# Prediction & modularity in complex dynamical systems

**Artemy Kolchinsky**
Indiana University, Bloomington
akolchin@indiana.edu
**Luis M. Rocha**
Indiana University, Bloomington
rocha@indiana.edu

Modularity is a concept of central importance in the study of complex systems. Broadly speaking, modular systems are composed of tightly-integrated subsystems, called *modules,* which are in turn weakly coupled to one another.

Despite its widespread use, the meaning and operationalization of modularity varies widely between scientific paradigms, fields, and processes of interest. We are interested in a general formalization of modularity applicable to multivariate dynamical systems, whether represented in the form of dynamical models or time series recordings. Such a formalization can be used not only for the analysis of system dynamics but also to compare notions of modularity found in disparate domains.

Information theory provides principled measures of information transfer and statistical dependence in distributed systems, and is well suited for quantitative analysis of modularity. Here, *multi-information* is a measure of the degree of statistical interdependence among a group of variables. Generalized to a dynamical setting, it is referred to as *stochastic interaction* [1], and measures the amount of information flowing between components of a system evolving dynamically through time.

One approach to modularity in dynamical systems attempts to partition system variables so as to minimize stochastic interaction between partition blocks. Low values of stochastic interaction correspond to partitions with little dynamical interdependence between blocks, reaching 0 when blocks represent completely independent subsystems. Stochastic interaction, however, cannot be a cost function *per se* because it generally selects partitions with larger blocks. For example, the partition with the largest block (that which includes the entire system) always has a stochastic interaction of 0. For this reason, weakly-coupled subsystems are sometimes identified by optimizing informationtheoretic measures normalized by ad-hoc terms [2].

Here, we look at modularity from the point of view of statistical inference, which leads to principled normalization terms as well as a variety of conceptual insights. We model a Bayesian learner who observes a certain number of transitions from a dynamical system of interest, and must then predict future transitions of that system. The learner assumes that the system may factorize into independent blocks, and selects the factorized model with the best posterior predictive performance.

In this case, stochastic interaction can be interpreted as a lower bound on the error of factorized statistical models. However, while partitions with smaller blocks generally have higher stochastic interaction terms, their smaller state-spaces induce dynamics that can be learned with fewer parameters. Because the predictive performance of a model depends on both the optimal possible fit (stochastic interaction) and the complexity of the parameter space, a trade-off between learning difficulty and predictive power is established. This trade-off can be controlled by varying the amount of training data provided: as it grows larger, simpler (more factorized) models gradually become less predictive than more complex ones. This trade-off can be used to generate a multi-scale decomposition of a dynamical system of interest.

The perspective of statistical learning can be used to define other measures useful for studying modularity. For example, we define the *total modularity* of a system as the improvement in predictive performance of optimally-factorized models above that of the largest (non-factorized) model, summed across the entire training process. High measures of total modularity indicate the presence of modular architectures that can be exploited by the learning process.

*State-dependent* and *causal* versions of modularity can also be quantified. The organization and degree of coupling between subsystems can vary between different parts of a dynamical system's state space. This can be inferred in a state-dependent manner by training and testing on dynamics sampled from a sub-region of the state space. In certain parts of the state space, however, correlations can obscure the presence of causal interactions between partition blocks. Causal modularity can detect such interactions by using a fully-supported 'interventional' distribution for training, while testing on a distribution preferentially weighted over a state space region of interest.

We argue that approaching modularity from the perspective of statistical modeling is both formally and conceptually fruitful (see also [3]). It provides principled normalization terms, can yield useful analyses of real-life dynamical systems (such as biochemical regulation models and recordings of neural dynamics), and sharpens intuitions regarding the role of modularity in scientific modeling.

## Bibliography


[1] Ay, N, and T Wennekers, "Dynamical properties of strongly interacting Markov chains", *Neural Networks* (2003).

[2] Balduzzi, D, and G Tononi, "Integrated information in discrete dynamical systems: motivation and theoretical framework", *PLoS Comput Biol* (2008).

[3] Kolchinsky, A, and L M Rocha, "Prediction and Modularity in Dynamical Systems", *ECAL 2011*, Submitted for publication.