The Problem of Piecemeal Induction

Conor Mayo-Wilson[†‡]

I argue that, in causal inference from many observational studies, the piecemeal collection of data can cause underdetermination, even if arbitrarily large amounts of reliable data are available. Two theorems reveal that, for any variable set V, there are causal theories over V that can be distinguished if and only if all variables are simultaneously measured. These results entail that, a priori, one cannot know which observational studies will be most informative with respect to the true causal theory describing V. Hence, scientific institutions may need to play a larger role in coordinating differing research programs.

1. Introduction. The piecemeal construction of theories is an essential part of scientific practice. Data sets are pooled so that scientists have larger numbers of observations from which to construct theories; mathematical models are generalized so that a multitude of equations can be derived from a common set of axioms; and theories are unified so that multiple phenomena can be explained by a single theory. In this way, the work of hundreds of scientists, perhaps living on different continents during different eras, can potentially all contribute to the construction of current theories.[1] The history of astronomy is a paradigmatic case of such piecemeal construction: observations from differing continents were pooled over cen-

†To contact the author, please write to: Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213; e-mail: conormw@andrew.cmu.edu.

‡Thanks to David Danks and Clark Glymour, who provided useful comments, corrections, and suggestions on several drafts of this article. The article also benefited greatly from discussions with participants at three conferences, namely, those sponsored by the Philosophy of Science Association, the British Society for Philosophy of Science, and the graduate students at Princeton and Rutgers.

1. Whewell writes, "The fact that science is capable of being resolved into separate processes of verification, is that which renders it possible to form a great body of scientific truth, by adding together a vast number of truths, of which many men, at various times and by multiplied efforts, have satisfied themselves" (Whewell 1859, 82).

Philosophy of Science, 78 (December 2011) pp. 864–874. 0031-8248/2011/7805-0013$10.00
Copyright 2011 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/662564 Published online by Cambridge University Press

turies, and numerous mathematical theories of planetary motion (e.g., Kepler's laws) were generalized and unified by Newton's law of gravity.

Nowhere is the piecemeal construction of scientific theories more prominent, however, than in modern social science and medical research. In the social sciences and medicine, scientists must frequently synthesize the results of many observational studies and/or randomized controlled trials in order to study the causes of a given phenomenon. For instance, there are literally thousands of papers that document the (causal) relationships between heart disease and factors such as dieting, smoking, and prescription drug use. Although these thousands of studies share a commonly measured variable (namely, incidence of heart disease), there are many variables that are measured in some studies but not in others (e.g., patient's diet). It is essential to have reliable techniques for making inferences about the causal relationships among all these variables. Similar remarks apply in the social sciences. For example, economic models of the causes of poverty incorporate many variables, including race, gender, marital status, parental income, education level, juvenile record, and so on. No observational study could feasibly measure all these variables, and yet it is crucial that we understand the intricate causal relations among all of them.

The piecemeal construction of scientific theories is generally thought to be desirable for at least two reasons. First, if theories can be developed by amalgamating the results of many smaller studies, then researchers can specialize in the study of particular phenomena. This specialization arguably increases the quality of the data gathered, as researchers can focus on learning appropriate measurement techniques, analysis methods, and so on. Second, if scientific theories can be constructed in a piecemeal fashion, then various studies can be conducted over time, by different researchers, allowing knowledge to be accumulated gradually. This latter advantage is also a practical necessity given financial and technical constraints on the number of variables that can be measured in a particular study.

The piecemeal construction of causal theories has been relatively unexplored in the philosophy of science, despite its central importance for the actual practice of science. Danks (2005) provided a rule for knowing that there is no causal relationship between two variables (i.e., neither variable is a cause of the other nor do they have a common cause) even when the two variables are never simultaneously measured. Tillman, Danks, and Glymour (2008) generalized this rule into a procedure for integrating arbitrarily many data sets to determine the full set of causal theories (over all variables under investigation) that are consistent with the data. Their integration procedure exhibits long-run reliability: as data are collected without bound, the procedure discovers all the causal information that could possibly be detected given that only particular subsets of the var-

https://doi.org/10.1086/662564 Published online by Cambridge University Press

iables are simultaneously measured.[2] This research focuses on what could be learned; it has left open the question of how much (if any) information is lost if one cannot measure all variables simultaneously.

This article argues that the piecemeal construction of causal theories can drastically increase underdetermination of theories by evidence. I state and explain two theorems that together show that, for any collection of relevant variables, there will be distinct causal theories T_1 and T_2 that can be distinguished if and only if all the relevant variables are simultaneously measured. That is, for any group of related variables, measuring only subsets of the group (no matter how many) can fail to reveal the full causal structure. The first of these theorems has an additional implication: namely, that it is impossible to choose a priori a sequence of small (in terms of number of variables) observational studies that will be most informative; given two different sequences of differing observational studies, there will be, in general, causal theories distinguishable by one, and only one, of the sequences.

The next section briefly outlines the basic commitments of the causal learning framework discussed here. Specifically, I state and explain two assumptions, the Causal Markov Condition (CMC) and the Causal Faithfulness Condition (CFC), both of which have been unwaveringly defended and staunchly criticized.[3] I will not enter into this debate, because my goal is to show that, even assuming the two principles, there will still be causal theories that are distinguishable only if all relevant variables are simultaneously measured. If the piecemeal construction of theories can increase underdetermination of theories even under the most advantageous plausible conditions for causal inference from observational data, then adopting weaker assumptions will not eradicate the problem. In section 3, I illustrate the problem of piecemeal induction using a medical case study. I show that, given the subsets of variables that were simultaneously measured in the case study, there are several distinct causal

2. Their algorithm is thus asymptotically reliable in the same manner as other causal learning algorithms. Specifically, as the amount of data grows without limit, the procedures produce (with probability one) all and only those causal theories that are, in principle, indistinguishable from the true one (without active experimentation). The Tillman, Danks, and Glymour algorithm thus shares the common flaw of not providing short-run information, though it does eliminate the requirement that all variables be simultaneously measured.

3. For defenses of the CMC, see Hausman and Woodward (2004a, 2004b) and Steel (2005). For criticisms of the Markov condition, see Cartwright (2002, 2007). For criticisms of faithfulness, see Freedman and Humphreys (1999) and Cartwright (2007). The case study involving birth control and thrombosis, often cited as a counterexample to CFC, is discussed in Hesslow (1976) and Cartwright (1989). Both the CMC and CFC are defended and employed in Spirtes, Glymour, and Scheines (2000).

https://doi.org/10.1086/662564 Published online by Cambridge University Press

theories that are not distinguishable (even in principle) but would have been had different observational studies been conducted. I then state and discuss the implications of the three central theorems about underdetermination from piecemeal science.

2. Causal Inference from Observational Data. In philosophy, machine learning, and statistics, two principles have been employed extensively in drawing causal conclusions from observational data:

> **Causal Markov Condition (CMC)**. Any variable is conditionally independent of its noneffects given its direct causes.
>
> **Causal Faithfulness Condition (CFC)**. No two variables are conditionally independent unless so entailed by the CMC.

To understand what the CMC and CFC imply, it is best to consider some examples. Suppose that smoking increases tar buildup in one's lungs, and tar buildup causes lung cancer. Smoking thus causes lung cancer only indirectly. These relationships can be represented graphically as shown below. In general, given a set of variables V, one can define a causal theory to be a directed graph of the sort pictured, which visually represents exactly which variables are causes of the others. I will use "causal graph" and "causal theory" interchangeably.[4]

$$\text{Smoking} \rightarrow \text{Tar} \rightarrow \text{Lung Cancer}$$

Knowing whether an individual smokes is clearly helpful in predicting whether the individual will develop lung cancer. However, such knowledge is irrelevant if one already has accurate information concerning tar buildup: information about smoking habits matters only because it helps one predict the amount of tar in the lungs. To see this point more clearly, imagine there is a drug that completely stops tar buildup. If an individual took such a drug, then smoking would no longer be a cause of lung cancer in that individual (assuming smoking only indirectly causes lung cancer by way of tar buildup).

This informal reasoning can be done precisely using the CMC: the variable "lung cancer" is (informationally and probabilistically) independent of its noneffect "smoking," given its direct cause "tar." The CMC, therefore, captures the important intuition that indirect causes are correlated with their effects but only when one fails to account for the direct

4. Of course, one often desires to know much more than simply which variables cause which others. Causal theories, for example, should ideally also tell one how strong the causal connection is between two variables. I focus on causal graphs because they represent the minimal amount of information that one generally desires when constructing causal theories for a set of variables.

https://doi.org/10.1086/662564 Published online by Cambridge University Press

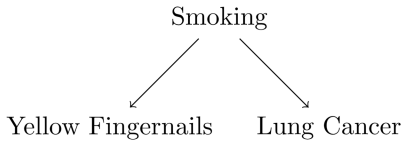


Figure 1.

causes through which the indirect ones exert influence. In contrast with our informal reasoning, the CMC (and actual instantiations of it) can all be expressed precisely using the language of (conditional) probabilities. In this article, that technical specification will not be relevant; what matters for our present purposes is simply the fact that everything discussed in section 3 can be done precisely in the language of (conditional) probabilities.

The CMC also captures important intuitions about common causes. Expanding the above example, suppose smoking also directly causes stained fingernails. Then knowing whether an individual has yellow fingernails would provide no additional evidence of his or her chances of lung cancer, if one knows whether the individual smokes. The variable "yellow fingernails" is (informationally and probabilistically) independent of its noneffect "lung cancer" given its direct cause "smoker" (fig. 1). The CMC uses causal structure to place constraints on the informational/probabilistic relationships between variables. The CFC essentially does the converse: informally, it says that variables that are (conditionally) independent are not (directly) causally connected. The CFC thus encodes and generalizes a completely standard principle of scientific inference: if variations in one factor are uncorrelated with variations in another, then (barring rich, domain-specific knowledge) the factors are not directly causally connected.

Given CMC and CFC, one can define two causal theories/graphs to be indistinguishable if they imply the same conditional independencies for a set of variables. Intuitively, two causal theories are indistinguishable if no amount of observational data could allow one to conclude which is correct, unless one employed domain-specific knowledge beyond the probabilistic relations among the variables. Returning to the fictitious example above, what causal theories are indistinguishable from the theory that smoking causes tar buildup, which in turn causes lung cancer? It turns out there are two more such theories, namely, ones that assert (1) lung cancer causes tar buildup, which in turn causes smoking, or (2) tar buildup is a common cause of both smoking and lung cancer. The three theories are depicted below. Of course, in this example, background knowledge is

https://doi.org/10.1086/662564 Published online by Cambridge University Press

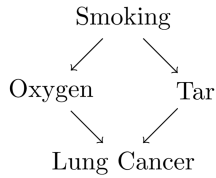


Figure 2.

sufficient to show that these rival theories are implausible, but, in general, such domain-specific knowledge may not be available.

A simple graphical criterion determines whether two causal theories are indistinguishable, but it requires two additional definitions. Say that two variables, X and Y, are adjacent if one is the cause of the other. Second, suppose X and Y both cause Z, but neither X nor Y causes the other. Then Z is called an unshielded collider with respect to X and Y. Pearl and Verma's theorem asserts the following:[5]

> **Theorem 1 (Pearl and Verma)**. Assuming CMC and CFC, two causal graphs are indistinguishable if and only if they have the same adjacencies and unshielded colliders.

Pearl and Verma's result is stunning. Although two indistinguishable causal theories might differ with respect to whether X is a cause of Y, or vice versa, they will agree that there is some direct causal link between the two. Conversely, differences in conditional probabilities imply differences in adjacency and unshielded collider structure. This theorem thus provides a ground (assuming CMC and CFC) for inferring significant, though often not complete, causal knowledge from observational data alone. Crucially, however, Pearl and Verma's theorem assumes that all variables can be simultaneously measured. For example, two theories might be distinguishable because one implies that X is independent of Y, given $Z_1, Z_2, \ldots, Z_n$, while the other does not. Yet, determining whether such a conditional independence holds requires simultaneously measuring $X, Y, Z_1, Z_2, \ldots, Z_n$. What happens if we cannot perform such a simultaneous measurement?

3. Piecemeal Causal Inference. Suppose that smoking is an indirect cause of lung cancer in two ways: by increasing the buildup of tar in one's lungs and also by decreasing the amount of oxygen supplied to one's lungs (fig. 2). In this case, even if we know about the tar buildup, an individual's smoking habits will not be independent of lung cancer because knowledge

5. Pearl and Verma (1991).

https://doi.org/10.1086/662564 Published online by Cambridge University Press

of smoking habits provides evidence of oxygen deprivation, which is assumed also to be a cause of lung cancer.[6] However, if one knew both how much tar occupied an individual's lungs and the supply of oxygen to the lungs, then smoking once again would be irrelevant to predicting whether an individual will develop lung cancer (given CMC and assuming that this is the true causal structure). More precisely, smoking is conditionally independent of lung cancer given both tar buildup and oxygen levels.

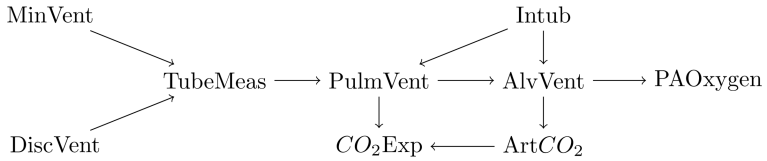
Consider what can be learned about the relationship between smoking and lung cancer in this fictitious example, if one could only measure any proper subset of the four variables, rather than all simultaneously. As shown above, smoking habits are not conditionally independent of lung cancer, given only tar buildup, and, by symmetric reasoning, the two variables are dependent, given only oxygen deprivation as well. And no other (three-variable) proper subset contains both of the relevant variables. Intuitively, it appears that one cannot rule out a direct causal link between smoking and lung cancer unless one can measure all four variables simultaneously.

This intuition is correct. Let T_1 be the true causal theory described above and T_2 be a theory just like T_1, except that T_2 asserts that smoking is also a direct cause of lung cancer. The CMC and CFC entail that T_1 and T_2 have exactly the same probabilistic relations involving three variables or fewer, and so T_1 and T_2 are indistinguishable if one cannot measure all four variables simultaneously. Of course, because T_1 and T_2 do not postulate the same direct causal links, they are distinguishable if all variables are simultaneously measured (by Pearl and Verma's theorem).

Although the above example is fictitious, the underlying problem that it illustrates is not. Danks (2005) considers the task of learning the causal relations among ventilators and blood oxygen saturation in an intensive care unit setting, assuming one can only measure certain subsets of those variables. The true causal structure is depicted in figure 3, and the three subsets of variables measured simultaneously are listed. Danks (2005) demonstrates that one can learn that neither MinVent nor Intub is the cause of the other, nor do they have a common cause, even if they are never simultaneously measured. This demonstration suggests that perhaps, with the proper care and insight, the results of many observational studies might be fused to yield all possible knowledge about all the variables. In fact, Danks (2005) used this case (and others) for optimism about the prospects for piecemeal causal inference. Such optimism must

6. If oxygen deprivation is more strongly tied to lung cancer than tar buildup, then smoking habits might be a better predictor of lung cancer than tar buildup, even after controlling for tar buildup. Indirect causes can be better predictors than direct ones when there are multiple pathways from the indirect cause to the effect.

https://doi.org/10.1086/662564 Published online by Cambridge University Press



Study 1	Study 2	Study 3
MinVent	PulmVent	TubeMeas
DiscVent	AlvVent	Intub
TubeMeas	ArtCO_2	AlvVent
PulmVent	CO_2	PAOxygen

Figure 3.

be tempered. Although one can rule out the existence of a causal link between MinVent and Intub, any amount of data (over these subsets) will underdetermine whether there is a direct causal link from MinVent to AlvVent or one from DiscVent to AlvVent (or both).[7] There are at least four distinct causal theories that are compatible with evidence and that would be distinguishable were all variables measured simultaneously (by Pearl and Verma's theorem). How common such underdetermination is in practice is an open question. But the theorems below prove that this underdetermination (or more precisely, its possibility) is not an artifact of the examples I have chosen but, rather, an intrinsic feature of piecemeal causal inference.

Let V be any collection of variables. For instance, V might be {Smoking, Lung Cancer, Tar}, or the collection of ventilator measurements, blood oxygen levels, and so on, in the medical case study. The following theorem states that the causal links are always underdetermined if one measures only proper subsets of V:

> **Theorem 2**. There exist distinct causal theories T_1 and T_2 (over V) with different direct causal links that are distinguishable (given sufficient data) if and only if every variable in V is simultaneously measured.[8]

This theorem implies that the piecemeal construction of causal theories from multiple observational studies always faces the threat of underdetermination even about adjacency information, regardless of the quantity and quality of data. Such failures, of course, can be of critical im-

7. The claim can be proven by exhaustive enumeration of the (conditional) independencies in the different theories.

8. Please contact the author for proofs.

https://doi.org/10.1086/662564 Published online by Cambridge University Press

portance, as without correct causal models government officials might institute policies with unforeseen consequences and doctors might prescribe medicines or surgeries without sufficient knowledge of how such treatments might affect patients' health. Pearl and Verma's theorem also implies that the orientations of some direct causal links can sometimes be learned, depending on whether there are unshielded colliders. The next theorem shows that this conclusion likewise depends on measuring all variables simultaneously.

> **Theorem 3**. Suppose that (*a*) one never simultaneously measures all variables in V, and (*b*) there is at least one $v \in V$ such that the full set of variables $V \backslash \{v\}$ is never simultaneously measured. Then there exist causal theories T_1 and T_2, such that T_1 and T_2 postulate the same adjacencies and are indistinguishable with respect to the observational studies in question, but T_1 contains an unshielded collider that T_2 does not.

Together, the two theorems show that the piecemeal construction of causal theories can drastically increase underdetermination. Important features of causal theories, which can be learned if one can conduct large observational studies in which all variables under investigation are simultaneously measured, become inaccessible when one attempts to patch together a series of smaller studies. It is equally important to recognize what the theorems do not say: they do not imply that one can never determine the unique causal structure (over all variables) using piecemeal causal inference. For some causal structures and particular collections of subsets, it is possible to recover identical information through piecemeal methods or through simultaneous measurement of all variables. This observation about the possibility of success, combined with the negative theorems, raises a crucial question: is it possible to plan a priori a series of observational studies on subsets that minimize the underdetermination of theories? More practically, could scientific institutions coordinate different research programs so as to minimize the amount of information that is lost by piecemeal inquiry?

Unfortunately, the answer is no. A priori, differing series of observational studies will generally be of incomparable value, in the sense that one series will distinguish between theories that the other does not, and vice versa. To make this precise, let U and U' be distinct collections of subsets of some variables V under investigation. For example, in the ICU case, U might contain the sets {MinVent, DiscVent, TubeMeas}, {TubeMeas, Intub, PAOxygen, AlvVent}, and {PulmVent, AlvVent, CO_2Exp, ArtCO_2}. Each set, U or U', thus represents a series of observational studies, rather than just a single study.

Say U is as strong as U' if any two causal theories that can be distin-

https://doi.org/10.1086/662564 Published online by Cambridge University Press

guished by studies specified by U' can be distinguished by conducting the studies specified by U. That is, anything that we can learn through U' can also be learned through U. Say they are equivalent just in case each is as strong as the other. Then theorem 2 implies the following:

> **Theorem 4**. U is as strong as U' if and only if every member of U' is contained in a member of U. In particular, if no member U_1 of U is a proper subset of any other member $U_2 \in U$ (and similarly for U'), then U and U' are equivalent if and only if $U = U'$.

The theorem has striking consequences. Suppose the National Institutes of Health (NIH) is considering funding a series of observational studies to determine the causes of schizophrenia. Because of the sheer number of variables that might be relevant to schizophrenia, such studies cannot measure all variables simultaneously, and so the NIH must pick and choose which observational studies to fund. The above theorem implies that, without further information, the only rule that would justify the NIH's preference for funding one study rather than another is "bigger is better." In other words, if one observational study will measure all the variables that another does, then there is reason to prefer the former to the latter. Otherwise, there is no a priori reason to prefer one series of observational studies rather than another; each will distinguish between only some series-specific sets of causal theories.

However, the qualification "without further information" is of critical importance. Given prior experiments and observational studies, there might, in fact, be principled ways to distinguish between different observational studies in terms of the informational value that they might possess, as suggested by Danks (2005). The present results imply, however, that the coordination of differing research programs must take place during inquiry; we cannot parcel out responsibilities for particular variable sets or phenomena in advance of knowing something about the structure of the world. Planning ahead is futile for this particular problem.

4. Conclusion. The piecemeal construction of theories is a central and unavoidable feature of scientific practice. Some recent work suggests that piecemeal causal inference can nonetheless lead to large-scale causal theories over many variables (Danks 2005; Tillman et al. 2008). I have argued that this optimism must be significantly tempered: in the construction of causal theories, piecemeal inquiry can drastically increase the number of distinct theories that are indistinguishable with respect to all evidence. Moreover, because a priori it is generally impossible to decide which types of observational studies will be most informative, scientific institutions might need to play a larger part in coordinating distinct research programs at successive stages of inquiry.

https://doi.org/10.1086/662564 Published online by Cambridge University Press

The arguments above leave two important questions unanswered. First, exactly what types of features of causal theories can be learned from different sets of observational studies? That is, even if information is lost by piecemeal inquiry, can one characterize how much one still might learn from measuring only proper subsets of a collection of variables? Second, is the piecemeal problem of induction a more general phenomenon? That is, in what types of inquiry does such piecemeal construction of theories increase underdetermination, and what is the extent of that underdetermination? This article, I hope, will spur future research into these important questions for philosophy of science and for scientific practice more generally.

REFERENCES

Cartwright, Nancy. 1989. *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

———. 2002. "Against Modularity, the Causal Markov Condition, and Any Link between the Two: Comments on Hausman and Woodward." *British Journal for the Philosophy of Science* 53:411–53.

———. 2007. *Hunting Causes and Using Them*. Cambridge: Cambridge University Press.

Danks, David. 2005. "Scientific Coherence and the Fusion of Experimental Results." *British Journal for the Philosophy of Science* 56:791–807.

Freedman, David, and Paul Humphreys. 1999. "Are There Algorithms That Discover Causal Structure?" *Synthese* 121:29–54.

Hausman, Daniel, and James Woodward. 2004a. "Manipulation and the Causal Markov Condition." *Philosophy of Science* 71 (Proceedings): 846–56.

———. 2004b. "Modularity and the Causal Markov Condition: A Restatement." *British Journal Philosophy of Science* 55:147–61.

Hesslow, G. 1976. "Discussion: Two Notes on the Probabilistic Approach to Causality." *Philosophy of Science* 43:290–92.

Pearl, Judea, and Thomas Verma. 1991. "A Theory of Inferred Causation." In *Principles of Knowledge Representation and Reasoning: Proceeding of the Second International Conference*, ed. J. A. Allen, R. Fikes, and E. Sandewall, 441–52. San Mateo, CA: Morgan Kaufmann.

Spirtes, Peter, Clark Glymour, and Richard Scheines. 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press.

Steel, Daniel. 2005. "Indeterminism and the Causal Markov Condition." *British Journal for the Philosophy of Science* 56:3–26.

Tillman, Robert E., David Danks, and Clark Glymour. 2008. "Integrating Locally Learned Causal Structures with Overlapping Variables." In *Advances in Neural Information Processing Systems 21*, ed. D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, 1665–72. Vancouver: NIPS.

Whewell, William. 1859. *History of the Inductive Sciences*. Vols. 1–2, 3rd ed. New York: Appleton.

https://doi.org/10.1086/662564 Published online by Cambridge University Press