TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI

Is Cause and Effect Transitive?

Dr. Ralph deLaubenfels

TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI

Teacher-Scholar Institute

Columbus, Ohio

2018

IS CAUSE AND EFFECT TRANSITIVE?

Ralph deLaubenfels

ABSTRACT. The answer is "yes" for deterministic (meaning no nontrivial probability distributions involved) cause and effect, but we will demonstrate that the answer is "no" for probabilistic analogues of cause and effect.

I. INTRODUCTION and TERMINOLOGY. See [1] for the definition of a *relation* from one set into another. The relations of interest in this short paper are between events, described in Terminologies 1.3, 1.4, and 1.5.

Definition 1.1. A relation is **transitive** if

$$[(A \text{ is related to } B) \text{ and } (B \text{ is related to } C)] \text{ implies that } (A \text{ is related to } C).$$

The relation of "greater than" from the real line to itself:

$$(a \text{ is related to } b) \text{ if } b > a$$

is transitive. Boxing victories are not transitive: In the early 1970s, Joe Frazier defeated Muhammad Ali and Muhammad Ali defeated George Foreman, but Joe Frazier did not defeat George Foreman.

The words "cause" and "effect" are bandied about so glibly that a dictionary definition is in order.

Dictionary Definition 1.2. From [2], here is the definition of cause and effect: "Noting a relationship between actions or events such that one or more are the result of the other or others." Paraphrasing the same dictionaries' definition of cause: If A and B are events, A is a **cause** of B if B is a result of A, that is, A is a producer of the **effect** B.

Terminology 1.3. In the language of sets, A is the cause of B if $A \subseteq B$; this is denoted

$$A \to B, \quad \text{or} \quad A \textbf{ implies } B, \quad \text{or} \quad \textbf{if } A, \textbf{ then } B \quad \text{or} \quad A \textbf{ causes } B.$$

For example, if A means living in Columbus, Ohio and B means living in Ohio, then $A \to B$; living in Columbus, Ohio implies living in Ohio.

Armed with taxonomy, if A is being a primate and B is being a mammal, then $A \to B$; every primate is a mammal.

Now consider a popular medical assertion "smoking causes cancer." This is easily demonstrated to be false, by producing people who smoked but did not get cancer.

Here is what people probably mean when they say "smoking causes cancer." Denoting S for smoking and C for cancer, studies apparently have shown that

$$P(C|S) > P(C);$$

in words, "smoking increases the probability of getting cancer."

Let's have suggestive terminology for this sort of relationship between smoking and cancer.

For events A, B, at least if we ignore sets of probability zero,

$$A \to B$$

is equivalent to

$$P(B|A) = 1.$$

For example, if smoking (denoted S) literally caused cancer (denoted C), so that $S \to C$, we would have

$$P(C|S) = 1.$$

Terminology 1.4. For $0 \leq \omega \leq 1$, denote

$$A \to B \text{ with probability } \omega$$

for

$$P(B|A) = \omega.$$

The closer ω is to one, the closer $A \to B$ with probability ω is to A (literally) causing B.

Terminology 1.5. We will say that A **encourages** B if $P(B|A) > P(B)$ and we will say that A **discourages** B if $P(B|A) < P(B)$.

Note that the only remaining possibility, $P(B|A) = P(B)$, is the definition of independence of A and B.

For example, although smoking does not literally cause cancer, it does apparently encourage cancer.

II. STATEMENT of TRANSITIVITY RESULTS. Proposition 2.1 states that the nonprobabilistic cause and effect as in Terminology 1.3 is transitive, while Propositions 2.2 and 2.3 state that the probabilistic versions of cause and effect in Terminologies 1.4 and 1.5 are not transitive.

Since $A \subseteq B$ and $B \subseteq C$ implies that $A \subseteq C$, we have the following obvious result (see Terminology 1.3).

Proposition 2.1. If $A \to B$ and $B \to C$, then $A \to C$.

In words, if A implies B and B implies C, then A implies C. This is called **transitivity** of the implication relation $\to$.

Proposition 2.2. Events A, B, and C can be chosen so that

$A \to B$ with probability one, $B \to C$ with probability arbitrarily close to one, but $A \to C$ with probability zero.

That is, for any δ strictly between 0 and 1, there are events A, B, and C so that

$$P(B|A) = 1, \;\; P(C|B) = (1 - \delta), \;\text{ and }\; P(C|A) = 0.$$

Proposition 2.3. Events A, B, and C can be chosen so that

$$A \text{ encourages } B, \;\; B \text{ encourages } C, \;\text{ but }\; A \text{ discourages } C,$$

$$A \to B \text{ with probability one}, \;\; B \to C \text{ with probability arbitrarily close to one},$$

but

$$A \to C \text{ with probability arbitrarily close to zero.}$$

That is, for any ϵ strictly between 0 and 1, there are events A, B, and C so that

$$P(B|A) > P(B), \;\; P(C|B) > P(C), \;\; P(C|A) < P(C),$$

$$P(B|A) = 1, \;\; P(C|B) > (1-\epsilon), \;\; \text{and} \;\; P(C|A) < \epsilon.$$

Discussion and Examples 2.4. The significance of the implication $A \to B$, even in the weaker probabilistic sense of Terminologies 1.4 and 1.5, if A is intrinsically harmless (that is, harmless when viewed in isolation from all possibly bad consequences) and B intrinsically bad, is that A inherits the badness of B, and may be pursued with the same vigor with which B is pursued. Stated more positively, we have the opportunity to diminish B before it appears, by attacking A. If A is smoking and B is cancer, $A \to B$ says that we may prevent many cases of cancer before their onset by cutting down on smoking. If A is reading comic books and B is juvenile delinquency (we give hints about our age here), $A \to B$ says that we may reduce juvenile delinquency by keeping comic books away from juveniles. Why wait for a crime to be committed, the responsible social engineer argues, when we can prevent crime before it is committed?

Here are other popular examples of $A \to B$, at least in one of the senses of Terminologies 1.3–1.5; which sense is plausible we leave to the reader.

getting wet $\to$ getting a cold.

rock and roll $\to$ violence.

playing with candy fake cigarettes $\to$ smoking

naughty words $\to$ naughty ideas.

obesity $\to$ heart attack.

cholesterol $\to$ high blood pressure.

high blood pressure $\to$ heart attack.

If we are truly intent on stopping bad things by stopping intrinsically harmless things that lead to bad things, here is a natural extension of our discussion so far. If A and B are intrinsically harmless, but C is bad, and

$$A \to B \to C,$$

shouldn't we stop, not only B, but A? For example, we already know that cigarettes $\to$ (in some sense) cancer; if we also know, as in our list above, that (candy fake cigarettes) $\to$ (smoking), so that we have a string

candy fake cigarettes $\to$ smoking $\to$ cancer,

should we (as apparently happened; we haven't seen any candy fake cigarettes since our childhood) get rid of candy fake cigarettes?

If, again embellishing some of the examples above,

advertising electric guitars $\to$ some people buy electric guitars $\to$ rock and roll is produced $\to$ violence,

should we outlaw both the advertising and purchasing of electric guitars?

More generally, if we can set up a string of dominoes of intrinsically harmless activities, concluding with a bad activity, with the toppling of one domino making all the others, up to the bad one, also topple, should we get rid of all the harmless dominoes?

Propositions 2.2 and 2.3 suggest that, unless we have the literal cause and effect of Proposition 2.1, the reasoning of the preceding three paragraphs is faulty. For example, if (electric guitars) cause (rock and roll) and (rock and roll) causes violence, in both cases as in Terminologies 1.4 and 1.5, it does *not* follow, just from that information, that (electric guitars) cause violence. Proposition 2.2 states that

$$P(\text{violence}|\text{electric guitars}) \text{ could equal zero;}$$

Proposition 2.3 states that (electric guitars) might *dis*courage violence.

III. CONSTRUCTION of TRANSITIVITY FAILURES. Throughout this section, A, B, and C are events, with probabilities of all possible intersections given in tables.

For any event E, "not E" (terminology appropriate for a "Dick and Jane" reading book) is the complement of E.

Demonstration of Proposition 2.2.

Intersections of A and B with C

	B	(not B)
A	0	0
(not A)	$(1-\delta)$	0

Intersections of A and B with (not C)

	B	(not B)
A	δ	0
(not A)	0	0

We leave it to the reader to calculate

$$P(B|A) = \frac{\delta}{\delta} = 1, P(C|B) = \frac{(1-\delta)}{1} = (1-\delta), \text{ and } P(C|A) = \frac{0}{\delta} = 0.$$

Demonstration of Proposition 2.3. In the following, δ could be any number strictly between 0 and $\frac{1}{6}$.

Intersections of A and B with C

	B	(not B)
A	δ^2	0
(not A)	$(1-\delta^2-4\delta)$	δ

Intersections of A and B with (not C)

	B	(not B)
A	δ	0
(not A)	δ	δ

We again leave it to the reader to calculate

$$P(B|A) = \frac{(\delta^2+\delta)}{(\delta^2+\delta)} = 1, P(C|B) = \frac{(1-4\delta)}{(1-2\delta)}, P(C|A) = \frac{\delta^2}{(\delta^2+\delta)} = \frac{\delta}{(\delta+1)},$$

$$P(B) = (1-2\delta), \text{ and } P(C) = (1-3\delta).$$

We need to verify that the assertions of Proposition 2.3 are true.

$$P(B|A) > P(B) \text{ and } P(B|A) = 1$$

are clear.

Since

$$P(C|A) = \frac{\delta}{(\delta+1)} < \frac{\delta}{1} = \delta,$$

$P(C|A) < \epsilon$ occurs whenever δ is chosen to be less than ϵ.

Also, since

$$\delta < (1-3\delta) \iff \delta < \frac{1}{4},$$

we also have

$$P(C|A) < \delta < (1-3\delta) = P(C)$$

whenever $\delta < \frac{1}{4}$.

$P(C|B) > P(C)$ is a result of the following calculation.

$$\frac{(1-4\delta)}{(1-2\delta)} > (1-3\delta) \iff (1-4\delta) > (1-2\delta)(1-3\delta) = 1-5\delta+6\delta^2 \iff \delta > 6\delta^2 \iff \frac{1}{6} > \delta.$$

Since

$$P(C) = (1-3\delta) > (1-\epsilon) \iff \delta < \frac{\epsilon}{3},$$

this also shows that $P(C|B) > (1-\epsilon)$, for δ sufficiently small.

REFERENCES

[1] Paul R. Halmos, *Naive Set Theory,* Springer-Verlag, 1974

[2] *The Random House College Dictionary,* 1973