

**INFORMS Philadelphia**  
**November 7, 1999**

# **Escaping the Monkey's Paw: The Value of Control Revisited**

**Ross D. Shachter**  
**Engineering-Economic Systems and Operations Research**  
**Stanford University**

joint work with  
**David E. Heckerman**  
**Microsoft Research**

**Why do decision analysts need causal models  
and what can they learn from them  
about the value of control?**

# Abstract

**Causal models are particularly useful for individuals making important decisions. Traditionally, decision analysts and other systems analysts have avoided using “causality” and other terms that suggest imprecision and informality. One concept from decision analysis that has gained limited acceptance is the “Value of Control (or Wizardry)”, the amount a decision-maker should be willing to pay to transform an uncertainty into a decision, that is, to choose its state. Although it is easy to calculate, it is not always clear how it should be interpreted. We examine it from a causal perspective to gain new insights and understanding. In particular, we formalize the conditions when the “Monkey's Paw” is not an issue, and when the “Value of Revelation” is equivalent to the Value of Control.**

# **Outline**

**A Decision Theoretic Approach to Causal Reasoning**

**The Value of Control**

**The Monkey's Paw**

**Avoiding the Monkey's Paw**

**Reframing the Value of Control**

**The Value of Revelation**

**Determinacy**

# **The Serenity Prayer**

**God grant me  
the Serenity to accept  
the things I cannot change,  
the Courage to change  
the things I can,  
and the Wisdom to know  
the difference.**

***from a poem credited to Reinhold Niebuhr (1892-1971)  
apparently inspired by Friedrich Oetinger  
and maybe Boethius, too***

# The Serenity Prayer

**God grant me  
the Serenity to accept  
the things I cannot change,  
the Courage to change  
the things I can,  
and the Wisdom to know  
the difference.**

Uncertainties

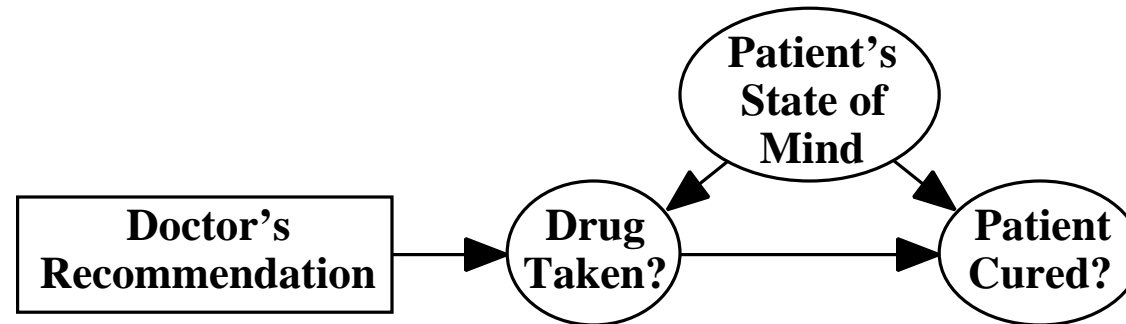
Decisions

Decision  
Analysis

*from a poem credited to Reinhold Niebuhr (1892-1971)  
apparently inspired by Friedrich Oetinger  
and maybe Boethius, too*

# Informal Preview to Causal Reasoning for Decision Analysis

Consider a physician choosing whether or not to recommend to her patient that he take a drug. Her beliefs are represented by the diagram:



Let  $U$  be the set of uncertainties

$$U = \{ \textit{Drug Taken?}, \textit{Patient's State of Mind}, \textit{Patient Cured?} \}$$

Let  $D$  be the set of decisions

$$D = \{ \textit{Doctor's Recommendation} \}$$

In our framework, we say that *Drug Taken?* is a cause of *Patient Cured?* with respect to *Doctor's Recommendation*.

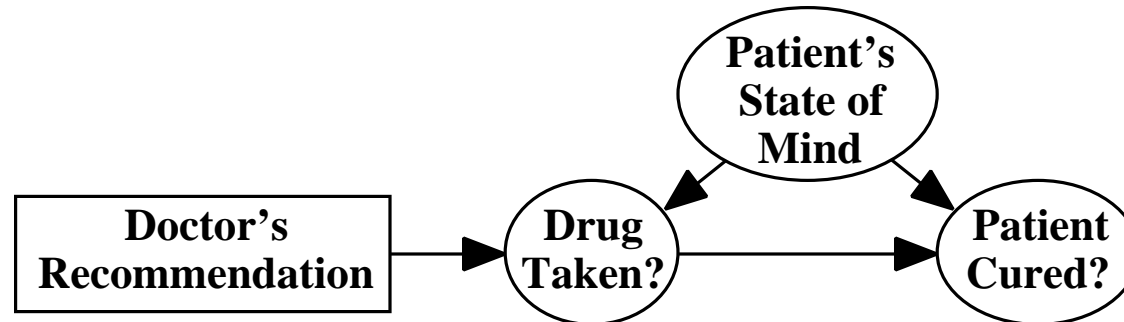
# Unresponsiveness

Uncertainty  $x$  is **UNRESPONSIVE** to decisions  $D$  if we believe that

$$x|D_1 = x|D_2$$

for any alternatives  $D_1$  and  $D_2$  for  $D$ ,  
in all states of the world.

Otherwise,  $x$  is said to be **RESPONSIVE** to decisions  $D$ .



*Patient's State of Mind* is unresponsive to *Doctor's Recommendation*.

*Drug Taken?* and *Patient Cured?* are responsive to *Doctor's Recommendation*.

If  $x$  is unresponsive to  $D$  then  $x$  and  $D$  are independent ( $\Pr\{x | D\} = \Pr\{x\}$ ), but not necessarily vice versa.

Unresponsiveness is a subjective judgement.

An uncertainty is responsive if and only if the decision maker has some control over it.

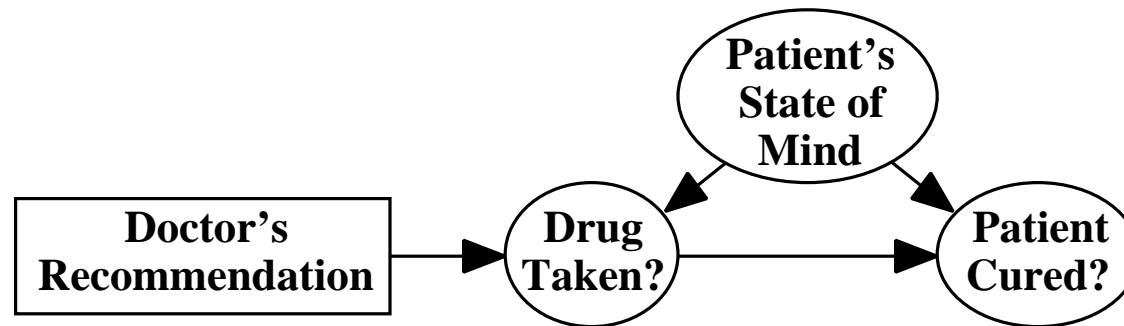
# Arc Reversal and Responsiveness

## Reversing Arcs Can Never Create Irrelevance

If A might be relevant to B given C in an influence diagram or belief network, arc reversals can never make A irrelevant to B given C in that diagram.

## Similarly, Reversing Arcs Can Never Create Unresponsiveness

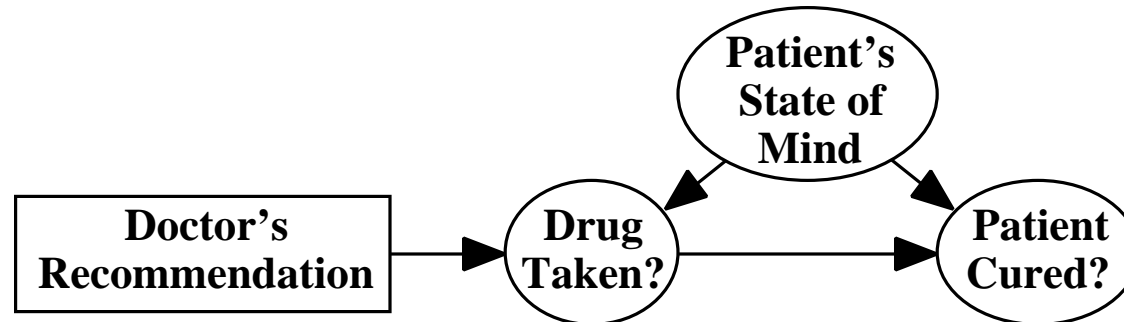
If A might be responsive to D in worlds limited by C, then arc reversals can never make A unresponsive to D in worlds limited by C in that diagram.



# Limited Responsiveness

Uncertainty  $x$  is **UNRESPONSIVE** to decisions  $D$  in **STATES LIMITED BY  $Y$**  if we believe that  $x|D_1 = x|D_2$  whenever  $Y|D_1 = Y|D_2$  for any alternatives  $D_1$  and  $D_2$  for  $D$ , in all states of the world.

Otherwise,  $x$  is **RESPONSIVE** to decisions  $D$  in **STATES LIMITED BY  $Y$** .



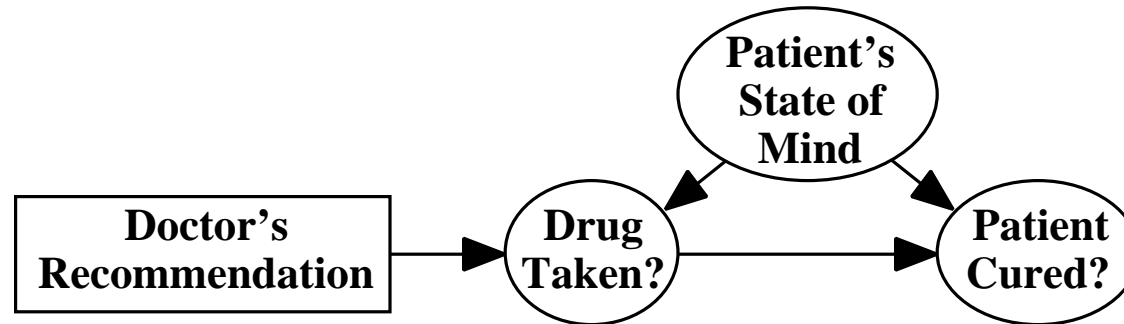
*Patient Cured?* is unresponsive to *Doctor's Rec.* in states limited by *Drug Taken?*.  
*Drug Taken?* is responsive to *Doctor's Rec.* in states limited by *Patient Cured?*.  
*Patient's State of Mind* is unresponsive to *Doctor's Rec.* in states limited by *Patient Cured?*.

Limited unresponsiveness does not imply conditional independence.

*Patient Cured?* is not independent of *Doctor's Rec.* given *Drug Taken?*.

# Cause

Given uncertainties  $U$ , decisions  $D$ , and uncertainty  $x \in U$ ,  
 $C \subseteq D \cup U \setminus \{x\}$  is said to be a CAUSE for  $x$  WITH RESPECT TO  $D$  if  
 $C$  is a minimal set of variables such that  
 $x$  is unresponsive to  $D$  in worlds limited by  $C$ .



*Drug Taken?* is a cause of *Patient Cured?* with respect to *Doctor's Rec.*  
*Doctor's Rec.* is a cause of *Patient Cured?* with respect to *Doctor's Rec.*  
*Doctor's Rec.* is a cause of *Drug Taken?* with respect to *Doctor's Rec.*  
 $\emptyset$  is a cause of *Patient's State of Mind* with respect to *Doctor's Rec.*  
(because *Patient's State of Mind* is unresponsive to *Doctor's Rec.*  
In other words, it is “exogenous.”).

All uncertainties in a cause are responsive to  $D$ .

It is conceivable to observe  $x$  if and only if  $x$  is unresponsive to  $D$

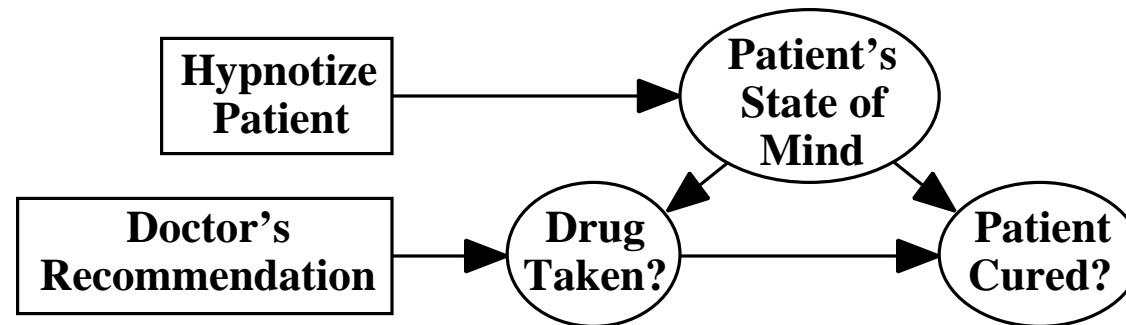
$\Rightarrow$  An uncertain effect cannot be observed before the decisions causing it are made.

# Why Define Causes With Respect to Decisions?

- Obtain an unambiguous definition based on narrow subjective judgements
- Concentrate on determining causes for uncertainties we can affect
- No need to assume all relevance relations are causal
- Avoid philosophical debate about the ultimate nature of causality

## Why Not Model Only Uncertainties and Introduce “Set” Decisions?

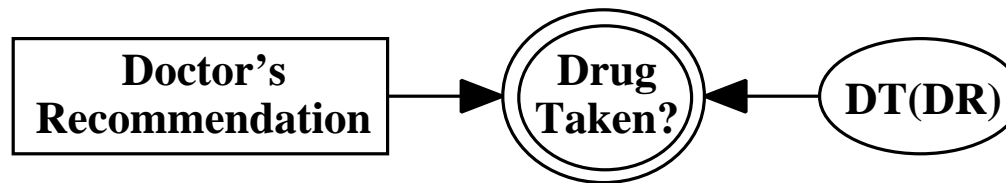
- We can recognize any “set” decision we can contemplate. For example, if we can contemplate controlling *Patient’s State of Mind*, perhaps through hypnosis, then a decision parent can be included for *Patient’s State of Mind*.



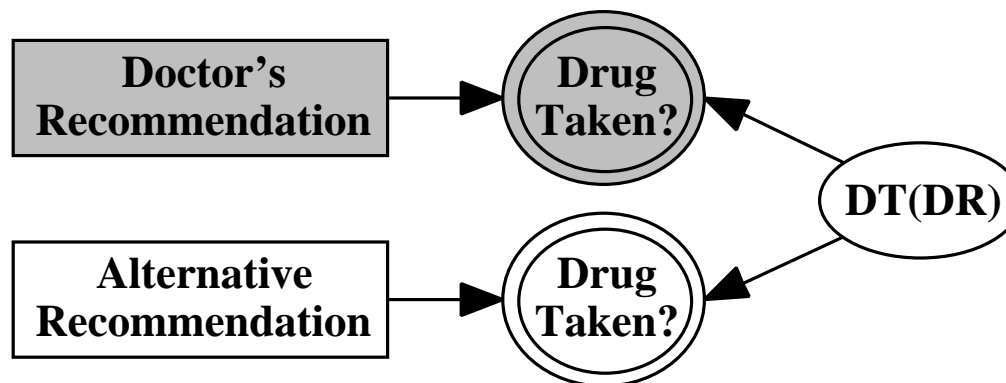
- We don’t need to consider set decisions for any uncertainty we don’t want to hypothesize controlling, or as a cause for another variable.
- A “set” decision for an uncertainty does not need to be unique.

# So, What Can We Do With This Causality?

- By explicitly modeling decisions, we are able to predict the effects of actions.
- We are thus able us to pick the best actions.
- We also gain insight into what new alternatives we might want to generate.
- If we are willing to do the extra work to decompose responsive uncertainties into a responsive part and an uncertain part, we get additional benefits:



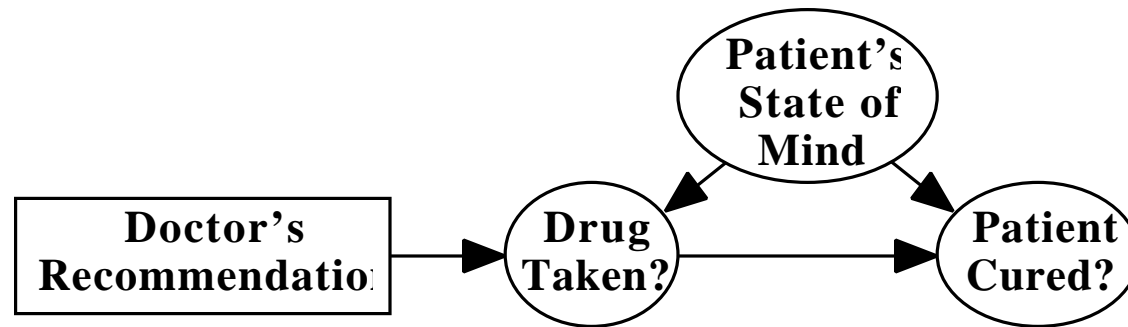
- We can now gather evidence about the "causal mechanism"  $DT(DR)$ .



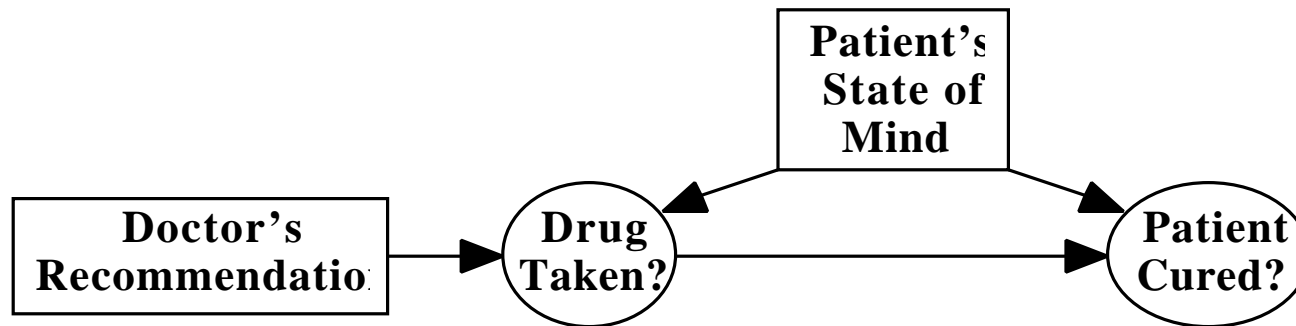
- We can also reason counterfactually, learning from observed responsive variables, to predict the effect of an alternative choice.

# The Value of Control

How much value is added to a situation if the decision maker can take control of some of the uncertainties and choose their states?



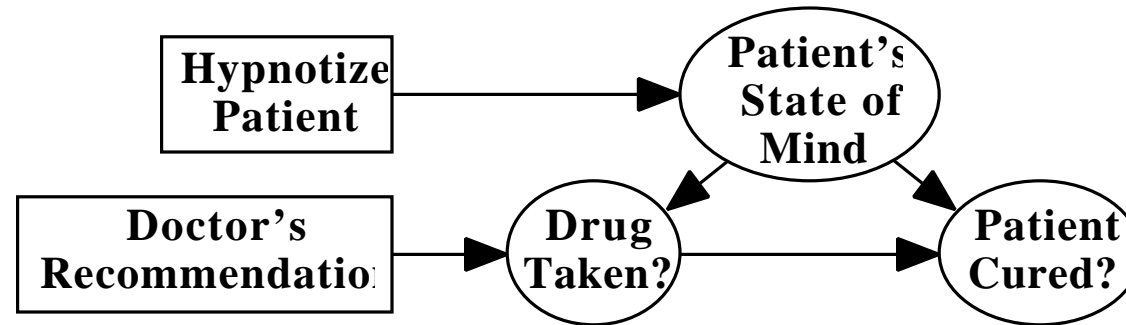
can be changed to



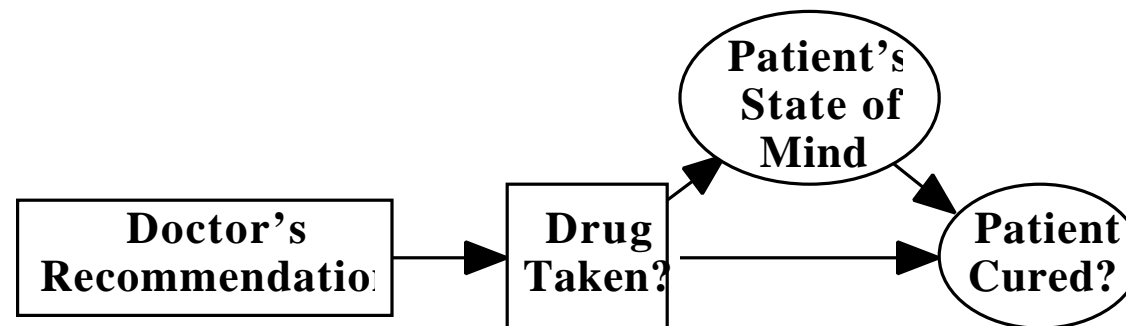
The difference can be computed exactly when utility is exponential (or linear) in dollars (or another numeraire), but what does it mean?

# The Value of Control (Continued)

It makes more sense to consider less perfect control:



It is even less clear how to interpret changes to variables that were not “exogenous:”



And, even more troubling, is the Monkey's Paw . . .

# **The Monkey's Paw**

*Jacobs, William Wyman. "The Monkey's Paw."  
Harper's Monthly, September 1902, 634-639.*

## **An Excerpt:**

"I should like to see those old temples and fakirs and jugglers," said the old man. "What was that you started telling me the other day about a monkey's paw or something, Morris?"

"Nothing," said the soldier, hastily. "Leastways nothing worth hearing."

"Monkey's paw?" said Mr. White, curiously.

"Well, it's just a bit of what you might call magic, perhaps," said the sergeant-major, off-handedly.

His three listeners leaned forward eagerly. The visitor absent-mindedly put his empty glass to his lips and then set it down again. His host filled it for him.

"To look at," said the sergeant-major, fumbling in his pocket, "it's just an ordinary little paw, dried to a mummy."

He took something out of his pocket and proffered it. Mrs. White drew back with a grimace, but her son, taking it, examined it curiously.

"And what is there special about it?" inquired Mr. White as he

took it from his son, and having examined it, placed it upon the table.

"It had a spell put on it by an old fakir," said the sergeant-major, "a very holy man. He wanted to show that fate ruled people's lives, and that those who interfered with it did so to their sorrow. He put a spell on it so that three separate men could each have three wishes from it."

His manner was so impressive that his hearers were conscious that their light laughter jarred somewhat.

"Well, why don't you have three, sir?" said Herbert White, cleverly.

The soldier regarded him in the way that middle age is wont to regard presumptuous youth. "I have," he said, quietly, and his blotchy face whitened.

"And did you really have the three wishes granted?" asked Mrs. White.

"I did," said the sergeant-major, and his glass tapped against his strong teeth.

"And has anybody else wished?" inquired the old lady.

"The first man had his three wishes, yes," was the reply. "I don't know what the first two were, but the third was for death. That's how I got the paw."

# The Monkey's Paw Lives On

The model itself depends on how the variables are set.

**“Magic Genie” problem: the mechanism isn’t specified**

- Does the patient believe he/she is receiving the treatment?
- Is medication administered covertly in his/her food?
- Is his/her jaw wired shut to prevent taking the treatment?
- Is the patient killed to prevent taking the treatment?

**“Fat Hand” problem: side effects possible even given a mechanism**

- Patient infected by dirty needle or other patients when treated
- Patient has to cut back on necessities to pay for treatment
- (Proximal components--heat in circuits, Piper Alpha)

**Bottom Line: the cause may depend on how the variables are set!**

**For example, we have a quite different causal model if**

- A hypochondriac patient is cured by believing he/she has received the treatment, even if it was just a placebo
- Treatment Taken? is set to False by killing the patient

# Avoiding the Monkey's Paw

There are several characteristics of the ideal intervention

Given an uncertain variable  $x$ , let  $x'$  be a decision.

We define  $x'$  as targeted for  $x$  if  
all other uncertain variables are unresponsive to  $x'$  in  
worlds limited by  $x$

We define  $x'$  as perfect for  $x$  if  
 $x'$  has an alternative corresponding to each possible state  
of  $x$  plus one more, "idle"  
when  $x'$  is set to a state then  $x$  takes on that state  
when  $x'$  is "idle", there is no manipulation of  $x$

We define  $x'$  as an atomic intervention on  $x$  if  
 $x'$  is both targeted and perfect for  $x$ .

There is no Monkey's Paw problem with an atomic intervention  
Note that when an intervention is perfect we "observe"  $x$  whenever  
 $x'$  is not "idle"

# **Avoiding the Monkey's Paw (Continued)**

**We have to be honest in how we model our interventions**

**Are our interventions really atomic?**

**Are they really targeted and perfect?**

**Magic Genie, Fat Hand are real issues**

**If we want to know the real effects of our intervention rather  
than the ideal,**

**we must model the real intervention and**

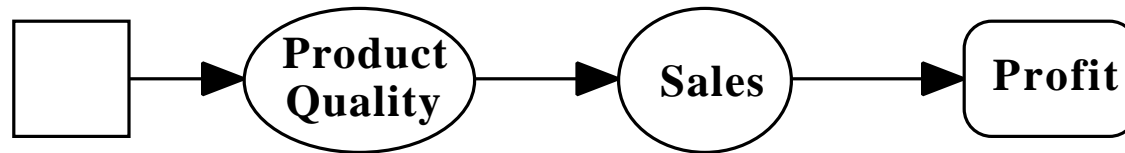
**take the care to recognize whatever “Monkey's Paws” might  
come with our interventions**

# Reframing the Value of Control

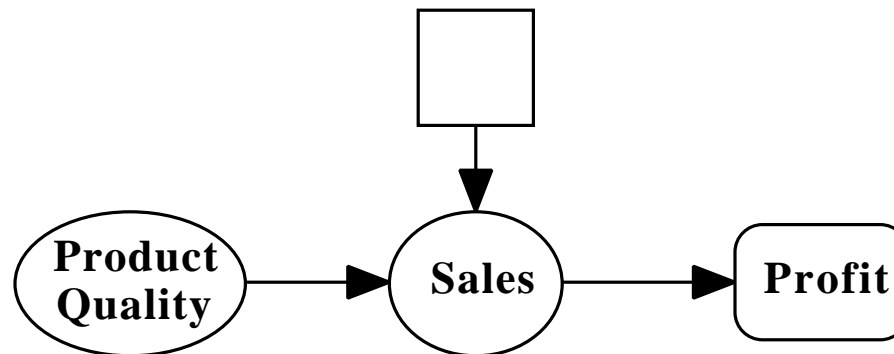
In light of these definitions, the Value of Control has normally been computed by assuming an atomic intervention:



becomes

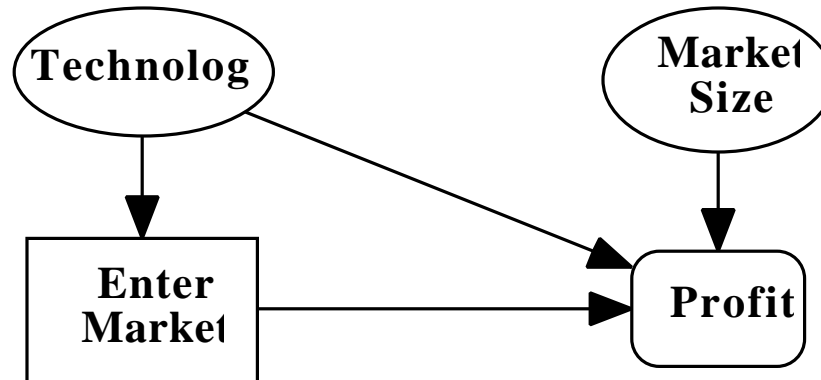


or

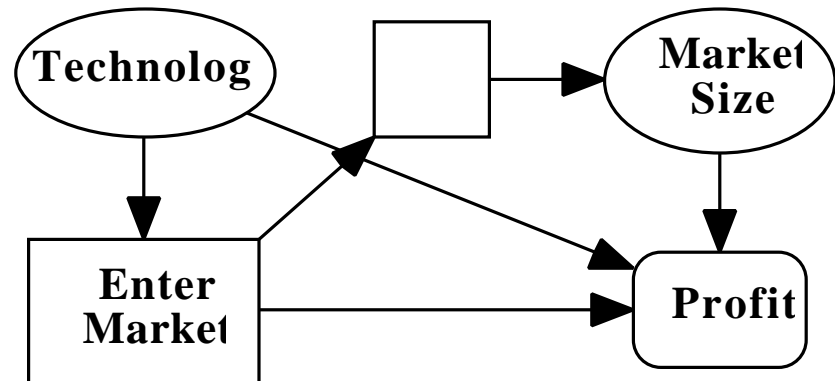
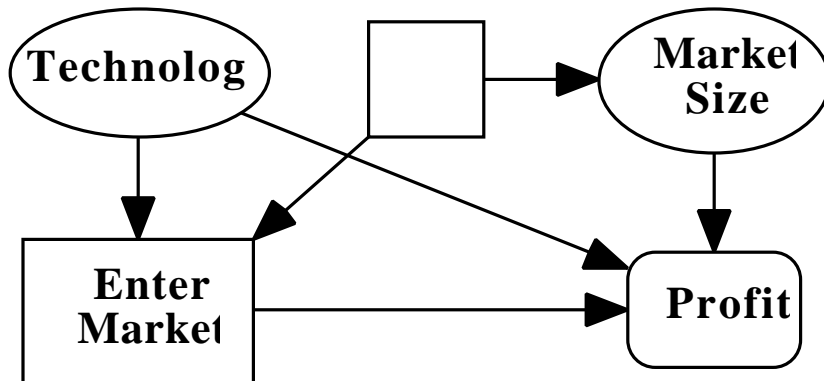


# Reframing the Value of Control (Continued)

But, in general, if we add a new decision to an influence diagram, we must also ask when the decision is made and what is known at the time.



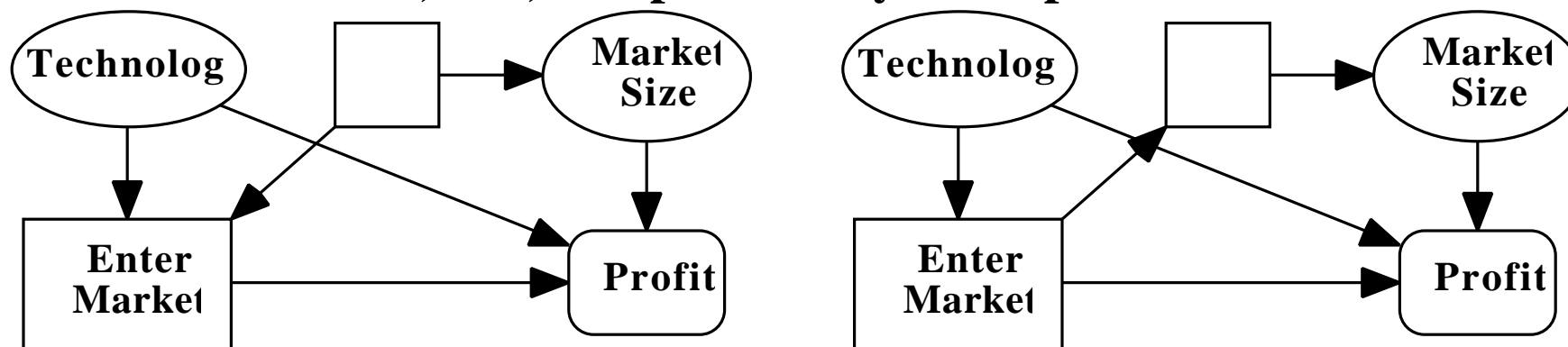
can become either of the two different situations below



Do we want to make a larger market for our competitor?

# Reframing the Value of Control (Continued)

Allow now, too, the possibility of imperfect controls.



Suppose we can affect, but not perfectly control the size of the market.

If we entertain the notion that the “Value of Control” lets us think about imperfect, untargeted interventions at times of our choice on the uncertainties we face, then we can reframe “Value of Control” as a fundamental paradigm of Decision Analysis.

Thus, even if we are thinking of perfect, targeted controls, when will be intervening and what we will know at the time?  
Furthermore, if the controls are not perfect, do we observe the results?

# The Value of Revelation

How much value is added to a situation if the decision maker can observe that some of the uncertainties have obtained particular states?



can be changed to either of the following



or

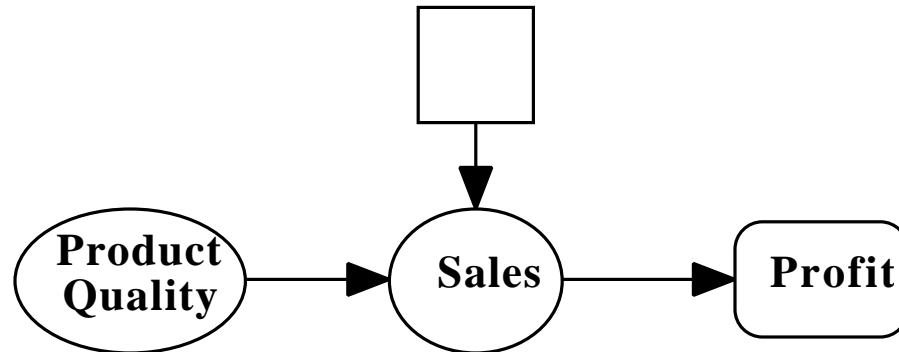


Technically, we are just interested in one term from the Value of Clairvoyance calculation.

This is always well-defined and interpretable, since it requires no intervention in the system.

When will a “peek” tell us what we would learn from a “poke”?

# The Value of Revelation (Continued)

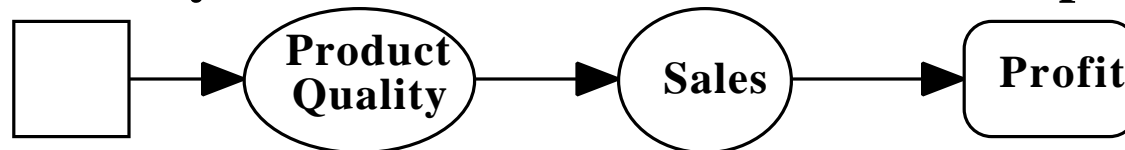


**The value of control is equal to the value of revelation for “Sales”**



**when there is no relevance through the parents of “Sales” to the value.**

**This is always true when the variable has no parents:**

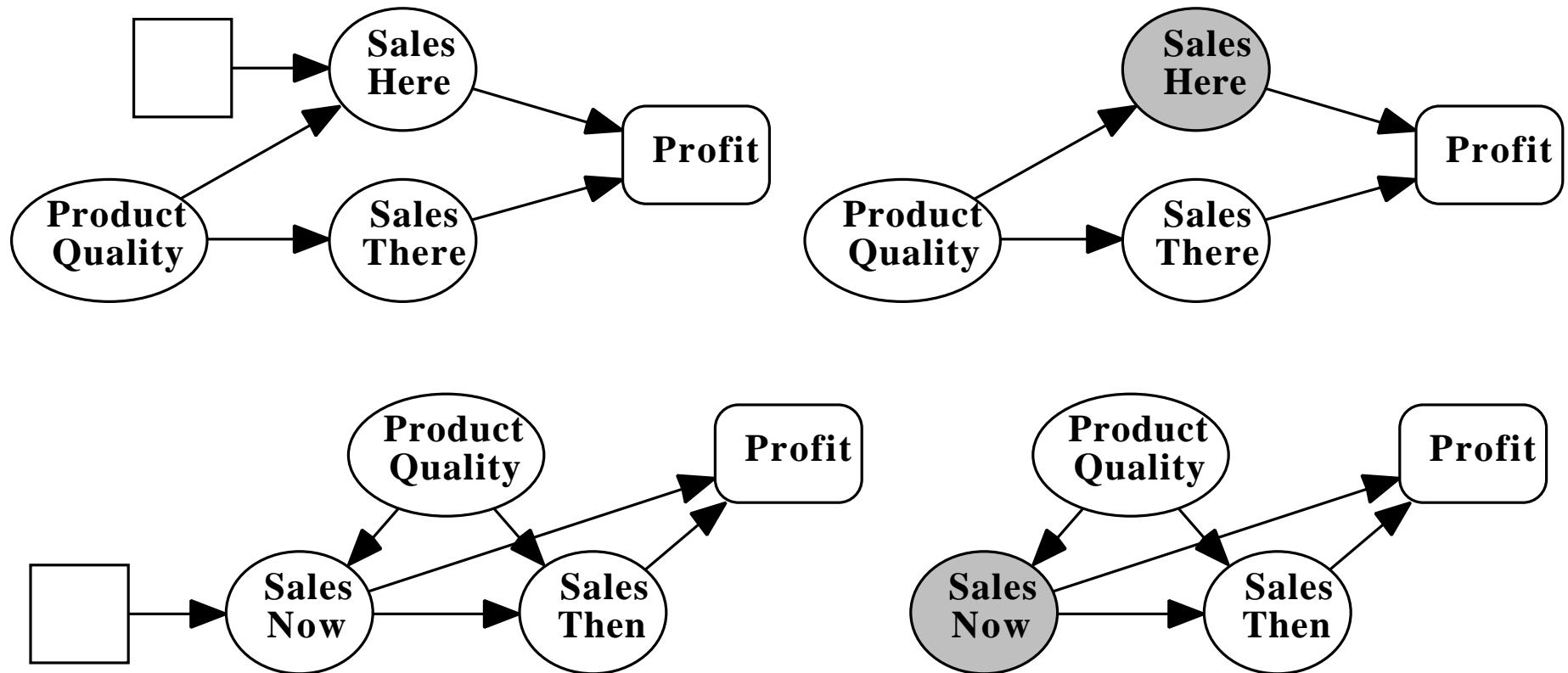


**The values of control and revelation are equal for “Product Quality”.**



# The Value of Revelation (Continued)

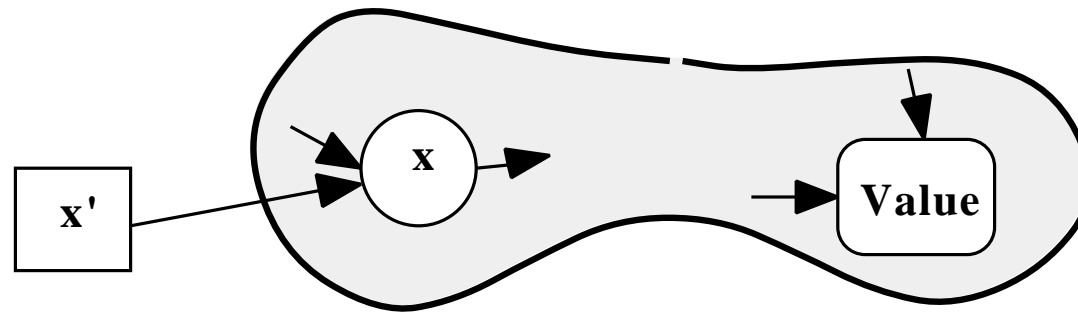
The values of control and revelation are not equal when there is relevance through the parents to the value.



This gets even more interesting when we start thinking of “when” (before which decisions) we get the revelation!

# Conclusions

When considering the value of control  $x'$  on an uncertainty  $x$



- **we have to ask: when we will exercise the control, how perfect it will be, and what we will be able to observe**
- **There is no “Monkey’s Paw” problem if  $x'$  is targeted for  $x$ , or if the value is unresponsive to  $x'$  in worlds limited by  $x$**
- **The value of revelation is equal to the value of control if the value is conditionally irrelevant to  $x'$  given  $x$**

**Thinking causally gives us valuable perspective about decision making!**