Neuroscience and AI Share the Same Elegant Mathematical Trap

Tsvi Achler, Eyal Amir

University of Illinois at Urbana-Champaign
201 N. Goodwin Ave, Urbana IL 61801, USA

Abstract

Animals display exceptionally robust recognition abilities to analyze scenes compared to artificial means. The prevailing hypothesis in both the neuroscience and AI literatures is that the brain recognizes its environment using optimized connections. These connections are determined through a gradual update of weights mediated by learning. The training and test distributions can be constrained to be similar so weights can be optimized for any arbitrary pattern. Thus both fields fit a mathematical-statistical framework that is well defined and elegant.

Despite its prevalence in the literature, it remains difficult to find strong experimental support for this mechanism within neuroscience. Furthermore this approach is not ideally optimized for novel combinations of previously learned patterns which typically form a scene. It may require an exponential amount of training data to achieve good precision.

The purpose of paper is to 1) review the difficulties associated with this approach in both neuroscience experiments and AI scenarios. 2) Direct the reader towards 'less elegant' mathematically-difficult inherently nonlinear methods that also address both literatures (better optimized for scenes and emulating experiments) but perform recognition without optimized weight parameters.

Introduction

Though modern day studies reveal important information about the brain, for example, which regions of the brain become active, the underlying circuits are still unknown.

The unparalleled robustness of brain processing serves as a motivation for AI. Hebb's seminal work led to the phrase 'what fires together wires together'. Since then, the predominant hypothesis is that the brain performs computations based on optimized connections determined by learning-dependent synaptic plasticity. It hypothesizes small incremental and distributed changes in the networks during learning of connection weights. The networks are not dependent on single individual connections and as a biological model are resistant to injury of single units. Moreover, the networks can learn any arbitrary function as long as the training distribution is similar to the test distribution. This requirement also facilitates probabilistic analysis, tying into probability theory and well-defined mathematical analysis. This approach elegantly connects neuroscience, probability theory and a practical ability to learn any arbitrary pattern. Subsequently it is the fundamental building block for the fields of machine learning and AI.

Searching Under the Streetlight

Despite the elegance of this mechanism, two problems persist: 1) The AI methods do not seem to scale to brain function. 2) Synaptic plasticity mechanisms are still mysterious after over a half-century of experiments.

Neuroscience. Synaptic plasticity in the context of learning is the most studied phenomenon in neuroscience. A search in the Pubmed database reveals 13,000 publications.

Synaptic plasticity is assumed to occur whenever a long-lasting change in communication between two neurons occurs as a consequence of stimulating them simultaneously. The changes are labeled Long Term Potentiation (LTP) or Long Term Depression (LTD), if the responses increase or decrease respectively.

Understanding these experiments' motivation and design enables the understanding of the variability of results.

In the simplest scenario, two electrodes are introduced into two neurons within brain tissue and activation of one electrode (neuron) will cause a response in the other (note: electrodes may be placed in numerous neurons before an appropriate pair is found). In these experiments, the electrode that is used to stimulate a response is the input electrode and the electrode that records the response is the output electrode. Once this configuration is found, an activation protocol is followed. Stimulation of the input electrode is adjusted until the output neuron fires 50% of the time. The electrical settings of the input electrode that satisfy this characteristic are labeled as the *Amplitude of 50%* (A_{50}).

'Synaptic change' is induced by stimulating the electrodes simultaneously and in rapid succession. After induction is complete, A_{50} is applied again to the input and any changes in the output electrode are observed. If the output activity is greater than the original 50% then LTP occurred. If the output activity is less than 50% then LTD occurred. These changes can last for hours or days (as long as the experimental preparation is viable).

A problem with this experimental design is that the high-frequency stimulation of Long Term Potentiation induction can affect many neurons, yet their contribution to the LTP phenomena is rarely considered. The synaptic plasticity hypothesis assumes that the only important interaction is between the input and output neurons. However this is highly variable suggesting a multi-cell mechanism. In experiments in brain regions that involve sensory processing, memory and logic, there are always more neurons present than a single input and output neuron.

Thus it is not surprising that it is difficult to determine under what induction patterns LTP occurs and under what induction patterns LTD occurs. Some studies find that if the input electrode spikes are activated within tens of milliseconds before the output electrode spikes, LTP is induced. The reversed order of firing results in LTD. In other studies, the electrode of the first spike or the last spike can determine the LTP or LTD and their magnitude. Yet other studies show that modification is frequency dependent. High-frequency bursts of spikes lead to LTP, regardless of the relative input-output electrode timing. Even other studies show that spikes are not necessary for the induction of LTP and LTD (Froemke, Tsay et al. 2006). Furthermore, these properties may change during development. Thus the criteria just to induce LTP or LTD are unclear.

In summary, reliable activity-dependent plasticity relations have not been determined in sensory regions, let alone more complex learning algorithms. Clearly, robust learning occurs in the brain. However learning rules that are based on connection weights and able to learn any arbitrary pattern may not be warranted biologically or even beneficial functionally.

AI. Learning algorithms require the training distribution and the test distribution to be the same. This limits AI because the distribution is violated in the natural environment such as a scene with many patterns. If a network is trained on pattern *A* and *B* separately and they are presented side-by-side simultaneously, this is outside the training distribution. Because of distribution limits, every pair of patterns possible (or triplets, quadruplets, etc.) must be trained. This type of problem is inescapable because it occurs with objects embedded within other objects (another common natural occurrence). This combinatorial explosion of training is long known as the 'superposition catastrophe problem' (Rosenblatt 1962; Rachkovskij and Kussul 2001). This problem is mostly due to the fundamental distribution assumptions in learning and has important implications for robustness. Thus, in order to implement AGI training requirements should be relaxed, and alternatives to connection parameter optimization explored. AGI algorithms should make intelligent structured inference onto different distributions.

Steps Forward

An example of a network that is less elegant in terms of solvability, but more powerful in handling situations outside of its training distribution is a self-regulatory feedback network (Achler, Omar et al. 2008).

This method was deduced from biological studies showing an overwhelming preponderance of feedback inhibition. Self-regulatory feedback, inhibitory feedback from outputs back to inputs (also known as pre-synaptic inhibition in neuroscience and negative feedback in engineering) are found in same numbers as feed-forward connections in the brain, especially in sensory processing regions.

Its functional characteristics are difficult to derive analytically (such connections are inherently nonlinear) but

are better optimized for novel combinations of previously learned patterns, scenes (Achler, Omar et al. 2008).

Though connections are determined by supervised learning, they are not trained in a conventional sense (i.e. through parameter optimization) since there are no connection weights to optimize.

However, this network implements a surprisingly robust multiclass classifier that can process simultaneous patterns (Achler, Omar et al. 2008) addressing the superposition catastrophe problem. Furthermore the networks make complex recognition decisions based on distributed processing (Achler and Amir 2008) addressing components of the binding problem. This structure requires minimal resources (less parameters) and training.

As importantly, given the scenario of an LTP experiment, the network behaves in a similar manner to that predicted by activity-dependent synaptic plasticity. However, no connection changes are needed (Achler 2008). This approach demonstrates that by relaxing training and distribution requirements there may be more benefits than difficulties.

Conclusion

To move forward towards AGI and better understanding of brain circuits, researchers must be willing to trade elegant mathematics for nonlinear methods that are difficult to solve, but function beyond the distribution limits.

References

Achler, T. (2008). Plasticity Without the Synapse: A Non-Hebbian Account of LTP. Society for Neuroscience. Washington DC.

Achler, T. and E. Amir (2008). "Input Feedback Networks: Classification and Inference Based on Network Structure." Artificial General Intelligence **1**: 15-26.

Achler, T., C. Omar, et al. (2008). "Shedding Weights: More With Less." Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IJCNN'08).

Froemke, R. C., I. A. Tsay, et al. (2006). "Contribution of individual spikes in burst-induced long-term synaptic modification." J Neurophysiol **95**(3): 1620-9.

Rachkovskij, D. A. and E. M. Kussul (2001). "Binding and Normalization of Binary Sparse Distributed Representations by Context-Dependent Thinning." Neural Computation **13**(2): 411-452.

Rosenblatt, F. (1962). Principles of neurodynamics; perceptrons and the theory of brain mechanisms. Washington,, Spartan Books.