Causal analysis of rule-based models by counterfactual reasoning

Jonathan Laurent[1], Jean Yang[1], and Walter Fontana[2]

[1]Carnegie Mellon University
[2]Harvard Medical School

Introduction

Rule-based modeling languages for molecular biology, such as Kappa [3] and BioNetGen [7], or organic chemistry, such as Mød [1], can be used to write mechanistic models of complex reaction systems. In these approaches, chemical transformations are represented by local graph-rewrite rules equipped with stochastic firing rates. In a dynamical simulation, rules induce a time series of events that might reach a state of interest in processes like the assembly of a molecular machine, the activation of a transcription factor, or the synthesis of a specific chemical compound. While rule-based models provide compactness, transparency, and the ability of handling combinatorial complexity, the perhaps most significant advantage lies in their suitability for causal analysis that takes into account the logically concurrent nature of interactions. The causal analysis [5, 3] of event series generated by such models provides a formal definition of "pathway" and a means for revealing the emergence of pathways from low-level interactions. These methods take advantage of rule structure to (i) compress a given simulation trace into a minimal subset of events that are necessary and jointly sufficient to replicate a phenomenon of interest and (ii) highlight the direct causal influences between events, exposing the extent of concurrency.

We propose a distinct but complementary approach based on *counterfactual reasoning* that improves causal explanations by (i) being more sensitive to kinetics and (ii) properly accounting for the causal impact from inhibition between events.

1 Motivating example

We illustrate the need for counterfactual reasoning on a toy example in Kappa. Consider a model with two types of agents, kinases K and substrates S, interacting according to the rules depicted in Figure 1.

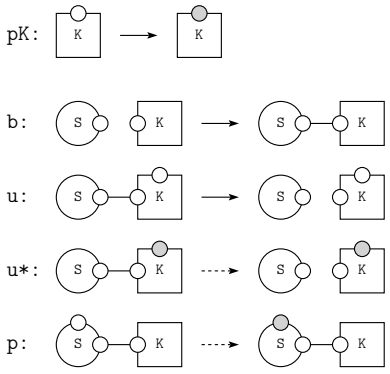

Figure 1: A motivating toy model. As usual in Kappa, sites not mentioned in a rule are left unchanged by it. Instead of naming sites, we here identify them by their position on an agent. Phosphorylated sites are shown in gray. *Slow* reactions are indicated by dotted arrows.

For the sake of simplicity, consider an initial mixture I with only a single kinase and a single substrate whose sites are free and unphosphorylated. We then ask: Starting from I, **how** is p triggered ? We are not merely looking for an account of reachability but rather for causal narratives, that is, collections of necessary events connected by causal influences.

A stochastic simulation [4] might produce the following trace (events are labelled by the rules that

induced them):

$$b, \quad u, \quad pK, \quad b, \quad p, \quad u^*, \quad \cdots$$

Figure 2 depicts the causal narrative explaining the occurrence of p according to existing techniques [5, 3]. The arrow between b and p is called an *activation arrow*, meaning that b modifies an aspect of state (by creating a link) that enables p to happen.

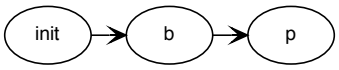

Figure 2: A causal explanation for p.

This narrative, however, is blind to the critical role of pK in the original trace. Looking at the rules in Figure 1 one notes that: (i) the phosphorylation rule p is slow (ii) the average time K and S remain bound depends on whether K is phosphorylated, as manifest in the two unbinding rules u (fast, if K is not phosphorylated) and u^* (slow, if K is phosphorylated). It seems reasonable to assert that p would probably not have happened had pK not happened, as the opportunity for p would have been cut short by a fast unbinding event. We therefore argue that pK, although it does not activate b or p directly, should be part of a causal narrative for p. Reasoning of this kind is *counterfactual* and can be deployed to define causality [9, 10].

In section 2, we give a rigorous semantics to this line of reasoning. In section 3 we show that counterfactual statements can be expressed using inhibition arrows, leading to the explanation shown in Figure 3.

2 Counterfactual simulation

Counterfactual statements are tricky because their truth is context-dependent. The statement "Had it not rained, the driver might have arrived earlier" can fail to be true in many ways. Intuitively, how easily the driver could have arrived earlier depends on how great a departure from actuality is required for it to be case [11]. This is why counterfactual reasoning is tied to modal logic. The standard approach is to require that the consequent in a counterfactual be true in some of those possible worlds (in which the antecedent holds) that *are most similar to the actual world*. If the counterfactual statement is true in all of these worlds, we can replace "might" with "would". We now operationalize this approach in the context of Kappa traces and interpolate between "might" and "would" using probabilities.

We start by formalizing the notion of an intervention. An intervention ι ("blocking pK" in our example) is a predicate $\mathsf{blocked}_\iota[t, e]$ that determines whether or not event e is blocked at time t. Given a predicate φ over traces, we write the proposition *"Had intervention ι happened in trace τ, φ would have been true with probability greater than $p \in [0, 1]$"* as:

$$\tau \models_p [\iota]\, \varphi.$$

To give an operational meaning to this statement, we invoke the continuous time Markov chain (CTMC) semantics of a Kappa model as defined and implemented in [4, 2]. For the present purpose it is conceptually useful to think of a CTMC abstractly in terms of the random realization of "potential events". A potential event is a pair (r, ξ) where r is a rule and ξ an injective mapping from local agents involved in r to global agents in a huge virtual mixture of many instances of all possible molecular species.[1] For every such potential event, we imagine a bell that rings at a time t drawn from an exponential distribution $\lambda_r \exp(-\lambda_r t)$, where λ_r is the stochastic rate constant of r. A simulation trace can be viewed as the realization of a random variable T determined by the set ω of ring times: Starting with an initial mixture, when a bell rings at t, its associated potential event (r, ξ) transforms the mixture according to r if ξ yields a valid embedding of the left hand side of r in the current mixture and time advances by t. Otherwise, time advances and nothing happens—a null event. Repeat on the resulting mixture.

We can extend this viewpoint to include interventions. For an intervention ι, we define the random variable $\hat{T}_\iota$ much in the same way as T, except that each time the bell rings, we require $\mathsf{blocked}_\iota[t, e]$ to be false for the potential event $e = (r, \xi)$ to be considered.

[1] For the sake of simplicity, we assume that no agent is created or deleted by a rule.

Counterfactual traces that are closest to the actual trace τ are then sampled by generating realizations of $\hat{T}_\iota$ that inherit, whenever possible, the subset of ω that made up τ. An efficient implementation of this specification for sampling the conditional random variable $\hat{T}_\iota \,|\, (T = \tau)$ is available at `https://github.com/jonathan-laurent/kappa-counterfactuals`. We refer to this natural extension of CTMC semantics as *counterfactual re-simulation* or *co-simulation* for short.

Using co-simulation, we operationalize the counterfactual statement $\tau \models_p [\iota]\, \varphi$ as:

$$\tau \models_p [\iota]\, \varphi \quad \Leftrightarrow \quad \mathbf{P}(\varphi(\hat{T}_\iota) \mid T = \tau) \geq p$$

3 Inhibition arrows

Returning to our example, we can use co-simulation to quantify the influence of pK on p by estimating the probability of p happening had pK not occurred. However, we can go further by using counterfactual traces to *explain* this influence using both activation and inhibition arrows (Figure 3).

Activation arrows are easy to define and identify in a trace. We say that an event e activates e' if e is the last event before e' that modifies some site to the value it is tested for by e'. Inhibition arrows are trickier because they must relate events that happened to events that did not. We use counterfactual traces to give a rigorous account of inhibition using arrows that connect events from the factual trace to events in the counterfactual trace and vice versa.

A *counterfactual experiment* is any triple (τ, ι, τ') for which there exists a random realization ω such that $\tau = T(\omega)$ and $\tau' = \hat{T}_\iota(\omega)$. Such triples are produced by co-simulation. Then, an event e that occurs at time t in τ is said to inhibit an event e' that occurs at time t' in τ' if all of the following hold: (i) $t < t'$ (ii) there exists a site s such that e is the last event in τ before t' that modifies the value of s away from what e' tests it for (iii) there are no events in τ' that modify s during the time interval (t, t').

Figure 3 shows the influence of pK on p based on a counterfactual experiment. Dotted nodes correspond to events proper to the counterfactual trace τ', thick nodes to events proper to the factual trace τ, and the remaining nodes correspond to events common to both traces. Activation arrows are depicted in black and inhibition arrows in red.

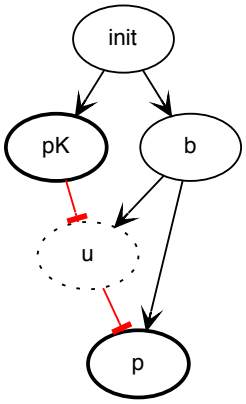

Figure 3: A better causal explanation for p. It is based on the counterfactual experiment (τ, ι, τ') where $\tau = (pK, b, p)$, ι blocks pK and $\tau' = (b, u)$.

The example illustrates the influence of pK on p mediated by the counterfactual event u. Such mediating events always exist, as stated by the following theorem.

Theorem. *Let (τ, ι, τ') be a counterfactual experiment and e an event that belongs only to τ. Then there exists an event $e_0 \in \tau$ that is blocked by ι and there is a path from e_0 to e with an even number of inhibition arrows.*

Conclusion

We have proposed a new way to generate causal explanations in Kappa based on counterfactual reasoning, in which causal explanation are augmented by including inhibition between events. By leveraging the Kappa stochastic simulator for co-simulation, we expect this technique to increase the sensitivity of explanatory accounts to kinetics.

Future work will investigate how this approach interacts with trace compression [5] and establish heuristics to determine which counterfactual interventions are worth attempting on a given reference trace.

References

[1] Jakob L. Andersen, Christoph Flamm, Daniel Merkle, and Peter F. Stadler. A Software Package for Chemically Inspired Graph Transformation. *Lecture Notes in Computer Science*, 9761:73–88, 2016.

[2] Pierre Boutillier, Thomas Ehrhard, and Jean Krivine. Incremental update for graph rewriting. In *Programming Languages and Systems - 26th European Symposium on Programming, ESOP 2017*, pages 201–228, 2017.

[3] V. Danos, J. Feret, W. Fontana, R. Harmer, and J. Krivine. Rule-Based Modelling of Cellular Signalling, invited paper. In L. Caires and V.T. Vasconcelos, editors, *Proceedings of the Eighteenth International Conference on Concurrency Theory, CONCUR 2007, Lisbon, Portugal*, volume 4703 of *Lecture Notes in Computer Science*, pages 17–41, Lisbon, Portugal, 2007. Springer, Berlin, Germany.

[4] V. Danos, J. Feret, W. Fontana, and J. Krivine. Scalable Simulation of Cellular Signaling Networks, invited paper. In Z. Shao, editor, *Proceedings of the Fifth Asian Symposium on Programming Systems, APLAS 2007, Singapore*, volume 4807 of *Lecture Notes in Computer Science*, pages 139–157, Singapore, 2007. Springer, Berlin, Germany.

[5] Vincent Danos, Jérôme Feret, Walter Fontana, Russell Harmer, Jonathan Hayman, Jean Krivine, Christopher D Thompson-Walsh, and Glynn Winskel. Graphs, Rewriting and Pathway Reconstruction for Rule-Based Models. In Deepak D'Souza, Telikepalli Kavitha, and Jaikumar Radhakrishnan, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2012*, volume 18 of *LIPIcs*, pages 276–288. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2012.

[6] JY Halpern and J Pearl. Causes and explanations: A structural-model approach. Part I: Causes. *The British journal for the philosophy of*, 2005.

[7] Leonard A. Harris, Justin S. Hogg, José-Juan Tapia, John A. P. Sekar, Sanjana Gupta, Ilya Korsunsky, Arshi Arora, Dipak Barua, Robert P. Sheehan, and James R. Faeder. BioNetGen 2.2: Advances in rule-based modeling. *Bioinformatics*, 32(21):3366—-3368, 2016.

[8] Florian Leitner-Fischer and Stefan Leue. Causality Checking for Complex System Models. In *VMCAI*, volume 7737, pages 248–267, 2013.

[9] David Lewis. Causation. *The journal of philosophy*, 70(17):556–567, 1974.

[10] David Lewis. Causation as influence. *The Journal of Philosophy*, 97(4):182–197, 2000.

[11] David K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, 1973.