## thesis

## A natural bias for simplicity

A 'map' in ordinary usage is a representation of one thing by another, usually simpler thing. Mathematicians define a map more formally as a function, *f*, taking the elements of one set into another. Using today's computing terminology, one might also think of inputs and outputs — a map acts on an input to produce an output. Maps of this kind run through all of science, from plasma physics to linguistics.

Scientists are experts in the maps of their own special areas. Earth scientists understand basic models of plate tectonics, just as biochemists have great familiarity with models of cell signalling. Yet one might wonder if all these different maps may share certain general properties, beyond simple properties such as continuities. Recent work suggests that the answer is yes — maps that arise across many fields of science appear to share a common and rather surprising bias toward simplicity in their outputs. That is, while we might naively expect nature's maps to reflect a random sample of all possible maps, they actually tend strongly toward those with many simple outputs. In this sense, nature really seems to be simpler than it could be.

This idea emerges from an application of the ideas of algorithmic information theory to mathematical mappings (K. Dingle, C. Q. Camargo & A. A. Louis, Nat. Commun. 9, 761; 2018). Algorithmic information theory aims to quantify the inherent complexity of a string of digits. In the case of binary digits, for example, a repeating string — 010101. — is inherently simple, whereas a highly erratic string is (in general, though not always) more complex; it contains more information and is more difficult to describe. A key idea in the field is the Kolmogorov complexity K(x) of a binary string x, defined (leaving aside some technicalities) as the length of the shortest programme able to generate *x* when run on a universal Turing machine.

This definition centres on a mapping. It goes from an output string x to another binary string giving the length of the shortest programme generating x. One interesting question is the following: what is the probability P(x) that a randomly chosen input will generate x? Mathematicians have developed estimates for this quantity, finding that it goes crudely in proportion to  $2^{-K(x)}$ . Hence, output strings become exponentially less likely with increasing Kolmogorov complexity. This is a general result for one particular mapping of importance in algorithmic information theory.

The question Dingle and colleagues address is whether anything similar can be established for maps in general. Any map can be considered to carry out a computation, and so can be considered in terms of the computer code that would be required to implement it. Following this approach, the authors derive an analogue of the classic result discussed above, but applying to any of a broad class of discrete maps. The result asserts that the probability P(x) a randomly chosen input will generate an output x is bounded by  $2^{-K(x)}$ . Again, there's a natural weighting toward simpler outputs.



Maps that arise across many fields of science appear to share a common and rather surprising bias toward simplicity in their outputs.

It may seem almost miraculous that one can derive expectations about maps in a general setting just by considering computational aspects of how maps process information. But it does seem to work, as Dingle and colleagues demonstrate in a number of real world examples. An example from biophysics is the map connecting RNA nucleotide sequences or genotypes to so-called RNA secondary structures — the structures these linear molecules naturally assume due to interactions between different bases. Dingle and colleagues examined the case n = 55 — RNAs of length 55 — and first used a popular software package to calculate the minimum free energy structure for a large number of random sequences. They then turned to a commonly employed shorthand for describing the secondary structures to turn each structure into a binary sequence. Finally, by estimating the complexity of each resulting output, they found that the results clearly showed a bias toward simpler sequences.

A more general example is ordinary differential equations, which yield discrete models when coarse-grained by discretizing input parameters and outputs, as happens in applications throughout science. Here the authors tested a representative model drawn from biochemisty and used in modelling the biochemical pathways controlling circadian rhythms. They first generated a

discrete mapping by discretizing the system of ordinary differential equations, which has 15 tunable parameters, and solving the resulting numerical model to get outputs as discrete time trajectories. Again, they found that the outputs were strongly biased toward simplicity, and the likelihood of outputs of different complexity could be accurately estimated using the output complexity only.

This work may take us part way to understanding why so many aspects of nature can be understood on the basis of relatively simple models, and indeed why the world seems understandable at all. It's a fundamental mathematical bias at work. Even so, there are many open questions. The authors note that a general map could conceivably have any distribution of complexities of its outputs. A map might be designed, for example, to have almost all complex outputs, and very few simple outputs. In this case, even if simpler outputs have a higher relative probability than more complex ones, the imbalance toward complexity in the sheer number of outputs will win out, leading mostly to complex outputs. But in the applications they've examined so far, this doesn't seem to happen. Most maps do seem to favour simpler output — although exactly why remains mysterious.

The authors also believe the work may help to explain some observed features of so-called deep learning neural networks, particularly why they seem to generalize so well and fit data well outside of the examples on which they've been trained, and despite the over parametrization of the models — the number of parameters they contain is also much larger than the training data. In a preprint (https:// arxiv.org/abs/1805.08522), Camargo and Louis, working with Guillermo Valle-Pérez, suggest this might be another example of the bias toward simpler outputs. Deep learning models may generalize well because they're biased in the same way that data from real world problems are — toward simple outputs.

The work of Dingle, Camargo and Louis reminds me of the related observation by Mark Transtrum and colleagues (preprint at https://arxiv.org/abs/1501.07668) concerning the surprising effectiveness of 'sloppy models' in so many areas of science. Why is the world understandable? We may never know for sure. But the mystery may be starting to shed its secrets.

## Mark Buchanan

Published online: 4 December 2018 https://doi.org/10.1038/s41567-018-0370-y