Project Proposal: Discovering Novel Memory Cell Arcitecutres for Recurrent Neural Networks Using Evolutionary Algorithms

Jared Murphy
Graduate Student
Department of Software Engineering
Rochester Institute of Technology
`jmm2188@rit.edu`

Travis Desell
Associate Professor
Department of Software Engineering
Rochester Institute of Technology
`tjdvse@rit.edu`

Goal and Implementation

The goal of this project is to discover new and useful memory cell architectures for improved performance in recurrent neural networks. The goal of this project will be achieved through carrying out the following steps:

1. Deconstruct current known memory cells such as GRU, LSTM, MGU and UGRNN into simple constituent parts, such as recurrent connections, activation functions etc.

2. Use simple constituent parts to replace/expand existing memory cell building blocks in the Evolutionary eXploration of Augmenting Memory Models (EXAMM) algorithm [1]

3. Evolve recurrent neural networks using EXAMM on large multivariate time series datasets from the National General Aviation Flight Information Database (NGAFID) [2] and a coal-fired power plant which has requested to remain anonymous.

4. Choose the best performing evolved models by comparing their performance to that of more traditional RNN architectures.

5. Examine network structures of chosen evolved models and identify possible novel memory cell architectures using EXAMM graphical interface.

6. Use discovered novel memory cells to create hidden layers in a traditional layered RNN, then train using the time series data and evaluate performance.

7. Compare performance to RNN's that use standard complex memory cells for hidden layers i.e. GRU, LSTM, MGU, and UGRNN

Background

Recurrent Neural Networks (RNN) are used to learn data that unfolds over time. For time series data prediction the number of possible RNN architectures is immense. Given that there is an ever increasing number of more complex memory cells to choose from, and that connections between RNN nodes can be made through time as well as between layers, doing a grid search for the optimal network architecture quickly becomes cost prohibitive.

Applied researcher expertise and heuristics are also inefficient ways to explore the design space, given that the possibilities are so vast. Automated processes that search for optimal RNN architecture are definitely needed.

In 2019, the Evolutionary eXploration of Augmenting Memory Models (EXAMM) algorithm was proposed [1]. Inspired by biological evolution, this algorithm presents a systematic way of exploring the search space of possible RNN architectures. At it's most basic level, EXAMM uses a genetic algorithm that calculates the fitness of an RNN, and then combines it with other RNNs based off their respective fitness', much like in biological natural selection. If an RNN has relatively low fitness it will be discarded, and if it performs well it will combine with other high performers so that combinations/mutations of only the best performing networks are created. In this way the search space for possible RNN architectures is explored efficiently. EXAMM allows for several complex memory cell structures to be used in network design, such as LSTM, GRU, MGU, UGRNN, and also allows for a simple node that applies an activation function to the sum of its inputs. Many different node, edge, and recurrent edge level mutations are defined for when RNN's are combined/mutated, and weights are initialized randomly from the start.

When EXAMM was used on data obtained by the National General Aviation Flight Information Database (NGAFID) [2] and a coal-fired power plant, an interesting result was obtained. It was found that when hidden node types were restricted to MGU nodes, the results of EXAMM were the overall worst. However, when simple nodes were allowed in combination with MGU nodes, the results were the overall best. In fact, it was found that allowing for simple neurons in addition to complex memory cells helped performance generally with few exceptions. With these results come two observations. The first is that many widely used complex memory cell structures are possibly incomplete and not capable of tracking some key dependencies[1]. The second is that EXAMM has the ability, in most cases, to drastically improve performance when given an RNN's most simple constituent parts as building blocks for designing a network.

These observations beg the question: If we increasingly do not trust researcher expertise and heuristics to best design the RNN as a whole, why do we trust it to design its more complex constituent parts, such as memory cells? Given that complex memory cell structures are just a collection of weights, biases, activation functions, connections, recurrent connections etc. would it not stand to reason that if provided these simple parts as building blocks, EXAMM could assemble them in ways that would perform quite well? If memory cells, such as LSTM [3], MGU, GRU and UGRNN were deconstructed and its parts given to EXAMM, is it possible that a new type of memory cell would emerge? It is the goal of this project to explore these questions and possibly answer them.

References

1. *Alexander Ororbia, AbdElRahman ElSaid, Travis Desell* Investigating Recurrent Neural Network Memory Structures using Neuro-Evolution. *GECCO '19, July 13–17, 2019, Prague, Czech Republic*

2. National General Aviation Flight Information Database (NGAFID). [Accessed Online 2023] https://ngafid.org. 2023

3. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in Neural Computation, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.