The Role of Conditional Independence in the Evolution of Intelligent Systems

Jory Schossau
Department of Computer Science & Engineering
Michigan State University
East Lansing, Michigan 48824, U.S.A.
jory@msu.edu

Larissa Albantakis
Department of Psychiatry
University of Wisconsin
Madison, Wisconsin 53719, U.S.A.
albantakis@wisc.edu

Arend Hintze
Department of Microbiology & Molecular Genetics
Michigan State University
East Lansing, Michigan 48824, U.S.A.
hintze@msu.edu

ABSTRACT

Systems are typically made from simple components regardless of their complexity. While the function of each part is easily understood, higher order functions are emergent properties and are notoriously difficult to explain. In networked systems, both digital and biological, each component receives inputs, performs a simple computation, and creates an output. When these components have multiple outputs, we intuitively assume that the outputs are causally dependent on the inputs but are themselves independent of each other given the state of their shared input [11]. However, this intuition can be violated for components with probabilistic logic, as these typically cannot be decomposed into separate logic gates with one output each. This violation of conditional independence on the past system state is equivalent to *instantaneous* interaction — the idea is that some information between the outputs is not coming from the inputs and thus must have been created instantaneously. Here we compare evolved artificial neural systems with and without instantaneous interaction across several task environments. We show that systems without instantaneous interactions evolve faster, to higher final levels of performance, and require fewer logic components to create a densely connected cognitive machinery.

CCS CONCEPTS

•**Computing methodologies → Cognitive science;** *Probabilistic reasoning;* Temporal reasoning; Spatial and physical reasoning;

KEYWORDS

Information Integration, Networks, Causality, Prediction, Learning

ACM Reference format:
Jory Schossau, Larissa Albantakis, and Arend Hintze. 2017. The Role of Conditional Independence in the Evolution of Intelligent Systems. In *Proceedings of the Genetic and Evolutionary Computation Conference 2017, Berlin, Germany, July 15–19, 2017 (GECCO '17),* 2 pages.
DOI: http://dx.doi.org/10.1145/3067695.3076033

DOI: http://dx.doi.org/10.1145/3067695.3076033

1 INTRODUCTION

Evolvable Markov Brains are networks of deterministic and probabilistic logic gates whose function and connectivity are genetically encoded. They are a useful model to study the evolution of behavior [10], cognitive properties [8], and neural-network complexity [1, 4], and can also be used as classifiers [3]. At each generation of evolution within a particular task environment, networks are selected based on their fitness and the populations adapt through random genomic mutations. The genome is sequentially processed with specific sites indicating the start of a gene. An individual gene encodes one Hidden Markov Gate (HMG), which specifies connections between network elements and also determines input-output logic [8]. These HMGs are generalized logic gates that encompass conventional logic gates such as XOR or NAND, whose logic table is typically a static mapping of two inputs to a single output (Figure 1A), but allow for more than the typical two-in-one-out format and can use a probabilistic mapping between input and output states. Here, HMGs could receive up to four inputs mapped to maximally four outputs. In this way, each gene may encode an entire logic module, as opposed to only a single logic function.

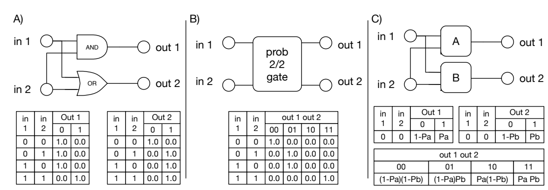

Figure 1: Examples of gate decomposition into combinations of simpler gates. Panel A shows two deterministic logic gates whose inputs are cross-wired so that both gates receive the same inputs. The tables below show their probabilities to output 0 or 1 respectively. These probabilistic logic boundary cases are effectively deterministic logic gates. Panel B shows a two-in-two-out logic gate that is functionally identical with the two gates depicted in Panel A. Panel C shows two probabilistic logic gates similarly connected like the deterministic gates from panel A. The logic tables below only show the case where both inputs are 0. The lower table shows replacing both probabilistic logic gates with a single two-in-two-out probabilistic logic gate (similar to panel B) and how the new probabilities for that gate are constructed from the individual probabilities of both gates.

Apart from introducing randomness into the Markov Brains, probabilistic HMGs also differ from deterministic HMGs in the way they can be represented by a collection of simpler logic gates. The outputs of a deterministic HMG are necessarily conditionally independent from each other: given the input state, an output is either on ('0') or off ('1'). Information about the state of other outputs is irrelevant, and as a consequence, a deterministic HMG can always be decomposed into several logic gates with one output each. For example, a two-in-two-out deterministic HMG (see Figure 1 panel B) can easily be decomposed into two independent two-in-one-out gates (see Figure 1 panel A). This decomposition works similarly for larger gates with more inputs and more outputs, requiring one logic gate per output.

Probabilistic HMGs, on the other hand, are not generally decomposable into separate logic functions for each output. In the case of a two-in-two-out probabilistic HMG, a probability table (for a detailed explanation see [8]) maps all four possible input states to all four possible output states. As a result, probabilistic HMGs are not necessarily always decomposable into smaller units, which sometimes results in the output wires having information, and thus violates conditional independence given the input state (cf. [7]) - which is also know as an *instantaneous interaction.* Without such conditional independence, the interactions between elements in a probabilistic Markov Brain cannot be represented as a directed acyclic *causal* graph [11]. This prohibits analyzing the causal composition of these Markov Brains, which means that the theoretical framework of integrated information theory (IIT) [2, 9], which assesses how sets of elements within a system causally constrain each other, cannot be applied to these Markov Brains.

While instantaneous interaction may be a curious phenomenon in evolvable Markov Brains, the question remains whether it hampers or helps Markov Brains to adapt. To explore this question, we implemented a *decomposable* version of the evolvable probabilistic HMGs (up to four-in-four-out) with conditionally independent outputs $\{\mathrm{out}_1, \ldots, \mathrm{out}_N\}$.

2 RESULTS

When evolving Markov Brains to solve three independent tasks: to forage, to perform active categorical perception [1, 8], and to integrate information to navigate [5, 6] we find that Brains using decomposable gates evolve faster (see Figure 2) and find generally better solutions (see Figure 3). In addition we find that the architecture of the evolved Markov Brains differs significantly (data not show). Markov Brains using only decomposable gates have a larger diameter, tend to use fewer gates, and always end up having a higher connectivity.

3 CONCLUSION

Using decomposable logic gates not only allows us to study integrated information in greater detail, but provides us with a new way to accelerate evolution. Our results suggest there is no apparent reason to include instantaneous interactions in Markov Brains. In future work we will explore which environments give the greatest benefit to decomposable gates, and if the tendency to evolve towards deterministic logic is the reason for their faster adaptation.

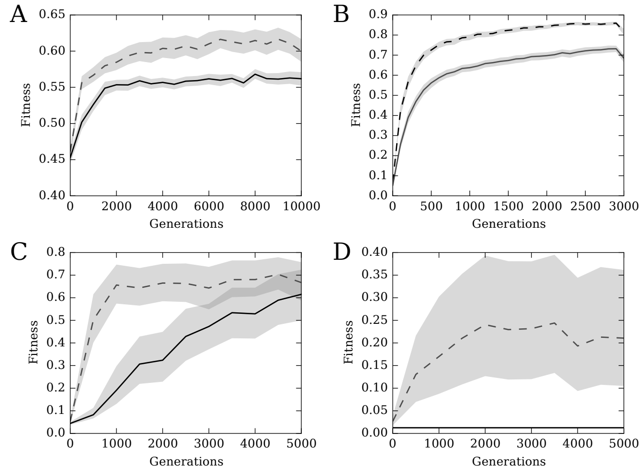

Figure 2: Average fitness of organisms on the line of descent in the spatial temporal (A) and foraging (B) environments, or association environment with no punishment (C) or a cost of punishment being 0.1 (D). The solid line represents average performance of agents restricted to conventional probabilistic HMGs (with instantaneous interactions), dashed lines represents average performance of agents restricted to decomposable HMGs (without instantaneous interactions). The y axes are normalized to show the fraction of maximally attainable fitness in each environment. The gray shadow indicates the bootstrapped 95% confidence interval of the mean.

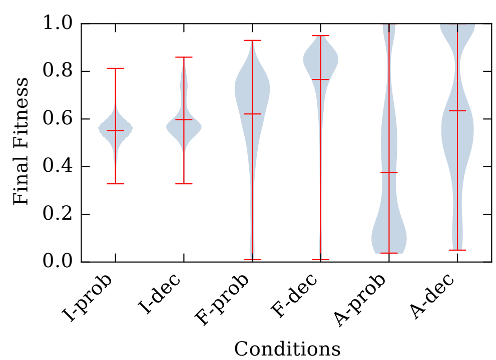

Figure 3: Distribution of performances at the end of evolution for each of two conditions ("prob" for using conventional- and "dec" for using decomposable logic gates) in each of three environments represented as I (temporal spatial integration task), F (foraging task), and A (association task with a punishment of 0.05). Red dashes indicate the mean and extrema. Fitness was normalized such that maximal theoretically attainable fitness is represented as 1.0. Significance was tested using the Mann-Whitney U test, with $p < 0.05$ for each environment ($p = 0.0$, $p = 0.0$, $p < 2.2 \times 10^{-112}$).