4L3-J-8-01

The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Connection-Aware Spectrum-Diversity for Neuroevolution

Danilo Vasconcellos Vargas[*1] Yuta Inoue[*2]

[*1] Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan
[*2] Graduate School of System Life Sciences, Kyushu University, Fukuoka, Japan

Spectrum-diverse Unified Neuroevolution Architecture (SUNA) is currently the most adaptive neuroevolution methods which is able to tackle different problems efficiently. This is possible by making use of (a) a unified neural model which allows a greater representation power together with (b) a new diversity metric called spectrum diversity which enables it to search in the huge search space created by the unified neural model. However, many questions remain unanswered regarding the feasibility of layers and other improved structures as well as improved diversity measures. Here we provide a study over a variation of the diversity measure. In other words, we create a connection-aware spectrum diversity. Experiments show that a connection-aware spectrum diversity allows for better results to arise over the course of evolution. This is justified by the fact that neural networks with a low number of connections are kept even when increasing the connections might improve slightly the results. Moreover, these networks themselves are easier to improve than ones with a high number of connections.

1. Introduction

Neuroevolution allow powerfull models to be evolved. This models are essentially not limited in any way. They can be non-differentiable, with circular feedback, chaotic dynamical systems, among others. In fact, SUNA recently proposed a unified neural model that unified most if not all of the features of state-of-the-art neural networks [Vargas and Murata2017]. I.e., an unified neural model with neuromodulation, activation/inhibition neurons, slow/fast neurons, feedback, circular feedback, memory states, different activation function, among other features. These set of features allowed the algorithm to find the best set of features as well as the topology and weights for a certain problem, allowing for fast and accurate adaptation.

Unified neuron models may have all the features that would able a network to better adapt to a specific scenario. However, a consequence from unified neuron models is a huge search space. Spectrum-diversity was proposed and shown capable of dealing with such a huge search space. This was possible using niching by spectrum and not allowing candidate solutions to compete with other solutions from different niches (In fact, this idea was called niched/relative fitness because the fitness is indeed separated and dependent on the niche (or subpopulation) [Vargas et al.2013], [Vargas et al.2013], [Vargas et al.2014], [Vargas et al.2015b]. Isolated subpopulations solve many conflicts which lie inside algorithms with a single population (panmictic populations)[Vargas et al.2015a].) However, the spectrum-diversity used did not take the number or types of connections into account. Thus it may create a tendency towards networks with a high number of connections.

In this work, a connection-aware spectrum diversity is investigated. Recently, spectrum-diversity showed that even a small population could deal with large search spaces if the spectrum of the candidate solution is considered instead the candidate solution itself. Here we investigate the addition of connection to the normalized equation of the spectrum.

Contact: Danilo Vasconcellos Vargas, Kyushu University, Fukuoka, vargas@inf.kyushu-u.ac.jp

1.1 Spectrum-diverse Unified Neuroevolution Architecture

The unified neural model (UNM) has allows for a great flexibility to deal with various problems. However, it also means that it has many dimensions and a huge search space.

Spectrum-diverse Unified Neuroevolution Architecture (SUNA) [Vargas and Murata2017] tackled the problem of search in such a huge search space by proposing a new diversity method called spectrum-diversity. Spectrum-diversity divide the candidate solutions into niches (subpopulations) based on their spectrum and disallow competition between niches (subpopulations). A spectrum is built by extracting from UNMs important information such as the numbers of neurons for each neuron types, number of neurons for each type of activation function, number of neurons with each adaptation speed, among other features. SUNA treat networks with only its spectrum not their structure to reduce dimensions of search space. Moreover, this also define the novelty in terms of approaches (which tools are used in the solution) rather than the solution (the detailed methods and steps). This type of novelty is a new paradigm for neuroevolution.

For evaluating the spectrum, SUNA uses the novelty map [Vargas and Murata2017]. Novelty map is a table to keep most novel elements. In this map, the novelty of a node is considered as the distance from other nodes. The process to add a new item is shown below. Consider P the size of the novelty map and $|P|_{max}$ is the maximum population of the map. The dynamic is described explicitly as follows:

1. If the current population number $|P| \leq |P|_{max}$, add the new item.
2. In the population $|P| > |P|_{max}$,
   (a) If the shortest distance between the new item and elements in the current map is longer than the

longest distance between elements in the current map, one of elements has the shortest distance is removed from the map and the new item is added to the map.

(b) Else then, the current elements is kept and not to add the new item to the map.

SUNA divides the candidate solutions into subpopulations using a novelty map [Vargas and Murata2017] which evaluate their uniqueness and classify into a given number of classes with respect to their difference in spectrum. In each evolution step, all candidate solutions are evaluated and for each subpopulation only the candidate solution with the highest fitness survive for the next generation. The surviving candidate solutions are selected to mutate for making the next population. In other words, the architecture judges the likenesses between networks and gives chances based on the spectrum, allowing different approaches to co-evolve even when there are gaps in fitness.

2. Connection-aware SUNA

In this section both the definition and the settings of the proposed method is discussed in detail.

2.1 Connection Spectra

In this paper, the addition of number of connections in the spectrum calculation of the original SUNA algorithm is considered. In the original SUNA, the spectrum is made of a histogram in which every bin is the number of nodes which share the same activation function. This simple spectrum contributes to preserving its diversity and is shown to work overall in the original paper. However, here we raise the hypotheses if this simple mechanism may not cause early on a high number of connections to appear. By adding connection count to spectra, it is possible to preserve the diversity regarding the number of connection as well. That is, we expect that this modification would allow simple networks to survive easierly and enable SUNA to find more compact solutions.

However, the careless addition of the number of connections to the formula causes the spectrum to become unbalanced, i.e., to easily become dominated by the huge number of connections which are usually more numerous than neurons. Consequently, when the difference between spectra is calculated, this would cause the diversity to focus on connections alone. In other words, some type of normalization is needed. To enable spectrum-diversity to work independent of the scale of connection and neurons the following normalization is added:

$$spectra^{ntype} = \frac{n^{ntype}}{N^{inModule}} \tag{1}$$

$$spectra^{connect} = \frac{n^{connect} - N^{connect}_{min}}{N^{connect}_{max}}, \tag{2}$$

where n^{ntype} is the number of neurons for each type of neuron (identity, threshold, random, sigmoid and control) as well as slow adapting neurons (i.e. neurons whose adaptation speed is more than one). $n^{connect}$ is the number of neurons of a type or connections and $N^{connect}_{min}$ and $N^{connect}_{max}$ are the minimum or the maximum number of the same item of n calculated over all of the individuals in the current population. In connection-aware SUNA, the spectrum of every agent has seven lengths totally and the given normalizations all values for each bin have a maximum of 1 and a minimum of 0.

3. Experiments and Discussion

Here two experiments will be run over the proposed connection-aware SUNA.

3.1 Experiment 1: Learning Caesar Cipher

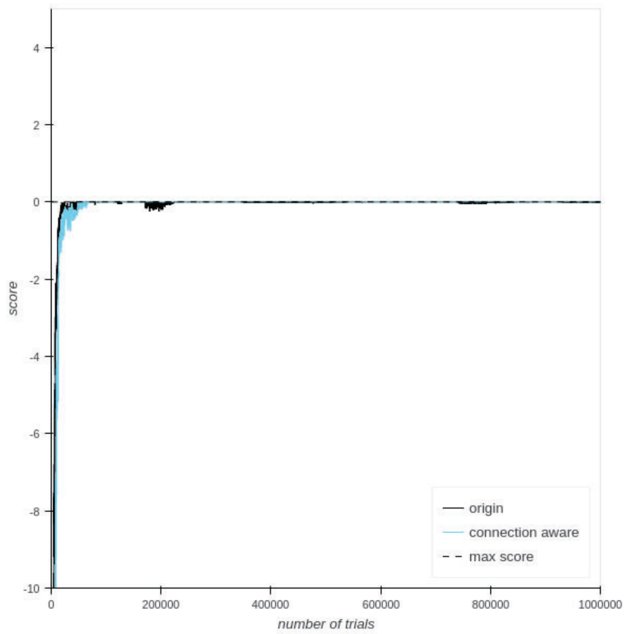

Figure 1: Result of origin, connection-aware models in Caesar Cipher. Lines are the mean of 100 trials. Variances are standard deviation.

Caesar cipher is the an encryption in which each original letter (message) is shifted by a certain number (key). The objective of the agent is to decode the cipher. In this environment, each agent receives as input both the original letter (message) coded as an integer and the key. The aim of the agent is to have an output that matches the output of the cipher.

The range of letter number and key number were from 0 to 25. Each trial consists of 10000 steps. The letter is randomly decided each step while the key is updated every 100 steps. In each step, reward was calculated based on the agent's action a as $-|a - (letter + key)|$. All of the agents were evaluated using the sum of reward in whole steps.

Regarding the results, both original and connection-aware SUNA reached the maximum score of 0 at almost same trials (Fig. 1). However, the original model presented some instability after the convergence in comparison with the connection-aware version. Here we raise the hypotheses that fewer connections may lead to less complex solutions and consequently more stable ones. In fact, it will be shown in the next Section that the preservation of networks with fewer connections is achieved by connection-aware SUNA while SUNA has a tendency for high number of connections.

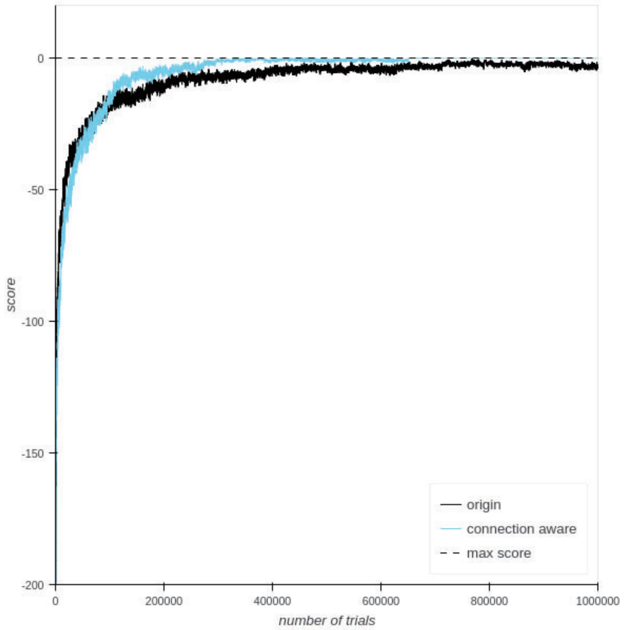

Figure 2: Result of origin, connection-aware models in Caesar Cipher. Lines are the mean of 100 trials. Variances are standard deviation.

3.2 Experiment 2: Learning Double Key Caesar Cipher

In this experiment, the number of the key used in the Caesar Cipher experiment was divided into two numbers with their multiplication resulting in the original key. The range of each key numbers was set from 1 to 5 (i.e., $\sqrt{25}$). The number of the message and other parameters was set to be the same as the Caesar Cipher experiment. All of the 100 individuals (agents) of the connection-aware model reached the maximum score 0. However, 11 agents of the original SUNA did not.

One of the simplest solutions can be seen in Figure 3. This is a simple and effective solution, however finding it requires a small number of neurons and connections. Therefore, it makes it harder for the original SUNA to reach such solutions in comparison with connection-aware SUNA. Moreover, fewer connections also force the number of functional neurons (i.e. connected neurons that have some influence in the network's behavior) to drop therefore it is also indirectly related.

4. Conclusions

This paper proposes a variation of SUNA which enables its diversity to be aware of connections. The experiments demonstrate that such a modification allows SUNA to preserve candidate solutions which have are a more compact

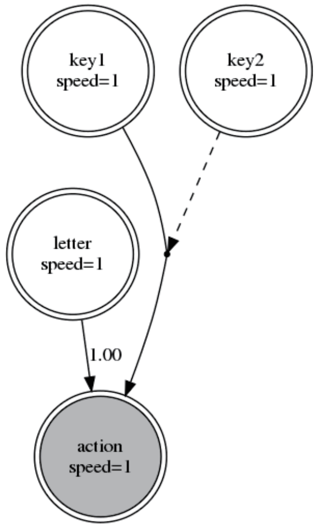

Figure 3: Typical topology of the network solves the Caesar Cipher with divided 2 keys. Key1 and key2 are multiplied as neuromodulation and added to message.

topology. At least for the experiments here, such a modification increased the quality of solutions as well as learning speed of SUNA.

References

[Vargas and Murata2017] Vargas, D. V. and Murata, J. (2017). Spectrum-diverse neuroevolution with unified neural models. *IEEE transactions on neural networks and learning systems*, 28(8):1759–1773.

[Vargas et al.2015a] Vargas, D. V., Murata, J., Takano, H., and Delbem, A. C. B. (2015a). General subpopulation framework and taming the conflict inside populations. *Evolutionary computation*, 23(1):1–36.

[Vargas et al.2013] Vargas, D. V., Takano, H., and Murata, J. (2013). Self organizing classifiers: first steps in structured evolutionary machine learning. *Evolutionary Intelligence*, 6(2):57–72.

[Vargas et al.2014] Vargas, D. V., Takano, H., and Murata, J. (2014). Novelty-organizing team of classifiers-a team-individual multi-objective approach to reinforcement learning. In *SICE Annual Conference (SICE), 2014 Proceedings of the*, pages 1785–1792. IEEE.

[Vargas et al.2015b] Vargas, D. V., Takano, H., and Murata, J. (2015b). Novelty-organizing team of classifiers in noisy and dynamic environments. In *Evolutionary Computation (CEC), 2015 IEEE Congress on*, pages 2937–2944. IEEE.