Evaluation Methods for Active Human-Guided Neuroevolution in Games

Igor V. Karpov, Leif Johnson, Vinod K. Valsalam and Risto Miikkulainen

Dept. of Computer Science, The University of Texas at Austin
1616 Guadalupe, Suite 2.408, Austin, TX, 78701 USA
{ikarpov,vkv,leif,risto}@cs.utexas.edu

Abstract

Machine learning (ML) games such as NERO incorporate human-guided ML methods such as real time neuroevolution (NE) as an integral part of the gameplay, i.e. by allowing the player to train teams of autonomous agents to compete with those trained by others. In order to improve human-guided ML, a way to systematically compare and validate new such methods is needed. To this end, this paper describes the results of a human subject study comparing human-guided ML methods and an online tournament validating them at scale. Additionally, this paper describes ongoing work to extend human-guided NE methods through active advice, examples and shaping and to combine these modalities into a more flexible and powerful overall system for agents in games.

A *machine learning game* relies on machine learning for part of its gameplay, allowing the user to train agents or teams of agents which then e.g. compete with those trained by other players. *Human-guided machine learning* (HGML) methods are methods that are well suited for such games because they allow natural and effective interaction with the human player during the training process. These methods can also be used by game developers to create desired controllers at game design time. One concrete example of a machine learning game is the NeuroEvolving Robotic Operatives (NERO) game, with human-guided machine learning via the real-time NeuroEvolution through Augmenting Topologies (rt-NEAT) method (Stanley, Bryant, and Miikkulainen 2005).

Human Subject Study

Machine learning games can be employed to systematically evaluate and validate human-guided ML methods in order to improve them. In previously reported work, the authors conducted a human participant study comparing several distinct modalities of guiding neuroevolution in ML games - namely by providing advice using a scripting language, by providing examples of desired behavior, and by modifying the environment to shape the learning process (Karpov, Valsalam, and Miikkulainen 2011). These three modalities (advice, examples, and shaping) were compared with two baselines–building behaviors through manual scripting and evolving
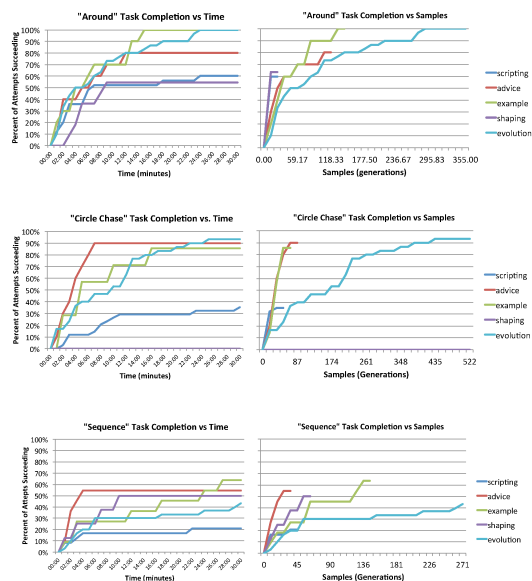
Figure 1: Performance of the human-guided neuroevolution methods, neuroevolution, and manual design on the three design tasks tested. In all cases, a human-guided NE method achieves success faster and with fewer samples than either manual design or hands-off NE. However, the modality that works best depends on the task.

them through hands-off neuroevolution. The comparisons were made on three different tasks which often constitute subtasks during organic NERO training. The results indicate that human-guided NE methods work better than either NE alone or manual design alone, but that the modality that works best depends on the task at hand (Figure 1).

Machine Learning Game Tournament

While such careful smaller-scale comparisons can provide insights into how to improve human-guided ML methods, the improved methods should also be validated on a larger scale in order to study the variety of approaches taken by human teachers and the natural interactions that occur during the training process. In order to do this, a special tournament format, a *machine learning game tournament*, was de-
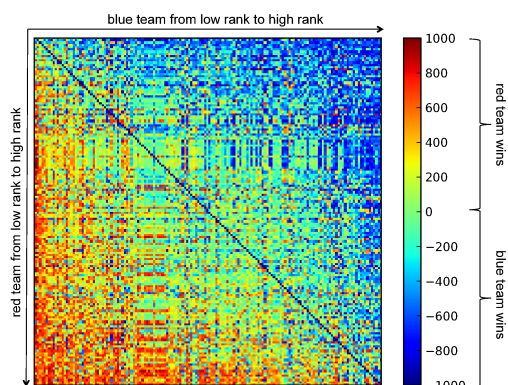
Figure 2: Results from the round-robin NERO tournament. Teams are sorted by average score differential over all matches. Rows and columns in the matrix represent teams in the tournament, and colors represent score differentials for the respective individual matches between two teams. Even the top ranked teams lose some matches, and the bottom ranked teams win some matches.

veloped and tested. In such a tournament, participants train their teams using an array of available human-guided machine learning methods and then submit their teams to compete in e.g. virtual combat tasks.

In the initial run of the tournament, about 85 students in an on-line AI course used an open source implementation of human-guided NE and a temporal difference reinforcement learning (RL) method to train and submit 156 teams of agents. The tournament submissions turned out to be varied both in the training strategies used by participants and in the resulting teams, and the task was challenging enough that no team was absolutely dominant or absolutely dominated (Figure 2).

Active Advice, Examples and Shaping

Ongoing work aims to extend the individual modalities with the ability to query the player for appropriate pieces of advice, examples or shaping. This approach is similar to active learning (Settles 2009), where the machine learning algorithm selects the pieces of data that need to be labeled by the oracle (human).

In active advice, partially filled out advice templates are generated, and the user is asked to fill in the blanks. The templates are generated based on the previously provided advice, the fitness of the individuals within the population using that advice, and the degree to which each piece of advice is incorporated by the population.

In active examples, a query is represented via a starting location for the human-controlled agent. The initial location of the example query is determined based on the state space explored by the current population and the amount of reward received in those locations. Because the problem setting and the NE method are both population based, it is also possible to direct the active selection of examples in areas of maximal disagreement between candidate solutions as in query by committee (Seung, Opper, and Sompolinsky 1992).
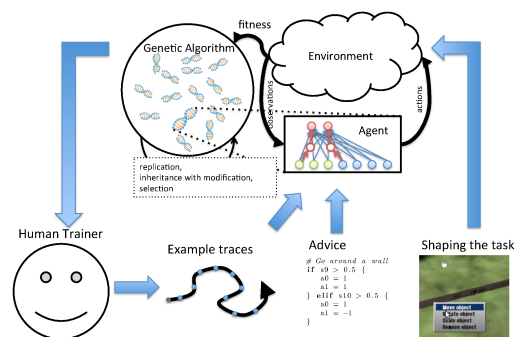

Figure 3: A diagram of the overall active neuroevolution system, showing the population of evolving agents, the different modalities of human training, and the active feedback.

In active shaping, the training system determines when learning is no longer taking place, and asks the player to provide it with the next step in the shaping process.

Integrated Active Neuroevolution

In order to provide the most flexible system for training game agents interactively, a combination of the training modalities (advice, examples, and shaping) can be used (Figure 3). Queries from individual modalities are viewed as actions in a Markov decision process, and the system learns to ask for the right kind of input via reinforcement learning.

Evaluation and Validation of Active Neuroevolution

The overall system will be evaluated in several ways. First, the active advice, example and shaping modalities of human-guided neuroevolution will be compared to the other human-guided, hands-off and manual design methods. Second, a tournament will validate the overall combined approach. During this tournament, user actions will be recorded during training, allowing for further improvements of the system by learning to ask queries that lead to successful solutions faster.

References

[Karpov, Valsalam, and Miikkulainen 2011] Karpov, I. V.; Valsalam, V. K.; and Miikkulainen, R. 2011. Human-assisted neuroevolution through shaping, advice and examples. In Krasnogor, N., and Lanzi, P. L., eds., *Proceedings of the 13th Annual Genetic and Evolutionary Computation Conference*, GECCO '11. Dublin, Ireland: ACM.

[Settles 2009] Settles, B. 2009. Active learning literature survey. Technical report, University of Wisconsin-Madison.

[Seung, Opper, and Sompolinsky 1992] Seung, H.; Opper, M.; and Sompolinsky, H. 1992. Query by committee. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 287–294. ACM.

[Stanley, Bryant, and Miikkulainen 2005] Stanley, K. O.; Bryant, B. D.; and Miikkulainen, R. 2005. Real-time neuroevolution in the NERO video game. *IEEE Transactions on Evolutionary Computation* 9(6):653–668.