The relevance of natural kinds to the computational discovery of novel scientific variables

Benjamin C. Jantzen, Virginia Tech

We set out to address two deeply related problems. The first is the recognition of what philosophers would call "natural kinds" [1] and what we call *dynamical kinds*, i.e., collections of physical systems that share enough of the causal structure that generates their behavior to draw accurate inferences about any one member of the collection from properties of the others. Put differently, the challenge is to determine whether two systems share an unknown set of governing equations, or to say it yet another way, whether they share physics. This is a central feature of scientific inference – the clustering of phenomena into classes for which a scientific law or theory may fruitfully be sought [2].

The second problem is the identification of novel scientific variables that appear to expand the degrees of freedom in terms of which phenomena are described [3]. The identification of electrical resistivity is one such instance in the history of science [4]. This sort of inference is distinct from the dimension-reducing feature selection of machine learning, and it depends upon recognition of variation in the law-like behavior of a type of system in different contexts. With resistivity for instance, the key observation is the change in relation between voltage and current as one wire is replaced with another. However, not all variation is indicative of a new degree of freedom in a single type of system. Variation across systems may be driven by a difference in physics altogether. For instance, internal combustion engines in operation are different kinds of things from a thermodynamic perspective than inactive, passively cooling engines, and it would be a mistake to suggest that the differences in their behavior correspond to differences in the value of a previously unrecognized variable or intrinsic property. Rather, the variations that matter for identifying novel variables are those that occur within a dynamical kind [5]. When two systems exhibiting distinct behaviors are nonetheless known to share causal structure, such variation is fruitfully interpreted as due to variation in the value of a new variable. Thus, an ability to identify whether two physical systems with unknown governing laws belong to the same dynamical kind is essential to solving both problems described above.

The first step in turning these qualitative discovery problems into computationally tractable problems is the construction of a proper metric that quantitatively tracks the similarity of causal structure or governing laws. There are indefinitely many ways of defining similarity among causal structures (see, e.g., [6]). But there is one that hews closely to the implicit tradition in scientific practice of treating as instances of a common kind all and only those systems whose dynamics are governed by differential equations of the "same form". This approach depends upon the concept of a *dynamical symmetry*, an intervention on the state of a system that commutes with its time evolution [2]. For example, a population of bacteria undergoing exponential growth manifests a scaling symmetry for every positive scaling coefficient, k. That's because one obtains the same final state whether one increases the population size by a factor of k and then allows it to evolve for time Δt, or lets it evolve for Δt and then scales the population by k. The collection of dynamical symmetries — and the group structure induced by their behavior under composition — is dictated by a system's detailed causal structure [7]. Thus, a collection of dynamical symmetries and an attendant algebra of composition picks out an equivalence class of causal structures — a dynamical kind.

While it is sometimes useful to consider the binary question of whether two systems belong to the same dynamical kind [8], many advantages accrue from defining a degree to which the dynamics of two systems in strictly different dynamical kinds diverge. By identifying a probability distribution characteristic of the dynamical symmetries of a given dynamical kind, one can construct a metric of dynamical similarity from an existing proper metric on probability distributions [9]. Called the dynamical distance, D_D, this metric vanishes for members of the same dynamical kind, and has been been shown to be sensitive to differences in causal skeleton (which variables cause which), the effective order, degree of nonlinearity, and the presence of chaos in the governing dynamics; the more two systems differ in any of these respects, the greater the distance between them as measured by D_D. Importantly, D_D can be estimated from observational timeseries of system states, and there are well-posed statistical tests for the significance of differences in D_D. All of this is true both for noisily sampled but deterministically evolving systems and for systems exhibiting genuinely stochastic dynamics.

With the ability to determine similarity of the unknown governing dynamics, it is possible to search for evidence of additional variables not represented in one's initial set. Explicitly, the inference engine is provided timeseries tracking the values of a set V of n variables for the physical system of interest as its state evolves unperturbed from some initial state, along with a discrete tag indicating the qualitative circumstances under which the timeseries was taken (e.g., one wire versus another). If two such timeseries belong to the same dynamical kind and yet represent divergent time evolution from the same initial conditions, then we infer a novel variable. That is, we add a variable v_1 to V and interpret this additional variable as reflecting a previously unrecognized causal property of the system. Furthermore, each distinct time evolution of the system from a fixed initial condition is taken to reflect a distinct value of this novel variable. From this point on, it is possible to consider timeseries in terms of the $n + 1$ variables of the expanded set V and seek both novel types of system—i.e., distinct dynamical kinds—and additional novel variables. This, of course, requires additional experiments and remains at present beyond the possibility of full automation. Rather, what we have is a cyclical discovery engine involving automated assessment of the similarity of dynamical kinds and the introduction of novel variables leading to experiments conducted by humans and designed to exhibit the causal role of these novel variables. These produce timeseries of expanded dimensionality that in turn feed another round of assessment and discovery.

[1] M. H. Slater, *Br. J. Philos. Sci.*, vol. 66, no. 2, pp. 375–411, Jun. 2015.

[2] B. C. Jantzen, *Synthese*, vol. 192, no. 11, pp. 3617–3646, 2014.

[3] W. W. Rozeboom, *Philos. Sci.*, vol. 28, no. 4, pp. 337–377, 1961.

[4] P. Langley, H. A. Simon, G. Bradshaw, and J. M. Zytkow, *Scientific discovery*. Cambridge, Mass: MIT Press, 1987.

[5] B. C. Jantzen, *Philosophies*, vol. 6, no. 4, p. 103, Dec. 2021. [Online]. Available: `https://www.mdpi.com/2409-9287/6/4/103`.

[6] B. Eva, R. Stern, and S. Hartmann, en, *Philos. Sci.*, vol. 86, no. 5, pp. 821–835, Dec. 2019.

[7] S. Roy and B. Jantzen, *Chaos*, vol. 28, no. 7, p. 075 305, Jul. 2018.

[8] B. C. Jantzen, *Proceedings of the UAI 2016 Workshop on Causation*, 2017, arXiv: 1612.04933. [Online]. Available: `http://ceur-ws.org/Vol-1792/paper2.pdf`.

[9] C. Shea-Blymyer, S. Roy, and B. Jantzen, *Entropy*, vol. 23, no. 9, p. 1191, Sep. 2021. [Online]. Available: `https://www.mdpi.com/1099-4300/23/9/1191`.