Précis of Causal Induction in Time

Tianwei Gong

Ph.D. 2023, Department of Psychology, The University of Edinburgh

1 Background (Chapter 1-3)

We all learn how to navigate the journey of life. Throughout history, humans have sought to uncover the secrets of hunting, farming, cooking, and healing. The challenge lies in discovering what truly "works", a challenge intricately intertwined with our understanding of the objective world's *causal structure*.

Humans began seeking causal conclusions long before the evolution of science or the advent of formal scientific education for individuals. Every day, we continue to collect experiences, with each passing moment bringing forth myriad occurrences. We might judge that a backfiring car startled some birds, a new food gave us indigestion, infer from a boiling kettle that someone was recently in the break room, from the heavy evening traffic that an accident has occurred, predict that you will be sore the day after the gym, or that a storm is coming after a pink sunrise. All these inferences leverage causal models linking events in virtue of their experienced and historical temporal proximity through the lens of our intuitive causal theories. *Time* is inherent to our understanding of the world, shaping how we link the things that happen around us and the actions we take. As such, understanding how individuals acquire knowledge of causal structures from temporal events becomes a core aspect of comprehending the cognitive process of causal learning.

Throughout the extensive history of causal cognition research, the field has established many theoretical and computational models aimed at describing human causal learning processes. Some notable examples include causal Bayesian networks, widely utilized not only in cognitive research (Griffiths & Tenenbaum, 2005; Rottman & Hastie, 2014) but also in scientific causal discoveries (Pearl, 2000). Additionally, classic learning theories like the Rescorla-Wagner rule (Rescorla & Wagner, 1972) and Power PC (Buehner et al., 2003; Cheng, 1997) are encompassed within this framework, along with recent process-level models that further consider the constraints of human cognitive resources (Bonawitz et al., 2014; Bramley et al., 2017; Davis & Rehder, 2020). However, most of these studies focus on scenarios where the evidence is helpfully "prepackaged" in the form of multiple (typically independent) trials or observations in which combinations of causal variables take different states (Allan, 1980; Cheng, 1997; Griffiths & Tenenbaum, 2005; Rescorla & Wagner, 1972). One common paradigm involves presenting participants with a set of independent samples in which putative cause and effect events are either present or absent. Cover stories have been used to frame this as data arising from experimental research in biology (Buehner et al., 2003; Lu et al., 2008), physics (Coenen et al., 2015; Lagnado & Sloman, 2004), and psychology (Rottman & Keil, 2012), since multiple independent trials are often the data that scientists collect under laboratory conditions.

While these settings put timing considerations to one side, they do not eliminate them (Pacer & Griffiths, 2015). Fundamental questions remain as to how to determine an appropriate time window to measure outcomes, how to ensure the observations are sufficiently independent to be aggregated (Gong & Bramley, 2023b), and how to determine the timing of interventions since some time-dependent factors (e.g. age) may also mediate the relationships between variables (Gong et al., 2023; Rottman, 2016). This implies we must to be able to litigate between competing causal explanations linking multiple events even as they occur and recur within a single ongoing data-stream, which remains relatively unexplored in the literature. Besides, many causal processes in the natural world are also cyclic (Malthus, 1872), and people frequently report causal beliefs that include feedback loops when allowed to do so in experiments (Kim & Ahn, 2002; Nikolic & Lagnado, 2015; Rehder, 2017; Sloman et al., 1998). Cyclic systems can involve both excitatory or inhibitory feedback, which can result in complex, periodic and chaotic behavior (Davis et al., 2020). Recognizing, predicting, explaining or

controlling the behavior of such cyclic causal systems is only possible if one properly represents the temporal dimension.

On the other side, empirical studies have shown that people adhere to several generic principles when processing delay information: They prefer causal explanations that suggest shorter delays between the causes and effects (Buehner & McGregor, 2006; Shanks et al., 1989), those that are more reliable (Bramley et al., 2018; Greville & Buehner, 2010), and those that involve more predictable delays (Hagmayer & Waldmann, 2002; Stephan et al., 2020). However, due to the lack of a computational framework, we cannot fully account for the rationale behind these preferences or the existence of multiple preferences within the same task (Greville & Buehner, 2010). Meanwhile, the empirical patterns were discovered non-systematically and mainly focused on systems that include only one evaluated cause and one effect. It remains unknown how people would judge when the causal system involves more than two variables, different kinds of causal links (generative vs. preventative; acyclic vs. cyclic), and when they need to make their own interventions.

This thesis explores computational theories and empirical knowledge regarding how people learn causal structures from events unfolding in continuous time. The precis is constructed using three key features of time:

Continuous Present: There exists an infinity of time points where events could occur. Chapter 4 provides a rational Bayesian framework capable of representing these dynamic and continuous features. It demonstrates how this framework can explain the principles of short, reliable, and predicted delays found in human reasoning. Additionally, this framework is shown to recover human judgments in seven causal learning tasks.

Intractable Past: Any event happening now could, in principle, be attributed to any event that occurred earlier. Chapter 5 explores how people learn in systematically manipulated causal learning tasks, encompassing causal systems with generative and preventative links. Chapter 6 delves into how people strategically intervene in time to enhance their understanding of causal structures. In both chapters, computational models are proposed to approximate the normative solution, considering the issue of intractability.

Unknown Future: Evidence is still forthcoming. Chapter 7 explores how people consider the unseen future and how different contextual implications about the future can even lead to opposing causal conclusions.

2 Chapter 4: Continuous Present

This chapter develops a rational framework that incorporates the role of time in guiding causal learning. I work within the Bayesian rational analysis tradition (Anderson, 1990; Marr, 1982), as this has proven successful in developing theories of atemporal causal induction (Griffiths & Tenenbaum, 2005; Rottman & Hastie, 2014). However, I depart from past analyses of causal inference by linking causal influence with dependence between events in continuous time (i.e. *contiguity*) rather than co-incidence of variable *states* across independent trials (i.e. *contingency*).

I here use a daily example to demonstrate the way of using Poisson-gamma family to capture temporal information to make causal inference. We suppose a fictional substance called 5-HTP is used to treat insomnia. Consuming an 5-HTP capsule can cause a person to sleep, representing a one-cause–one-effect scenario. Here, temporal information is embedded within the delay between the causative event of pill consumption and its effect event of falling asleep. The causal delay can vary across different mechanisms and hence follow distributions with different shapes (see Figure 1a). The gamma distribution provides a variety of shapes to capture different temporal mechanisms. It can be codified with a shape parameter α along with the rate parameter β:

$$P_d(t|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} \tag{1}$$

We might also model the same effect less granularly in terms of the pill's production of Melatonin particles over time (Figure 1b), forming as one kind of one-cause–many-effect scenarios. In this case, instead of focusing on the relationship between a cause and individual effect events, it may be more practical to think at a rate-level about how many additional events we expect it be generated by the cause per time unit and how this rate change to be spread over time. This requires reasoning about the functional form of the event's causal influence on the effect's rate over time, such as, a potential

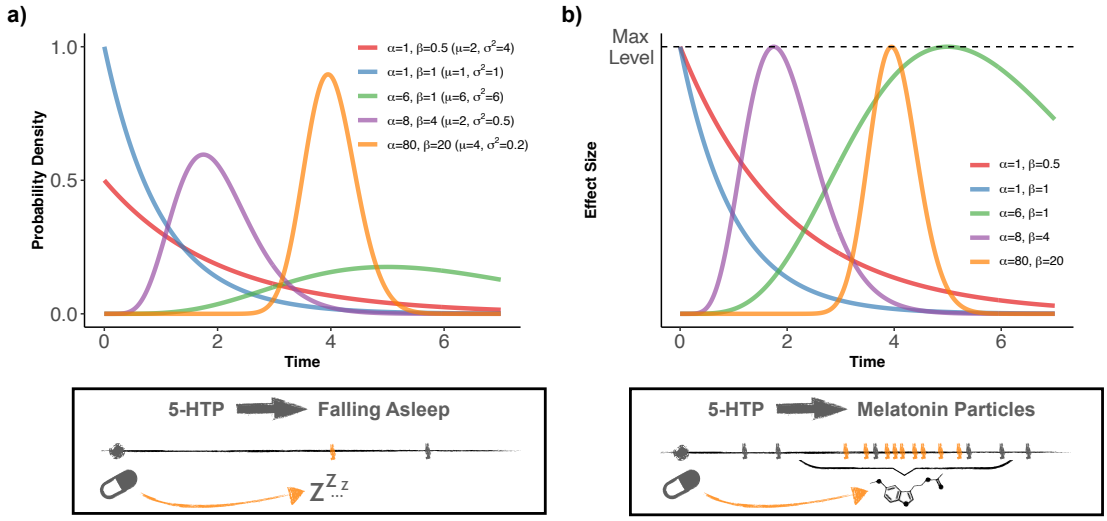

Figure 1: Examples of two types of function that could be used to model cause-effect delays and causal influences, respectively. Illustrative example relates a drug "5-HTP" and sleep. (a) a gamma probability density function capturing the delay between taking a drug and falling asleep and (b) scaled gamma density function capturing the rate of melatonin production after drug is administered. The purple distribution is the ground truth generative distribution. The orange effects in the timeline are those in fact generated by the drug while the gray effects are the base rate effects.

incubation period, peak, and a decay process. The scaled Gamma distributions are used to represent how the rate changes $\lambda = f(\lambda_{max}, t)$ across time (Figure 1b). The Poisson process is used to model the probability $P_r(k|\lambda)$ of observing a particular quantity k of such independent events (the number of events per time unit) given their assumed $r$ate λ:

$$P_r(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{2}$$

Challenging our initial 5-HTP example, everyday continuous-time evidence is often more complicated: Events can occur at any time point, and different potential causes can overlap in time leading to pervasive credit assignment questions. Given two levels of granularity shown above, I propose two schemes to solve the problem. The event-based scheme uses the concept of token-level "actual causation" to map each event to its possible causes (Halpern, 2016), identifying which of several candidate events actually caused the observed outcome (Gerstenberg et al., 2021; Stephan et al., 2020). One can consider various possible causal pathways that could produce the observed events, depending on the underlying causal mechanisms (Bramley et al., 2018; Valentin et al., 2022).

The rate-based scheme models causes that temporarily affect the rate of occurrence of some effect. For a generative cause like "5-HTP", we expect the rate of its effect to temporarily increase from its base rate, and intuitively expect such rate increases to be additive (unless there are also interactions between the base rate causes and the focal cause). That is, an independent generative cause is something that adds extra events to the timeline without affecting those that would have been there anyway. For example, we might think of a large gathering causing Covid rates to spike by contributing additional infection events. We can model this even though it is infeasible to establish a one-to-one relationships between individual cases caused by base rates and those caused by the gathering.

This rational approach anticipates the ceteris paribus human preference for causal explanations that posit shorter, more reliable and more predictable causal influences (Buehner & McGregor, 2006; Greville & Buehner, 2010; Hagmayer & Waldmann, 2002; Shanks et al., 1989). Besides, the framework shows how to uncover the underlying causal structure all manner of complex, mechanistically specified continuous-time settings. I showed our framework can explain behavioral patterns across a range of learning tasks from the last 20 years (see Table 1 for a summary). These tasks encompass a variety of scenarios, including extended episodes containing many events and episodic evidence in which the same events recur. They also vary in the size of the hypothesis space, the involvement of background activity,

Table 1: The features of datasets modeled in Chapter 4.

Name	Reference	Base Rate	Prevention	Cycle	Delay Prior
Continuous timeline, effect specified:					
Earthquake	Lagnado and Speekenbrink (2010)	✔	✗	✗	✗
Device: Prevention	Gong and Bramley (2023a)	✔	✔	✗	✔
Continuous timeline, effect unspecified:					
Device: Active Learning	Gong et al. (2023)	✗	✗	✔	✔
Episodic evidence, effect specified:					
Bacteria	Greville and Buehner (2007)	✔	✔	✗	✗
Future Bacteria	Gong and Bramley (2023b)	✔	✔	✗	✗
Episodic evidence, effect unspecified:					
Computer Virus	Lagnado and Sloman (2006)	✗	✗	✔	✗
Device: Chain or Fork	Bramley et al. (2018)	✗	✗	✗	✗

preventative connections, and the presence of cyclic dynamics, as well as how much information learners had about the relevant causal delay distributions and base rates. I find a good degree of consistency between judgments from the rational framework and judgments from people (e.g. Pearson correlation: 0.68-0.96). Thus, people are not only capable of utilizing temporal information in diverse causal learning situations but also doing so in systematic, predictable, and to some extent, rational ways.

3 Chapter 5: Intractable Past [1]

Few studies have examined learning and reasoning about systems with events unfolding in continuous time. Among these, none have explored learning about preventative causal influences. How do people use temporal information to infer which components of a causal system generate or prevent activity in other components? In what ways do generative and preventative causes interact to shape the behavior of causal mechanisms and their learnability? What algorithms might people employ to process temporal dynamics for making causal inferences, particularly when they need to solve problems in real time?

In this chapter, I explore human causal structure learning within a space of hypotheses that combine generative and preventative causal relationships. Participants observe the behavior of causal devices as they are perturbed by fixed interventions and subject to either regular or irregular spontaneous activations (see Figure 2). The normative solution to this question, as mentioned in Section 2, relies on detailed considerations about the token-level causation giving rise to the observable evidence (i.e. which particular event actually caused which particular effect). While the enumerative approach achieves benchmark performance by inverting the generative model, exhaustively considering pathways linking all observed events, it makes unrealistic demands on memory storage and computing power compared to what could plausibly be involved in humans. I propose a family of more cognitively plausible algorithmic approximations. It is based on the simulation-and-summary-statistic idea which is an important approach in Approximate Bayesian Computation in statistics (Blum et al., 2013; Sunnåker et al., 2013; Ullman et al., 2018). The proposed approximation algorithms consider three bounded features that are often highlighted in cognitive psychology: mental simulation — people make inferences by comparing their observations to mental simulations of what *kind of pattern* they expect to happen (Battaglia et al., 2013; Gerstenberg et al., 2021); (structurally) local computation — people make causal attributions at the level of individual links without accommodating the full space of global causal models (Davis et al., 2020; Fernbach & Sloman, 2009); and (temporally) local evidence — people abstract cues to help reasoning by segmenting the evidence encountered across an extended observation (Bramley et al., 2017; Harman, 1986).

In Experiment 1, I provided an initial empirical demonstration that people can use real-time temporal information to detangle the influences of generative and preventative causes and identify causal structures involving combinations thereof. Besides, it is found that the base rate regularity matters as participants better identified preventative connections when the effect otherwise activates regularly. The type of neighboring connections matters as participants better identified a connection when

[1] The research from this chapter appears in the literature as: Gong, T., & Bramley, N. R. (2023a). Continuous time causal structure induction with prevention and generation. *Cognition*, *240*, 10553. The task demo is available here.

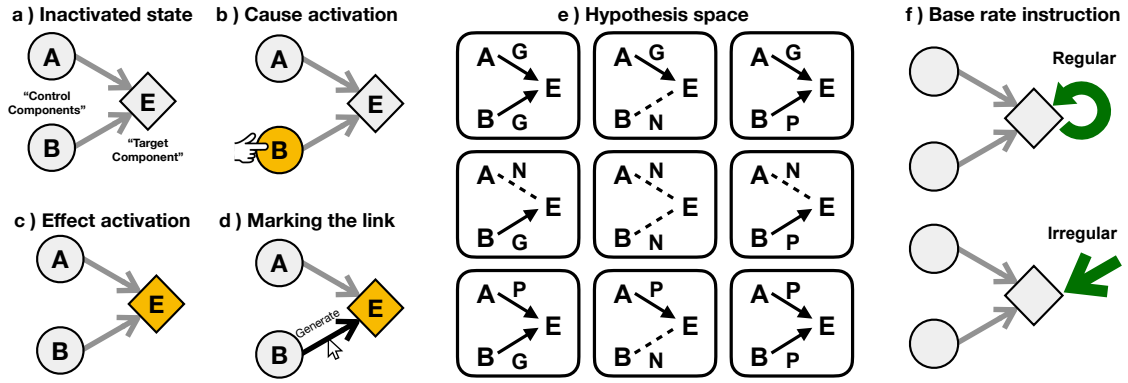

Figure 2: Causal devices tested in Chapter 5. a-d) Experimental interfaces. Participants were instructed to the control components and target components in the causal devices and observed how the system reacted to pre-set interventions. They marked their answers of the role of each connection during or after the observation. e) The response hypothesis space (all possible pairwise combinations of generative (G), non-causal (N), and preventative (P) connections). f) The illustrations shown to participants in the regular (periodic) vs. irregular (exogenous) base rate condition.

it was paired with a generative neighbor. The timing and sequence of interventions matter as participants identified generative relationships better when the interventions were clustered rather than interleaved. These rich empirical patterns, as well as the quantitative judgments, were better captured by the summary-statistic approach rather than the normative approach.

In Experiment 2, I further constructed two special types of stimuli for which two models have different dominant answer. They are based on the two locality principles driving the summary-statistic model: (1) Local computation indicates a failure to account for the influence of the other connections in the system, and (2) Local evidence indicates a failure to take into account whatever happened before their current observation window. In the first type of scenarios, the learner needs to identify a generative target cause that is paired with a preventative cause. This presents a challenge for local computation because the preventative cause can block the generative causes' influence and mislead a local learner into believing the target connection is a non-causal connection. In the second type of scenarios, a non-causal target is paired with a generative neighboring component. For a local learner who only focuses on a small time window after each intervention, the generative influences can easily spill over to the observation window during which the learner is focused on the target non-causal component and leading to statistics more typical of generative causation. Participants' judgments for both types were consistent with the summary-statistic learners rather than the normative learners. In sum, this project demonstrates a quantitative account of how people manage to learn causal structure on the basis of real-time continuous temporal dynamics.

4 Chapter 6: Making Interventions in Continuous Time [2]

The ability to predict, plan, and control events in the world demands a sophisticated representation of the world's causal structure. Learning such a causal model requires gathering causal evidence through interventions (Pearl, 2000). However, learning causal structure in general, and selecting interventions in particular, are computationally challenging problems even under idealized conditions (Bramley et al., 2017). In everyday life, this challenge is compounded by the need to interact with the causal environment in real time, bringing computational constraints to the fore (Griffiths et al., 2015; Simon, 1982). How do people actively learn about causal structure in real time?

In this chapter, I constructed a causal learning task in which participants interact with causal devices in real time, deciding when and where to intervene in order to gather information about how the device works. There are at least two typical ways to intervene in a causal system: (1) By activating components, thus potentially setting in motion a new sequence of events; and (2) by

[2] Gong, T., Gerstenberg, T., Mayrhofer, R., & Bramley, N. R. (2023). Active causal structure learning in continuous time. *Cognitive Psychology*, *140*, 101542. The task demo is available here.

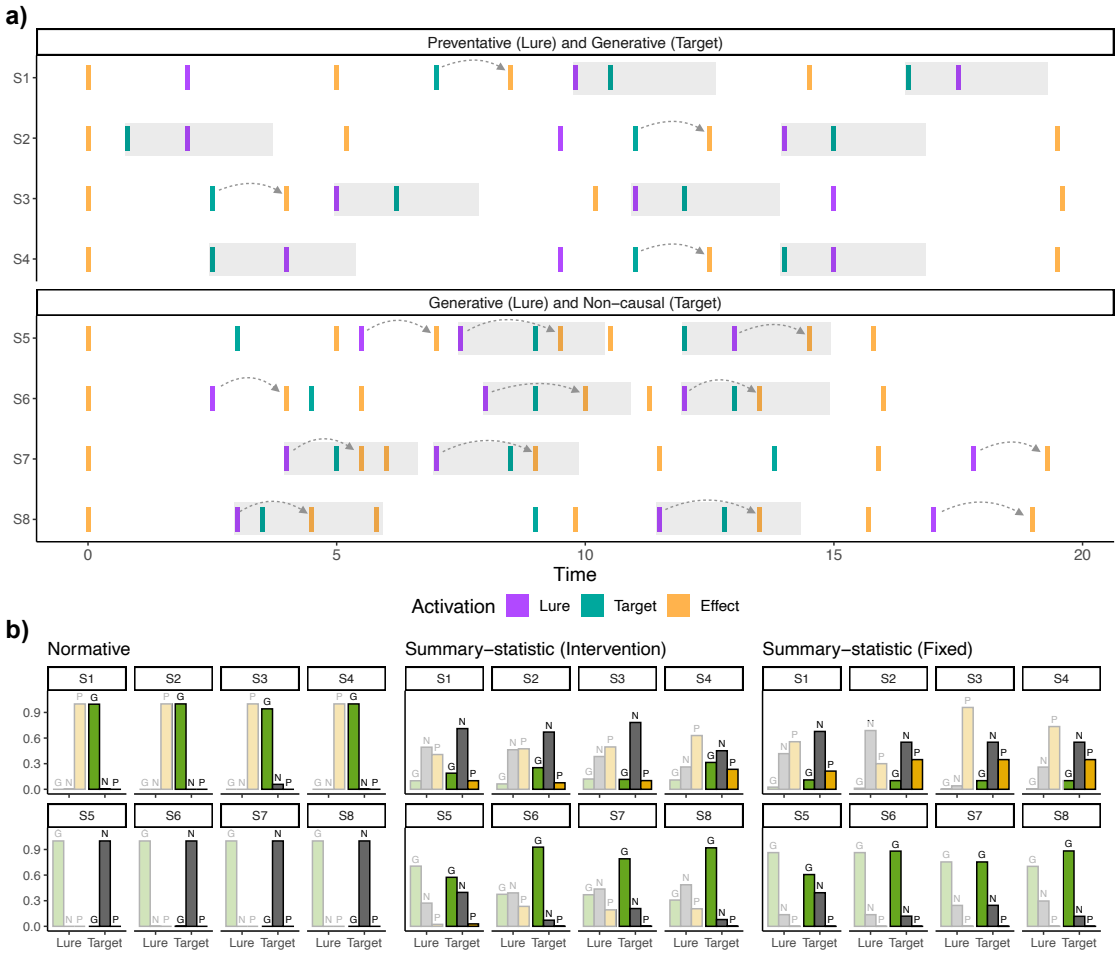

Figure 3: Stimuli and model predictions in Experiment 2 of Chapter 5. a) Stimuli. Curved arrows indicate the true underlying generative process. b) Judgment predictions from different models. The normative and summary-statistic models particularly differ in their judgments about the target components, with opaque bars used to highlight where the modal response shifts between normative and summary statistic models.

blocking components, thus preventing that component both from being activated and from activating any other components until it is unblocked again. In two experiments, I included a range of acyclic and cyclic causal structures. In Experiment 1, I allowed participants to activate components while in Experiment 2, I also allowed them to block components.

Two experiments showed that people are able to infer causal structure through active intervention in a challenging continuous-time learning setting. Participants had different error patterns from an ideal observer model, in particular making more accurate judgments about acyclic structures than cyclic structures while the ideal observer had the reverse pattern. It was also found that differences in accuracy across conditions were associated with differences in the character of the evidence. The informativeness of evidence predicted participants' performance in acyclic structures, but the complexity of evidence appeared to dominate participants' performance in cyclic structures where it was generally higher. Participants who were able to generate evidence that was both informative but not overly complex tended to perform best overall (see Figure 4).

Intervention choices were also partly shaped by a drive to control computational demands. In terms of when to intervene, participants performed fewer activating interventions and waited longer between them on cyclic structures that tended to produce more events. They also tended to perform more activating interventions on four-node structures yielding a similar number of events as for three-node structures but presumably responding to the greater initial uncertainty (larger space of structure

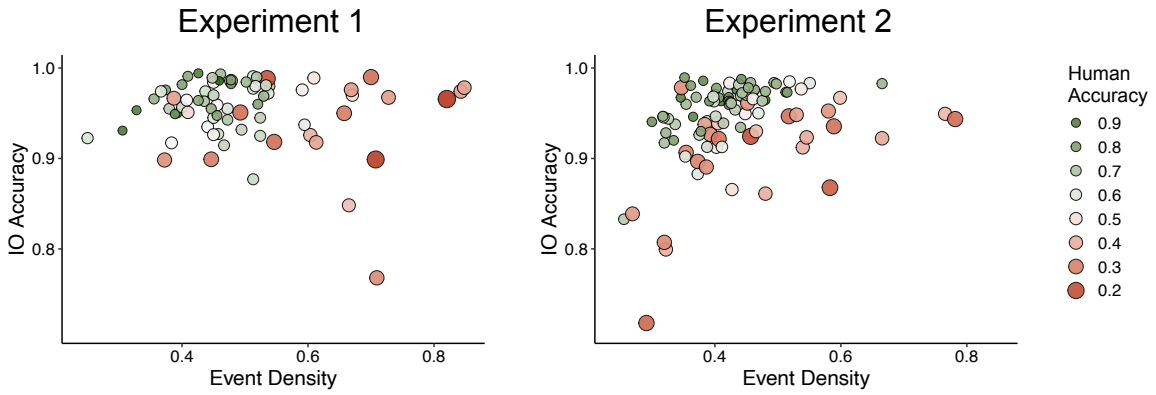

Figure 4: Scatterplots of average final IO accuracy (indexing evidence informativeness) and event density (indexing evidence complexity) for each participant with color and size indicating that participants' judgment accuracy in Chapter 6. Participants with higher accuracy generated evidence that was both more informative and less complex (the upper left area).

possibilities). They blocked more often in cyclic than acyclic devices and did so when many events could be expected to occur in the near future. When the expected upcoming evidential complexity was already high, participants were more likely to wait or block rather than activate another component to produce more events.

The empirical results support the central idea that managing computational cost plays an important role in interventional decisions and success in the real-time causal learning setting. I then model the role of complexity in shaping participants' causal judgments and intervention choices. The judgment model assumes that human causal judgments $q \in \mathbf{Q}\{X \rightarrow Y, X \leftarrow Y, X \leftrightarrow Y, X \varnothing Y\}$ are a noisy version of the ideal observer's posterior marginalized across connections IO_q, where the noise degree depends on the density, and hence complexity. I capture this with a dynamic softmax function (Luce, 1959):

$$P(\text{judgment} = q) = \frac{\exp\left(\mathrm{IO}_q / (\tau_1 N + \tau_2)\right)}{\sum_{q' \in \mathbf{Q}} \exp\left(\mathrm{IO}_{q'} / (\tau_1 N + \tau_2)\right)} \tag{3}$$

Here N denotes a trial's event density (average number of events per second). The judgment temperature component is thus a linear function of events $f(N) = \tau_1 N + \tau_2$ with two parameters $\tau_1, \tau_2 \in (0, +\infty)$ that are constant across trials, while N varies across trials depending on what interventions are performed and how the system reacts to them.

The intervention model is based on the resource-rational framework (Lieder & Griffiths, 2020), suggesting human minds discover solutions that trade off efficiently between the costs of computation and its rewards in greater accuracy or performance. Accordingly, the expected utility of an action $\mathbb{E}[U(i^*)]$ to a bounded learner balances expected reward EIG (i.e. expected information gain) and cost of computation ECC (i.e. expected computational cost; measured by the expected event density):

$$\mathbb{E}[U(i^*)] = \sum_{t=t_x}^{t_y - 1} R(t) \cdot \left[\mathrm{EIG}[i^*]_t^{t+1} - \omega \cdot \mathrm{ECC}[i^*]_t^{t+1}\right] \tag{4}$$

Here ω scales the cost component to align it with the epistemic reward scale of bits, the sum aggregates the expected future gains and costs over future seconds up until t_y, with $R(t)$ as a discount function which diminishes the utility of information and the dis-utility of computational costs the further into the future they occur.

Both models proposed above can best capture participants' judgments and interventions among a set of different computational models (see Figure 5 for a demostration). The experiments and modeling show that participants' causal judgments depend on not just the informativeness but also the complexity of the evidence they gather, and that they adapt their actions to the ongoing event

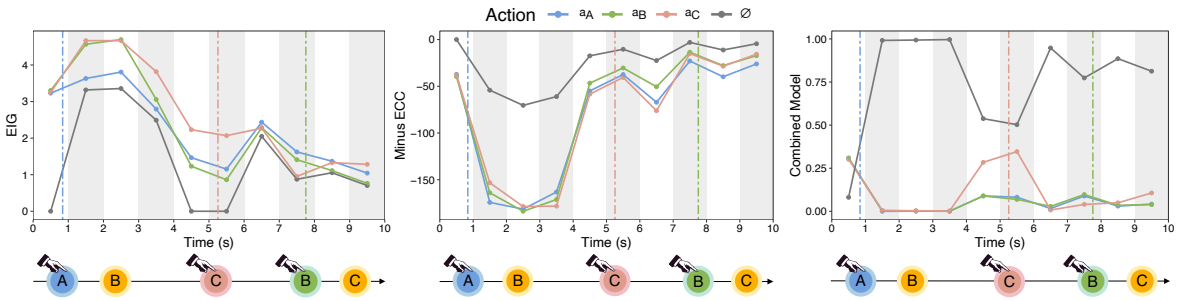

Figure 5: Example of real-time model prediction for a participant in reliable condition of Experiment 1 in Chapter 6 facing $A \rightarrow B \rightarrow C$ structure. Lines and points show instantaneous value for each potential intervention (colors) or non-intervention (black). Dashed vertical lines show participants interventions. Model takes earlier interventions and observations into consideration and predicts value of intervention choices for each 1-second window (marked by vertical white/gray shading). Parameters of the combined model based on EIG + local polynomial cost model fit to this individual. Model fit is the product of likelihoods of the chosen action or non-action in each window.

dynamics during learning so as to strike a balance between expected information gain and anticipated inferential complexity. Participants' ability to do this was presumably limited by their information processing capacity, leading to a kind of "less is more" phenomenon (Gigerenzer & Todd, 1999) in which simpler evidence was often more valuable to them even when less normatively informative. These findings contribute to our understanding of causal inference in continuous time, incorporate a new dimension to the study of human active learning and offer new directions for research into human learning.

5 Chapter 7: Unknown Future [3]

Discovering and measuring causal effects is of central interest for both individual cognition and scientific practice. Unfortunately, even with good quality experimental data and a well matched control group this can still be challenging, because genuine causal influences can take complex forms and our measurements of them are inevitably incomplete: Some effects might occur instantly and dissipate rapidly (such as from electric shocks or adrenaline injections), but others might peak later (paracetamol) grow or compound over minutes, days or years (perhaps lockdowns on covid rates, or European membership decisions on GDP). This highlights a central challenge for causal induction: To estimate the strength and direction of a novel cause, we need to decide when best to measure it. But to the extent that a treatment is truly novel, we are likely to lack the necessary mechanistic understanding to make this choice and so be forced into guesswork based on our inductive biases and whatever measurements we have. How do people make causal inferences when they only have data that is "un-prepackaged" with the future open?

I demonstrate the research design with the following scenario (Greville & Buehner, 2007): Imagine a biotechnology lab examines the effect of several types of radiation treatment on the survival of bacterial cultures. Bacterial cultures die naturally after a number of days, but the treatment might promote the survival of bacterial cultures (be beneficial) or kill them prematurely (be harmful). In the example shown in Figure 6a, are Sigma-Rays harmful or beneficial to the survival of AB-loop bacteria?

There are competing perspectives in the literature: (1) Contingency provides no straightforward answer here since both groups have experienced the same total number of deaths by the end of the observation. (2) The contiguity principle (Greville & Buehner, 2007) would predict that the treatment seems to be beneficial, potentially postponing the death of bacteria, as there are fewer deaths in the observations on days 1–3. (3) If the learner notices the trend, they might rather suspect the treatment will ultimately prove harmful since the experimental condition has a worryingly increasing trend and most of the forty samples are still alive on Day 5 (see Figure 6b). To test whether people hold these

[3] Gong, T., & Bramley, N. R. (2023b). Evidence from the future. *PsyArXiv* (In press at *Journal of Experimental Psychology: General).*

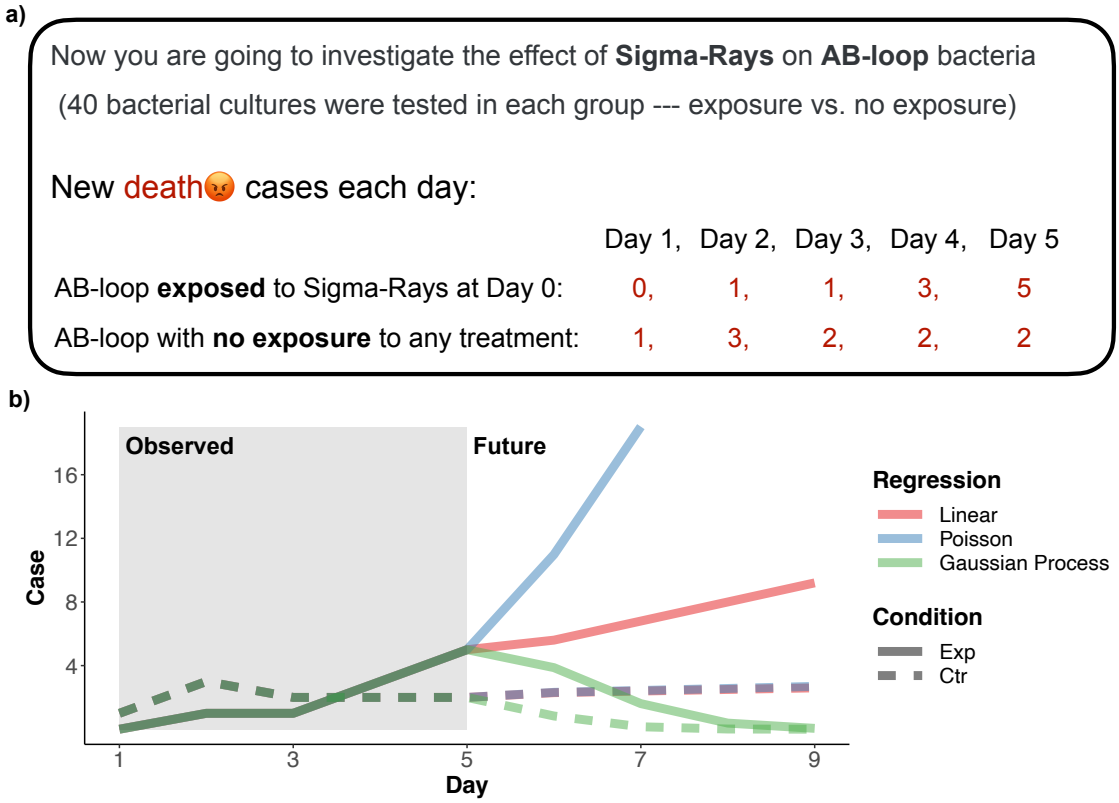

Figure 6: An example stimulus material of the study in Chapter 7 (a) and the corresponding extrapolation results of how the new case will be in the future given different regression models (b). The Poisson regression would predict the experimental case as 0 at Day 9 due to the cumulative cases have exceeded the max sample size. The Gaussian process regression was based on RBF kernel (Schulz et al., 2017).

three different mindsets and may apply different principles in different contexts, I manipulated in three experiments what participants are told about the experimenter's stopping rule (Experiment 1), the display format (Experiment 2), and the sample size (Experiment 3).

Across three experiments, I constructed trajectories in which new death cases after treatments increased or decreased over time. I found that participants robustly used the contingency information (Buehner et al., 2003; Cheng, 1997; Griffiths & Tenenbaum, 2005). Beyond this, they used the temporal information and used it in a malleable way. Participants judged a treatment to be more harmful if more samples died in the early days in the experimental condition, consistent with the contiguity principle found in previous studies (Greville & Buehner, 2007; Pacer & Griffiths, 2012). However, this only happened when participants saw the data in a static format and were either told that the observation had finished (Experiment 1) or that the total sample size was so small that they had seen the most of the potential data by day 5 (Experiment 3). In contrast, participants relied on the trend when they were informed that the observation had not ended (Experiment 1) or experienced a dynamic format where the data were revealed sequentially (Experiment 2). These effects consistently occurred regardless of whether the contingency information suggested the cause to be harmful, beneficial, or non-causal.

The findings indicate that when utilizing temporal information, people are sensitive to the wider context (here cued by the cover story, presentation format and sample size). Whether strength judgments reflected generalization beyond the data depended on the extent that the context and the available measurements implied that all the relevant causality had been captured in the provided observations. Besides the theoretical implication, one practical implication of this study is its demonstration that instructional framing influences how people interpret the data they are shown. This means that providing accurate context as well as data is vital for accurate scientific communication (Soyer & Hogarth, 2012).

6 In Summary

In sum, this thesis explores the process of human causal structure induction from events in continuous time. It builds a bridge between the normative and process levels in causal reasoning, and provide quantitative predictions about human judgments in various situations. By understanding the mechanisms behind people's rapid and efficient learning with limited resources, this thesis contributes to our understanding of natural cognition while also offering insights into the quest for more human-like algorithms. In our daily lives, we encounter not only expected or surprising events but also ponder their connections to the past and future. Incorporating a formal framework for temporal causal inference into the causal theory, would be crucial, for our understanding of human reasoning processes.

References

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*(3), 147–149.

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332.

Blum, M. G., Nunes, M. A., Prangle, D., & Sisson, S. A. (2013). A comparative review of dimension reduction methods in approximate bayesian computation. *Statistical Science*, *28*(2), 189–208.

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating bayesian inference. *Cognitive Psychology*, *74*, 35–65.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, *124*(3), 301–338.

Bramley, N. R., Gerstenberg, T., Mayrhofer, R., & Lagnado, D. A. (2018). Time in causal structure learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1880–1910.

Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1119–1140.

Buehner, M. J., & McGregor, S. (2006). Temporal delays can facilitate causal attribution: Towards a general timeframe bias in causal induction. *Thinking & Reasoning*, *12*(4), 353–378.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405.

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133.

Davis, Z., Bramley, N. R., & Rehder, B. (2020). Causal structure learning in continuous systems. *Frontiers in Psychology*, *11*, 244.

Davis, Z., & Rehder, B. (2020). A process model of causal reasoning. *Cognitive Science*, *44*(5), e12839.

Fernbach, P. M., & Sloman, S. A. (2009). Causal learning with local computations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 678–693.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, *128*(5), 936–975.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press.

Gong, T., & Bramley, N. R. (2023a). Continuous time causal structure induction with prevention and generation. *Cognition*, *240*, 105530.

Gong, T., & Bramley, N. R. (2023b). Evidence from the future. *PsyArXiv (In press at Journal of Experimental Psychology: General)*.

Gong, T., Gerstenberg, T., Mayrhofer, R., & Bramley, N. R. (2023). Active causal structure learning in continuous time. *Cognitive Psychology*, *140*, 101542.

Greville, W. J., & Buehner, M. J. (2007). The influence of temporal distributions on causal induction from tabular data. *Memory & Cognition*, *35*(3), 444–453.

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology: General*, *139*(4), 756–771.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334–384.

Hagmayer, Y., & Waldmann, M. R. (2002). How temporal assumptions influence causal judgments. *Memory & Cognition*, *30*(7), 1128–1137.

Halpern, J. Y. (2016). *Actual causation*. MIT Press.

Harman, G. (1986). *Change in view: Principles of reasoning*. The MIT Press.

Kim, N. S., & Ahn, W.-k. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General*, *131*(4), 451–476.

Lagnado, D. A., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 856–876.

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 451–460.

Lagnado, D. A., & Speekenbrink, M. (2010). The influence of delays in real-time causal learning. *The Open Psychology Journal*, *3*(1), 184–195.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, 1–60.

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review*, *115*(4), 955.

Luce, R. D. (1959). *Individual choice behavior*. Wiley.

Malthus, T. R. (1872). *An essay on the principle of population*. Reeves & Turner.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Nikolic, M., & Lagnado, D. A. (2015). There aren't plenty more fish in the sea: A causal network approach. *British Journal of Psychology*, *106*(4), 564–582.

Pacer, M., & Griffiths, T. L. (2012). Elements of a rational framework for continuous-time causal induction. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual conference of the cognitive science society* (pp. 833–838).

Pacer, M., & Griffiths, T. L. (2015). Upsetting the contingency table: Causal induction over sequences of point events. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the cognitive science society* (pp. 1805–1810).

Pearl, J. (2000). *Causality*. Cambridge University Press (2009 reprint).

Rehder, B. (2017). Reasoning with causal cycles. *Cognitive Science*, *41*, 944–1002.

Rescorla, R. A., & Wagner, A. R. (1972). A theory on pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning ii: Current theory and research* (pp. 64–99). Appleton Century Crofts.

Rottman, B. M. (2016). Searching for the best cause: Roles of mechanism beliefs, autocorrelation, and exploitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(8), 1233–1256.

Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*(1), 109–139.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*(1-2), 93–125.

Schulz, E., Tenenbaum, J. B., Duvenaud, D., Speekenbrink, M., & Gershman, S. J. (2017). Compositional inductive biases in function learning. *Cognitive Psychology*, *99*, 44–79.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgement of causality by human subjects. *The Quarterly Journal of Experimental Psychology*, *41*(2), 139–159.

Simon, H. A. (1982). *Models of bounded rationality: Empirically grounded economic reason*. MIT press.

Sloman, S. A., Love, B. C., & Ahn, W.-K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*(2), 189–228.

Soyer, E., & Hogarth, R. M. (2012). The illusion of predictability: How regression statistics mislead experts. *International Journal of Forecasting*, *28*(3), 695–711.

Stephan, S., Mayrhofer, R., & Waldmann, M. R. (2020). Time and singular causation–a computational model. *Cognitive Science*, *44*(7), e12871.

Sunnåker, M., Busetto, A. G., Numminen, E., Corander, J., Foll, M., & Dessimoz, C. (2013). Approximate bayesian computation. *PLoS Computational Biology*, *9*(1), e1002803.

Ullman, T. D., Stuhlmüller, A., Goodman, N. D., & Tenenbaum, J. B. (2018). Learning physical parameters from dynamic scenes. *Cognitive Psychology*, *104*, 57–82.

Valentin, S., Bramley, N. R., & Lucas, C. G. (2022). Discovering common hidden causes in sequences of events. *Computational Brain & Behavior*, 1–23.