James Thorniley
(University of Sussex)

The strength of weak coupling:
The misapplication of information theory to causal strength

Bradford Hill's eponymous list of criteria for causal inference gives *strength* as the first consideration. Recent years have seen a proliferation of proposed information theoretic analyses of causal strength in complex systems, with a view to applications in neuroscience, ecology and climate science, amongst others. The theory of Granger causation is a popular basis for this work, resulting in measures like Schreiber's *transfer entropy*. However these measures, especially when interpreted as causal strength, seem to contradict the conclusions of dynamical systems-based analysis of complex systems. I argue that the most coherent interpretation of this phenomenon is that transfer entropy *is* a justifiable part of an epistemological approach to causation, but it should not be viewed as a measure of causal strength. I ask what we want from a measure of causal strength, and conclude that perhaps it is something very much like the familiar coupling parameters of dynamical systems models.

Transfer entropy, like Granger causality, is known to be confounded by various aspects of complex system dynamics. For example, a pair of coupled dynamical systems ("chaotic synchrony" being a typical example) will tend to exhibit low transfer entropy for both low coupling (where the relationship is random) and high coupling (where the relationship is stable), with a peak in the middle where coupling is weak and the overall dynamic is complex. As a result some have suggested that transfer entropy does not measure causation, but some form of complexity, or in the words of one study, "emergent computation". Others have developed alternative statistics that they hope better reflect strength of causal influence.

I suggest that such attempts confuse the epistemological role of statistics. In fact, transfer entropy is a justifiable *inferential statistic* for causation, at least provided that the Reichenbachian assumption that all correlations are a result of causal relationships is met. It is akin to an indicator of what Deborah Mayo calls *severity*: a post-data assessment of the inferential warrant for a given hypothesis. From this perspective, a statistical result only supports a hypothesis to the extent that it would be very unlikely to reach the value obtained if the hypothesis was false. If conditions are right, high transfer entropy is very unlikely if there is no causal relationship, and an experiment resulting in high transfer entropy is thus a *severe* test for the presence of that causal relationship. However, low transfer entropy is not necessarily evidence of *no* causal relationship, on account of the fact that low transfer entropy may be likely even when there is a causal relationship.

This asymmetry mirrors a debate around the fundamental assumptions used in causal Bayes net theories. A set of variables with no non-causal statistical dependencies are termed Markovian by Judea Pearl. Such a Markovian set will satisfy the *causal Markov condition* (CMC), namely that any subset of those variables will be statistically independent of any other subset when conditioning on a third subset that is "blocking" or "d-separating" according to the causal graph. Transfer entropy can be seen to reflect this logic if a plausible causal graph of a dynamical system is introduced. Though the Markovian assumption is subject to various objections and caveats, it is the "converse" assumption, *stability*, that is more problematic in the current context. Stability states that whenever variables are *not* d-separated according to a causal graph one can expect a corresponding statistical dependence. Loosely speaking, where the CMC says that all correlations must be explained by causes, stability says that all causes will lead to correlations (note that these are logically distinct claims).

What stability does not say, though perhaps it suggests it, is that *stronger* causes will lead to *stronger* correlations in a monotonic fashion. Let us call this extra claim *strength-stability*. Transfer entropy is a generalized non-linear equivalent to conditional correlation. Synchronized systems appear to offer a clear counter-example against strength-stability - as dynamical coupling (which I argue is a reasonable measure of causal influence) increases, transfer entropy first increases with it, then decreases. Thus consideration of strength-stability in the context of complex systems extends existing critiques (e.g. by Cartwright) of the standard stability criterion.

A fundamental question that remains unanswered is what are we looking for in a universal quantifier of causal influence strength, and why? Hill is rather ambiguous on just what strength *is*, giving only a few examples concerning mortality rates. In any case, he seems to regard it as a means to an end, a heuristic for evaluating the confidence in the very existence of a causal relationship. For this purpose, there is nothing wrong with transfer entropy as it is, though it is only one of many available tools.

The parameters of dynamical models give relatively unproblematic characterizations of causal influence strength in the cases where they are relevant. Information theory offers tools for statistical inference that also potentially serve as descriptors of complexity. It appears unlikely that there is a universal measure that serves both the inferential and strength-quantifying functions. Some scientific programmes, especially in the neurosciences, appear to demand a generalised quantifier of causality (rather than of complexity) - for these my argument may have negative implications. Yet in other contexts a clearer view of the role of information theoretic statistics should be overwhelmingly positive.