Veli-Pekka Parkkinen
(University of Oslo)

# Causal modularity and the explanation of complex systems

An influential account of causal explanation, the interventionist theory (Woodward 2003), requires that an explanation in terms of many component causes must satisfy a condition of modularity. Modularity states that it must be possible to intervene to change the properties of individual components without changing the functional (causal) connections between the components. This condition has been contested on the grounds that complex biological systems do not afford actual interventions on their subsystems that would leave all connections between the subsystems intact (Mitchell 2008 & 2009). Instead, when an intervention changes one subsystem, the other subsystems might shuffle their connections so that the overall functionality of the system remains unchanged.

In this talk I will present an inferential reading of the modularity condition, and defend it against the complex systems critique. I will argue that decomposing a system into functionally differentiated subsystems is useful explanatory strategy insofar as the decomposition allows us to make inferences about the properties of the system under interventions and does this in a particular way that is characterized by modularity.

As a background I assume a view according to which explanation is about providing information that furnishes inferences concerning what would happen to a target system if it were intervened on. This ability to make what-if -inferences is what constitutes understanding (Ylikoski & Kuorikoski 2010). A representation of a system that decomposes it in some way explains insofar as it provides understanding of the properties of the system.

I argue that we must distinguish between an ontological notion of modularity that applies to causal systems as concrete things in the world, and an epistemological notion of modularity that applies to explanations as representations of real-world systems. Pending an agreed-upon analysis of the former notion, I focus on the role of modularity as a property of explanations. I distinguish modularity's role in Woodward's theory as a criterion that can be used to select from many possible decompositions of a system the ones that support correct inferences about outcomes of actual manipulations: A representation of a system as a collection of components is modular to the extent that it allows us to infer outcomes of local interventions on the system by representing them as interventions on the components in the representation, without us having to reconsider the description of dependencies among the components anew with respect to each intervention. In such a way, a modular decomposition of a system provides understanding in a particularly parsimonious manner.

Understood this way, modularity comes in degrees. Different decompositions of a system can have different modularity properties; a decomposition may permit correct what-if inferences about the outcomes of some interventions on some properties of a target system, but not others. For instance, a particular decomposition might furnish correct what-if –inferences about qualitative properties of a system, but not about its quantitative properties. Different decompositions provide understanding of those properties of a system with respect to which they are modular to some degree. This does not require, even in principle, that there exists a privileged decomposition that is perfectly modular, so that it would correctly predict changes in all the properties of a system with respect to all conceivable interventions. Instead, the modularity-condition suggests that in trying to understand complex systems where properties of individual subsystems heavily depend on their connections to each other, we often have to consider various decompositions of the same system. This is in accordance to actual scientific practice. For instance, in studying living systems – a paradigmatic case of complex systems – biologists employ various different decompositions of the same system depending on their primary explanatory interests. I provide an example from research on genetics of circadian rhythms that illustrates how biologists consider decompositions both on the level of individual genes and on the level of clusters of interacting genes, in order to answer qualitative and quantitative questions about the circadian system.