TECHNICAL REPORT
R-283b
February 2001

Comments on Nozer Singpurwalla's "On Causality and Causal Mechanisms"

Judea Pearl
Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles, CA 90024
judea@cs.ucla.edu

Professor Singpurwalla is anxious to prove that "the calculus of probabilities, endowed with a time dynamic, is indeed the calculus of causality".

As I have explained in my comments on Lindley's review, the effort to embed causality within probability theory cannot succeed and, even if successful, would not yield desirable or useful results. I have also explained why probability calculus need be extended with new notation to cover causal relationships. Interestingly, the idea of adopting new notation for expressing causal relationships has been traumatic to most persons trained in statistics; partly because the adaptation of a new language is difficult in general, and partly because statisticians have been accustomed to assuming that all phenomena, processes, thoughts, and modes of inference can be captured in the powerful language of probability theory.[1]

I will here demonstrate that:

1. Professor Singpurwalla's exercise of Section 4.1, which he presents as being entirely "within the calculus of probability", invokes in fact elements of a brand new calculus, foreign to the syntax of standard probability theory, and very akin to the causal calculus that Singpurwalla deemed unnecessary.

2. Time dynamic, combined with probability calculus are not sufficient for representing causal information.

[1] Commenting on my $do(x)$ notation a leading statistician wrote: "Is this a concept in some new theory of probability or expectation? If so, please provide it. Otherwise, 'metaphysics' may remain the leading explanation." (e-mail correspondence)

1. My humble engineering background may disqualify me from officially challenging authorities in questions of philosophical foundations, but, as a computer scientist, I am quite sensitive to questions of notation and syntax. Armed with such sensitivities, I am on notice that Professor Singpurwalla's exercise in Section 4.1 invokes two connectives: the usual conditioning bar, as in $P(y|x)$, and the *semicolon*, as in $P(y|x; H)$. The semicolon, to the best of my knowledge, is not part of standard probability theory; it is conspicuously absent from Kolmogorov's axioms as well as other axiomatic theories of probability that I have studied. Moreover, unlike the conditioning operator, one cannot express $P(y; x)$ in terms of the joint probability $P(y, x)$ or of the joint probability of all variables in the space.

True, the semicolon is not foreign to statistical writings; it is used frequently to distinguish parameters from variables, and the expression $P(x; \alpha, \beta)$ is rampart in the literature on multivariate distributions. Yet, in this context, the semicolon serves merely as an index to distinguish one density function from another and, as an index, the semicolon can be excused from intruding upon the traditional syntax of probability calculus.

Singpurwalla is using the semicolon in a drastically more liberal way, characterized by two new features. First, in addition to parameters, the semicolon now embraces ordinary events from the probability space, such as the event $X = x$ in $P(U, Y; x, H)$ (Eq. (4.2)). Second, certain events are permitted to cross over from one side of the semicolon to the other (as in Eq. (4.4)), as well as to be deleted from the right hand side of the semicolon (as in (4.3)). Such unprecedented transformations are not normally associated with indexical symbols, and the questions arise: (1) What are the conditions that legitimize such transformations? and (2) Can these conditions be expressed in terms of probabilities and temporal information? A formal answer to the first question will mark the birth of a new calculus, the "calculus of semicolon". A positive answer to question (2) would establish the new calculus as a legitimate machinery within the framework of probability theory, paralleling (and supplementing) the traditional calculus of conditioning. But a negative answer to question (2) would place the new calculus outside the framework of probability theory.

Singpurwalla explains the conditions that legitimize the transformations in Eqs. (4.3) and (4.4), but his explanations are mostly verbal and informal, involving new primitives such as "histories", "control", "knowledge", "observations", "in light of H", and other nonstandard relationships. It is not clear, for example, what variables may enter the "history" part of the expressions, H. Singpurwalla says that if "X were to be controlled, or selected at some value, say x. Then x should become a part of the history." But if this is the sole criterion for becoming a part of the history, then the entire past should enter "history", regardless of whether events in the past were controlled, observed or left unknown. Consequently, it is not clear why observing the barometer drop would generate different predictions about the weather from, say, squeezing the barometer and causing it to drop; in both cases the barometer reading would enter the "history" part of the formula with the same value.

The desiderata of showing that "the calculus of probabilities, endowed with a time dynamic, is indeed the calculus of causality" requires that such ambiguities be eliminated and that all considerations regarding the semicolon symbol be reduced to formal sentences involving probabilities of time-indexed variables, nothing else.

I have no doubt that some formal explication of these considerations is feasible, because the transformations in Eqs. (4.3) and (4.4) are special cases of Rules 2 and 3 of *do*-calculus (with the semicolon representing the *do* operator) where they are given formal conditions of applicability (*Causality*, page 85, Theorem 3.4.1). However, in *do*-calculus, these applicability conditions are expressed in terms of properties of a directed acyclic graph which, in turns, encode knowledge of causal mechanisms. Singpurwalla aspires to formulate these conditions in terms of a probability distribution alone, enriched with temporal information. This aspiration, as the next example demonstrates, is untenable.

2. Consider a switch X that turns on two lights, Y and Z, and assume that, due to differences in location, Z turns on a split second before Y. This example represents functional relationships between the three variables, which, according to Singpurwalla's Section 2.1, is sufficient to characterize Z as a (spurious) cause of Y. Consider now a variant of this example where the switch X activates Z, and Z, in turns, activates Y. This case is identical in all respects to the previous one, because both the functional and temporal relationships are identical. Yet few people would perceive the causal relationships to be the same in the two situations; the latter represents cascaded process, $X \longrightarrow Z \longrightarrow Y$, while the former represents a branching process, $Y \longleftarrow X \longrightarrow Z$. Intervening on Z would affect Y in the cascaded case, but not in the branching case. Singpurwalla's "semicolon calculus" would be ade-

quate for causality if it were to entail the relations $P(Y=1; Z=1, H)=1$ in the cascaded case and $P(Y=1; Z=1, H)=P(Y=1; H)$ in the branching case. Not surprisingly, it does not.

The preceding example illustrates the impossibility of defining causation in terms of temporal-functional relationships without attending to the mechanisms that sustain those relationships. In the branching case, for example, although all three variables are symmetrically constrained by the functional relationships: $X=Y$, $X=Y$, $Z=Y$, these relationships in themselves do not reveal the information that the three equalities are sustained by only two mechanisms, $Y=X$ and $Z=X$, and that the first equality would still be sustained when the second is modified. A set of mechanisms, each represented by an equation, is not equivalent to the set of algebraic equations that can be assembled from those mechanisms. Mathematically, the latter is defined as *one* set of n equations, whereas the former is defined as n separate sets, each containing one equation. These are two distinct mathematical objects that admit two distinct types of solution-preserving operations. The calculus of causality deals with the dynamics of such modular systems of equations, where the addition and deletion of equations represent interventions.

In summary, Professor Singpurwalla attempt to show that "the calculus of probabilities, endowed with a time dynamic, is indeed the calculus of causality" is deficient on three accounts: (1) the rules for his "calculus of semicolon" are not explicated *formally*, (2) these rules are not expressed in terms of probabilistic and temporal information *alone* and, finally, these rules *cannot* be formulated in terms of probabilistic and temporal information alone; knowledge of causal mechanism must be invoked.

Singpurwalla's reluctance to accepting new ideas in statistics is understandable in light of his low opinion of other professions (e.g., engineering), but his resistance to enriching probability calculus with new capabilities is enigmatic, for it stands in glaring contradiction to the conception of mathematics as "an expression of the human mind." (Courant and Robbins (1981) "*What is Mathematics?*") The human mind speaks cause and effect – mathematics should echo this language.