Combining Data-Driven and Knowledge-Guided Methods to Induce Interpretable Physiological Models

Pat Langley
Institute for the Study of Learning and Expertise
2164 Staunton Court, Palo Alto, CA 94306

Will Bridewell
Center for Biomedical Informatics Research
Stanford University, Stanford, CA 94305

Abstract

In this paper, we review the paradigm of inductive process modeling and examine its application to human physiology. This framework represents models as a set of interacting processes, each with associated differential or algebraic equations that express causal relations 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 their associated parameters, and thus identify quantitative models that explain time-series data. We present an initial process model for aspects of human physiology, consider its uses for health monitoring, and discuss the induction of such models. In closing, we consider related efforts on physiological modeling and our plans for collecting data to evaluate our framework in this domain.

Introduction

Computational methods for machine learning and discovery have become widely used for automatically constructing models from observations, but they typically generate descriptive models that utilize formalisms developed in AI or statistics. In contrast, computational techniques for scientific modeling have long supported the explanation of complex phenomena in terms of first principles, but domain experts typically construct and tune them manually. This divide between data-intensive and knowledge-guided modeling has led to two distinct communities with little common ground.

However, these two paradigms are not antithetical, and each stands to benefit from the techniques and insights of the other. In this paper, we present a framework for model induction that combines time-series data with domain knowledge to produce interpretable models stated in formalisms familiar to scientists. More specifically, our approach encodes models as sets of processes that incorporate differential and algebraic equations, simulates these models' behavior over time, utilizes background knowledge to generate model structures, calculates differences between observed and predicted behavior to tune parameters, and ranks the discovered models by their fit to data.

We have applied this computational framework to automatically construct interpretable models in a number of fields, including ecology, environmental science, and biochemistry. In more recent work, we are exploring its use in developing computational models of human physiology, with the aim of supporting health monitors that track an individual's bodily state and detect cases of anomalous behavior. We are aiming for shallow but broad models that relate variables like heart rate, respiratory rate, and body temperature to exogenous various like activity level, external temperature, and humidity.

In this paper, we review the basic techniques that induce process models from data and background knowledge, using examples from ecology to illustrate the framework. After this, we explain how we can adapt this approach to human physiology, presenting an abstract process model that, when provided with additional detail, could support both generic and personalized health monitoring. In closing, we briefly discuss related work on quantitative physiological models and our plans for collecting time-series data from wearable sensors to use in monitoring and model construction.

Inductive Process Modeling

Historically, most scientific models have focused on explaining observed phenomena in terms of underlying processes or mechanisms. In many fields, these models incorporate mathematical equations that make quantitative predictions about changes over time. Unfortunately, constructing such models is time consuming and error prone, which suggests there is value in automating this effort. However, most work on model induction within artificial intelligence and machine learning has emphasized descriptive models that make little contact with domain concepts and that lack an explanatory character (e.g., Lavrac & Džeroski, 1994; Langley et al., 1987). Automating the construction of explanatory models requires a marriage of knowledge-guided and data-driven methods that draws on the strengths of both frameworks. In this section, we review one approach to this challenge known as *inductive process modeling* (Bridewell et al., 2008).

This framework incorporates a formalism that combines the notion of processes, which specify causal connections between two or more variables, with differential and algebraic equations, which provide quantitative details about these relations. These *quantitative process models* are mod-

Table 1: A process model of a predator–prey interaction between populations of foxes (f) and rabbits (r). The notation $d[X, t, 1]$ indicates dX/dt.

model Predation;
entities $r\{prey\}, f\{predator\}$;
process rabbit_growth;
 equations
  $d[r.pop, t, 1] = 1.81 * r.pop * (1 - 0.0003 * r.pop)$;
process fox_death;
 equations
  $d[f.pop, t, 1] = -1 * 1.04 * f.pop$;
process predation_holling_1;
 equations
  $d[r.pop, t, 1] = -1 * 0.03 * r.pop * f.pop$;
  $d[f.pop, t, 1] = 0.30 * 0.03 * r.pop * f.pop$;

ular in that one can easily remove a process, add a new one, or replace one process with another. Table 1 presents a process model for a simple, two-species predator–prey system. Here there are three processes that characterize prey growth, predator death, and predation. Each process contains equation elements that specify how one or more variables change over time. For example, predation_holling_1 states how the density of the prey and predator populations combine to influence each other. When coupled with the effects of fox_death via addition, the resulting equation, $d[f.pop, t, 1] = -1{*}1.04{*}f.pop{+}0.30{*}0.03{*}r.pop{*}f.pop$, provides an account of predator population dynamics. One simulates the model by compiling it into a set of differential and algebraic equations, then calling a numerical solver (Cohen & Hindmarsh 1996). The resulting trajectories are available for analysis and comparison with observations.

Over the past few years, we have developed several systems that construct quantitative process models from a mixture of domain knowledge and time-series data. The most recent instance (Bridewell & Langley, 2010) takes as input knowledge in the form of generic processes, entities, and constraints, along with measurements of variables to be modeled. The background knowledge primarily consists of generic processes such as

generic process holling_1:
 variables $S1\{prey\}, S2\{predator\}$;
 parameters $ar[0.01, 10], ef[0.001, 0.8]$;
 equations
  $d[S1.pop, t, 1] = -1 * ar * S1.pop * S2.pop$;
  $d[S2.pop, t, 1] = ef * ar * S1.pop * S2.pop$;

which is essentially a template for the predation process in Table 1. Here numeric ranges replace the parameters and typed identifiers replace the entities (i.e., $S1$ and $S2$ for f and r). Generic entities are named types that contain a combination of variables and parameters, whereas model constraints place limits on how one can combine the generic components. For instance, if there are multiple potential predation processes, the constraints could state that only one such relationship may exist between any pair of species. The data are continuous-valued time-series that are tied to variables in the model. Importantly, some variables may be hidden or theoretical, with no associated measurements.

Given background knowledge and data, an inductive process modeler will produce an account that both explains the observations and offers predictions. The model induction phase involves two interleaved stages. First, the system generates candidate structures by binding entities to generic processes and combining these components in ways that observe the constraints. Second, it estimates the parameters of these structures using a local optimization routine (Bunch et al., 1993) coupled with random restarts. Finally, the system ranks the structures and returns the best explanations based on their fit to the time-series data.

In previous work we have demonstrated the power of inductive process modeling in a variety of scientific domains. The earliest applications dealt with simple population dynamics and battery performance (Langley et al., 2003), with later work turning to the hydrodynamics of fjords (Asgharbeygi et al., 2006) and more complex ocean ecosystems (Bridewell et al., 2008). We have also used the framework to address problems in systems biology and biochemistry (Langley et al., 2006).

Process Models of Human Physiology

Given our previous success with inductive process modeling, it seems natural to consider applying the paradigm to physiology. Many models in this field are already stated as sets of differential equations, which further suggests it as a promising approach. However, most physiological models described in the literature focus on one particular function, such as the cardiac or respiratory subsystem, in considerable depth. Although we could replicate these accounts within the process modeling framework, we believe there is also need for high-level models that characterize interactions among variables from different subsystems.

For instance, consider a process model that relates respiratory rate, heart rate, blood oxygenation, tissue oxygenation, and activity level. Figure 1 depicts the causal structure of one such model with four distinct processes. These include mechanisms that:

- decrease tissue oxygenation as a monotonic function of the activity level;
- increase tissue oxygenation as a combined monotonic function of blood oxygenation and heart rate, as well as decrease the blood oxygenation;
- increase the blood oxygenation as a monotonic function of respiratory rate; and
- decrease respiratory rate rate as a monotonic function of tissue oxygenation.

The figure does not specify the structure or parameters of equations associated with each process, or even whether they take the form of algebraic or differential equations, since we are still exploring these alternatives. However, the signs on links indicate whether the output variables increase or de-

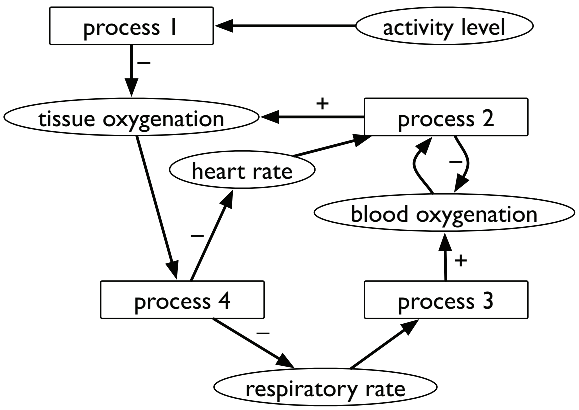

Figure 1: Abstract depiction of a process model that relates heart rate, respiratory rate, blood oxygenation, tissue oxygenation, and activity level. The signs indicate the direction of change in output variables as input variables increase.

crease with the inputs.[1] The model cannot make quantitative predictions about how attributes change over time without specific mathematical formulas, but this version still offers an abstract account of their interactions.

Walking through the model should clarify how it would operate if additional details were available. An increase (decrease) in the activity level, an exogenous variable, uses oxygen in the tissues and thus causes the tissue oxygenation to decrease (increase), other things being equal. This is offset by another process that increases tissue oxygenation as a combined function of blood oxygenation and heart rate, which determines how much oxygen is transported to the tissues. However, this process also decreases blood oxygenation by an equal amount as the tissue absorbs the oxygen.

A third process raises (lowers) the blood oxygenation as the respiratory rate increases (decreases). Finally, a fourth process leads to a decrease (increase) in respiratory rate and heart rate with an increase (decrease) in levels of blood oxygenation. These last two mechanisms introduce feedback into the system, letting it attempt to maintain blood oxygenation and tissue oxygenation at nominal levels. We should note that even this abstract model is almost certainly flawed, but it should clarify how one can apply process modeling to physiological phenomena.

Clearly, we might expand this abstract model to incorporate additional variables. These could include endogenous attributes like body temperature and blood glucose level, as well as exogenous variables like external temperature, humidity, and food intake. Appropriate processes would account for other phenomena like temperature and glucose regulation. The modular character of process models would let us introduce these separately or in concert, although different combinations would predict different behaviors.

Once we provide such a process model with parameterized equations, we can use it to simulate the how variables change over time. When combined with data obtained from measuring instruments, we can also use the resulting trajectories to monitor a person's physiological state. We envision two approaches to monitoring:

- predicting that blood oxygenation, heart rate, or some other variable will go below or above a critical value based on exogenous variables like activity level and oxygen pressure; and
- detecting anomalous observations that suggest a person is entering an unmodeled region of the physiological state space.

The first alternative will be useful in well-understood situations like oncoming heat stroke or hypoxia, where informed action (e.g., reducing activity or giving oxygen) can mitigate health consequences. The second case will be appropriate for situations in which the model is inaccurate or incomplete; it may indicate the physiological system is behaving abnormally, which would be a cause for medical concern.

Naturally, we can also compare the simulated trajectories to data obtained from measuring instruments to drive inductive process modeling, as described earlier. This will let us:

- guide search through the space of model structures and thus improve our understanding of physiological processes; and
- estimate parameter values for the processes that determine physiological behavior of average humans; and
- estimate personalized parameter values that account for individual differences in human physiology;

The resulting models could, in turn, be used in medical monitoring systems, as just outlined. Clearly, models that have been fit to physiological data will be more useful to this end than ones that are crafted by hand, especially if their parameters are tuned to explain individual differences.

Concluding Remarks

In this paper, we reviewed inductive process modeling, a framework for representation, simulation, and construction of quantitative models that explain the behavior of dynamical systems. We presented an example of a quantitative process model from ecology, described the simulation of such models to produce trajectories, and discussed ways to search the space of model structures and their associated parameters. We also explained how one can apply this framework to human physiology, illustrating these ideas with an abstract model from the domain, and how to use quantitative process models for the purpose of health monitoring.

Of course, we are not the first researchers to develop mathematical models of physiology. One mature framework, known as Archimedes (Schlessinger & Eddy 2002), incorporates a number of "first principles" physiological models. A key distinction is that Archimedes is a monolithic, multiscale model stated as an opaque Java program, rather than in a declarative formalism like our process models. Another is that scientists and medical experts set parameters manually based on evidence from the literature, whereas our approach estimates them from fits to measured time-series data.

[1] This model is only valid when the activity level is above normal, as during exercise; other physiological regimes would require additional processes.

In other work, Gribok, Buller, and Reifman (2008) have compared shallow models of body temperature produced by autoregression with deeper models developed manually. They note that the former models assume normal environmental conditions and homogeneous subjects to generalize across individuals, whereas the deeper models could apply to a range of conditions and to varied individuals. On the other hand, their shallow models were computer generated, while they claim that deeper models require hand construction and tuning. Inductive process modeling holds potential for automatically creating deep physiological accounts and tuning parameters to account for individual differences.

However, we must still demonstrate our framework's potential for physiological modeling. This will involve devising quantitative processes that move beyond the abstract structures in Figure 1 and producing a library of plausible generic processes for this domain. We hope to use wearable sensors for collecting data on physiological and environmental variables, which we can then use to guide search through the space of process model structures, estimate parameters for different individuals, and monitor subjects' physiological states as they change over time.

Acknowledgements

This material is based on research sponsored by ONR under agreement N00014-11-1-0107. The U. S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are the authors' and should not be interpreted as representing the official policies or endorsements, either expressed on implied, of ONR or the U. S. Government. We thank Mark Buller, Sandeep Gupta, and Wei-Nchih Lee for informative discussions about models of human physiology.

References

Asgharbeygi, N., Bay, S., Langley, P., and Arrigo, K. 2006. Inductive revision of quantitative process models. *Ecological Modelling* 194: 70–79.

Bridewell, W., Langley, P., Todorovski, L., and Džeroski, S. 2008. Inductive process modeling. *Machine Learning* 71: 1–32.

Bridewell, W., and Langley, P. 2010. Two kinds of knowledge in scientific discovery. *Topics in Cognitive Science* 2: 36–52.

Bunch, D., Gay, D., and Welsch, R. 1993. Algorithm 717: Subroutines for maximum likelihood and quasi-likelihood estimation of parameters in nonlinear regression models. *ACM Transactions on Mathematical Software* 19: 109–130.

Cohen, S., and Hindmarsh, A. 1996. CVODE, a stiff/nonstiff ODE solver in C. *Computers in Physics* 10: 138–143.

Gribok, A. V., Buller, M. J., and Reifman, J. 2008. Individualized short-term core temperature prediction in humans using biomathematical models. *IEEE Transactions on Biomedical Engineering*. 55: 1499–1487.

Langley, P., George, D., Bay, S., and Saito, K. 2003. Robust induction of process models from time-series data. In *Proceedings of the Twentieth International Conference on Machine Learning*, 432–439. Menlo Park, CA: AAAI Press.

Langley, P., Shiran, O., Shrager, J., Todorovski, L., and Pohorille, A. 2006. Constructing explanatory process models from biological data and knowledge. *AI in Medicine* 37: 191–201.

Langley, P., Simon, H. A., Bradshaw, G., and Żytkow, J. M. 1987. *Scientific Discovery*. Cambridge, MA: MIT Press.

Lavrač, N., and Džeroski, S. 1994. *Inductive Logic Programming: Techniques and Applications*. New York, NY: Ellis Horwood.

Schlessinger, L., and Eddy, D. M. 2002. Archimedes: A new model for simulating health care systems—the mathematical formulation. *Journal of Biomedical Informatics* 35: 37–50.