: Independence of Counterfactuals.

The answer is yes. The single-world intervention graph (SWIG) [21, 22], which is a simple graphical theory unifying causal directed acyclic graphs (DAGs) and potential (aka counterfactual) outcomes via a node-splitting transformation, is proposed for answering such questions. Here we showwe are able to answer such question elegantly with info intervention and related tools.

Since the info intervened causal graph $G^{\sigma(A=a,B=b)}$ factorizes, we can use d-separation to read conditional independence from the info intervened graph Figure 4b:

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

which means

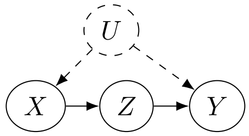
$$Y^{\sigma(a,b)} \perp\!\!\!\perp B^{\sigma(a,b)}|\{A^{\sigma(a,b)}, Z^{\sigma(a,b)}\}$$

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

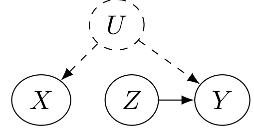
Similarly, we have the front-door criterion for causal inference with an elegant proof.

Theorem 5.5 (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,*

$$P(y|\sigma(\tilde{x})) = \sum_z P(z|\tilde{x}) \sum_x P(y|z, x)P(x) \tag{2}$$



(a) G.



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

Figure 4: Front-door Criteria.

Proof. The causal graph with hidden variables U of this front-door problem is given by Pearl [17] shown in Fig. 4a. By the factorization property of the info causal DAG in Fig. 4:

$$P(y, z, x, u|\sigma(\tilde{x})) = P(y|u, z)P(z|\tilde{x})P(x|u)P(u) \tag{3}$$

Then we have

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

□

6 Discussion and Conclusion

We interpret causality as information transfer and propose the info intervention to highlight the information view of causal modeling. We have not only addressed the logic behind this concept, but also invented the $\sigma(\cdot)$ operator to formulate causal queries and 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 [7]. 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] Tineke Blom, Stephan Bongers, and Joris M Mooij. Beyond Structural Causal Models: Causal Constraints Models. may 2018.

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

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

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

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

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

[7] Bernard Feltz, Marcus Missal, and Andrew Cameron Sims. *Free Will, Causality, and Neuroscience*. Koninklijke Brill NV, 2019.

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

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

[10] 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.

[11] 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.

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

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

[14] 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.

[15] 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.

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

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

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

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

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

[21] 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.

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

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

[24] 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.