26

Pearl Causality and the Value of Control

Ross Shachter and David Heckerman

1 Introduction

We welcome this opportunity to acknowledge the significance of Judea Pearl's contributions to uncertain reasoning and in particular to his work on causality. In the decision analysis community causality had long been "taboo" even though it provides a natural framework to communicate with decision makers and experts [Shachter and Heckerman 1986]. Ironically, while many of the concepts and methods of causal reasoning are foundational to decision analysis, scholars went to great lengths to avoid causal terminology in their work. Judea Pearl's work is helping to break this barrier, allowing the exploration of some fundamental principles. We were inspired by his work to understand exactly what assumptions are being made in his causal models, and we would like to think that our subsequent insights have contributed to his and others' work as well.

In this paper, we revisit our previous work on how a decision analytic perspective helps to clarify some of Pearl's notions, such as those of the *do* operator and *atomic intervention*. In addition, we show how influence diagrams [Howard and Matheson 1984] provide a general graphical representation for cause. Decision analysis can be viewed simply as determining what interventions we want to make in the world to improve the prospects for us and those we care about, an inherently causal concept. As we shall discuss, causal models are naturally represented within the framework of decision analysis, although the causal aspects of issues about counterfactuals and causal mechanisms that arise in computing the value of clairvoyance [Howard 1990], were first presented by Heckerman and Shachter [1994, 1995]. We show how this perspective helps clarify decision-analytic measures of sensitivity, such as the value of control and the value of revelation [Matheson 1990; Matheson and Matheson 2005].

2 Decision-Theoretic Foundations

In this section we introduce the relevant concepts from [Heckerman and Shachter 1995], the framework for this paper, along with some extensions to those concepts.

Our approach rests on a simple but powerful primitive concept of *unresponsiveness*. An uncertain variable is unresponsive to a set of decisions if its value is unaffected by our choice for the decisions. It is unresponsive to those decisions in worlds limited by other variables if the decisions cannot affect the uncertain variable without also changing one of the other variables.

We can formalize this by introducing concepts based on Savage [1954]. We consider three different kinds of distinctions, which he called *acts*, *consequences*, and *possible states of the world*. We have complete control over the acts but no control over the uncertain state of the world. We might have some level of control over consequences, which are logically determined, after we act, by the state of the world. Therefore, a consequence can be represented as a deterministic function of acts and the state of the world, inheriting uncertainty from the state of the world while affected, more or less, by our choice of action.

In practice, it is convenient to represent acts and consequences with variables in our model. We call a variable describing a set of mutually exclusive and collectively exhaustive acts a *decision*, and we denote the set of decisions by D. We call a variable describing a consequence *uncertain*, and we denote the set of uncertain variables by U. At times we will distinguish between the uncertain variables that serve as our objectives or *value variables*, V, and the other uncertain variables which we call *chance variables*, $C = U \setminus V$. Finally, in this section we will use the variables S to represent the possible states of the world. As a convention we will refer to single variables with lower-case (x or d), sets of variables with upper-case (D or V), and particular instances of variables with bold ($\mathbf{x}$ or $\mathbf{D}$). In this notation, the set of uncertain variables X takes value $X[\mathbf{S}, \mathbf{D}]$ deterministically when $\mathbf{D}$ is chosen and $\mathbf{S}$ is the state of the world.

DEFINITION 1 (Unresponsiveness). Given a decision problem described by uncertain variables U, decision variables D, and state of the world S, and variable sets $X \subseteq U$ and $Y \subseteq D \cup U$, X is said to be *unresponsive to* D, denoted $X \not\hookrightarrow D$, if we believe that

$$\forall \mathbf{S} \in S, \mathbf{D_1} \in D, \mathbf{D_2} \in D : X[\mathbf{S}, \mathbf{D_1}] = X[\mathbf{S}, \mathbf{D_2}]$$

and, if not, X is said to be *responsive to* D.
Furthermore, X is said to be *unresponsive to* D *in worlds limited by* Y, denoted $X \not\hookrightarrow_Y D$, if we believe that

$$\forall \mathbf{S} \in S, \mathbf{D_1} \in D, \mathbf{D_2} \in D : Y[\mathbf{S}, \mathbf{D_1}] = Y[\mathbf{S}, \mathbf{D_2}] \Longrightarrow X[\mathbf{S}, \mathbf{D_1}] = X[\mathbf{S}, \mathbf{D_2}]$$

and, if not, X is said to be *responsive to* D *in worlds limited by* Y.

The distinctions of unresponsiveness and limited unresponsiveness seem natural for decision makers to consider. Unresponsiveness is related to independence, in that any uncertain variables X that are unresponsive to decisions D are independent of D. Although it is not necessarily the case that X independent of D is unresponsive to D, that implication is often assumed [Spirtes, Glymour, and Scheines 1993]. In contrast, there is no such general correspondence between limited unresponsiveness and conditional independence.

To illustrate these concepts graphically, we introduce influence diagrams [Howard and Matheson 1984]. An *influence diagram* is an acyclic directed graph G with

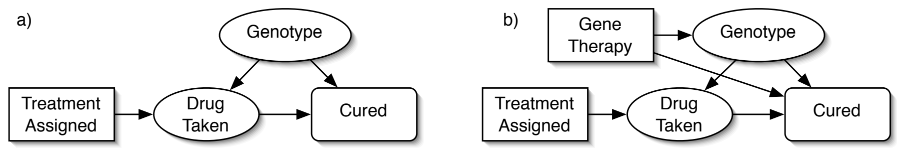

Figure 1. The treatment assignment only cures the patient if it affects whether the drug is taken, but genotype does not have a causal effect unless it is responsive to decisions.

nodes corresponding to the variables, rectangles for decisions, ovals for chance variables, and rounded rectangles for value variables. Arcs into chance and value nodes, are *conditional*. For each uncertain variable x there is a conditional probability distribution for x given its parents, $Pa(x)$. If the distribution is a deterministic function, we represent that in the graph by a double oval or double rounded rectangle. Arcs into decisions are *informational*, representing that the parent variables will be observed before the decision is made. Although there are significant issues involving informational arcs, we will focus primarily on models in which there are no informational arcs and all of the decisions could be made in any order, before any of the uncertain variables are observed.

We allow multiple value nodes, all with no children, assuming that their values will be summed. We assume that the criterion for making decisions is either the total value or an increasing exponential utility function of the total. This simplifies the valuation of a proposed change to a decision problem because the most a decision maker should be willing to pay for the change is the difference in the values of the diagrams with and without the proposed change.

Although we have defined unresponsiveness without regard to a graphical representation, there is an intuitive graphical interpretation (with some technical exceptions described in Heckerman and Shachter [1985]). The uncertain descendants of decisions are usually responsive to them, and the other uncertain variables are usually unresponsive. Also, X is usually unresponsive to D in worlds limited by Y if all of the directed paths from D to X include nodes in Y. When these rules of thumb are all satisfied, we say that an influence diagram is causal.

DEFINITION 2 (Causal Influence Diagram). An influence diagram with graph G and decision nodes D, chance nodes C, and value nodes V, is said to be *causal* if we believe that uncertain variables $X \subseteq C \cup V$ are unresponsive to decisions D, $X \not\hookleftarrow D$, whenever there is no directed path from D to X, and X is unresponsive to decisions D in worlds limited by Y, $X \not\hookleftarrow_Y D$, whenever every directed path from D to X includes a node from Y.

Consider the influence diagram shown in Figure 1a which we believe is causal. In this case, we believe that *Drug Taken* and *Cured* are responsive to *Treatment Assigned* while *Genotype* is unresponsive to *Treatment Assigned*. We also believe that *Cured* is unresponsive to *Treatment Assigned* in worlds limited by *Drug Taken*. Note that *Treatnent Assigned* is *not* independent of *Genotype* or *Cured* given *Drug Taken*.

The concept of limited unresponsiveness allows us to define how one variable can cause another in a way that is natural for decision makers to understand.

DEFINITION 3 (Cause with Respect to Decisions). Given a decision problem described by uncertain variables U and decision variables D, and a variable $x \in U$, the set of variables $Y \subseteq D \cup U \setminus \{x\}$ is said to be a *cause for x with respect to D* if Y is a minimal set of variables such that $x \not\hookleftarrow_Y D$.

Defining cause with respect to a particular set of decisions adds clarity. Consider again the causal influence diagram shown in Figure 1a. With respect to the decision *Treatment Assigned*, the cause of *Cured* is either {*Treament Assigned*} or {*Drug Taken*}, while the cause of *Genotype* is {}. Because we believe that *Genotype* is unresponsive to *Treatment Assigned* it has no cause with respect to D. On the other hand, we believe that *Cured* is responsive to *Treatment Assigned* but not in worlds limited by *Drug Taken*, so {*Drug Taken*} is a cause of *Cured* with respect to D.

Consider now the causal influence diagram shown in Figure 1b, in which we have added the decision *Gene Therapy*. Because *Genotype* is now responsive to D, the cause of *Genotype* is {*Gene Therapy*} with respect to D. If the gene therapy has some side effect on whether the patient is cured, then {*Gene Therapy*, *Drug Taken*} but not {*Genotype*, *Drug Taken*} would be a cause of *Cured* with respect to the decisions, because *Cured* is unresponsive to D in worlds limited by the former but not the latter.

The concept of limited unresponsiveness also allows us to formally define direct and atomic interventions. A set of decision I is a direct intervention on a set of uncertain variables X if the effects of I on all other uncertain variables are mediated through their effects on X.

DEFINITION 4 (Direct Intervention). Given a decision problem described by uncertain variables U and decision variables D, a set of decisions $I \subseteq D$ is said to be a *direct intervention on $X \subseteq U$ with respect to D* if (1) $x \hookleftarrow I$ for all $x \in X$, and (2) $y \not\hookleftarrow_X I$ for all $y \in U$.

In a causal influence diagram every node in I has children only in X and there is a directed path from I to every node in X. In the causal influence diagram shown in Figure 1b, *Treatment Assigned* is a direct intervention on *Drug Taken*, and the set of decisions is a direct intervention on all three uncertain variables. Note that whether a decision is a direct intervention depends on the underlying causal mechanism. If the gene therapy had no side effect then *Gene Therapy* would be a direct intervention on *Genotype*, but regardless whether there is a side effect, *Gene Therapy* is a direct intervention on {*Genotype*, *Cured*}.

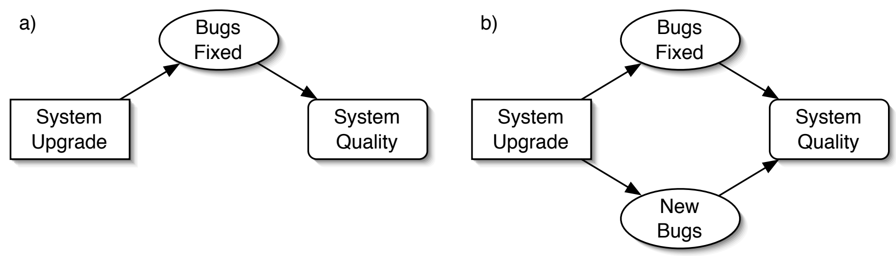

Figure 2. We believe that a system upgrade will affect system quality by fixing bugs unless new bugs are introduced in the process.

DEFINITION 5 (Atomic Intervention). Given a decision problem described by uncertain variables U and decision variables D, a decision $do(x) \in D$ is said to be a *atomic intervention on* $x \in U$ *with respect to* D if (1) $do(x)$ is a direct invention on x with respect to D, and (2) $do(x)$ has precisely the instances (a) **idle**, which corresponds to no intervention, and (b) $\mathbf{do(x)}$ for every instance $\mathbf{x}$ of x, where $x = \mathbf{x}$ whenever $do(x) = \mathbf{do(x)}$.

This is precisely the atomic intervention described without definition in Pearl [1993]. The assumptions underlying it are quite strong. The causal influence diagram shown in Figure 2a assumes that we can upgrade our system and improve the quality by fixing the bugs, but the diagram shown in (b) illustrates the all too familiar situation when new bugs are introduced in the process, compromising system quality. In that case, *System Upgrade* is not a direct intervention on *Bugs Fixed* and $\{$*Bugs Fixed*$\}$ is not a cause of *System Quality* with respect to D. Although the system upgrade was intended to be an atomic intervention, it can have unintended and undesirable consequences.

We can now represent the relationship between an uncertain variable x and other variables Y, such as its parents in a causal influence diagram. We consider the uncertain function $x(Y)$ as a variable, and now x is a deterministic function of Y and $x(Y)$. In fact, if Y is a cause of x with respect to D then $x(Y)$ must be unresponsive to D.

DEFINITION 6 (Mapping Variable). Given a decision problem described by uncertain variables U and decision variables D, $x \in U$ and variables Y such that for every $y \in Y \cap U$ there exists an atomic intervention $do(y) \in D$, the *mapping variable* $x(Y)$ is the chance variable that represents all possible mappings from Y to x.

Finally, we have developed the machinery to characterize a *Pearl causal model* and structural equations [Pearl 1993]. Given uncertain variables U, suppose the decisions D comprise an atomic intervention $do(x)$ on every $x \in U$. Given a graph

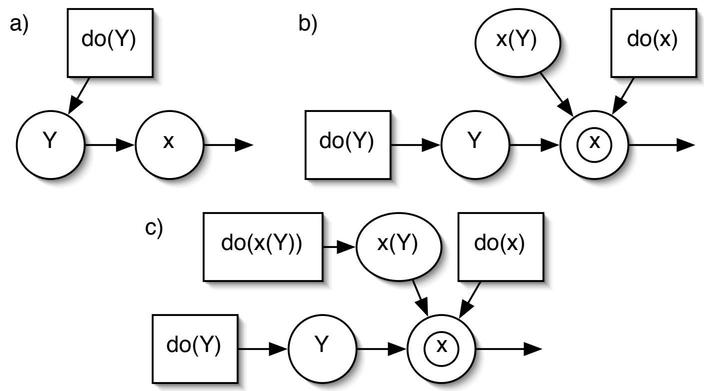

Figure 3. The partial influence diagram for x in a causal model, shown in (a) with parents Y, becomes the diagram shown in (b) explicitly representing the structural equation for x, and, when Y is nonempty, the diagram shown in (c) with an explicit atomic intervention on the mapping variable.

G with nodes U, such that $Pa(x) \cup \{do(x)\}$ is a cause for x with respect to D. Then

$$x = f_x(Pa(x), do(x), x(Pa(x)))$$

for all $x \in U$ where f_x is a deterministic function such that $x = \mathbf{x}$ if $do(x) = \mathbf{do}(\mathbf{x})$.

We can extend this to allow manipulation of a mapping variable for $x \in U$ with parents to obtain a *Pearl causal model with an atomic intervention for mapping variable* $x(Pa(x))$. The decisions D now also include a atomic intervention $do(x(Pa(x)))$. As a result, $Pa(x) \cup \{do(x), do(x(Pa(x)))\}$ is now a cause for x with respect to D and $x(Pa(x)) = \mathbf{x}(Pa(x))$ when $do(x(Pa(x))) = \mathbf{do}(\mathbf{x}(Pa(x)))$.

The causal model is represented by the partial influence diagrams shown in Figure 3 with $Y = Pa(x) \subseteq C$ in the graph G. We assume in (a) that there are atomic interventions $do(y)$ on each $y \in Y$ represented as $do(Y)$. The diagram shown in (b) explicitly represents the structural equation for x as a deterministic function of Y, an atomic intervention, $do(x)$, and the mapping variable, $x(Y)$. The influence diagram is causal, showing that $Y \cup \{do(x)\}$ is a cause for x with respect to D. We can extend the model by adding an atomic intervention for the mapping variable, $do(x(Y))$. If Y is empty then nothing needs to be added, as $do(x)$ is the same atomic intervention as $do(x())$, but otherwise we obtain the diagram shown in (c). Now $Y \cup \{do(x), do(x(Y))\}$ is a cause for x with respect to D.

An influence diagram is said to be in *canonical form* if each uncertain variable responsive to a decision is a descendant of that decision and represented as a deterministic node. Each decision, including atomic interventions, is explicit. Each uncertain variable that is responsive to D is a deterministic function of its parents, including any decisions that are direct interventions on it, and a mapping variable. As an example, the influence diagram shown in Figure 3b is in canonical form.

In the next section we apply these concepts to define and contrast different measures for the value to a decision maker of manipulating (or observing) an uncertain variable.

3 Value of Control

When assisting a decision maker developing a model, sensitivity analysis measures help the decision maker to validate the model. One popular measure is the *value of clairvoyance*, the most a decision maker should be willing to pay to observe a set of uncertain variables before making particular decisions [Howard 1967]. Our focus of attention is another measure, the value of control (or wizardry), the most a decision maker should be willing to pay a hypothetical wizard to optimally control the distribution of an uncertain variable [Matheson 1990], [Matheson and Matheson 2005]. We consider and contrast the value of control with two other measures, the value of do, and the value of revelation, and we develop the conditions under which the different measures are equal.

In formalizing the value of control, it is natural to consider the value of an atomic intervention on uncertain variable x, in particular $\mathbf{do}(\mathbf{x}^*)$, that would set it to $\mathbf{x}^*$ the instance yielding the most valuable decision situation, rather than to **idle**. We call the most the decision maker should be willing to pay for such an intervention the *value of do* and compute it as the difference in the values of the diagrams.

DEFINITION 7 (Value of Do). Given a decision problem including an atomic intervention on uncertain variable $x \in U$, the *value of do for x*, denoted by $VoD(\mathbf{x}^*)$, is the most one should be willing to pay for an atomic intervention on uncertain variable x to the best possible deterministic instance, $\mathbf{do}(\mathbf{x}^*)$, instead of to **idle**.

Our goal in general is to value the optimal manipulation of the *conditional* distribution of a target uncertain variable x in a causal influence diagram, $P\{x|Y\}$, and the most we should be willing to pay for such an intervention is the *value of control*. The simplest case is when $\{do(x)\}$ is a cause of x with respect to D, $Y = \{\}$, so the optimal distribution is equivalent to an atomic intervention on x to $\mathbf{x}^*$, and control and *do* are the same intervention. Otherwise, the *do* operation effectively severs the arcs from Y to x and replaces the previous causal mechanism with the new atomic one. By contrast, the control operation is an atomic intervention on the mapping variable $x(Y)$ to its optimal value $\mathbf{do}(\mathbf{x}^*(Y))$ rather than to **idle**.

DEFINITION 8 (Value of Control). Given a decision problem including variables Y, a mapping variable $x(Y)$ for uncertain variable $x \in U$, and atomic interventions $do(x)$ and $do(x(Y))$ such that $Y \cup \{do(x), do(x(Y))\}$ is a cause of x with respect to D, the *value of control for x*, denoted by $VoC(\mathbf{x}^*(Y))$, is the most one should be willing to pay for an atomic intervention on the mapping variable for uncertain variable x to the best possible deterministic function of Y, $\mathbf{do}(\mathbf{x}^*(Y))$, instead of to **idle**.

If $Y = \{\}$, then $do(x)$ is the same atomic intervention as $do(x(Y))$, and the values of do and control for x are equal, $VoD(\mathbf{x}^*) = VoC(\mathbf{x}^*())$.

In many cases, while it is tempting to assume atomic interventions, they can be cumbersome or implausible. In an attempt to avoid such issues, Ronald A. Howard has suggested an alternative passive measure, the *value of revelation*: how much better off the decision maker should be by observing that the uncertain variable in question obtained its most desirable value. This is only well-defined for variables unresponsive to D, except for those atomic interventions that are set to **idle**, because otherwise the observation would be made before decisions it might be responsive to. Under our assumptions this can be computed as the difference in value between two situations, but it is hard to describe it as a willingness to pay for this difference as it is more passive than intentional. (The value of revelation is in fact an intermediate term in the computation of the value of clairvoyance.)

DEFINITION 9 (Value of Revelation). Given a decision problem including uncertain variable $x \in U$ and a (possibly empty) set of atomic interventions, A, that is a cause for x with respect to D, the *value of revelation for* uncertain variable $x \in U$, denoted by $VoR(\mathbf{x}^*)$, is the increase in the value of the situation with $d = \mathbf{idle}$ for all $d \in A$, if one observed that uncertain variable $x = \mathbf{x}^*$, the best possible deterministic instance, instead of not observing x.

To illustrate these three measures we, consider a partial causal influence diagram including x and its parents, Y, which we assume for this example are uncertain and nonempty, as shown in Figure 4a. There are atomic interventions $do(x)$ on x, $do(x(Y))$ on mapping variable $x(Y)$, and $do(y)$ on each $y \in Y$ represented as $do(Y)$. The variable x is a deterministic function of Y, $do(x)$ and $x(Y)$. In this model, $Y \cup \{do(x), do(x(Y))\}$ is a cause of x with respect to D. The dashed line from x to values V suggests that there might be some directed path from x to V. If not, V would be unresponsive to $do(x)$ and $do(x(Y))$ and the values of do and control would be zero.

To obtain the reference diagram for our proposed changes, we set all of the atomic interventions to **idle** as shown in Figure 4b1. We can compute the value of this diagram by eliminating the idle decisions and absorbing the mapping variable into x, yielding the simpler diagram shown in (b2). To compute the value of do for x, we can compute the value of the diagram with $\mathbf{do}(\mathbf{x}^*)$ by setting the other atomic interventions to **idle**, as shown in (c1). But since that is making the optimal choice for x with no interventions on Y or $x(Y)$, we can now think of x as a decision variable as indicated in the diagram shown in (c2). We shall use this shorthand in many of the examples that we consider. To compute the value of control for x, we can compute the value of the diagram with $\mathbf{do}(\mathbf{x}^*(Y))$ by setting the other atomic interventions to **idle**, as shown in (d1). But since that is making the optimal choice for $x(Y)$ with none of the other interventions, we can compute its value with

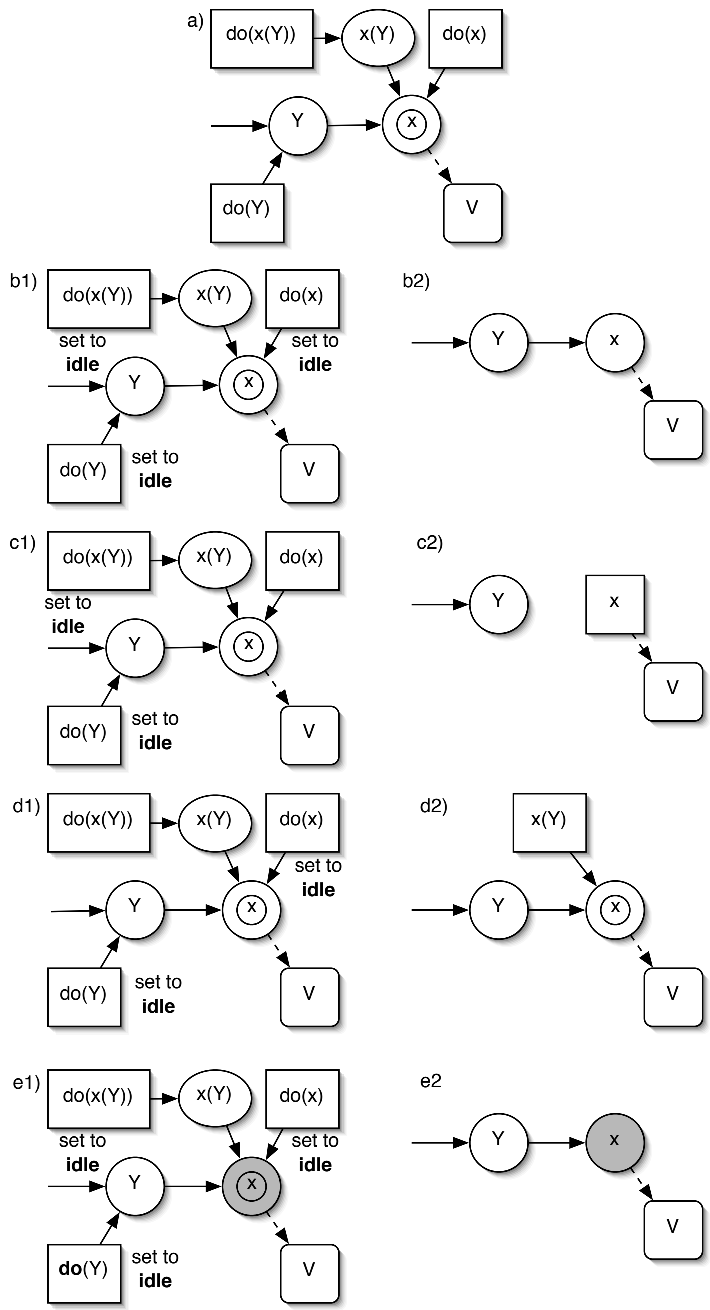

Figure 4. Partial causal influence diagrams to compute the values of do, control, and revelation for x when Y is nonempty.

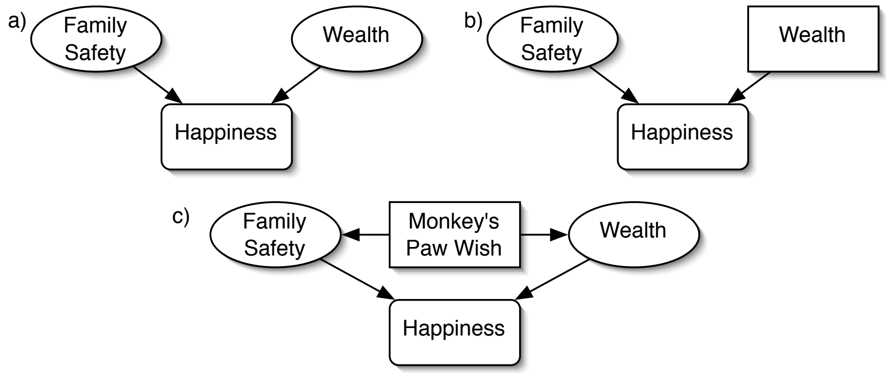

Figure 5. Unless the intervention is direct there can be disastrous side effects.

the simpler influence diagram shown in (d2), again using our shorthand. Finally, to compute the value of revelation for x, we can compute the value of the diagram with $x = x^*$ and all of the atomic interventions **idle**, as shown in (e1). The observation is well-defined because all of the interventions are **idle**, but that also means that we can compute its value with the simpler influence diagram shown in (e2).

Each of the three measures requires evaluation of two influence diagrams to determine its value, the reference diagram with all of the atomic interventions set to **idle** and a revised one, a diagram with either an atomic intervention or an observation. The values of these diagrams can be computed using simpler influence diagrams, with either one new decision, an atomic one made with no observations, or a new observation made before any decision, and the simpler diagram for the reference value has neither new decisions nor observations. These simpler diagrams are well-defined even if there are other decisions elsewhere and some observations prior to some of the other decisions [Shachter 1986]. Note that care must be taken in computing the value of control because there can be an exponential number of instances for the mapping variable.

The assumption of a direct intervention is crucial. Matheson and Matheson [2005](refer to it as "pure" and to an atomic intervention as "perfect".) There is a classic horror story of a man granted three wishes on a monkey's paw [Jacobs 1902]. He chooses to be wealthy and his wish is granted, tragically, through the death of his son. This corresponds to the causal influence diagrams shown in Figure 5. The value of his situation with no intervention is represented by the diagram in (a). The atomic intervention on *Wealth* he intends would yield the same value as a diagram in which *Wealth* is a decision as in (b), but the value with his intervention actually equals the value of the diagram shown in (c). The wish decision he actually made was not the direct intervention on *Wealth* he desired. The lesson is clear: in

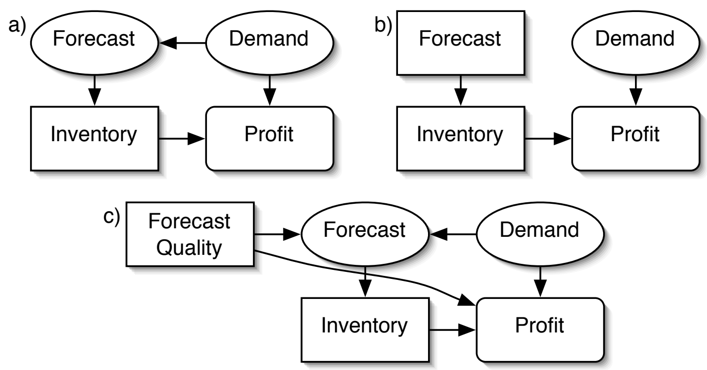

Figure 6. When we intervene on a forecast, we want to improve its quality, rather than to obtain a single most desirable instance.

manipulating our situation, we must beware of the unintended consequences.

Suppose the uncertain variable is being used to provide information, such as a forecast. Consider the causal influence diagram shown in Figure 6. This situation corresponds to one in which inventory decisions must be made before demand is observed, but a forecast relevant to demand will be observed before choosing inventory as shown in (a). Alas, an atomic intervention setting the forecast to our most desirable value ("highest demand") as in (b) does not improve profit since it tells us nothing about the real demand. What we would like to manipulate is the quality of the forecast, having it represent the best possible signal about demand as in (c). In this case, the value of do for *Forecast* is zero, but the value of control for *Forecast* should be positive. In fact, if there are as many instances for *Forecast* as there are for *Demand*, the highest quality forecast possible is clairvoyance on the demand, and the value of control would be equal to the value of clairvoyance. In the diagram *Forecast Quality* might not be an atomic intervention, both because there might only be a choice among imperfect information sources, and because there might be different costs associated with those different information sources.

Consider the causal influence diagrams shown in Figure 7, in which we believe that *Product Quality* is unresponsive to direct interventions (not shown) on *Sales* or *Profit*. We would like to understand how much we would improve our profit by manipulating our product quality. The diagram shown in (a) treats quality and sales as uncertain with its atomic interventions set to **idle**, and its value is the reference for any changes. The diagram shown in (b) has the same value as an atomic intervention on *Product Quality* to its optimal instance, and because that intervention is the cause of *Product Quality* with respect to D, the difference in values of this diagram relative to the one in (a) is both the value of do and the value of control for *Product Quality*. Alternatively, in (c) if we observed that *Product*

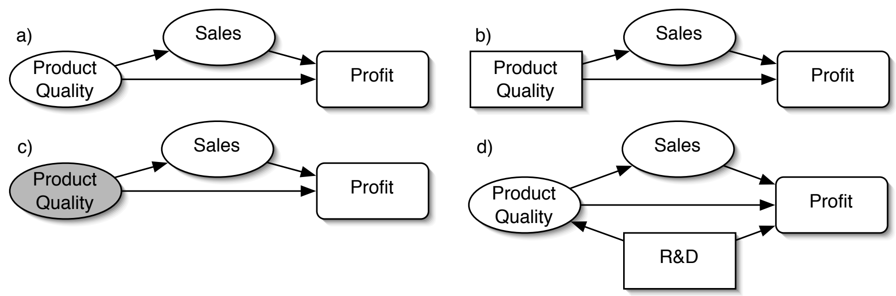

Figure 7. In this causal influence diagram the values of do, control, and revelation for *Product Quality* are equal.

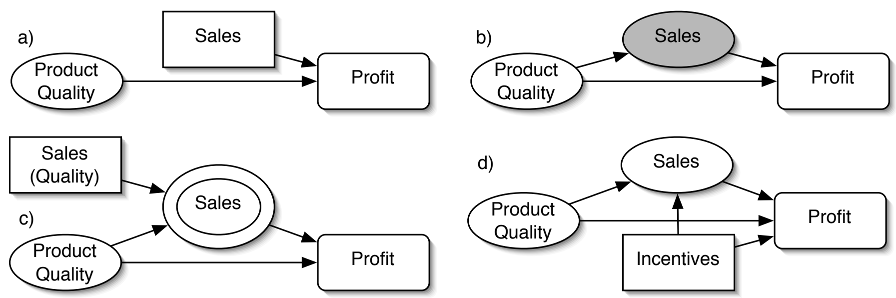

Figure 8. The values of do, control, and revelation for *Sales* might not be equal.

Quality takes the best possible value, this diagram has the same value as the one in (b). As a result, the value of revelation is equal to the other two values. Finally, in (d) we could contemplate a research and development effort that might lead to higher product quality. Because the diagram in (d) is causal, {*Product Quality*} is a cause of *Sales* with respect to D.

Now consider the causal influence diagrams shown in Figure 8, in which we are manipulating sales rather than product quality to improve our profit. We obtain the diagram shown in (a) by assuming that *Product Quality* is unresponsive to an atomic intervention on *Sales*. In (b) we could observe that *Sales* takes that same value, but this observation updates our belief about the *Product Quality*, and the value of this diagram might not be equal to the value of the diagram in (a). We obtain the diagram shown in (c) by an atomic intervention on the mapping variable for *Sales*, not determining sales but rather how it depends on quality (assuming that there is an atomic intervention on *Product Quality*). In this situation the values of do, control, and revelation could all be different! Finally, in (d) we consider offering incentives to boost sales, recognizing that it might affect our profits both directly

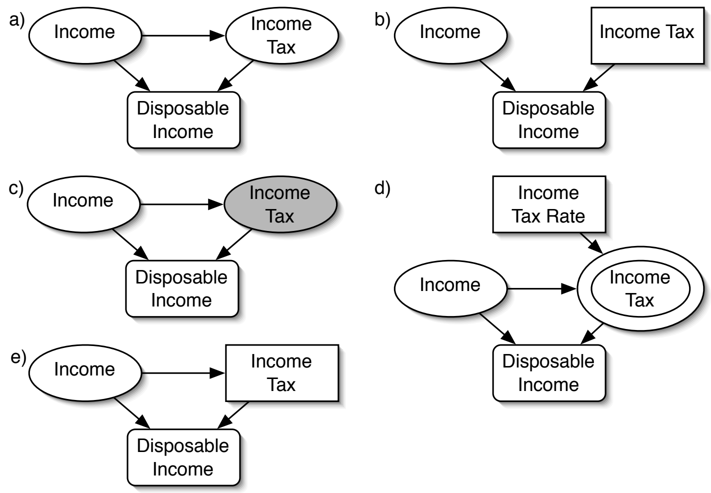

Figure 9. The values of do, control, and revelation are different for *Income Tax*.

and indirectly.

There can be a significant difference between passive observation of uncertain variable x and intervention on x. Consider the causal influence diagrams shown in Figure 9 representing disposable income after taxes. We believe that *Income* is unresponsive to a direct intervention on *Income Tax*, but *Income Tax* might be responsive to a direct intervention on *Income*. However, the value of do, the difference between the values of the diagrams in (b) and (a), is quite different from the value of revelation based on (c) and (a). Being able to choose not to pay any tax is quite different from learning that you will pay no tax, since it is more likely in the latter case that you have lost your job. Alternatively, we can consider setting the income tax rate as shown in (d), which would lead to the value of control. In this case, we can simplify the calculation in (d) that searches all possible mapping variable instances, to a simpler decision shown in (e), recognizing that in this case there is no interaction among the components of the mapping variable, and therefore we can independently search for the best possible instance for *Income Tax* for each possible instance of *Income*.

The correspondence between passive observation and intervention has been studied, primarily to identify causal effects from observational data [Robins 1986], [Pearl 1993] and [Spirtes, Glymour, and Scheines 1993]. In our framework, a set of variables Y is said to satisfy the *back door condition* for x if Y is unresponsive to $do(x)$ while $do(x)$ is d-separated from V by $\{x\} \cup Y$. When Y satisfies the back door condition, there is a correspondence among the values of do, control and revelation, in that

$$P\{V|\mathbf{Y},\mathbf{x}^*\} = P\{V|\mathbf{Y},\mathbf{do}(\mathbf{x}^*)\} = P\{V|\mathbf{Y},\mathbf{do}(\mathbf{x}^*(Y))\}.$$

However, in valuing the decision situation we do not get to observe Y and thus $P\{V|\mathbf{x}^*\}$ might not be equal to $P\{V|\mathbf{do}(\mathbf{x}^*)\}$. Consider the diagrams shown in Figure 9. Because *Income* satisfies the back door criterion relative to *Income Tax*, the values of do, control and revelation on *Income Tax* would all be the same if we observed *Income*. But we do not know what our *Income* will be and the values of do, control, and revelation can all be different.

Nonetheless, if we make a stronger assumption, that Y is d-separated from V by x, the three measures will be equal. The atomic intervention on x or its mapping variable only affects the value V through the descendants of x in a causal model, and all other variables are unresponsive to the intervention in worlds limited by x. However, the atomic interventions might not be independent of V given x unless Y is d-separated from V by x. Otherwise, observing x or an atomic intervention on the mapping variable for x can lead to a different value for the diagram than an atomic intervention on x.

We establish this result in two steps for both general situations and for Pearl causal models. By assuming that $do(x)$ is independent of V given x, we first show that the values of do and revelation are equal. If we then assume that Y is d-separated from V by x, we show that the values of do and control are equal. The conditions under which these two different comparisons can be made are not identical either. To be able to compute the value of revelation for x we must set to **idle** all interventions that x is responsive to, while to compute the value of control for x we need to be ensure that we have an atomic intervention on a mapping variable for x.

THEOREM 10 (Equal Values of Do and Revelation). *Given a decision problem including uncertain variable $x \in U$, if there is a set of atomic interventions A, including $do(x)$, that is a cause of x with respect to D, and $do(x)$ is independent of V given x, then the values of do and revelation for x are equal, $VoD(\mathbf{x}^*) = VoR(\mathbf{x}^*)$.*

If $\{do(x)\}$ is a cause of x with respect to D, then they are also equal to the value of control for x, $VoC(\mathbf{x}^*()) = VoD(\mathbf{x}^*) = VoR(\mathbf{x}^*)$.

Proof. Consider the probability of V after the intervention $\mathbf{do}(\mathbf{x}^*)$ with all other interventions in A set to **idle**. Because x is determined by $\mathbf{do}(\mathbf{x}^*)$, and $do(x)$ is independent of V given x,

$$P\{V|\mathbf{do}(\mathbf{x}^*)\} = P\{V|\mathbf{x}^*,\mathbf{do}(\mathbf{x}^*)\} = P\{V|\mathbf{x}^*\} = P\{V|\mathbf{x}^*, do(x) = \mathbf{idle}\}.$$

If $\{do(x)\}$ is is a cause of x with respect to D then the values of do and control for x are equal. □

COROLLARY 11. *Given a decision problem described by a Pearl causal model including uncertain variable $x \in U$, if $Pa(x)$ is d-separated from V by x, then the* *values of do and revelation for x are equal,* $VoD(\mathbf{x}^*) = VoR(\mathbf{x}^*)$. *If x has no parents, then the the values of do, control, and revelation for x are equal,*

$$VoD(\mathbf{x}^*) = VoC(\mathbf{x}^*()) = VoR(\mathbf{x}^*).$$

THEOREM 12 (Equal Values of Do and Control). *Given a decision problem described by an influence diagram including uncertain variable $x \in U$, and nonempty set of variables Y. If there are atomic interventions $do(x)$ for x, $do(y)$ for every $y \in Y \cap U$, and $do(x(Y)))$ for the mapping variable $x(Y)$, $Y \cup \{do(x), do(x(Y))\}$ is a cause of x with respect to D, and Y is d-separated from V by x, then the values of do and control are equal,*

$$VoD(\mathbf{x}^*) = VoC(\mathbf{x}^*(Y)).$$

Proof. We know that $Y \cup \{do(x), do(x(Y))\}$ is independent of V given x, because otherwise Y would not be d-separated from V by x. Because $do(x)$ is an atomic intervention on x and $do(x)$ is independent of V given x, as in Theorem 10, $P\{V|\mathbf{do}(\mathbf{x}^*)\} = P\{V|\mathbf{x}^*, \mathbf{do}(\mathbf{x}^*)\} = P\{V|\mathbf{x}^*\}$. Now consider the probability of V after the intervention $\mathbf{do}(\mathbf{x}^*(Y))$. Because $x = \mathbf{x}^*(\mathbf{Y})$ is determined by $\mathbf{do}(\mathbf{x}^*(Y))$ and $\mathbf{Y}$, and $Y \cup \{do(x(Y))\}$ is independent of V given x,

$$\begin{aligned} P\{V|\mathbf{do}(\mathbf{x}^*(Y)), \mathbf{Y}\} &= P\{V|x = \mathbf{x}^*(\mathbf{Y}), \mathbf{do}(\mathbf{x}^*(Y)), \mathbf{Y}\} \\ &= P\{V|x = \mathbf{x}^*(\mathbf{Y})\}, \end{aligned}$$

The optimal choice of $x(Y)$ does not depend on Y, $\mathbf{x}^*(Y) = \mathbf{x}^*$, yielding

$$P\{V|\mathbf{do}(\mathbf{x}^*(Y)), \mathbf{Y}\} = P\{V|\mathbf{x}^*\}.$$

As a result,

$$\begin{aligned} P\{V|\mathbf{do}(\mathbf{x}^*(Y))\} &= \sum_{\mathbf{Y}} P\{V, \mathbf{Y}|\mathbf{do}(\mathbf{x}^*(Y))\} \\ &= \sum_{\mathbf{Y}} P\{V|\mathbf{do}(\mathbf{x}^*(Y)), \mathbf{Y}\} P\{\mathbf{Y}|\mathbf{do}(\mathbf{x}^*(Y))\} \\ &= \sum_{\mathbf{Y}} P\{V|\mathbf{x}^*\} P\{\mathbf{Y}|\mathbf{do}(\mathbf{x}^*(Y))\} \\ &= P\{V|\mathbf{x}^*\} \sum_{\mathbf{Y}} P\{\mathbf{Y}|\mathbf{do}(\mathbf{x}^*(Y))\} \\ &= P\{V|\mathbf{x}^*\} \end{aligned}$$

□

COROLLARY 13. *Given an uncertain variable $x \in U$ with parents in a decision problem described by a Pearl causal model with an atomic intervention for mapping*

Figure 10. The values of do, control, and revelation are equal for each uncertain variable.

variable $x(Pa(x))$*, if* $Pa(x)$ *is d-separated from* V *by* x*, then the values of do, control, and revelation for* x *are equal,* $VoD(\mathbf{x}^*) = VoC(\mathbf{x}^*(Pa(x))) = VoR(\mathbf{x}^*)$.

Consider the causal influence diagrams shown in Figure 10, concerning a communicable disease, for which we believe that *Exposure* is unresponsive to any direct intervention on *Infection*, and both of them are unresponsive to any direct intervention on *Health*, but all of the uncertain variables might be responsive to a direction intervention on *Exposure*. Because *Exposure* has no parents, the values of do, control, and revelation for it will be equal. Furthermore, in this case, even though *Infection* has a parent, the values of do, control, and revelation for it will be also equal, because *Exposure* is independent of *Health* given *Infection*. Likewise, there will be equal values of do, control, and revelation for *Health*.

4 Conclusions

We have sharpened the distinctions underlying the value of control and related value of revelation and value of do, and shown that they are equivalent when the target variable x in a causal influence diagram either has no parents, or its parents, $Pa(x)$ are d-separated from the value V by x.

The general problem, which have only touched upon, permits multiple decisions and information sets at those other decisions. In that case, there is a question of how to recognize when $Pa(x)$ in d-separated from V by x. We can address this in general by either constructing the normal form diagram [Bhattacharjya and Shachter 2007] or by building a policy diagram, iteratively substituting deterministic policies for decisions starting with the latest decision [Shachter 1999]. These approaches exploit the causal structure and the separable value function represented in the influence diagram.

References

Bhattacharjya, D. and R. Shachter (2007). Evaluating influence diagrams with decision circuits. In R. Parr and L. van der Gaag (Eds.), *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*, pp. 9–16. Oregon: AUAI Press.

Heckerman, D. and R. Shachter (1995). Decision-theoretic foundations for causal reasoning. *Journal of Artificial Intelligence Research 3*, 405–430.

Heckerman, D. E. and R. D. Shachter (1994). A decision-based view of causality.

In R. Lopez de Mantaras and D. Poole (Eds.), *Uncertainty in Artificial Intelligence: Proceedings of the Tenth Conference*, pp. 302–310. San Mateo, CA: Morgan Kaufmann.

Howard, R. (1967). Value of information lotteries. *IEEE Transa. Systems Sci. Cybernetics SSC-3*(1), 54–60.

Howard, R. A. (1990). From influence to relevance to knowledge. In R. M. Oliver and J. Q. Smith (Eds.), *Influence Diagrams, Belief Nets, and Decision Analysis*, pp. 3–23. Chichester: Wiley.

Howard, R. A. and J. E. Matheson (1984). Influence diagrams. In R. A. Howard and J. E. Matheson (Eds.), *The Principles and Applications of Decision Analysis*, Volume II. Menlo Park, CA: Strategic Decisions Group.

Jacobs, W. W. (1902, September). The monkey's paw. *Harper's Monthly 105*, 634–639.

Matheson, D. and J. Matheson (2005). Describing and valuing interventions that observe or control decision situations. *Decision Analysis 2*(3), 165–181.

Matheson, J. E. (1990). Using influence diagrams to value information and control. In R. M. Oliver and J. Q. Smith (Eds.), *Influence Diagrams, Belief Nets, and Decision Analysis*, pp. 25–48. Chichester: Wiley.

Pearl, J. (1993). Comment: Graphical models, causality, and intervention. *Statistical Science 8*, 266–269.

Robins, J. (1986). A new approach to causal inference in mortality studies with sustained exposure results. *Mathematical Modeling 7*, 1393–1512.

Savage, L. (1954). *The Foundations o Statistics*. New York: Wiley.

Shachter, R. D. (1986). Evaluating influence diagrams. *Operations Research 34*(November-December), 871–882.

Shachter, R. D. (1999). Efficient value of information computation. In *Uncertainty in Artificial Intelligence: Proceedings of the Fifteenth Conference*, pp. 594–601. San Francisco, CA: Morgan Kaufmann.

Shachter, R. D. and D. E. Heckerman (1986). A backwards view for assessment. In *Workshop on Uncertainty in Artificial Intelligence*, University of Pennsylvania, Philadelphia, pp. 237–242.

Spirtes, P., C. Glymour, and R. Scheines (1993). *Causation, Prediction, and Search*. New York: Springer-Verlag.