1

An Introduction to Time and Causality

Samantha Kleinberg

1.1 Why Time and Causality?

Time is fundamental to how we perceive and reason about causes. It lets us immediately rule out the sound of a car crash as its cause, and learn through experience that alcohol may lead to hangovers. Our first instinct when asking why something happened is to look backward, aiming to identify an earlier event that may have been responsible. If you develop food poisoning and want to know how you became ill, you will likely start by enumerating all the things you recently ate. It probably seems obvious to you that you should only consider foods you consumed before you got sick and ignore those you ate afterward. At the same time, you would not bother trying to remember what foods you ate a week ago, last month, or last year. A cause needs to be earlier – but not too early. There is a range of times where we believe a food can plausibly lead to illness. If we want to explain a particular stock market crash, we again look backward to what happened leading up to the crash. The important events must still be earlier, but now we might consider a different window of time than for food poisoning, based on how we believe factors may influence the market and whether we believe this process is faster or slower than that for food-based illness.

Our reliance on time is not limited to explanation of specific events (token causality), but is central to how we learn about general relationships, such as what causes heart disease (type-level causality). To find causes of disease, we often examine what is different about people who become ill and those who do not. Once again, if people with heart disease tended to take up smoking after diagnosis, we would not consider this a cause of the illness, as we believe that only factors that happened before the onset of the disease could have caused it.

The order of events is one of the key features we use to separate potential causes from the large set of things we measure or observe. We are able to split items into those that are potentially relevant and those that are not, solely based on when they happen. Of course, not every earlier event is responsible for every later event, but this belief about the importance of time provides a key heuristic for locating potential causes.

This intertwined relationship between time and causality may seem so obvious that it does not warrant examination, but it is not a given that time is such an important clue toward causality. There are many other factors we could use to weed out things that are not causes. We could consider, for example, which events have the capacity to cause others, which are nearby in space, or which are often responsible for a particular effect. For instance, when one billiard ball hits another we may suggest that the moving ball has momentum that it transfers to the second one or that the ball closest to the second is the one that caused its motion. While this is true, we immediately know, that even without any understanding of physics, the ball that was moving first caused the second ball to move. It is obvious to us what happened based solely on the order of events. We could similarly determine which foods are likely to be causes of food poisoning by identifying those that are most frequently responsible for food poisoning. This requires much more background knowledge and thought, though, than an initial filtering by time does.

Even though time is one of the key clues for humans trying to identify causal relationships, and even though it has been part of definitions of causality dating back centuries, it can also be misleading. It is such a strong clue to causality that we may believe one event caused another just because it happened earlier. This is known as the *post hoc ergo propter hoc* (after therefore because) logical fallacy. When we push a button at a crosswalk and the light immediately changes, we may think the timing is too much of a coincidence for the button to not be the cause of the light changing (Miele, 2016) or we may receive a flu shot and believe that a case of the flu shortly after was caused by the shot. The downside of time having such a strong influence on our understanding of causality is that this information may drown out other conflicting signals, such as knowledge that the flu shot contains an inactive form of the virus.

Psychologists have done experiments designed to create such a conflict between our temporal perception and our knowledge of how things must work. In cases where participants had full knowledge of a system and could predict how long it would take to operate, even though the real cause was slow, children mainly relied on temporal perception and claimed that the event happening immediately before the effect was its cause. Adults, on the other hand, were able to reason that such an event did not have time to produce the effect (Schlottmann, 1999). Yet in cases where there is not a fully observed mechanism, adults too find it less plausible that a cause is responsible for an effect as the time delay between them increases (Shanks, Pearson, and Dickinson, 1989). These errors may persist even when participants are told there may be delays and that the order in which events are observed may not be the same as the order in which they happen (Lagnado and Sloman, 2006).

In addition to erroneous perceptions of causality by humans, time may also lead to incorrect causal inferences by algorithms. Since the mid-1900s US gross domestic product (GDP) has grown exponentially. Many other things have grown at similar rates, such as smartphone usage and the number of Starbucks stores. Simply applying standard algorithms used to identify causal relationships to this raw data would likely result in finding any pair of these variables to be causally related. There is no relationship, though, except that these time series are all nonstationary. This means that key factors such as their mean or variance are changing over time. Because this trend overwhelms the smaller day-to-day or year-to-year changes, we can find spurious connections such as smartphone usage causing Starbucks to open more stores.

Nonstationarity is one of the fundamental challenges for inference from time series. In addition to finding spurious causal relationships, it can also make it more difficult to find true ones. For instance, patients change over time in the hospital as they recover, and interventions to treat illness may lead to changes in causal relationships such as how a medication affects physiology. Similarly, macroeconomic variables are influenced by policy changes, which may change behavior, and the locations being studied may change as country boundaries are not static. Even policies affecting human behavior, such as laws requiring individuals to wear seatbelts, may lead to fundamental differences in how people drive or what risks they take. In contrast to the case of nonstationarity leading to spurious inferences, here the causal relationships governing how these systems behave actually are changing over time. This leads to a moving target for inference, and many methods can only find the relationships that are stable across time, leading to missed inferences. Yet, finding specifically which relationships are changing – and how – is often more important than identifying those that are constant.

Even if a time series is stationary, there are other challenges for causal inference (finding causal relationships from data). A major problem is when the timing associated with data is at the wrong level of granularity, making events seem simultaneous when they are not. If causal influence is happening across time, and this order of events is important for finding causes, then inference becomes even more challenging when we cannot determine which event happened first. For example, if we measure GDP and inflation once a year, then if influence between these variables happens faster than yearly, we will have no idea which change came first and will only be able to find a correlation between the two time series.

In more challenging cases, different factors may be measured with different delays, leading to the events appearing out of order. Surprisingly, this situation is not uncommon. Continuous glucose monitors provide regular measurements of glucose for use in managing diabetes. These devices record glucose values every 5 minutes, but they measure glucose in fluid between the cells rather than in the blood stream. As a result, the measurements are delayed relative to blood glucose, on the order of 5–15 minutes (Bequette, 2010). Further the signals are not just shifted in time but rather the signal may be distorted in a way that depends on the current glucose value and its trajectory (Cobelli et al., 2016). Thus, it is possible that we may observe what seems like a meal and injection of insulin happening before a rise in glucose, solely because one variable (glucose) is measured with a delay that is not present for the others. This could lead to counterintuitive findings like insulin causing an increase in glucose if the delay is not taken into account or if too fine a timescale is used. Data that is changing, unreliable, or where the order of events cannot be trusted poses many challenges for learning causal models.

Despite these issues, a cause happening before its effect has been a core, and often unquestioned, part of how we describe causality, both in everyday life and in scientific research. However, understanding the relationship between time and causality is now an increasingly significant focus of research across many disciplines. This shift is driven in part by the availability of new datasets and measurement methods. Large observational datasets such as from medical records and social media enable us to ask new types of questions on causality. They also bring new challenges in identifying causal relationships, as we face missing variables, noise, and causal relationships that change over time. These challenges are common to research in many fields, and their solutions may in turn provide insight into core philosophical problems. For example, we can now examine risk factors for disease using electronic medical records going back decades, from much larger populations than were previously possible. Yet the meaning of variables changes over time (what was considered light smoking in the 1970s may not mean the same thing today), as do populations and their behaviors. Similarly, more sensitive measurement techniques may now let us determine whether events thought to be simultaneous actually have a time order. These in turn may shed light on questions such as whether a cause coming before an effect is necessary or just the most likely order of events. Given these many opportunities and challenges, it is an ideal time to examine the relationship between time and causality in depth.

1.2 Interdisciplinary Perspectives on Time and Causality

This book brings together research on all facets of how time and causality relate across the sciences. Through these interdisciplinary perspectives, we explore the role of time in philosophical theories of causality, how time affects our perception and judgment of causality, and new methods for inferring causes from temporal data. This section provides historical context for the three main segments of the book (philosophy, sciences, and models and algorithms) and brief overviews of the chapters, with a focus on discussing how they illuminate different aspects of the same core problems.

1.2.1 Philosophy

In the eighteenth century, David Hume (1748, Section VII, Part II; emphasis original) defined a cause as "*an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second*. Or, in other words, *where, if the first object had not been, the second never had existed.*" More than two hundred years later, these definitions are the foundation of two main philosophical approaches to defining causality, as well as the computational methods for finding it that will be discussed later in this book. The first sentence says causality is about regularities, where the effect routinely follows the cause. The second elaboration provides the basis for the counterfactual approach to causality, where we examine what would have happened had the cause not occurred (see Lewis, 1973a). That is, rather than a repeated sequence of events, here the effect must make a difference to whether or not the cause happens. While there are many differences between the two definitions, a key commonality is that both assume that there is a first event (the cause) and a second (the effect). This requirement that a cause be earlier than its effect is known as temporal priority. Hume also stipulated that causally related events must be contiguous (nearby in space and time) and that the cause is required for the effect to occur (necessary connection). In this book we focus primarily on the temporal aspect of this definition.

In practice, effects may not always follow their causes. While smoking can cause lung cancer, not everyone who smokes develops cancer, and some non-smokers also develop cancer. One source of uncertainty may come from our observations (perhaps if we knew enough about what caused cancer, we could specify the cause such that every time it was present cancer resulted), while the other is that some causes are at their core non-deterministic. That is, in some cases even with complete knowledge about the scenario, we cannot be sure whether the cause will produce the effect. As a result, research stemming from the regularity theory has aimed to specify these conditions more fully to distinguish causal regularities from correlations (Mackie, 1974/1980), or incorporate probability into the definitions of causality.

In the 1950s, Hans Reichenbach (1956) attempted to define causality in terms of probability and, uniquely, to learn the direction of time from the direction of causality. That is, in his theory a common cause that fully explains its effects must be earlier than the effects, and by finding this explanatory event we can determine which way time flows – rather than using the direction of time to determine which events may be potential common causes. Reichenbach's theory is unique in this respect, using the asymmetry of probabilistic relationships between cause and effect to identify causes without knowing which is earlier. However, without the use of temporal information, one must make other, stronger, assumptions. Thus, Reichenbach proposed definitions for a common cause of two effects based on their probabilities, with the stipulation that the effects are simultaneous. If there is no earlier common cause, this ensures we will not find one event as a cause of the other. For example, if two lamps go out at the same time, we will not erroneously find one lamp's state causing the other's. Yet for time series data that is nonstationary we may find two series with similar trends where neither is a cause of the other and yet there is no common cause (violating Reichenbach's common cause principle).

While Reichenbach's theory influenced many later probabilistic theories, work on probabilistic causality returned to the use of temporal priority as a distinguishing feature of causality. Patrick Suppes' influential work on distinguishing mere events that raise the probability of their effects from genuine causes assumed we know the ordering of events and then defined ways in which earlier events may make later ones spurious (Suppes, 1970). Similarly, while Ellery Eells (1991) developed measures for probabilistic causal relevance that consider factors happening temporally in between the cause and effect, he again only considers causes that are strictly earlier than their effects. Note, though, that while Hume's initial goal was to uncover both the metaphysics (what causes are) and epistemology (how we can learn about causes), these probabilistic theories are primarily focused on epistemology. This is an important distinction when examining the relationship of time and causality, since it is possible that temporal priority is an effective clue toward identifying causes, even if it is unnecessary for something to actually be a cause.

Thus, while time is a core part of how people think about causes, is it essential for a definition of causality? According to Hume it is a part of causality itself, but it is also possible that while temporal order is a useful heuristic for people to identify causes, it may not be a fundamental part of what it means to be a cause. As discussed later in this book, physics has provided seeming counterexamples to this stipulation with cases where the future may potentially affect the past. For example, Price (2012b) argues that some quantum theories must allow retrocausality. Reichenbach argued that we can instead identify the order of time from the direction of causality. While less extreme than causality flowing backward in time, there has been compelling new evidence of simultaneous causation in psychology and neuroscience (Vernon, 2015).

The first segment of the book addresses the philosophy of time and causality, exploring what this new evidence means for both the metaphysics and epistemology of causes – what they are and how we can find them. One of the fundamental questions about the relationship between causality and time is whether causes must precede their effects. Even without backwards causality (causes affecting events that happen earlier than they do), we may ask whether all causes are indeed strictly earlier than their effects and whether this temporal precedence is essential to a theory of causality. In Chapter 2, Thomas Lodewyck and Bert Leuridan begin with a detailed examination of the role of time in all major theories of causality, providing a road map for all later chapters. They argue for an account of causation that is independent of time. Notably, the core of this argument is that existing philosophical theories that aim to identify causal direction from temporal direction or vice versa cannot adequately handle scientific phenomena. Phil Dowe, in Chapter 3, focuses on the relationship between time and metaphysics and the puzzle of how we can answer what it means for one event to cause another in so many different ways. Dowe shows that many answers are lacking and discusses what is needed to fully answer this question. While many arguments exist for why temporal (and even spatial) ordering is necessary for understanding causality, what about the reverse: must we have an understanding of causality to have a concept of space and time? That is, can there be worlds that proceed temporally but without causation? In Chapter 4, Victor Gijsbers argues that causality and time are so interlinked that they cannot be analyzed separately, suggesting that a new approach to both causality and time is required. In Chapter 5, Jenann Ismael discusses how recent advances in physics, computer science, and decision theory have provided new insights into these old philosophical problems, and prompted new questions. In particular, she ties together the philosophical study of causality, evidence from physics, and computational methods for inferring causes. A common theme throughout the volume is the use of methods and findings from other fields to advance understanding across disciplines.

1.2.2 Sciences

Hume and others have provided the foundations for how we understand the metaphysics and epistemology of causality. Many scientific disciplines, though, are primarily concerned with the methodology used to find causes or support causal claims. That is, regardless of what a cause actually is, what are the tools we can use to identify such relationships? Even though it is not always explicit, causality is at the core of most scientific pursuits, which aim to learn how things work. For example, biologists have asked questions such as what happens during the cell cycle, what functions different genes have, how organisms have evolved, and so on. The answers to these questions may involve uncovering specific labels for processes (e.g., mitosis is one phase of the cell cycle, and is when the cell divides) or links between events (e.g., a particular gene is necessary for a specific eye color), rather than explicitly causal language. Yet the goal is still uncovering why and how things happen, which are both causal questions (Mayr, 1961).

All of these processes unfold over time, though this is often implicit in the explanations given in biology. While there may be regularities at work, the way causes are described in the biological sciences bears more similarities to mechanistic theories of causality. Rather than determining whether or not there exists a causal link between events, these approaches instead focus on identifying the processes by which the cause brings about the effect (Glennan, 1996a; Machamer et al., 2000). A mechanism can be thought of as a set of parts that interact to regularly produce a change. Time is implicit here, as the process must have some duration at the end of which the effect is present.

In biomedical sciences, such as epidemiology and public health, the distinction between causality and correlation not only leads to deeper understanding of phenomena, but is fundamental for making policies and choosing medical interventions. A chemical may be found in cigarettes (and thus associated with cancer) but may not be a carcinogen on its own, and so different determinations about its safety may be made. Even though causality is critical for decision-making here, it is much more challenging to establish than in biology since it is rarely possible to conduct controlled experiments, many factors interact to affect health, these relationships are rarely deterministic, and often they unfold over long time periods. As a result, there has been much effort toward clarifying how causal claims can be made. Austin Bradford Hill (1965) proposed a set of factors to consider when evaluating causal claims, though these have been frequently misrepresented as criteria for causality. One of the factors proposed is temporality (does the cause happen before the effect?), yet Hill notes that in epidemiology this feature may not be as straightforward as it seems. For example, individuals with chronic heart failure have reduced exercise tolerance. Yet, given that heart failure takes time to develop (and we rarely see it in its early stages), we may observe the impact on exercise before heart failure is actually apparent and diagnosable – leading to a mistaken temporal impression and the idea that reduced exercise is causing heart failure.

The second part of this volume delves into the relationship between time and causality in specific disciplines, including psychology, biology, and physics. Work in psychology has examined many facets of our understanding of causality, including when this understanding develops in children, whether we are able to perceive causality directly, and how causality and moral considerations interact in assigning blame. There has been a large amount of work on how we learn about causes, and in particular how being able to experiment on a system (intervene) affects causal learning, and how temporal cues are used. In Chapter 6, Neil Bramley and others propose a unification, showing that it is difficult to separate time and causality both conceptually and experimentally and discussing how both affect our ability to learn in complex situations. This research into what affects human judgments of causality may provide insight that can guide algorithms that discover causal relationships from data. In biology, time is a key factor in understanding the mechanisms by which processes occur, yet time is often implicit. Yin Chung Au, in Chapter 7, argues that even when it is not explicitly mentioned, time plays a critical role in the search for biological mechanisms. Using the specific example of apoptosis (programmed cell death) studies from the 1970s to the present, the chapter examines both how biological mechanisms are identified using temporal information and the role of time in diagrams used for communicating and explaining these mechanisms. Much like biology, work in physics illustrates how pervasive the assumption of a temporal order is. Inge de Bal and Erik Weber, in Chapter 8, provide new insights into how time can be used to identify causes in ways that may translate across disciplines. They investigate the relationship between mechanistic and correlational evidence in reasoning about physical causal claims. In particular, they show that mechanistic evidence can explain how we go from symmetric physical laws to asymmetric causal claims. While de Bal and Weber build on analogies from the social sciences, Jonathan Livengood and Karen R. Zwier show, in Chapter 9, that time is often not explicitly modeled in the social sciences. This chapter shows how this can lead to models where causes do not precede their effects and proposes criteria to both preserve the intended ordering and remain consistent with the data analyzed. Using examples from sociology, political science, and education research, they show that these criteria are non-trivial and that researchers must carefully consider the commitments of the models they develop even – or especially – when time is not explicit.

1.2.3 Models and Algorithms

While uncovering causality has been a core goal of many sciences and a topic of inquiry for centuries, it was only in the 1980s that the computational methods for modeling causal relationships and inferring them from data emerged. A major advance was the work of Judea Pearl (2000) and Peter Spirtes, Clark Glymour, and Richard Scheines (Spirtes et al., 2000) who showed that probabilistic causal relationships could be represented using graphs called Bayesian networks. They defined what makes a graphical model a causal one, and what we need to assume to draw causal conclusions. Further, they created algorithms for learning these models from data. While these methods are highly general, focused on identifying relationships between variables, (e.g., smoking causes lung cancer) they do not define the timing of these relationships (e.g., does smoking cause lung cancer in 10 years or 20 years?) or what the meaning of a variable is (e.g., what does it mean for smoking to be true – smoking for a month, a year, 10 years? What quantity must one smoke?). While the methods described could use temporal data to orient edges in the graph from causes to effects, they had no way of learning the timing for each relationship, and further had to make many assumptions that are not widely true in time series data, such as that relationships are stable over time. More recent work has developed dynamic Bayesian networks, which extend these methods to include a notion of time, using a sequence of Bayesian networks connected across time to show how factors at one time influence those at a later time (Murphy, 2002).

Nevertheless, there has been significantly less attention paid to the much more challenging temporal case. The challenges are both conceptual and practical, including how to identify causes at different levels of temporal granularity and developing efficient methods given the increased complexity of inference in the temporal case. Despite limited computational work, understanding causal relationships over time is of significant practical importance. For example, to develop effective economic policies we must be able to compare both the cost of various strategies as well as how long they will take to achieve their goals. Further, many domains produce time series data. As a result, one of the earliest formulations of how to identify potential causal relationships in time series came from the economist Clive Granger (1980). Granger causality provides a way of identifying relationships between