Emergence of Modular Structure in the Recurrent Neural Network During Incremental Multi-task Learning

Yuhang Wu (202221080817@uestc.edu.cn)
School of Computer Science and Engineering
University of Electronic Science and Technology of China, Chengdu, China

Shi Gu (gus@uestc.edu.cn)
School of Computer Science and Engineering
University of Electronic Science and Technology of China, Chengdu, China

Abstract

Biological brain networks universally exhibit distinct simplicity: they are composed of modular components that may function relatively independently. Yet, there's currently no consensus on the origins of modularization. In this study, we trained single recurrent neural networks on multiple cognitive tasks requiring working memory, decision-making, classification, and inhibitory control, thereby simulating real-world challenges. Our findings reveal that under conditions of constrained network size, multitasking promotes greater modularity compared to scenarios involving fewer tasks. This implies that modularity may arise as an adaptation in models required to handle multiple tasks when the units available for computation are limited. Additionally, we compared the learning processes of dynamically evolving networks, which form new connections periodically, with those of statically fixed networks, where connections are pre-established at the start of training. We found that models growing sequentially lead to higher modular structures across all wiring rules. Our study proposes that functional demands consequently influence structural formation, offering new insights into neuroscience.

Keywords: modular neural network; computational neuroscience; structural-functional interaction; cognitive multitasking

Introduction

The biological brain exhibits formidable computational abilities, efficiently processing complex and dynamic information to perform diverse and challenging tasks. Contrary to what might be expected from such complexity, the underlying structure of the brain's network is not densely disordered or chaotic. Instead, it demonstrates remarkable simplicity: the network is composed of modular components that function relatively independently, and it features the recurrent use of specific neural circuit patterns, known as network motifs. The origins of these structural characteristics, however, remain a subject of inquiry. Previous research has investigated various contributing factors, including modularly varying goals ((Kashtan & Alon, 2005)), constraints related to wiring costs ((Clune, Mouret, & Lipson, 2013)), and adaptations to minimize catastrophic forgetting ((Ellefsen, Mouret, & Clune, 2015)).

In this work, we introduce a novel hypothesis positing that the functional demands of tasks inherently sculpt the structural organization of the brain. Through extensive experimentation, we show that a modular architecture may emerge as a necessary adaptation for the biological brain to manage multiple tasks efficiently under computational constraints. Moreover, our findings indicate that dynamically evolving networks yield structures with greater modularity and pronounced performance disparities, reflecting the brain's inclination towards highly organized and individually distinct differences.

Methods

To investigate how the adaptation to multitasks helps functionally shape the modular structure, we trained a single RNN model (Fig.1) to perform 20 inter-related cognitive tasks following the settings of (Yang, Joglekar, Song, Newsome, & Wang, 2019).

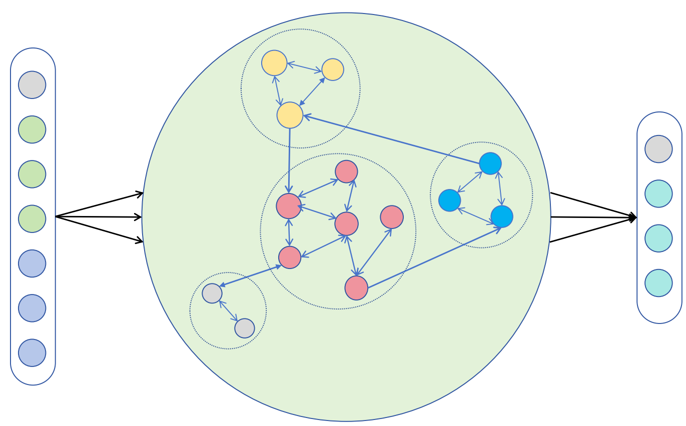

Figure 1: A trained RNN shows a modular structure to support the learning of multitaks.

During the training period, tasks were randomly selected from a set, and batches of trials for the selected task were generated and fed to the RNN model.

To characterize the learning behavior of the RNN through training, we regularly employed a widely used community discovery technique on the recurrent weight matrix to calculate its modularity (Leicht & Newman, 2008), which can be written

$$Q = \frac{1}{m}\sum_{ij}\left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]\delta(c_i, c_j) \tag{1}$$

where A_{ij} is defined to be the absolute value of W_{ij} (an element of the recurrent weight matrix of the RNN model), δ is the Kronecker delta symbol, and c_i is the label of the community to which vertex i is assigned.

Results

Multitask adaption

We study RNNs with varying sizes of recurrent layers, presenting how modularity relates to the number of tasks (Fig.2).

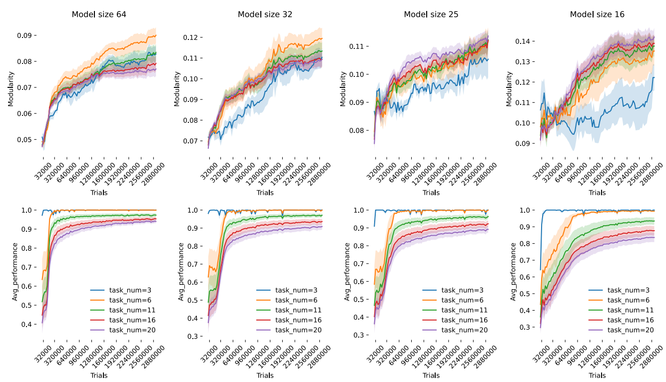

Figure 2: Modularity and the number of tasks exhibit a significant positive relationship as the model size decreases. Dark lines represent the mean of 5 runs, and light areas indicate the standard error.

We observed a significant and intriguing phenomenon: when the model scale is large, the network's capacity is sufficient to support multitasking learning (as indicated by the higher average performance in Fig.2 with model sizes of 64 and 32), and there is no apparent relationship between the network's modularity and the number of tasks. However, as the model size decreases and the units available for computation correspondingly diminish, the network's capacity drops to a level where it struggles to support multitasking learning (as shown by the lower average performance in Fig.2 with model sizes of 25 and 16). In these cases, modularity and the number of tasks exhibit a significant positive relationship. Particularly notable is the trajectory of the blue curve in Fig. 2, which corresponds to a task number of three. While this curve intertwines with others at a model size of 64, it distinctly diverges and drops below the others when the model size is reduced to 16, creating a significant margin.

This phenomenon suggests that under the constraint of limited computational units, models spontaneously develop highly modular structures to adapt to the demands of multitasking, which potentially allows for the efficient reuse of limited neural circuits, providing a greater survival advantage.

Evolving networks

To model the learning process in an evolving network, we first trained RNNs featuring 84 hidden nodes devoid of initial synapses, as illustrated by the add_conn curves in Figure 3. Synaptic connections were incrementally introduced at a rate of 10 per 500 batches, achieving a sparse topology of 800 connections by training's end. For comparative purposes, we also examined RNNs of identical scale but with 800 preformed connections, detailed in the fix_conn curves of Fig.3.

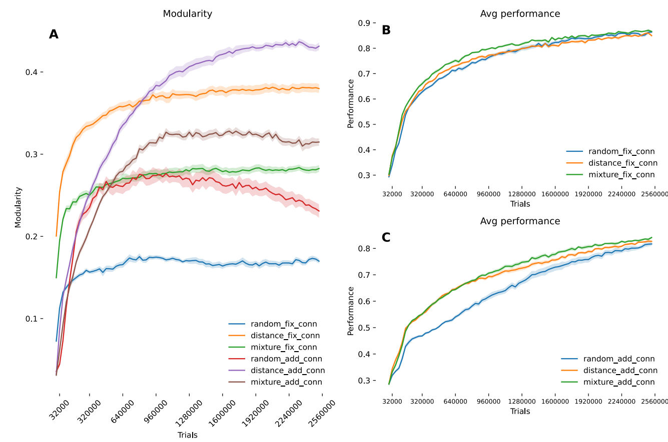

Figure 3: (A) Models that grow sequentially show higher modularity across all wiring rules. (B) Fixed models display similar performance curves. (C) Growing models are more sensitive to the degree of modularity in terms of performance. Dark lines represent the mean of 10 runs, while light areas indicate the standard error.

All connections are established according to probabilistic wiring rules. 'Distance' dictates that the connection probability is proportional to the negative power of the Euclidean distance between two nodes, utilizing a distance matrix derived from human cortical data. 'Random' ensures uniform distribution of connections. 'Mixture' denotes a hybrid approach that combines these two rules.

Discussion

In our study, we produce a novel explanation that links the emergence of modularization in neural network structures to functional demands and dynamic network growth. By demonstrating that modularity emerges as a strategic adaptation to multitasking in constrained environments, our research contributes novel insights into the interplay between the structure and function of brain networks. It posits that functional requirements reciprocally influence structural configurations, whereas the conventional paradigm, which typically regards structural attributes as the fundamental scaffold for functional capacities. The experiments on evolving networks assumed that all synapses were dynamically formed during the task-learning process. This setup universally induced a higher level of modularity and a pronounced performance disparity associated with the degree of modularity. Such findings may partly reveal the brain's propensity for high sparsity and modularity, as well as significant individual differences, under the costly constraints of synaptic growth. Moving forward, future research should explore how these principles can be applied more broadly across biological and artificial systems, enriching our understanding of neural network evolution and potentially informing the development of more sophisticated models with limited computational resources.

References

Clune, J., Mouret, J.-B., & Lipson, H. (2013). The evolutionary origins of modularity. *Proceedings of the Royal Society b: Biological sciences*, *280*(1755), 20122863.

Ellefsen, K. O., Mouret, J.-B., & Clune, J. (2015). Neural modularity helps organisms evolve to learn new skills without forgetting old skills. *PLoS computational biology*, *11*(4), e1004128.

Kashtan, N., & Alon, U. (2005). Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences*, *102*(39), 13773–13778.

Leicht, E. A., & Newman, M. E. (2008). Community structure in directed networks. *Physical review letters*, *100*(11), 118703.

Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X.-J. (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nature neuroscience*, *22*(2), 297–306.