Elements of a rational framework for continuous-time causal induction

Michael Pacer (mpacer@berkeley.edu)
Thomas L. Griffiths (tom_griffiths@berkeley.edu)
Department of Psychology, University of California, Berkeley, Berkeley, CA 94720

Abstract

Temporal information plays a major role in human causal inference. We present a rational framework for causal induction from events that take place in continuous time. We define a set of desiderata for such a framework and outline a strategy for satisfying these desiderata using continuous-time stochastic processes. We develop two specific models within this framework, illustrating how it can be used to capture both generative and preventative causal relationships as well as delays between cause and effect. We evaluate one model through a new behavioral experiment, and the other through a comparison to existing data.

Introduction

Causal induction plays a key role in human cognition, allowing people to identify the causal relationships that structure their environment. Recent work in cognitive science has resulted in many successful models of how people infer causal relationships from contingency data (Anderson, 1990; Anderson & Sheu, 1995; Cheng, 1997; Griffiths & Tenenbaum, 2005) and events that unfold in discrete time (Wasserman, 1990; Greville & Buehner, 2007). However, relatively few models have explored events that occur in continuous time. And yet, people regularly and easily reason about causal phenomena that evolve in continuous time (Michotte, 1963; Griffiths & Tenenbaum, 2009). Our understanding of causal inference would thus benefit from a framework capable of explaining human continuous-time causal inferences.

In this paper, we address this challenge by undertaking a rational analysis of continuous-time causal induction, in the spirit of Anderson (1990) and Marr (1982). We formalize the abstract problem posed by continuous-time causal inference, identifying a set of desiderata that a solution to this problem needs to incorporate. We then outline a framework that satisfies these desiderata, based on rational statistical inference over continuous-time stochastic processes. Our framework makes it possible to define both generative and preventative causes that unfold in continuous time, and to take into account delays between causes and effects.

With this framework in hand, we present two case studies from experimental psychology on human continuous-time causal inference. The first case study involves a novel experiment based on an experiment conducted by Griffiths and Tenenbaum (2005), allowing us to show how our framework can be used to infer whether a cause prevents events from occurring. The second case study is a re-analysis of an experiment on the effects of temporal information on human causal inference that was originally conducted by Greville and Buehner (2007). This second case study demonstrates the value of being able to use delay distributions to characterize how the effect of a cause changes over time.

We begin with a brief overview of previous work on the role of time in human causal inference, focusing on Griffiths and Tenenbaum (2005) and Greville and Buehner (2007). We then lay out a set of desiderata for a computational framework for human continuous-time causal inference. We go on to describe formally how we implement these desiderata in our proposed framework. Following this, we apply the framework to our two case studies, evaluating models that use preventative causes and delay distributions. Finally, we conclude and suggest directions for future work.

Continuous-time causal induction

Studying the role of time in causal induction has a long history in cognitive science. One of the earliest established findings in the study of human causal inference is our ability to perceive causal relations in collisions, which is highly dependent on precise timing (Michotte, 1963). More recently, Buehner and colleagues have been very active in studying causal inference as it interacts with temporal information (e.g., Buehner & May, 2003; Greville & Buehner, 2007). However, while studies of causal induction often present events to participants in continuous time, they are typically analyzed using the discrete trial structure which the researchers used to design the stimuli (e.g., Anderson & Sheu, 1995; Wasserman, 1990).

Nonetheless, it may be enlightening to treat events as if they occurred over continuous time. Models considered by Griffiths and Tenenbaum (2005, 2009) take this approach, treating events that occur in continuous time as existing in a continuous dimension, or analyzing summaries of events as if they had occurred during a continuous interval. Even stimuli that are explicitly designed to convey information in discrete time (e.g., Greville & Buehner, 2007) can be analyzed in terms of continuous time by integrating over the time intervals. In the remainder of this section we summarize results from two studies on causal induction with temporal information, providing context for our later analyses.

Causal induction from rates

Griffiths and Tenenbaum (2005) showed that people are capable of reasoning about causes that increase the rate at which events occur over continuous time, and their judgments are in close accordance with the predictions of a computational model engaging in continuous-time causal inference. In their experiments, participants observed a series of results that they were told came from physics experiments studying whether different electrical fields cause different radioactive compounds to emit more or fewer particles (the compound always released particles at some rate). For each "experi-

ment", participants were told how many particles were emitted during one minute when the electrical field was on and one minute when the field was off. Participants then indicated the degree to which they endorse the claim that the field caused the compounds to emit more particles on a scale of 0 (the field definitely does not cause the compound to decay) to 100 (the field definitely does cause the compound to decay).

Causal induction from tabular displays

Greville and Buehner (2007) demonstrated that the temporal distribution of event occurrences will alter people's causal judgments, even if the relative frequencies of the occurrence of the effect in the presence or absence of a cause is held constant. Their purpose was to show that "temporal regularity" influences people's judgments above and beyond mere contingency information. Their experiments used a tabular format to display events that unfolded over five days (split up into five segments of one day each), reporting in which day events occurred. This discretization allows the use of traditional models of causal inference which infer causes on the basis of contingency information.

In each condition of Greville and Buehner's experiments, participants were shown two groups of 40 bacterial cultures, one group which was exposed to radiation and one which was not. Participants were shown (in tabular format) on which of 5 days each batch of bacteria died (if they died). Participants were asked to rate the effect of the radiation on a scale of -100 to 100 where -100 meant that the treatment was very effective at killing the bacteria, while 100 meant that the treatment was very effective at preventing the bacteria from dying (a rating of 0 meant that the treatment had no effect).

Greville and Buehner asked each participant about 18 pairs of tables, which differed in the frequency and distribution of times of death. In particular, Greville and Buehner varied the number of cultures dead by day five and the distribution over the times at which the bacteria died. They first fixed the number of deaths that would occur in each table. In all conditions, the time distribution for the bacteria not exposed to radiation was such that each of the deaths occurred with equal probability in any of the five days. However, for the bacteria exposed to radiation there were three time-of-death distributions: "strong contiguity", in which bacteria death was more likely in the first few days after the radiation treatment; "weak contiguity", in which bacteria died more often later in five day period; and "random", in which bacteria death was uniformly distributed among the five days. Contingency information was held constant while varying contiguity. The results of the experiments showed that temporal information dramatically affects human causal inference.

Defining desiderata for the framework

The studies discussed in the previous section illustrate some of the great variety in the phenomena to be considered by a framework for continuous-time causal induction. Thus, it will be helpful to identify the most vital features for allowing the framework to capture a wide class of these cases. The following sections detail an important set of these properties.

Intervention. The framework should capable of considering interventions in the sense meant in causal graphical models (Pearl, 2000). That is, an intervened upon node is said to be rendered independent of its parent nodes.

Instantaneous and interval causes. The framework should include both causes that exist instantaneously as well as over intervals of time.

Generative and preventative causal relations. It is vitally important when modeling human causal inference to distinguish between causes that generate effects and causes that prevent effects (Griffiths & Tenenbaum, 2005, 2009). People make dramatically different predictions based on which type of relationship they are looking for. Thus, we would want the framework to be capable of doing the same. In discrete time, Griffiths and Tenenbaum (2005) used the Noisy-OR and Noisy-ANDNOT logic gates to represent a cause that generates or prevents effects with reference to a background rate of the effects' occurrence. Because these discrete time parameterizations will not hold in continuous time, we will have to redefine what we mean by a generative and a preventative relation for continuous time.

Delay distributions. In most models of causation that work in discrete time or over trials in which events occur simultaneously, a cause can only influence an effect if and only if that cause is present. This is undesirable if we are to develop a framework for continuous-time causal inference. Not only would it be useful to track how a cause's influence changes over time, instantaneous events occur for only an infinitesimal period of time. Thus, in order for such events to have any effect on other variables they must be able to exert influence even after they are no longer present. Thus, we will need to characterize *delay distributions*, which define how a cause's influence on its effects changes over time.[1]

A framework based on Poisson processes

To form a rational framework encompassing these desiderata, we draw from the wide literature in statistics and computer science on continuous-time stochastic processes. In particular we pay attention to one class of continuous-time stochastic processes: Poisson processes. Poisson processes provide an excellent starting ground for generalizing causal graphical models (and hence intervention) as they define a series of independent random variables indexed over continuous time, being the continuous analogue of the independent Bernoulli events that take place on a series of discrete trials in many causal graphical models (Griffiths, 2005).

In its simplest sense, a Poisson process is a stochastic process (i.e., a series of random variables) that defines the rate at which instantaneous events occur over continuous time. That rate is determined by the rate function $\lambda(t)$. If a set

[1]This notion of change over time is not meant to capture that described in Rottman and Ahn (2009) where the change occurs over successive presentations of the cause, but change associated with temporal distance to one presentation of the cause.

of events are produced by a Poisson process, the probability that a certain number of events (k) occurred in a time interval ($[t_0, t_1],\ 0 \leq t_0 < t_1$) is,

$$P[(N(t_1) - N(t_0)) = k] = \frac{e^{-\lambda_{t_0,t_1}}(\lambda_{t_0,t_1})^k}{k!},$$
$$\text{where } \lambda_{t_0,t_1} = \int_{t_0}^{t_1} \lambda(t)dt\ .$$

The rate function defines the distribution of waiting times between events. For example, the waiting time before the first event (τ_1) is distributed $P(\tau_1 = t) = \lambda(t)e^{-\lambda_{0,t}}$.

Poisson processes have several desirable properties. If you have two independent Poisson processes with rates $\lambda_1(t)$ and $\lambda_2(t)$, you can take the union of their event sets and this produces another Poisson process with rate $\lambda_1(t)+\lambda_2(t)$. This is a "superposition" of Poisson processes. Now, suppose the existence of some Poisson process PP_0, which has rate $\lambda_0(t)$. Suppose also that you have another function with the same support on t called $\pi_1(t)$, the range of which is a subset of $[0, 1]$. Then, for an event produced by PP_0, cancel that event (i.e., treat it as if it had never occurred) with probability $\pi_1(t)$. This procedure is called "thinning" the Poisson process PP_0, and the resultant Poisson process has rate $\lambda(t)(1 - \pi(t))$.

Poisson processes have been used to model aspects of continuous-time causal induction, including both causes that occur instantaneously and over intervals (Griffiths, 2005; Griffiths & Tenenbaum, 2005, 2009). However, this work has focused on generative causes and only explored a limited class of delay distributions. In the remainder of the paper, we show that this framework can address more of these desiderata. First, we define preventative causes within the Poisson process framework, evaluating the resulting model through a new behavioral experiment based on Griffiths and Tenenbaum (2005). We then introduce an extremely general approach to handling delay distributions, which we evaluate using the results of Greville and Buehner (2007).

Generative and preventative causes

The properties of the Poisson process – specifically invariance of the form of the stochastic process under the superposition and thinning transformations – can be used to characterize generative and preventative causal relations. Suppose that there are i generative causes($\{C_i\}$) and j preventative causes ($\{C_j\}$), and they exist over intervals of time. That is $\forall C_a \in \{C_i\} \cup \{C_j\},\ \exists T_a(C_a = 1) \subset \mathcal{T}$ where $\mathcal{T}$ is the set of all non-measure-zero time intervals and $T_a(C_a = 1)$ is the set of intervals during which C_a occurs. Let the Poisson process PP_0 be a background rate of effect occurrence with an unknown time-invariant rate function $\lambda_0 > 0$. Causes assert their influence by altering the base-rate of the effect.

Generative causes will superpose themselves onto the background process, thereby increasing the rate of effect occurrence. That is, we can think of a generative cause C_i as producing a series of effects on its own, thereby inducing a Poisson process PP_i with parameter $\lambda_i(t)$, where we assume that the cause only exhibits a non-zero effect when it is present (i.e., $t \in T_i(C_i = 1)$). That is, when C_i is present, the rate will be $\lambda_0 + \lambda_i$, and otherwise the rate will be λ_0. This is equivalent to a continuous-time version of the Noisy-OR logic gate, used in models of discrete-time causal inference (see Griffiths, 2005; Simma et al., 2008).

We will assume preventative causes will thin all Poisson processes that generate effects including both the background and generative processes. A preventative cause C_j will have thinning parameter π_j which affects the generative processes if and only if cause is present. Thus, if $\lambda_{\text{total}}(t)$ is the total rate, when C_j is absent, the rate be $\lambda_{\text{total}}(t)$, but when C_j is present the rate will become $\lambda_{\text{total}}(t)(1 - \pi_j)$. This is equivalent to a continuous-time version of the Noisy-ANDNOT logic gate, which in the discrete-time setting defines the probability that an event will be canceled when the cause is present.

In the case where there are many causes, we will presume that they are independent and thus can be composed with one another, such that you will have a summation of the rates for the generative causes and a product of $1-$ the thinning parameters for the preventative causes. The rate function for a case with background rate λ_0 and i generative causes and j preventative causes is defined as

$$\lambda(t) = \Big(\lambda_0 + \sum_i \lambda_i \int_{T' \in T_i(C_i=1)} \delta(t, T')dT'\Big) \prod_j (1 - \pi_j \int_{T' \in T_j(C_j=1)} \delta(t, T')dT'). \quad (1)$$

where $\delta(\cdot, \cdot)$ is the Dirac delta function, which has an infinite spike where the two arguments agree and 0 elsewhere. Note, this does not include a prior for the causes; i.e., we treat these causes as continuous-time interventions.

The situation used as a cover story by Griffiths and Tenenbaum (2005) – determining whether electrical fields change the rate at which radioactive compounds emit particles – involves a system that can be analyzed using this model, where it has one effect (particle emissions) with a background rate and (possibly) one generative cause (the electrical field, C_i). Griffiths and Tenenbaum presented participants with information summarizing the number of effect occurrences (particle emissions) that occurred during one minute with the cause on and one minute with the cause off. For each compound, participants rated on a scale of 0 (the electric field definitely does not cause the compound to decay) to 100 (the electric field definitely does cause the compound to decay) their belief regarding whether C_i was indeed a cause.

To model the participants' predictions, Griffiths and Tenenbaum (2005) treated the problem as one of model selection between a graphical model G_0 where the cause had no effect (i.e., $\lambda_i(t) = 0, \forall t$) and a graphical model G_1 where the cause did have an effect (i.e., $\lambda_i(t) > 0, \exists t$). They parameterized G_0 and G_1 as we have above, as Poisson processes with different rate functions, where generative causes are treated

as we have treated them above. The quantity used to predict human judgments, termed "Causal Support", was the log likelihood ratio in favor of G_1, integrating over the values of all of the parameters of the Poisson process. This model performed well at predicting the mean judgments of the participants, with a scaled correlation of $r = .978, \alpha = .35$.[2] Other models considered by Griffiths and Tenenbaum (2005) also performed well, with the raw difference in rates ΔR (Anderson & Sheu, 1995) giving $r = .899$, $\alpha = .05$, a variant on the Power-PC theory (Cheng, 1997) giving $r = .845$ $\alpha = .06$, and a modified χ^2 score giving $r = .980$, $\alpha = .01$.

Griffiths and Tenenbaum (2005) considered only generative causes, creating the opportunity to use the same paradigm to evaluate whether the treatment of preventative causes outlined above is effective. We ran a new experiment to address this question. Considering only one preventative cause, we used nearly identical materials to Griffiths and Tenenbaum (2005), only changing the word "increases" to "decreases" and using the following $(N(c^-), N(c^+))$ pairs (where $(N(c^-)$ and $N(c^+))$ are the number of particles that were emitted during the minute when, respectively, the cause was absent and was present): $(52, 2)$, $(60, 10)$, $(100, 50)$, $(12, 2)$, $(20, 10)$, $(60, 50)$, $(4, 2)$, $(12, 10)$, $(52, 50)$.

We recruited 18 participants through Amazon Mechanical Turk to participate in our study online. We asked each participant to make the following judgment about each of the nine cases: "Does this field decrease the rate at which this compound emits particles?" Participants responded on a scale ranging from 0 (the electrical field definitely does not decrease the rate of particle emissions) to 100 (the electrical field definitely does decrease the rate of particle emissions).

Following Griffiths and Tenenbaum (2005) we modeled this task as a model selection problem between two graphs G_0 where the cause has no effect and G_1 where the cause has a (preventative) effect on the rate of particle emissions. We used the model defined in Equation 1 with one potential preventative cause with parameter π_1 and a background rate λ_0 to define the likelihood functions for G_0 and G_1. We assumed that in G_0, π_1 is constrained to be equal to 0. To obtain the log likelihood ratio, we need to provide likelihoods in terms of the graphical models (i.e., $P(D|G_0)$ and $P(D|G_1)$ for the observed data D). However, as they stand, the Poisson processes associated with these graphical models assume that the parameters λ_0 and π_1 are known, which is not the case. We thus need to define prior distributions over these parameters. With defined prior distributions, we can use Monte Carlo integration to obtain our marginal likelihoods, corresponding to the probability of the data given just the graphical model. We defined the prior for π_1 as $U(0, 1)$, i.e., uniformly distributed in the interval $[0, 1]$. Griffiths and Tenenbaum (2005) used an improper prior for λ_0, with $\lambda_0 \sim \frac{1}{\lambda_0}$. We approximated the previously used prior by sampling $v_0 \sim U(\log(10^{-6}), \log(10^6))$ and letting $\lambda_0 = e^{v_0}$. [3]

[2] As is usual in these studies, the authors scaled their model's values with the non-linear transformation $y = \text{sign}(x)|x|^\alpha$ where α is chosen to maximize the linear correlation r.

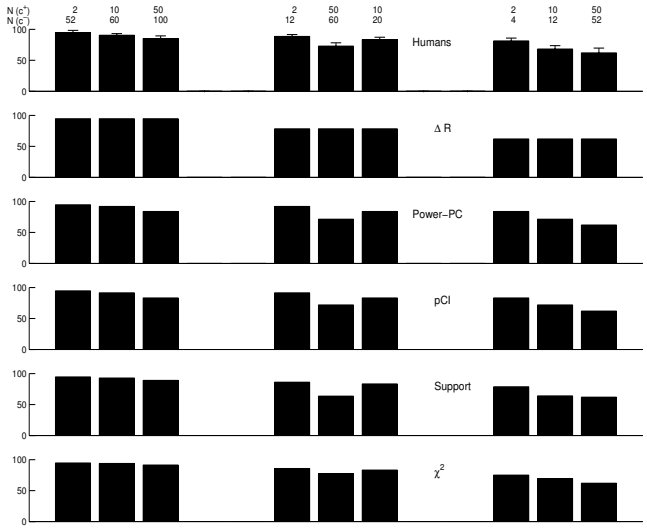

Figure 1: Preventative data for particle emissions: human responses and scaled model predictions. Support is the model that results from our framework.

Using the log likelihood ratio in favor of the hypothesis that $\pi_1 \sim U(0, 1)$ (G_1) over $\pi_1 = 0$ (G_0) as our predictor of mean human judgments, we see a high scaled correlation with the results of the experiment: Causal Support gives $r = 0.963$, $\alpha = 0.23$ (see Figure 1). We also evaluated the models tested by Griffiths and Tenenbaum (2005), which showed similarly high performance, ΔR : $r = 0.780, \alpha = 1.95 \times 10^-4$; Power PC: $r = 0.986, \alpha = 0.45$; and χ^2 : $r = 0.942, \alpha = 1.95 \times 10^-4$. Our purpose is not to claim that the model we have defined is the best model of human inference, but to demonstrate that the assumptions we have made about handling preventative causes in our framework are reasonable. Future work will hopefully clarify whether this model outperforms the other models in cases where their predictions diverge more dramatically.

Delay distributions

We will now describe how we implement delay distributions in our framework and apply the resultant model to modeling the results of Greville and Buehner (2007) – i.e., the bacteria death studies. We assume that generative and preventative causes have the same representation as above. We will assume that delay functions define what proportion of a cause's influence remains an arbitrary amount of time after it occurs, where a base parameter defines the maximum influence of the cause.

Let $f_i^g(\cdot, \cdot; \gamma_i)$ indicate the delay function with unknown parameters γ_i for generative cause C_i, and $f_j^p(\cdot, \cdot; \theta_j)$ indicate the delay function with unknown parameters θ_j for preventative cause C_j . Let the set $\{t_k\}$ be the set of times that instantaneous cause C_k occurs and $\{[t_{l,0}, t_{l,1}]\}$ be the set of

[3] To see the approximation, note that $v_0 = \log(\lambda)$ and $v_0' = \frac{1}{\lambda}$ and use a change of variables to find $f_{\lambda_0}(\cdot)$.

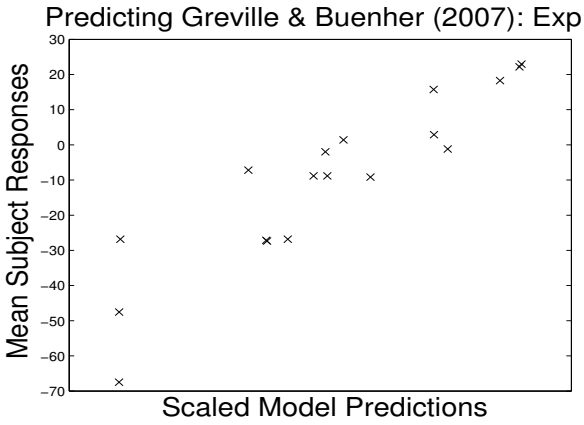

Figure 2: Model predictions for Greville and Buehner (2007), Experiment 1.

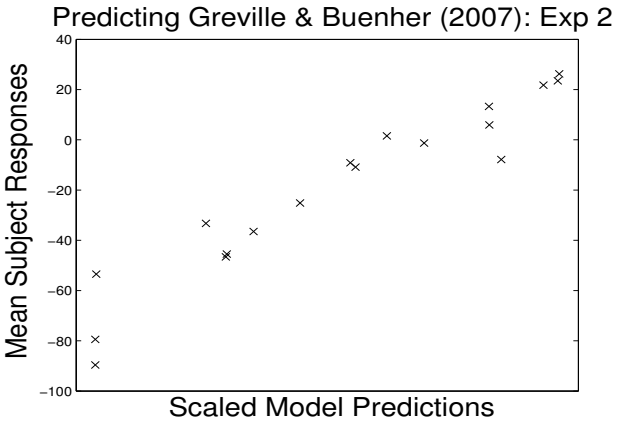

Figure 3: Model predictions for Greville and Buehner (2007), Experiment 2.

times over which interval cause C_l occurs. We set λ_0 to indicate the underlying rate of an effects' occurrence, and λ_i and π_j to indicate the maximum value of a cause's influence.

The effects of delay distributions can be accommodated by defining a Poison process with rate

$$\lambda(t) = \Bigg(\lambda_0 + \sum_{C_i^g} \Big(\sum_{t' \in \{t_i\}} \lambda_i f_i^g(t, t'; \gamma_i) + \sum_{[t_0', t_1'] \in \{[t_{i,0}, t_{i,1}]\}} \lambda_i f_i^g(t, [t_0', t_1']; \gamma_i) \Big) \Bigg) \prod_{C_j^p} \Big(\prod_{t'^* \in \{t_j\}} (1 - \pi_j f_j^p(t, t'; \theta_j)) \prod_{[t_0', t_1'] \in \{[t_{j_0}, t_{j_1}]\}} (1 - \pi_j f_j^p(t, [t_0', t_1']; \theta_j)) \Big).$$

where $f(t, [t_0, t_1]; \cdot)$ is the convolution of $f(t, x; \cdot)$ with the boxcar function on $[t_0, t_1]$ (i.e., the function that takes the value 1 for all $x \in [t_0, t_1]$ and 0 otherwise). This allows us to keep the expressivity needed to capture our first findings, while allowing greater generality in the types of delay distributions applicable to interval causes.

Modeling the studies in Greville and Buehner (2007) requires further formal specification. Because events in these experiments were deaths they happen only once. As such, we only consider a bacterium to have died on the first arrival in a Poisson process defining the rate of death (i.e., $p(\tau_1 = t) = \lambda(t) e^{-\lambda_{0,t}}$, where $\lambda_{a,b} = \int_a^b \lambda(t) dt$). But, we do not know the precise time at which the bacterium died, merely the day on which it died. Therefore, the likelihood that bacterium i died on day $t_{i,1}$ is $\int_{t_{i,0}}^{t_{i,1}} \lambda(t)\, e^{-\lambda_{0,t}} dt = e^{-\lambda_{0,t_{i,0}}} - e^{-\lambda_{0,t_{i,1}}}$, where $t_{i,0}$ is the day before $t_{i,1}$. Finally, we model the 80 bacterial cultures in each condition as 80 conditionally independent Poisson processes given an underlying graph, i.e.,$p(D|G) = \prod_{i=1}^{80} \int_{t_{i,0}}^{t_{i,1}} \lambda(t)\, e^{-\lambda_{0,t}} dt$.

Because Greville and Buehner (2007) asked participants to respond on a scale of -100 (the radiation definitely causes death) to 0 (the radiation has no effect) to 100 (the radiation definitely prevents death), we have effectively three graphs to choose from: G_g, the generative graph (where $\pi_1 = 0$ and $\lambda_1 \in \mathbb{R}^+$); G_p, the preventative graph (where $\lambda_1 = 0$ and $\pi_1 \in [0, 1]$); and G_0 the null graph (where $\pi_1 = \lambda_1 = 0$). As in Griffiths and Tenenbaum (2009), we modeled participants' mean responses in each condition as $P(G_p|D) - P(G_g|D)$, assuming all three graphs are a priori equally likely.

We assumed a scaled exponential decay function with parameter ϕ_1 is used for both generative and preventative causes (i.e., $f_i(t, t'; \gamma_i) = f_j(t, t'; \theta_j) = e^{-\phi_1(t-t')}$ where t' is the time that a cause occurs). Because the radiation is only applied to the bacteria once at the beginning of the five days, for G_g and G_p, the only occurrence of the cause is instantaneous and appears at $t = 0$. Thus, for the generative graph,

$$\lambda_{0,t} = \int_0^t \lambda_0 + \lambda_1 f_1(s, 0; \gamma_1) ds = t\lambda_0 + \frac{\lambda_1}{\phi_1}(1 - e^{-t\phi_1}),$$

and for the preventative graph,

$$\lambda_{0,t} = t\lambda_0 (1 - \frac{\pi_1}{\phi_1}(1 - e^{-t\phi_1}))$$

Similar to before, as a prior for λ_0 we used $v_0 \sim U(\log(10^{-1}), \log(10^1))$ and set $\lambda_0 = e^{v_0}$. The remaining priors were defined as $\pi_1 \sim U(0, 1), \lambda_1 \sim \Gamma(1, \lambda_0)$, and $\phi_1 \sim \Gamma(1, \lambda_0)$, where the priors are defined in terms of λ_0 such that they inherit the scale defined by λ_0.

Using Monte Carlo integration, we calculated our model's judgments $P(G_p|D) - P(G_g|D)$ for the data in Greville and Buehner (2007). Because the experiments used slightly different methods we evaluated our model predictions separately for each experiment but concurrently for all 18 conditions within each experiment. Our model has a scaled correlation of $r = .910$ ($\alpha = 2.74$) with mean participant responses in Experiment 1 and a scaled correlation of $r = .957$ ($\alpha = 1.72$) with mean participant responses in Experiment 2. Since submitting this paper, we learned of another model which outperforms our own – namely that described in Buehner (2006).

This was used to analyze the same data and had an excellent linear fit for the two experiments, $r = .97$ and $.953$ (Buehner, 2006). In these cases, our models make very similar predictions thus, we will need to explore more complex experimental scenarios (e.g., trials with multiple exposures to the cause at different times). This would put predictions from these models in starker contrast.

Conclusions and future directions

Continuous-time causal induction is so pervasive that we often go about not even noticing that we are engaging in it. The richness of temporal information surely aids people as they infer causes in their everyday life. Here we have developed a rational framework that makes use of that same wealth of information. The framework is based on an extension to causal graphical models to include continuous-time stochastic processes – specifically Poisson processes. We demonstrate in two case studies that our framework is capable of accurately predicting human judgments in tasks that require reasoning about preventative causes and reasoning about delay functions. This extends previous work on Poisson processes as rational models of continuous-time causal induction.

Continuous-time stochastic processes are a very rich class of mathematical objects, and we expect that the formal framework we have outlined will grow more powerful as further tools are added to it. Fortunately, there are currently many tools being crafted. Our hope is to develop inference algorithms that allow our framework to consider large, complex networks of causal variables and the relations between them, in the vein of Pearl (2000) and Simma and Jordan (2010). Additionally, it will add an additional layer of generality to develop an account for how instantaneous events can alter the states of events that occur over intervals of time. Currently we take the parameters of causes as fixed at all times – it is merely the influence of the cause on the effect that wanes. However, people are very capable of reasoning about causal relations that change their form over time (Rottman & Ahn, 2009), and as such explaining data of that sort may be essential for capturing the full range of human causal reasoning.

Developing a rational framework for continuous-time causal induction could potentially provide insight into other phenomena of human causal judgment. One of the advantages of taking a Bayesian approach to causal induction is its ability to form strong inferences from very small amounts of data. This feature may be essential in explaining why perceptual causality (where one makes a causal inference from a single piece of data) is extremely sensitive to subtle differences in timing (Michotte, 1963; Newman, Choi, Wynn, & Scholl, 2008). Finally, one of the central motivations for studying time in causation is a pervasive belief that the timing of events can unveil the direction of the underlying relationship (i.e., what causes what; Rottman & Keil, 2012). If people do believe this (even implicitly), then characterizing the role of continuous-time in causation would be absolutely necessary if we are to understand the full extent of the human mind's capacity and propensity for causal inference.

Of course, all of this belies the fact that there are many phenomena on human causal reasoning that have yet to be studied. What any computational-level framework offers is the ability to develop new questions out of the formal principles that originally drove the design – even if those questions did not exist when the framework was formulated. If such an event occurs in the near future, it would be a pleasant thought to think that it could have been the effect of the work presented here; but whether that will occur – only time will tell.

Acknowledgments. This work was supported by a Berkeley Fellowship awarded to MP and grant number FA-9550-10-1-0232 from the Air Force Office of Scientific Research.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory & Cognition*, *23*, 510-524.

Buehner, M. (2006). A causal power approach to learning with rates. In *Proceedings of the 28th annual conference of the cognitive science society.*

Buehner, M., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology Section A*, *56*(5), 865–890.

Cheng, P. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405.

Greville, W., & Buehner, M. (2007). The influence of temporal distributions on causal induction from tabular data. *Memory & Cognition*, *35*, 444453.

Griffiths, T. L. (2005). *Causes, coincidences, and theories*. Unpublished doctoral dissertation, Stanford University.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*, 354-384.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological review*, *116*(4), 661.

Marr, D. (1982). *Vision*. San Francisco, CA: W. H. Freeman.

Michotte, A. (1963). *The perception of causality*. New York: Basic Books.

Newman, G., Choi, H., Wynn, K., & Scholl, B. (2008). The origins of causal perception: Evidence from postdictive processing in infancy. *Cognitive psychology*, *57*(3), 262–291.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, UK: Cambridge University Press.

Rottman, B. M., & Ahn, W. (2009). Causal learning about tolerance and sensitization. *Psychonomic Bulletin and Review*, *16*(6), 1043-1049.

Rottman, B. M., & Keil, F. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*(1), 93–125.

Simma, A., Goldszmidt, M., MacCormick, J., Barham, P., Black, R., Isaacs, R., et al. (2008). Ct-nor: representing and reasoning about events in continuous time.

Simma, A., & Jordan, M. (2010). Modeling events with cascades of poisson processes. In *International conference on uncertainty in artificial intelligence.*

Wasserman, E. A. (1990). Detecting response-outcome relations: Toward an understanding of the causal texture of the environment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, p. 27-82). San Diego, CA: Academic Press.