Theories of Causal Ordering:
Reply to de Kleer and Brown

Yumi Iwasaki and Herbert A. Simon

Department of Computer Science

Carnegie-Mellon University

Theories of Causal Ordering:
Reply to de Kleer and Brown

Yumi Iwasaki and Herbert A. Simon
Carnegie-Mellon University

Abstract

In their reply to our paper, *Causality in Device Behavior*, de Kleer and Brown (*Comments*) seek to establish a clear product differentiation between the well-known concepts of causal ordering and comparative statics, on the one side, and their "mythical causality" and qualitative physics, on the other. Most of the differences they see, however, are invisible to our eyes. Contrary to their claim, the earlier notion of causality, quite as much as the later one, is qualitative and "derives from the relationship between the equations and their underlying components which comprise the modeled system." The concepts of causal ordering and comparative statics offer the advantage of a formal foundation that makes clear exactly what is being postulated. Hence, they can contribute a great deal to the clarification of the causal approaches to system analysis that de Kleer and Brown are seeking to develop.

In this brief response to the *Comments*, we discuss the source of the structural equations in the causal ordering approach, and we challenge more generally the claim that there are inherent differences (e.g., in the case of feedback) between the "engineer's" and the "economist's" approach to the study of system behavior.

Introduction

We will focus, as de Kleer and Brown do, on two issues, referring to their paper as *Comments* and organizing our remarks under the same section headings that they use. The first issue concerns the relation between the formal (syntactical) definition of causality and the empirical (semantic) causal relations among the components of the system being modeled. The second concerns the operation of causality in the presence of feedback.

Relation of Modeled Situation to Model. As de Kleer and Brown (*Comments*) observe, "a standard engineering approach to analyzing a device starts with its components and connections," or in the language of causal ordering, its underlying mechanisms. From a knowledge of these, a set of equations is formulated, *each member of the set representing the working of a distinct mechanism, hence of a specific connection among components.* In the literature on causal ordering, equations that correspond, individually, to mechanisms are called *structural equations.*

At the syntactical level, we can speak of causal relations among variables, defined by the structural equations; at the semantic level, we can speak of corresponding causal relations among the components, defined by the mechanisms. This is as true of the "economic" as of the "engineering" approach, and there is absolutely no difference between causal ordering and mythical causality on this score. Any question on this point as it applies to economic modeling, can be laid to rest by consulting the literature, for example, pages 51-52, 67-70, 85-86, 88-90, and 116-124 of Simon's essays on causal ordering referred to in our article (1986) and reprinted in his (1977). There is no "down playing" of modeling in the causal ordering literature. Hence, the first of the two "central differences" that de Kleer and Brown mention simply doesn't exist.

As corollary to this (non-existent) difference, de Kleer and Brown (*Comments*) claim that the modeling and interpretation steps in the causal ordering approach (See their Figure 1) "are insufficiently formalized to be automated," while these same steps can be automated in the qualitative physics approach. But the basis for their formalization is "a description of the physical situation in terms of components and their interconnections." Any situation so

described can be automated in the same way – – e.g., the evaporator example in our paper. Which approach is used has nothing to do with how formal a description of the situation is provided.

Nor need the system to be modeled be "physical." Consider the system: amount of rainfall – – –> size of wheat crop – – –> price of wheat. To build a model of this system, we begin with three components: the weather, the wheat-growing component, the consumption component. The weather is taken as an exogenous variable, whose causal mechanism is external to the system. There is an "agronomical" connection between wheat growing and the weather, and a marketing connection between the amount of wheat offered on the market and its price. Any formal system that will automate the modeling and interpretation of the evaporator will automate the modeling and interpretation of this simple set of mechanisms.

On the basis of these considerations, we must reject the claim that there is a difference in the relation of model to system being modeled between the two approaches to causality.

Qualitative Relations versus Differential Equations. It is further claimed (*Comments*) that the causal ordering approach "requires solving a set of conventional differential equations" rather than propagating qualitative disturbances through a structure. It is hard to see on what this claim rests, especially since most of the examples of causal ordering in the literature (including our evaporator) involve algebraic relations describing a system's steady state, and not differential equations at all.

Moreover, the only properties of the equations that are used in arriving at the causal ordering are the presence or absence of particular variables in particular equations. No use is made of the algebraic properties of the equations (including their linearity); only combinatorial properties are considered. How then can it be claimed that we "require conventional algebra to determine the qualitative behavior?" In fact, our determination of causal orderings is no less, and no more, qualitative than the method of propagation used with mythical causality. (See Sect. 3.43, pp. 26-27, of de Kleer and Brown, 1984.)

But perhaps de Kleer and Brown are confusing the techniques for determining causal ordering with our later discussion of feedback loops and the use of comparative statics to

determine the signs of perturbations. We separate these two issues since they are, in fact, quite independent of each other; a distinction not made in Qualitative Physics. Clearly they are speaking of comparative statics when they state our method "only determines the initial response of a system to a disturbance," while "qualitative physics and conventional physics produce a description of the behavior of the system *over time.*" Even this is not quite correct, for what is determined by comparative statics is not the initial response but the net response upon return to (static or steady state) equilibrium. This is why the question of dynamic stability -- -- whether the system will in fact return to equilibrium -- -- is so important.

Instead of the distinction between determination of causal ordering and determination of the signs of the effects of a disturbance, a distinction is made in Qualitative Physics between inter-state and intra-state behavior. However, this distinction is an artifact of the way rules are written and does not correspond to instantaneity and non-instantaneity in real time. This fact is obscured by de Kleer and Brown's description of intra- and inter-state behavior as behavior over time and instantaneous behavior (or behavior over "mythical time"), respectively. No causal effect actually propagates instantaneously from one variable to another. The only case in which a change in one variable should be reflected "instantaneously" in another variable is when the relation between the two variables is not causal but definitional, where one variable is *defined* in terms of the other. However, it is obvious that such relations have nothing to do with causality.[1] .

In *Comments* the claim is made, in contrast to this supposed limitation of comparative statics, that "qualitative physics provides a general procedure for solving a set of possibly nonlinear qualitative differential equations." This is a very strong claim, and we have not seen anything in the authors' papers that substantiates it. In particular, their intra-state propagation technique is not a method for solving differential equations that describe the general movement of a dynamic system over time, but a method for studying the displacement of a system from one equilibrium or steady state to another. For a non-linear system it will work only in the neighborhood of a local equilibrium, that is, in the domain where a linear

---

[1] See our later discussion in the section on Feedback in Conduit

approximation to the system can be employed. The causal ordering approach has the same limitation – – and no more.

Propagation of Constraints

In the second section of *Comments*, de Kleer and Brown review their method of propagating constraints, introducing in the process the unnecessary neologism, "plunking." As their example shows, what they have provided is a convenient algorithm for solving linear equation systems whose matrices are sparse, and which therefore don't require the heavy artillery of the Gauss-Seidel method. In comparing propagation of constraints with Gaussian elimination in Section 2.1 of *Comments*, de Kleer and Brown seem to regard the method of causal ordering as being equivalent to Gaussian elimination, but this is simply not the case.

Now what is the relation of their algorithm to causal ordering? If there is a simple causal ordering among the variables (as there is not in the example), then the propagation algorithm will solve for variables in the order of their causal precedence, starting with the exogenous variables. No "plunking" (elimination of a variable by solving for it in terms of other variables and removing it from the equations) is called for.

If the equation system is self contained and consistent, but not simply causally ordered, then variables must be eliminated by substitution along the way, and the selection of these variables is arbitrary (essentially, in any minimal self-contained subset, any one of the variables may be retained, and the others must be removed). Hence, as Figure 2 in de Kleer and Brown's *Comments* shows, the order of solution need not correspond to the causal ordering.

We must observe here that the constraint propagation process described in (De Kleer and Brown, 1984) is not exactly equivalent to the constraint propagation process with plunking described in *Comments* in at least one significant respect. In the former there is no notion of a system being underdetermined, because there is no explicit notion of a system boundary and consequently no checking for the system's being underdetermined; while in the latter, when plunked variables remain unsolved, it becomes obvious that the system is underdetermined. Since the process as described in (De Kleer and Brown, 1984) of

propagating constraints and making assumptions using heuristics continues until all the variables are assigned new values, underdetermination does not become an issue. When the system is underdetermined, the process simply produces *all* interpretations that are consistent.[2] When the system is self-contained, an assumption introduced (or a plunked variable) serves to break a feedback loop to impose a direction around the loop. If the system is underdetermined, in the propagation process as described in (De Kleer and Brown, 1984) assumptions are introduced to make the process go through when it gets stuck. Adding these assumptions by the use of heuristics is in fact equivalent to defining exogenous variables and holding them constant.

But if causal ordering is to be more than a syntactical exercise, the selection of variables to be treated as exogenous must be more than a matter of algorithmic convenience – – it must reflect the components and connections, the mechanisms, of the underlying real-world system. The "plunking" heuristics that are described in de Kleer and Brown's comment do not reflect this concern with the mechanisms, but only with computation. This seems inconsistent with their general desire to reflect the real system in their mythical causality.

Comparison of the Approaches to Causality

In the section of *Comments* comparing the differing approaches to causality, the erroneous claim is repeated that "we [de K. and B.] link causality directly to the structure of the system while Iwasaki and Simon link causality to the form of the equations describing the system." We have already seen that neither half of this claim is correct. On the one hand, their propagation algorithm responds to the form of the equations; on the other hand, the causal ordering approach insists that the equations of the model must be structural equations, each describing a mechanism.

The authors now draw six consequences from this incorrect claim. We agree with the first: "If no feedback is present, then both approaches specify the same causal ordering."

---

[2]This is also the reason that constraint propagation described in (De Kleer and Brown, 1984) works even when equations are redundant. The process does not detect the redundancy, and we suspect that all such redundancies will be interpreted as presence of feedback, provided the equations are consistent.

The claim that we specify no causal ordering within a feedback loop is also correct. Ordering the variables in a loop is like choosing between the causation of chickens by eggs and of eggs by chickens, and seems pointless.

The third claim asserts that causal ordering "requires a global analysis" while mythical causality "specifies none." Since both methods are based on the combinatorial properties of the equation system, the claim seems surprising. That it is wrong, can be seen from de Kleer and Brown's example, which shows that the computation depends on locating subsets of equations with appropriate numbers of variables, and that the "plunking" heuristics are precisely concerned with choosing variables within subsets. No more globality than this is required for causal ordering.

Similarly, we have shown that there is no basis for the claim that the one method is more "algebraic" than the other.

With respect to constraints on the equations, assumptions of linearity play no role in the causal ordering approach. With respect to redundancy, the method of mythical causality will produce consistent results only if the equations are not overdetermined.

Finally, it is not true that the causal ordering method "requires the original quantitative model." The causal ordering can be determined simply from a knowledge of which variables appear in which equations – – precisely the knowledge required to apply the propagation algorithm.

Thus, of the six consequences drawn, four are incorrect, while the other two do disclose one difference between the two approaches.

Feedback in the Conduit

What this section of the *Comments* chiefly illustrates is that there may be more than one way of writing structural equations for a system, and that one person's intuitions about the appropriate formulation may be another's perplexities. The authors provide no reasons why a conduit with a load has to be represented by two "narrow conduits" (one of them imaginary) especially since they add that "the narrow conduits are not the ideal conduits of qualitative physics . . . but are components." Note that on page 30 of de Kleer and Brown (1984) they

introduce a load, without introducing an additional mythical narrow conduit, in exactly the way we do. Their principles of qualitative physics seem here to be quite flexible and wholly informal.

The equation (I2) in our formulation simply represents the load mechanism (or "component" if you please) as a causal response of the output pressure to the flow rate. Thus, equation (I1) describes the conduit, which determines the flow rate as a function of pressure drop; the second equation (I2) describes the load, which determines output pressure as a function of flow rate; the third (I3) declares the input pressure to be an exogenous variable. These equations provide a straightforward representation of the structure of the conduit with its load.

Now consider the alternative system: (F1) through (F6). If we use the definition of P_{io} supplied by equation (F6) to substitute in (F1), then (F1) becomes identical with (I1). By noticing at once that the quantity of flow is conserved through the system, the subscripts on the Q's can be dispensed with, and the equations (F3) through (F5) can be removed from the system. (F2) is then identical with (I2). In place of explicitly declaring P_i to be exogenous (Equation I3), the alternative system defines an input disturbance (V1).

Thus, the two sets of equations are fully equivalent, and must describe the same system behavior. The only difference between them is one of "grain size." The equations (F1) through (F6) first identify the different volumes of liquid at different points in the conduits, then declare these volumes all to be equal; in the system (I1) through (I3), the equality is assumed from the start. Finally, it is curious that an equation like (F6), which is purely definitional, is included in a system that is supposed to represent physical components and connections between them. We avoided such an artifact by introducing the real physical pressures, instead of their defined difference, in (I1).

De Kleer and Brown object that (I2) [this must be a typo for (I1)] "is not a component confluence, being a combination of a connection law (F6) [really, as we have seen, a definition] and component law (F1)." But our version makes clear that if we take the conduit as the component (which fits our intuitive notion of the physics of the situation), then (I1) is a

confluence describing that precise component.

It is claimed that "Although confluences (I1') – (I3') have the same solutions as confluences (F1) – (F6), they are not causally interchangeable. Each of the confluences (F1) through (F6) directly derive from the structure of the device through a modeling process. . . Iwasaki and Simon ignore modeling, and obscure the underlying physical structure." The previous paragraphs should provide an adequate reply to this incorrect claim. Obscurity of physical structure appears to be more in the eye of the beholder than in the physical system itself.

Since no actual conduit is an "ideal" conduit with no pressure drop, the decision as to whether a piece of pipe should be modeled as such or as a "constrict" with a pressure drop is up to the person building the model. The decision may be based on additional information, such as the actual physical shape of the pipe, what the pipe is meant to do, or even the whole purpose of building the model. While de Kleer and Brown protest that "we must never abstract away the components," this *is* abstraction. Although it violates the no-function-in-structure principle, such abstraction is essential in constructing any device model. No representation of the components of a system is possible without it.

Next, it is claimed that we follow a "wrong tack" in considering "impossible interpretations" of the disturbances. However, these interpretations are impossible only if we have already accepted the universal legitimacy of the confluence heuristic. (We will have more to say about this when discussing dynamic stability.) But de Kleer and Brown themselves call attention to the (desirable) ambiguity of their procedure, which permits all alternative interpretations to be explored. Why is this not legitimate here?

The difficulty we find in de Kleer and Brown's analysis of the conduit is that they appeal to an "intuitive physics" for which they supply no formal foundation. Without such a foundation, we are continually confronted with explanatory qualifications that seem wholly ad hoc. For example, in Footnote 3 of *Comments*, positive feedback is ruled out because "from an engineering perspective it is easy to see that the system cannot have positive feedback." Quite so, and we show it to be so with our dynamic analysis; but how does it follow from the

formal analysis of de Kleer and Brown? Where do we find the axiom that states that if all components of a system are linear and passive, it cannot have positive feedback? In the absence of a formalization of what "qualitative physics" is, we must plead not guilty to the charge of misusing it.

Comparative Statics

We have little to say about sections 5.1-5.3 of the *Comments*. They are mainly a restatement of our own brief remarks on dynamics. However, we do note that, again, the authors introduce ad hoc assumptions whenever the argument requires them, for example, in such statements as "It is well known that positive feedback (usually) leads to instability," or "Negative resistance is rare and counterintuitive," or "As negative resistance is so rare, the causal heuristics presume positive resistance unless evidence to the contrary." It is our belief that such assertions are not properly part of a general causal argument, but are specific physical assumptions that need to be introduced explicitly when dealing with the particular kinds of systems to which they validly apply. They should not be hidden in general heuristics.

With respect to section 5.4, De Kleer and Brown are quite correct in saying that, in our treatment of comparative statics, we sometimes make use of properties of equation systems, some of them derived from considerations of dynamic stability, that are stronger than the properties they use in qualitative physics. It should come as no surprise, then, that we also are sometimes able to deduce stronger results. However, we must point out that the properties we use fall far short of a complete specification of the equation system, nor is anything more than algebra required to draw the inferences. Specifically, we retain the literal coefficients of variables in the equations, and we sometimes use stability conditions to infer inequalities between functions of these coefficients. These inequalities are, in general, invariant under arbitrary monotonic transformations of the variables, hence involve quantitative assumptions only in a very weak sense of that term.

Ambiguity and Feedback

Since this section of *Comments* largely reviews issues already discussed, our remarks can again be brief. De Kleer and Brown state that their causal heuristics cannot be used to

identify exogenous variables. We discussed our view on this point earlier in the section on Propagation of Constraints.

With respect to minimal complete subsets of several variables (section 6.2), we know of no reason why systems of structural equations should not contain such subsets (as in their example of section 2). If they do, then it is meaningless to talk about the causal ordering, mythical or otherwise, among variables in such a subset. Hence the alternative orderings produced by the propagation technique have no operational or physical significance if the description of the system by its equations is correct.

The discussion in the *Comments* of no-function-in-structure appears to accept most of our strictures. We still confess our inability to understand why our light switch model violates the no-function-in-structure principle. In particular, we do not understand what could be meant by the statement that voltage is a "pathway" for current. Those of us who would like to understand what the no-function-in-structure principle means, and how to apply it, would be greatly assisted if de Kleer and Brown would provide us with a formal description of it, which they have nowhere done.

The *Comments* make much of the ambiguities that are generated by the propagation algorithm in qualitative physics. These ambiguities stem from: (1) the qualitative nature of the rules, (2) the local nature of the rules, and (3) inability to distinguish between causes and effects among events. (See de Kleer, 1979.)

The first two sources of ambiguity are present in both the mythical causality and causal ordering approaches. As we pointed out earlier, we make no more use of the quantitative properties of equations than does qualitative physics, nor do we depend on global properties in different ways than qualitative physics does. Given alternative assumptions about exogenous variables, our approach produces different causal orderings. And, given different assumptions about signs of derivatives, the method of comparative statics produces different evaluations of the effects of disturbances.

The third source of ambiguity arises computationally as non-determinism in the constraint propagation process that is resolvable in different ways by application of different

11

causal heuristics. While qualitative physics provides three heuristics to remove the ambiguity, our procedures recognize the ambiguity by assigning mutually interacting variables to a single subset having no internal causal ordering.

The *Comments* assert that, "perhaps the single key advantage of qualitative reasoning is that it is ambiguous." But the ambiguity, recognized quite as fully by our procedures as by theirs, derives not from the analysis but from the phenomena being analyzed. It is simply a consequence of the incompleteness of the information provided about the system being modeled, and the existence of many alternative assumptions that might be made to fill the gap.

Conclusion

We share with de Kleer and Brown the hope that our exchange will be a step towards clearing up misunderstandings about our several approaches to the analysis of causal relations and to a qualitative calculus for analyzing the behavior of physical and other systems described in terms of their structural relations. If we have not been able to reach full agreement, perhaps a public discussion of this kind will help other readers to find their way through the literature on this topic, and to share in the task of clarifying these important and difficult issues. In this way we may be able to combine the insights of the earlier literature on causal ordering and comparative statics with those that have been gained in the search for a qualitative physics.

We do hope, however, that our reply to the *Comments* of de Kleer and Brown makes clear that both groups are trying to explicate the *same* concept of causality and not distinct and separate concepts. It would be unfortunate if anyone were to conclude that our disagreements could be waved away as due simply to "the fundamental differences between the approaches of engineering and economics." Both engineering and economics are concerned with understanding complex systems in terms of their components and the mechanisms that connect those components with each other. We have found no reasons to call upon different methods of causal analysis for these two classes of systems. Indeed, as our examples of the evaporator and the conduit indicate, methods of analysis developed

initially with particular reference to economic systems are just as readily applied to the systems of the physical world, and detect the same causal orderings as are detected by the methods of mythical causality.

References

De Kleer, J. The Origin and Resolution of Ambiguities in Causal
    Arguments. *Proceedings Sixth International Joint Conference
on Artificial Intelligence*, Tokyo, Japan, August, 1979

De Kleer, J., & Brown, J. S. A qualitative physics based on
    confluences. *Artificial Intelligence*, 24:7-83 (1984)

-----. Theories of Causal Ordering. *Artificial Intelligence* (1986)
    this volume.

Iwasaki, Y., & Simon, H.A. Causality in Device Behavior. *Artificial
    Intelligence* (1986)

Simon, H.A. *Models of Discovery*, Dordrecht, Holland: D. Reidel, 1977