Mind & Society (2019) 18:43–56
https://doi.org/10.1007/s11299-019-00216-1

Scientific discovery, causal explanation, and process model induction

Pat Langley[1]

Published online: 30 May 2019

Abstract
In this paper, I review two related lines of computational research: discovery of scientific knowledge and causal models of scientific phenomena. I also report research on quantitative process models that falls at the intersection of these two themes. This framework represents models as a set of interacting processes, each with associated differential equations that express influences among variables. Simulating such a quantitative process model produces trajectories for variables over time that one can compare to observations. Background knowledge about candidate processes enables search through the space of model structures and associated parameters to find explanations of time-series data. I discuss the representation of such process models, their use for prediction and explanation, and their discovery through heuristic search, along with their interpretation as causal accounts of dynamic behavior.

Keywords Scientific discovery · Process models · Causal models · Explanation

1 Computational scientific discovery

Science is one of civilization's crowning achievements, which makes the cognitive processes that underlie it eminently worthy of study. The scientific enterprise differs from other intellectual endeavors not only by relying on formal theories, laws, and models to explain and predict observations, but also by using such observations to construct, revise, and evaluate its formal accounts. Together, these activities produce a closed loop that drives research communities toward better understanding of their chosen phenomena. Many of these activities have been studied by philosophers of science for over a century, but most of their efforts focused on the 'logic of justification', which characterized how data can support or refute laws, models, or theories. In contrast, the community largely

✉ Pat Langley
patrick.w.langley@gmail.com

[1] Institute for the Study of Learning and Expertise, 2164 Staunton Court, Palo Alto, CA 94306, USA

avoided the 'logic of discovery', which would explain how observations could lead to their creation.

A common claim was that discovery required some 'creative spark', which might plausibly be studied by psychologists but which could never be analyzed in rational or logical terms. For example, Popper (1961) wrote:

> The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it…My view may be expressed by saying that every discovery contains an irrational element, or creative intuition…

He was not alone in taking this position; Hempel (1966) and many other philosophers of science also maintained that discovery was inherently irrational and beyond formal understanding. Nevertheless, scientific discovery remained a fascinating topic, and a deeper understanding of its mechanisms would have important implications, both theoretical and practical.

Fortunately, advances made during the 1950s in two other fields—cognitive psychology and artificial intelligence—suggested a way forward. The key insight came from Herbert Simon, a co-founder of AI who also developed some of the earliest computer models of human thinking. Simon (1966) proposed that scientific discovery, rather than depending on some unknown mystical ability, is actually a variety of problem solving that involves search through a space of problem states generated by applying mental operators and guided by heuristics to make search tractable. Heuristic search had already been implicated in many cases of human cognition, such as proving theorems and playing chess. At the time, this was a radical idea and few researchers in philosophy, psychology, or computing were willing to adopt it. Yet it was fully consistent with empirical studies of problem solving in humans, and there were already many AI systems that had demonstrated the power of heuristic search. Yet only Simon had the interdisciplinary expertise and the audacious research style to propose applying this computational framework to discovery of scientific knowledge.

Moreover, heuristic search offered not only a path to understanding how scientific discovery has operated historically, but also ways that we might automate this creative process. For my dissertation research with Simon, I adapted his ideas in this area to develop a computer program (Langley 1979, 1981) that:

- Carried out search in a problem space of theoretical terms;
- Using operators that combined old terms into new ones;
- Guided by heuristics that noted regularities in data; and
- Applied these recursively to formulate higher-level relations.

I named the system *Bacon*, after Sir Francis Bacon, the early champion of empirical scientific method, because it adopted a data-driven approach to discovery. Computational experiments showed that the system could rediscover a variety of laws from the history of physics, including Kepler's third law of planetary motion, the ideal gas law, Coulomb's law, and Ohm's law. Extended versions

found additional numeric relations like Snell's law of refraction, Archimedes' law of displacement, conservation of momentum, and Black's law of specific heat (Bradshaw et al. 1980), as well as results from early chemistry (Langley et al. 1983).

Bacon's success inspired the creation other many AI systems for discovering numeric laws. Researchers who reported results in this area included Falkenhainer and Michalski (1986), Kokar (1986), Żytkow et al. (1990), Schaffer (1990), Nordhausen and Langley (1990), Moulet (1992), Gordon et al. (1994), Murata et al. (1994), Džeroski and Todorovski (1995), Washio and Motoda (1997), Bradley et al. (2001), Koza et al. (2001), and, more recently, Schmidt and Lipson (2009). These relied on different methods but also induced explicit mathematical laws from observations. This task is sometimes referred to as *equation discovery*.

But the endeavor to understand discovery did stop there. Inspired by earlier work on the DENDRAL system (Feigenbaum et al. 1971), which inferred structural models of organic molecules from their mass spectra, the community developed additional systems that created models of other scientific phenomena (e.g., Langley et al. 1987; Valdés-Pérez 1996). The broader movement has come to be known as *computational scientific discovery*, and it has produced both accounts of many historical breakthroughs and novel results that have contributed to the scientific literature. Efforts in this paradigm differ from mainstream work in data mining and machine learning by producing content stated in established scientific formalisms (Langley et al. 2002a, b), ranging from componential models in particle physics (Kocabas 1991) to reaction pathways in chemistry (Valdés-Pérez 1994) to regulatory models in genetics (King et al. 2004). Collections edited by Shrager and Langley (1990) and by Džeroski and Todorovski (2007) present snapshots of the field at different points in its trajectory.

2 Causal models in science

An important recurring theme in science is the notion of a *causal model*. Intutively, we can say that a *variable x causally influences a variable y if a change in x's value produces a change in y's value provided other variables are held constant*. This definition of causality does not state that x is the only causal influence on y, specify the functional form of the relation, or state whether the effect is deterministic or stochastic, but it agrees with most scientists' understanding of the term. A given variable may causally influence one or more one other terms and it may be causally impacted by one or more other variables. An example of the latter is the equation $d = 2xy + 3z$, where x, y, and z each exert a causal influence on d. Such a relation may or may not describe reality, but we can still state a hypothesis that it holds.

Many of the numeric laws produced by equation discovery systems have causal interpretations, but this has not been an emphasis of research in the area. Moreover, it has focused on finding individual equations rather than collections of linked expressions that constitute full scientific *models*. The latter are a collection of causal hypothesizes that connect a set of variables, some that play the role of dependent terms in certain relations and independent factors in others. *Exogenous* terms exert

a causal influence on other variables but are not so affected themselves, whereas *endogenous* terms are causally influenced by one or more others, although they may also affect different variables. We can say that a causal model *explains* variations in observed values for endogenous variables in terms of other variables' values.

Herbert Simon and his colleagues also did insightful work on the nature of causal models. For instance, Iwasaki and Simon (1986) indentified the importance of causal ordering in accounts of device behavior. They clarified that it is not enough for a model to specify equations that relate variables; they must also state which terms serve as causes and which serve as effects. Iwasaki and Simon (1994) extended this analysis to sets of differential equations that characterize a system's dynamic behavior, as well as models that abstract away from details by aggregating over subsets of variables. These results dealt with the representation of causal models and ways of reasoning over them qualitatively, but Simon (1954) also made earlier groundbreaking contributions to their induction. He determined the statistical conditions—based on relations between partial corrections—under which one can infer the presence or absence of causal links between variables, and in some cases their direction, based on purely observational data, with no experimental control.

Decades later, Glymour et al. (1987) extended Simon's framework to handle more complex relations among partial correlations and incorporated the results into a computer system that discovered causal models from nonexperimental data. Early techniques focused on identifying structural equation models—sets of linear causal equations—but more recent results have generalized to broader types of accounts. A parallel line of research has dealt with induction of Bayesian networks (Darwiche 2009), some but not all with causal interpretations. Both frameworks have seen wide use in the sciences, with gene regulation being an especially popular application (e.g., Bay et al. 2003). These efforts also count as important contributions to the literature on computational scientific discovery.

3 Rate-based process models

One drawback of most computational research on causal models is that the formalisms take a very abstract form and make little contact with established scientific concepts. This holds for both structural equation models and Bayesian networks, two frameworks that have been widely adopted. An important excepton is Forbus' (1984) work on *qualitative process models*, in which each process specifies a set of causal variables that influence one or more effect variables, which are often derivatives. Each such influence is qualitative, specifying only whether an increase in the causal term leads the effect term to increase or decrease. This makes them similar to mental models that people use to reason informally about physical situations, but they are not detailed enough for many scientific settings.

Scientists and engineers in many disciplines state their models in terms of differential equations that describe changes in variables as functions of current values. These provide a formal specification of how variables interact over time, and computational researchers have developed numerous algorithms for simulating sets of differential equations both accurately and efficiently. They have used these techniques

to develop and simulate dynamic models of phenomena that range from weather patterns and aircraft turbulence, which involve a few equations repeated many times, to ecology and biochemistry, which involve a diverse set of relations among many variables. Differential equations have strong mathematical underpinnings and efficient computational methods for simulating their behavior, but they do not describe the *context* that surrounds many scientific models, they do not provide underlying *explanations* for observations, and they provide few constraints on model *discovery*. This suggests the need for an alternative framework that addresses these issues.

Fortunately, we can combine this approach with key ideas from Forbus' framework. Explanations in science and engineering are often stated in terms of *processes*, but only in informal terms, using natural language, and his notation treats processes as first-class structures. Bridewell et al. (2008) showed that one can restate differential equation models in terms of processes that govern system dynamics and variables that are changed by those processes. Variables describe the current state of the system, whereas processes relate these variables and specify how they evolve over time. This content imposes qualitative structure on the quantitative aspects of the model, which has advantages that will become apparent shortly. In this framework, a *process model* is a collection of linked processes and variables that explain the behavior of a dynamic system. For example, a model of aquatic population dynamics might include variables for phytoplankton and nitrogen concentrations, along with processes for nutrient absorption and remineralization. Each process specifies a fragment of one or more algebraic or differential equations. Taken together, a set of processes constitute an operational model that remains modular and interpretable.

In recent work, Langley and Arvay (2015) introduced three further constraints on the structure of process models:

- Each process specifies the *rate* of change for one or more dependent variables;
- Each rate[1] is determined by an algebraic combination of independent variables;
- The derivative of each dependent variable is proportional to the process's rate.

Together, these assumptions impose much stronger limits on the space of candidate models that should improve the ability to interpret and discover them. Chemical reactions offer a good example of such processes. A given reaction always involves the same substances, but its rate of operation can vary over time. This rate is an algebraic function of the concentrations of chemicals involved. For instance, consider a chemical reaction that combines two substances, with concentrations C_1 and C_2, to produce a third chemical as output, which has the rate $C_1 \times C_2$. Each reaction has one or more associated derivatives of the form $dC/dt = k \times R$, where R is the rate and k is a constant parameter. Our example would involve three such derivatives, two for the input substances, with negative coefficients, and one for the output substance, with a positive coefficient. However, this does not mean that these variables

[1] This rate is always positive and its values are inherently unobservable, so we can adopt any measurement units that we find convenient.

Table 1 (a) A rate-based process model for an aquatic ecosystem that relates concentrations of phytoplankton, nitrogen, and detritus. Each process specifies a rate expression and a set of derivatives proportional to this rate, which changes over time. (b) A set of linked differential equations that produce the same dynamic behavior as the process model

(a)

organism_loss[phyto, detritus]		
	rate	$r = phyto$
	equations	$d[phyto, t] = -0.05 * r$
		$d[detritus, t] = 0.05 * r$
nutrient_uptake[phyto, nitro]		
	rate	$r = phyto * nitro$
	equations	$d[phyto, t] = 0.5 * r$
		$d[nitro, t] = -0.005 * r$
remineralization[detritus, nitro]		
	rate	$r = detritus$
	equations	$d[detritus, t] = -0.04 * r$
		$d[nitro, t] = 0.04 * r$

(b)

$$d[phyto, t] = -0.05 * phyto + 0.5 * phyto * nitro$$
$$d[nitro, t] = -0.005 * phyto * nitro + 0.04 * detritus$$
$$d[detritus, t] = 0.05 * phyto + -0.04 * detritus$$

always increase or decrease over time, as other processes may also influence their values.

Table 1(a) shows a simple model for an aquatic ecosystem with three variables: phytoplankton, nitrogen, and detritus. This includes three distinct processes, one for phytoplankton loss, one for uptake of nitrogen by phytoplankton, and another for remineralization of nitrogen from detritus. The variables *phyto* and *nitro* refer to the concentrations of phytoplankton and nitrogen, respectively. Each process has an associated rate expression, one specifying that the rate equals the product of two variables and the others stating that it equals a single variable. Each process also includes two associated derivatives that are proportional to the rate, with parameters detailing this functional dependence. Table 1(b) translates these processes into a set of differential equations, one per variable, with each term on the right-hand sides mapping onto an equation fragment in some process. The two notations produce the same dynamic behavior, but the first one has a higher-level organization.

4 Using rate-based process models

We can use such a quantitative process model by compiling it into a set of differential equations. For each endogenous variable *v*, one collects all equation fragments from processes in which *v* appears on the left-hand side. The differential equation for that variable has the sum of these fragments as its right-hand side. This produces a set of linked equations that one can provide to a standard differential equation solver like CVODE (Cohen and Hindmarsh 1996) to simulate the behavior of each variable over time. For this purpose, we must provide not

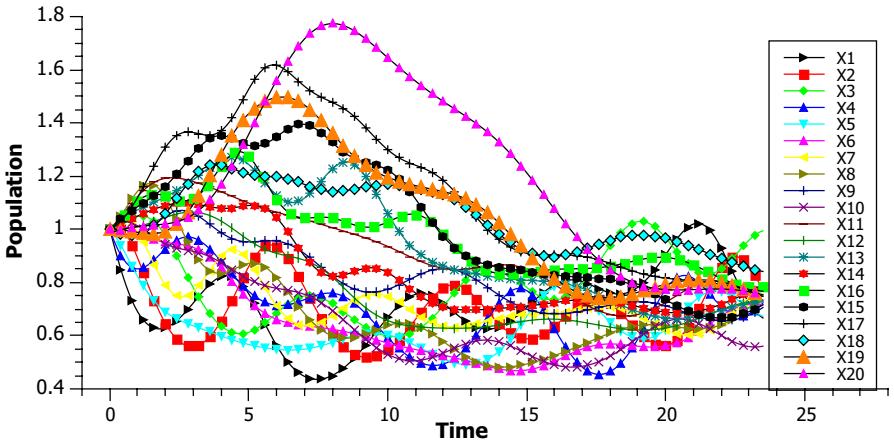

Fig. 1 Trajectories for a 20-variable predator–prey system in which the organisms are organized in a linear food chain. Each entity participates in two distinct processes. Taken from Arvay and Langley (2016)

only the parameters for each equation, but also the initial values for each endogenous variable. For models that include exogenous terms reflecting influences of the external environment, one must also specify their values on each time step, as they cannot be simulated.

There are two distinct but related uses for such simulations. The first involves predicting trajectories of future values for each endogenous variables. This is directly analogous to standard applications of both differential equations and statistical time-series models. Here the metric for success is the degree to which the values predicted for each variable match the values that are observed later. For example, Fig. 1 plots the simulated populations for a 20-variable predator–prey system generated by the differential equations that result from compiling a model with 21 distinct processes. These included 19 predation relations between pairs of organisms with the same rate expressions but different coefficients. Two other processes involved reproduction for the organism lowest on the food chain and loss for the one at the top of the chain. Some of the trajectories involve cyclic behavior, but nothing as regular as observed in classic two-organism systems. This complexity results from the fact that most variables take part in two feedback loops whose activity is influenced by others up or down the chain.

The second use of process models is more distinctive, as they provide an *explanation* of time-series data not only in terms of hidden terms like rates, but because they refer to unobserved processes that are familiar to domain specialists. Processes provide an overlay on standard differential-equation models that specify additional context about which terms must appear together. For example, we know that a predation process causes the population of the predator to increase and the population of the prey to decrease. Thus, the 21-process model mentioned earlier does more than generate the trajectories in Fig. 1; it also explains them in terms of such familiar relations. As we will see, this knowledge also provides important constraints for the discovery of such models.

5 Discovering rate-based process models

Although we can create quantitative process models manually, they also lend themselves to automated construction by computer. We will refer to the task of constructing such an account from data and knowledge as *inductive process modeling*. Observations take the form of time series for a subset of model variables, while knowledge takes the form of *generic* processes that specify possible relations among entities in terms of functional forms with bounded parameters. The output is a ranked list of interpretable models—collections of processes and associated differential equations—that explain the observations and make predictions about unseen data. Figure 2 depicts this computational discovery task in graphical terms.

Typical data sets for inductive process modeling are observational rather than experimental in character. One is provided with a multivariate trajectory over time, such as a scientist might collect for an aquatic ecosystem. Most variables (e.g., organism and nutrient concentrations) are endogenous, in that they are targets for prediction and explanation. The data set may also include exogenous variables (e.g., temperature and sunlight) that change over time and that can influence other measures, but that are not themselves targets for prediction. In most cases, only one trajectory is available and experimental control of variables is not possible. Table 2 presents three generic processes that underlie the model in Table 1(a). These take roughly the same form as concrete processes, but they refer to types of variables (e.g., organisms) rather than to specific ones (e.g., phytoplankton) and they contain functional forms that relate variables rather than particular equations. The latter do not specify coefficients but instead refer to parameters and constraints on their values. Generic processes serve as building blocks from which to construct quantitative process models.

Early methods for inductive process modeling (e.g., Bridewell et al. 2008) carried out search at two distinct levels. The first explored the space of model structures.

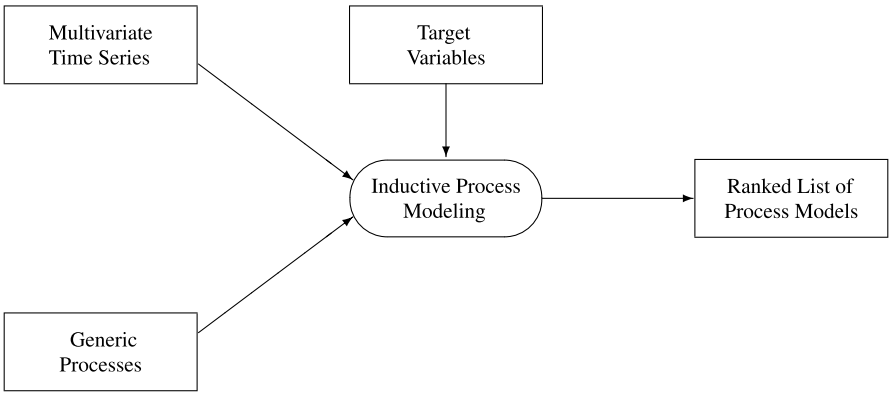

Fig. 2 The task of inductive process modeling, which is given a set of continuous variables, observed time series for their values, and generic domain knowledge about candidate processes. The result is a ranked list of parameterized model structures that fit the observations and explain them in terms of inferred processes

Table 2 Three generic processes for aquatic ecosystems with type information about the variables involved in each process and bounds on their parameter values

organism_loss[S, D]	
types	$S\{organism\}, D\{detritus\}$
parameters	a $[0, 1]$
rate	$r = S$
equations	$d[S, t, 1] = a * r$
	$d[D, t, 1] = -1 * a * r$
nutrient_uptake[S, N]	
types	$S\{organism\}, N\{nutrient\}$
parameters	b $[0, 1]$, c $[-1, 0]$
rate	$r = S * N$
equations	$d[S, t, 1] = b * r$
	$d[N, t, 1] = c * r$
remineralization[N, D]	
types	$N\{nutrient\}, D\{detritus\}$
parameters	e $[0, 1]$
rate	$r = D$
equations	$d[N, t, 1] = e * r$
	$d[D, t, 1] = -1 * e * r$

This involved finding all ways to instantiate the known generic processes with specific entities and variables, which provided elements for candidate models. Search started with the empty model and, on each step, added a new process instance provided the result did not exceed a user-specified maximum. For each such model structure, a second procedure carried out search through a parameter space. This started from values sampled at random from within the parameters' ranges and used gradient descent to converge on values, complemented by random restarts to guard against finding local optima. The objective function that guided search was the accuracy of a model's simulated trajectory against observations. The output was a list of parameterized model structures ranked by error on the training data.

This approach to process model induction has been applied successfully in a variety of settings. These have included fields as diverse as ecology, hydrology, and biochemistry (Asgharbeygi et al. 2006; Bridewell et al. 2008; Langley et al. 2006). Extensions to the basic framework have included adding the ability to organize process knowledge hierarchically (Todorovski et al. 2005), handle missing data using iterative optimization (Bridewell et al. 2006), reduce overfitting by combining models into ensembles (Bridewell et al. 2005), find models that involve partial differential equations (Park et al. 2010), draw on constraints to limit search (Bridewell and Langley 2010), and even learn these constraints from the performance of sample models (Todorovski et al. 2012). Each extension improved either the accuracy of induced models or the efficiency of discovering them.

Despite this progress, early methods for process model induction suffered from three important drawbacks. First, they evaluated only complete model structures by comparing a parameterized model's predictions with observed time series. This approach did not support heuristic search through the space of model structures and thus did not scale, as the number of structures is an exponential function of the number of variables and the number of generic processes. Second, they relied on

repeated simulation of each model structure to estimate its parameters. They initialized each parameter randomly, simulated parameterized model, calculated a gradient based on the resulting error, and iterated until convergence. Hundreds of iterations were not uncommon, so that simulations for parameter estimation often took 99.99% of the CPU time. Third, to avoid halting at local optima, they required repeated restarts from different initial random values. This further added to the computational burden and, again, it was borne separately for each distinct model structure.

More recent approaches to inductive process modeling instead carry out heuristic search through the space of model structures. To this end, they take advantage of the decomposition of processes into rate expressions and proportional equations. They also make two additional assumptions: that rate expressions are parameter free and that they include only observable variables. As before, they first enumerate candidate processes by instantiating the generic processes in all possible ways that satisfy type constraints. Next they use the observed time series to calculate rates for each such process instance. The assumption that rates are parameter-free algebraic functions of observed variables lets them make this calculation for each candidate process instance on each time step. The result is a trajectory of the rate for each such process. After this, we estimate the derivatives for each dependent variable X on each time step t as the average of $X(t) - X(t-1)$ and $X(t+1) - X(t)$, using what is commonly known as the 'center difference' method.

A heuristic search mechnism now iterates through the dependent variables. For each one, this attempts to find a differential equation that predicts its observed changes as a linear combination of process rates. For each variable D, it considers as candidate predictors only process instances that include D as one of their dependent terms. One can invoke multiple linear regression to find candidate equations for each dependent term, but this requires a set of rate expressions as independent factors. Langley and Arvay's (2015) RPM system first considers individual rates, then pairs, and so on until it finds an equation with acceptable fit. Arvay and Langley (2016) describe a more tractable approach that uses repeated feature selection to induce more complex differential equations. Both systems ensure that, if a rate expression appears in the equation for one dependent variable (e.g., a predator), it also appears in those for other derivatives in that process (e.g., the prey). Equations must also satisfy parameter constraints (e.g., negative for inputs to a chemical reaction and positive for outputs).

Together, these constraints reduce substantially the number of model structures that are entertained during the discovery effort. They make the difference between tractable heuristic search and the intractable exhaustive methods used in previous systems. For instance, Langley and Arvay (2015) report that RPM ran 83,000 more rapidly on a three-process modeling task than an earlier system that carried out exhaustive search through the structure space, yet it still reconstructed the target model far more reliably than its predecessor. Successors have induced accurate and plausible models for data on aquatic ecosystems, an eight-reaction chemical pathway, and the 20-organism predator–prey system shown in Fig. 2. Experiments with synthetic time series have also revealed that they scale well with the number of variables and the number of processes in the model.

6 Causal interpretation of process modeling

Now that I have reviewed the paradigm of inductive process modeling, I can examine its relationship to causal accounts in science. I will organize the discussion into three subtopics that parallel previous sections of the paper. The first concerns the causal status of rate-based processes and the models comprising them. The second issue addresses the forms of causal reasoning that such process models support. The final one deals with the ability to discover these models from background knowledge and observations of dynamic systems.

Earlier, I defined a causal model as a collection of hypotheses, each of which makes a claim about how changes in the values of some variables influence those of another. Clearly, rate-based process models satisfy this definition. Each process encodes a set of hypothesized influences that relate causal variables to effect variables. For example, the first process in Table 1, for nitrogen uptake, specifies four causal influences, two for effects of nitrate concentration and two for effects of detritus concentration on the derivatives of each variable. However, process models are more constrained than classic causal accounts, such as sets of structural equations or some Bayesian networks, because they make stronger representational assumptions. They also deal with temporal phenomena, making them related to dynamic Bayesian networks, but they encode derivatives rather than values on successive time steps. A more important difference is that effects are mediated through a rate term that can be a function of multiple factors and can influence one or more derivatives. In other words, processes combine multiple causal links that must stand or fall together.

Process models also support two important types of causal reasoning. First, they allow counterfactual thought experiments because one can alter the initial values for some variables and use simulation to determine how the trajectories would differ. Second, they provide a causal explanation of observed behavior in terms of familiar domain processes. Traditional causal frameworks, like structural equation models, Bayesian networks, and linked differential equations, also offer explanations, but they provide more shallow accounts. For instance, one can characterize a reaction pathway as a set of ordinary differential equations, but this account would not refer explicitly to the reactions, which are central concepts in chemistry. Other examples come from the field of ecology, which uses differential equation models but which also postulates processes like predation and nutrient absorption that, like reactions, have multiple effects on participating variables.

These techniques search a space of candidate process models and attempt to discover ones that fit the data and explain them in causal terms. I will not claim that these are the only models consistent with the trajectories; in fact, Arvay and Langley (2016) report cases in which multiple process models reproduce the data equally well. Neither will I claim that the discovered models are correct in the sense that they are guaranteed to generalize well, but this is a common limitation of induction. This is especially challenging given observational rather than experimental data. Even Simon's (1954) arguments for inference of causal relations in linear models, and their extension in Glymour et al.'s (1987) work, are statistical

in character and offer no ultimate guarantees. However, this does not negate the fact that the computational methods I have described can discover plausible and interesting causal explanations.

7 Concluding remarks

In this paper, I reviewed research on computational approaches to both scientific discovery and causal modeling. I also introduced the notion of quantitative process models that account for multivariate time series in terms of a set of component processes. Each such element specifies an algebraic rate expression and one or more derivatives that are proportional to this rate. A derivative can be influenced by multiple processes and a variable can participate in different rate expressions. One can compile a process model into a set of linked differential equations, and then simulate it to predict future trajectories and explain available observations. Moreover, this notation directly supports the discovery of explanatory scientific models. Given a multivariate time series and domain knowledge stated as a set of generic processes, one can search a space of candidate models to find candidates that explain and fit the observations. Implemented systems have used this approach to rediscover dynamic models in ecology, chemistry, and other domains, with heuristic search methods scaling well to increases in the number of variables and model complexity.

I also argued that rate-based processes encode hypothesized causal influences, which means that any collection of them constitutes a causal model. Such accounts incorporate more constraints than traditional causal accounts, in that they require some influences to stand or fall together, but these reflect implicit assumptions in many scientific domains. Process models also support both counterfactual reasoning about different situatons and explanation in terms of familiar concepts, with the latter distinguishing the framework from more abstract causal variants. Techniques for inductive process modeling offer ways to construct such explanations, but, like other inductive methods, they cannot guarantee that models will be correct or generalize well. Nevertheless, the framework covers an important class of causal models that are highly relevant to many fields of science.

Despite the progress to date, that we future research should address some important limitations that remain. This should extend existing methods to handle data sets with unobserved variables and deal with rate expressions that contain parameters, possibly by drawing on techniques for iterative optimization to estimate unknown values. The augmented approaches should also be able to postulate entirely new processes, which would make them less dependent on user-provided background knowledge. Finally, an extended version of the paradigm should take advantage of controlled experimentation, which would aid substantially the ability to infer causal influences in complex dynamic systems. Taken together, these will offer a more complete account for the discovery of scientific models that provide causal explanations for observations in terms of underlying processes.

Acknowledgements The research reported here was supported by Grant No. N00014-11-1-0107 from the US Office of Naval Research, which is not responsible for its contents. The results draw on joint work

with Adam Arvay, Will Bridewell, Saso Džeroski, Ljupčo Todorovski, and others over years of shared effort on computational scientific discovery.

References

Arvay A, Langley P (2016) Selective induction of rate-based process models. In: Proceedings of the 4th annual conference on cognitive systems, Evanston

Asgharbeygi N, Bay S, Langley P, Arrigo K (2006) Inductive revision of quantitative process models. Ecol Model 194:70–79

Bay SD, Shrager J, Pohorille A, Langley P (2003) Revising regulatory networks: from expression data to linear causal models. J Biomed Inf 35:289–297

Bradley E, Easley M, Stolle R (2001) Reasoning about nonlinear system identification. Artif Intell 133:139–188

Bradshaw GL, Langley P, Simon HA (1980) Bacon.4: the discovery of intrinsic properties. In: Proceedings of the 3rd biennial conference of the Canadian society for computational studies of intelligence, Victoria, pp 19–25

Bridewell W, Langley P (2010) Two kinds of knowledge in scientific discovery. Top Cogn Sci 2:36–52

Bridewell W, Bani Asadi N, Langley P, Todorovski L (2005) Reducing overfitting in process model induction. In: Proceedings of the 22nd international conference on machine learning, Bonn, pp 81–88

Bridewell W, Langley P, Racunas S, Borrett SR (2006) Learning process models with missing data. In: Proceedings of the 17th European conference on machine learning. Springer, Berlin, pp 557–565

Bridewell W, Langley P, Todorovski L, Džeroski S (2008) Inductive process modeling. Mach Learn 71:1–32

Cohen S, Hindmarsh A (1996) CVODE: a stiff/nonstiff ODE solver in C. Comput Phys 10:138–43

Darwiche A (2009) Modeling and reasoning with Bayesian networks. Cambridge University Press, New York

Džeroski S, Todorovski L (1995) Discovering dynamics: from inductive logic programming to machine discovery. J Intell Inf Syst 4:89–108

Džeroski S, Todorovski L (eds) (2007) Computational discovery of communicable scientific knowledge. Springer, Berlin

Falkenhainer BC, Michalski RS (1986) Integrating quantitative and qualitative discovery: the ABACUS system. Mach Learn 1:367–401

Feigenbaum EA, Buchanan BG, Lederberg J (1971) On generality and problem solving: a case study using the DENDRAL program. In: Meltzer B, Michie D (eds) Machine intelligence, vol 6. Edinburgh University Press, Edinburgh

Forbus KD (1984) Qualitative process theory. Artif Intell 24:85–168

Glymour C, Scheines R, Spirtes P, Kelly K (1987) Discovering causal structure: artificial intelligence, philosophy of science, and statistical modeling. Academic, San Diego

Gordon A, Edwards P, Sleeman D, Kodratoff Y (1994) Scientific discovery in a space of structural models: an example from the history of solution chemistry. In: Proceedings of the 16th annual conference of the cognitive science society. Lawrence Erlbaum, Atlanta, pp 381–386

Hempel CG (1966) Philosophy of natural science. Prentice-Hall, Englewood Cliffs

Iwasaki Y, Simon HA (1986) Causality in device behavior. Artif Intell 29:3–32

Iwasaki Y, Simon HA (1994) Causality and model abstraction. Artif Intell 67:143–194

King RD, Whelan KE, Jones FM, Reiser PGK, Bryant CH, Muggleton SH, Kell DB, Oliver SG (2004) Functional genomic hypothesis generation and experimentation by a robot scientist. Nature 427:247–252

Kocabas S (1991) Conflict resolution as discovery in particle physics. Mach Learn 6:277–309

Kokar MM (1986) Determining arguments of invariant functional descriptions. Mach Learn 1:403–422

Koza JR, Mydlowec W, Lanza G, Yu J, Keane MA (2001) Reverse engineering of metabolic pathways from observed data using genetic programming. Pac Symp Biocomput 6:434–445

Langley P (1979) Rediscovering physics with Bacon.3. In: Proceedings of the 6th international joint conference on artificial intelligence. Morgan Kaufmann, Tokyo, pp 505–507

Langley P (1981) Data-driven discovery of physical laws. Cogn Sci 5:31–54

Langley P, Arvay A (2015) Heuristic induction of rate-based process models. In: Proceedings of the 29th AAAI conference on artificial intelligence. AAAI Press, Austin, pp 537–544

Langley P, Bradshaw GL, Simon HA (1983) Rediscovering chemistry with the BACON system. In: Michalski RS, Carbonell JG, Mitchell TM (eds) Machine learning: an artificial intelligence approach. Morgan Kaufmann, San Mateo

Langley P, Simon HA, Bradshaw GL, Żytkow JM (1987) Scientific discovery: computational explorations of the creative processes. MIT Press, Cambridge

Langley P, Sanchez J, Todorovski L, Džeroski S (2002a) Inducing process models from continuous data. In: Proceedings of the 19th international conference on machine learning. Morgan Kaufmann, Sydney, pp 347–354

Langley P, Shrager J, Saito K (2002b) Computational discovery of communicable scientific knowledge. In: Magnani L, Nersessian NJ, Pizzi C (eds) Logical and computational aspects of model-based reasoning. Kluwer, Dordrecht

Langley P, Shiran O, Shrager J, Todorovski L, Pohorille A (2006) Constructing explanatory process models from biological data and knowledge. Artif Intell Med 37:191–201

Moulet M (1992) ARC.2: linear regression in Abacus. In: Proceedings of the ML 92 workshop on machine discovery, Aberdeen, pp 137–146

Murata T, Mizutani M, Shimura M (1994) A discovery system for trigonometric functions. In: Proceedings of the 20th national conference on artificial intelligence. AAAI Press, Seattle, pp 645–650

Nordhausen B, Langley P (1990) A robust approach to numeric discovery. In: Proceedings of the 7th international conference on machine learning. Morgan Kaufmann, Austin, pp 411–418

Park C, Bridewell W, Langley P (2010) Integrated systems for inducing spatio-temporal process models. In: Proceedings of the 24th AAAI conference on artificial intelligence. AAAI Press, Atlanta, pp 1555–1560

Popper KR (1961) The logic of scientific discovery. Science Editions, New York

Schaffer C (1990) Bivariate scientific function finding in a sampled, real-data testbed. Mach Learn 12:167–183

Schmidt M, Lipson H (2009) Distilling free-form natural laws from experimental data. Science 324:81–85

Shrager J, Langley P (eds) (1990) Computational models of scientific discovery and theory formation. Morgan Kaufmann, San Francisco

Simon HA (1954) Spurious correlation: a causal interpretation. J Am Stat Assoc 49:467–479

Simon HA (1966) Scientific discovery and the psychology of problem solving. In: Colodny R (ed) Mind and cosmos. University of Pittsburgh Press, Pittsburgh

Todorovski L, Bridewell W, Langley P (2012) Discovering constraints for inductive process modeling. In: Proceedings of the 26th AAAI conference on artificial intelligence. AAAI Press, Toronto, pp 256–262

Todorovski L, Shiran O, Bridewell W, Langley P (2005) Inducing hierarchical process models in dynamic domains. In: Proceedings of the 20th national conference on artificial intelligence. AAAI Press, Pittsburgh, pp 892–897

Valdés-Pérez RE (1994) Human/computer interactive elucidation of reaction mechanisms: application to catalyzed hydrogenolysis of ethane. Catal Lett 28:79–87

Valdés-Pérez RE (1996) Computer science research on scientific discovery. Knowl Eng Rev 11:57–66

Washio T, Motoda H (1997) Discovering admissable models of complex systems based on scale types and identity constraints. In: Proceedings of the 15th international joint conference on artificial intelligence. Morgan Kaufmann, Yokohama, pp 810–817

Żytkow JM, Zhu J, Hussam A (1990) Automated discovery in a chemistry laboratory. In: Proceedings of the 8th national conference on artificial intelligence. AAAI Press, Boston, pp 889–894