Creative Agent Reasoning through Adaptive Neural Networks

Alysson Ribeiro da Silva[1], **Luís Fabrício Wanderley Góes**[1]

[1]Pontifícia Universidade Católica de Minas Gerais (PUC MINAS)
Caixa Postal 1.686 – 30.535-901 – Belo Horizonte – MG – Brazil

alysson.silva@sga.pucminas.br,lfwgoes@pucminas.br

Abstract. *In this research we proposed the Honing Adaptive Resonance Process (HARP) to enable Adaptive Neural Networks to output creative decisions or artifacts. The HARP is mainly the fusion of The Honing Theory (HT), which describes how the human brain generates creative ideas, and The Adaptive Resonance Theory (ART), that describes how the human brain stores and retrieves information. It is composed of a set of Adaptive Neural Networks (ANN) we design to handle creativity as described by the HT. Each proposed ANN is primarily based on the ART, and they enable an agent to generate valuable strategies to solve complex problems with temporal learning techniques such as the Q-Learning and reactive models. In order to enable an agent to perform creative actions, we primarily created a symbolic computational model of the HT, using the Bayesian surprise, which makes the ANNs learn not only valuable strategies, but also new ones. The proposals were deployed in agents we created, called HoningStone and HearthBot, that are able to generate creative card combos and play the digital collectible card game HearthStone, respectively. Our results show that HoningStone was able to generate combos more creative than a handcraft heuristic. Moreover, the HearthBot was evaluated in terms of win rate performance when playing against the Monte Carlo Tree Search (MCTS) and a Board Control Greedy (BC-Greedy) heuristics, overcoming them and also its own non-creative baseline versions.*

1. Introduction

In the recent years, neural networks have been used successfully to deploy reasoning into agents [Silver et al. 2016]. However, little effort is dedicated to deploy on them the ability to develop creative behavior. Most of the solutions that deploy it on automated systems, such as agents, are not based on how the human brain works and are more commonly related to fields such as fashion, painting, and culinary [Amorim et al. 2017]. Differently from the aforementioned works, this research aims at enabling agents to output creative decisions or artifacts. This creative behavior is desired to solve complex problems in uncertain environments.

Agent control is composed by a set of tasks that are hard to achieve, since an agent needs to sense and model its reality, organize what it is perceiving in a search space, search for a solution, apply or perform it, adapt to uncertainties, care on how to perform its movements, and deal with constraints [Silva and Machado 2016]. Those tasks are usually solved by optimization techniques, such as the Q-Learning, to extract valuable strategies from a search space modeled with states in a Partially Observable Markov Decision Process (POMDP). Neural network models can be used to help in solving those tasks, such as Deep Neural Networks [Silver et al. 2016]. However, Deep Learning approaches tend to possess an extremely high computational cost, demanding more energy, time to process and learn information. On the other hand, Adaptive Neural Networks (ANN)

[Teng and Tan 2015], can solve control problems, helping machines to develop advanced cognition, through a Fusion Architecture for Learning COgnition and Navigation (FALCON), in a faster and more stable fashion, and tend to cost much less in terms processing power and computation time. They are based on the Adaptive Resonance Theory (ART), that describes how the human brain stores and retrieves information, in which mostly of this research proposals were inspired.

In order to equip a machine with the ability to be creative is a two-fold problem since defining what is creativity is hard but also unveiling what are the mechanisms inside the brain that promotes it is even harder. In the first front, much progress has been made in the field of Computational Creativity, in which scientists now tend to agree that a creative idea has to be novel but also useful. Quite a few general purpose metrics have been proposed to evaluate novelty, such as the Bayesian surprise, and also many context-specific ones to measure usefulness in art, music, culinary, games. In the second effort, many theories have been proposed to explain how creativity occurs, from all of them, we would like to highlight The Honing Theory (HT), created by Liane Gabora [Gabora 2010], which proposes that a creative idea is the result of an alternating process between associative and analytic modes within the brain. This alternation of modes, alongside with the world perception, eventually leads to a creative idea.

The main proposal of this research is called the Honing Adaptive Resonance Process (HARP), and it aims in enabling agents to output creative decisions or artifacts to solve complex problems. The HARP is an emergent neural process made of the fusion of the HT (Honing Theory) and the ART (Adaptive Resonance Theory). It is composed of a set of ANNs we design to handle creative thinking. Each proposed ANN is primarily based on the ART, and they enable an agent to generate valuable strategies to solve complex problems with the reactive and Q-Learning temporal learning techniques. As accomplished by [Teng and Tan 2015], a FALCON topology was also used to handle the agent's control problems modeled as a POMDP. To incorporate creative thinking, as described by the HT, we firstly created a symbolic computational model of it, what was never attempted before. This model also utilizes the Bayesian surprise to enable the proposal to evaluate or search for novelty. Finally, our the HT proposal was further extended and incorporated into the HARP.

To evaluate our proposals, they were deployed in agents we created, called HoningStone and HearthBot, that are able to generate creative card combos and to play the digital collectible card game HearthStone, respectively. Our results shows that HoningStone was able to generate combos that are more creative if compared to a handcrafted heuristic based on randomized decisions. Those results were reinforced as being more creative by [Ramos et al. 2016], which focused in evaluating them through human computation. On the other hand, the HearthBot was primarily evaluated in terms of win rate performance when playing against the Monte Carlo Tree Search (MCTS) and a Board Control Greedy (BC-Greedy) heuristics. Our results show that it was able to overcome the MCTS, the BC-Greedy, and non-creative versions of itself by obtaining a superior win rate performance for most of the cases.

2. Contributions

The main contributions of this research are summarized as follows.

- The first symbolic model of the HT (Honing Theory) based on Knowledge-Based Systems inspired on a Greedy Randomized Adaptive Search Procedure (GRASP).

This contribution was published in [Góes et al. 2016] and contributed to three other Master's Thesis which their results were published in [Junior et al. 2016], [Ramos et al. 2016], and [Amorim et al. 2017].

- The HARP (Honing Adaptive Resonance Process), composed by the proposed ANNs described in Section 3, is the fusion of the ART (Adaptive Resonance Theory) and the HT (Honing Theory), and it enables an agent to develop creative strategies to solve complex problems.
- The Expectation ART ANN. It enables agents to sense the novelty of a decision when exploring a search space. It incorporates the Bayesian surprise into its dynamics, to hold temporal information considering its experience, to perform advanced cognition and decision making that can lead to novel solutions.
- The Proximity ANN. Published in [Silva and Góes 2017], it enabled our agents to model and store information in a compact way, about its environments, with a precise categorization method. It is mainly used to enable the HARP to spend less space to store received stimulus from an agent's environment.
- The Unstructured Areas Multi-channel Adaptive Neural Network (UAM-ANN). It is a novel ANN able to represent semantic information hierarchically, in a similar fashion to Deep Convolution Neural Networks, and it was primarily designed to handle the large amount of information provided by the Expectation ART.
- Behavioral and non-behavioral feature and action models for the digital collectible card game HearthStone. It is proposed in this research handcrafted and atomic feature representation for ANNs. Moreover, are also proposed ways to represent its reward functions used by the Q-Learning and reactive temporal optimization methods.
- A benchmark evaluation of the Metastone simulator, where it was used in this research to evaluate all deployed agents when learning and playing HearthStone.
- Deployment and creation of HearthStone agents, able to learn and play HearthStone, simulated by Metastone, and used to evaluate all proposals.

Our secondary contributions include a performance and complexity analysis for the proposals, an analysis for the Bayesian surprise behavior for decision making when controlling agents, and practical examples of the ART when deployed in computational systems.

3. The Honing Adaptive Resonance Process

The HARP (Honing Adaptive Resonance Process) enables an agent to generate creative strategies, novel and valuable, to solve complex problems. Value is obtained through a trail reinforcement mechanism performed by the temporal learning method, that helps in distinguishing which strategies, sequence of POMDP states, are good or bad to follow. In contrast, novelty is achieved by forcing the temporal learning method to reinforce states that were less visited, which is achieved with the Bayesian surprise. Creative strategies emerge from the HARP as a consequence of a cyclic trail reinforcement and novelty exploration phases iteration. Due to its nature, the HARP can explore better a search space, and yet be able to achieve value due to its trail reinforcement mechanisms.

As exemplified in Fig. 1, the HARP is formed by our proposed ANNs. Its main structure is formed by the UAM-ANN, where it is internally divided into many semantic hierarchies of neuron clusters. They are exemplified as the big, red and blue, rectangles and are used to store information about POMDP states. Each neuron cluster is organized inside activation areas depicted as smaller rectangles. They are organized in a tree

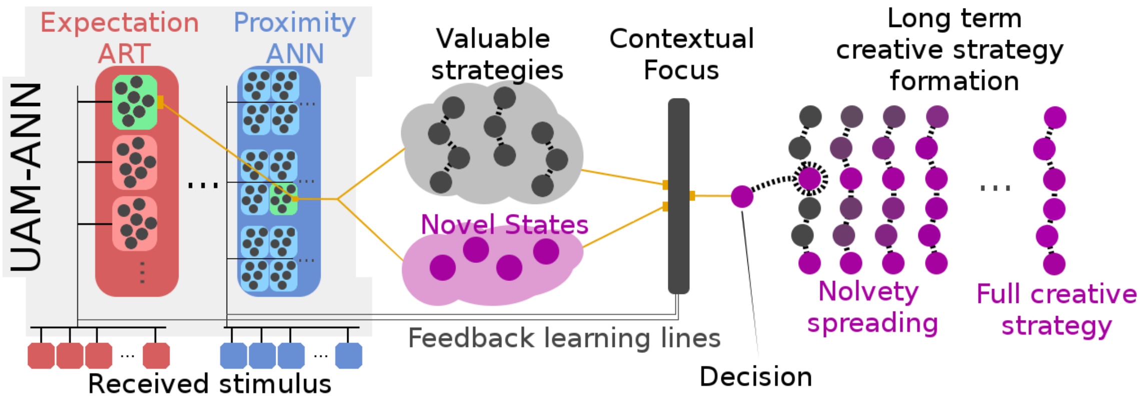

Figure 1. The HARP creative strategies formation scheme.

structure, created by links between hierarchies, thus enabling information sharing, which is essential to handle large amounts of information. To generate a solution, the HARP receives external stimulus for each hierarchy, depicted as small, blue and red, squares, received by an agent. By performing its routines, an activation path, depicted in yellow, is generated through all the UAM's semantic hierarchies activating the green areas. In its basic form, two predictions are achieved as a response. The first prediction, extracted from the full activation of the UAM, is part of a valuable strategy, and the second one, extracted only from the Expectation ART activation, is novel. Both predictions pass through a module called Contextual Focus, described by the HT as the ability to decide if the full valuable strategy will be followed or a novel state will be explored instead, and only one of them is used as a final response. This response is finally used by an agent to interact with its world.

In a long-term view, creative strategies are formed by the HARP, as a consequence of the temporal learning method iteration on the newly generated ones, after performing a certain number of responses obtained from the Contextual Focus. Since the HARP can decide wetter to be novel or efficient, hybrid valuable strategies are consequently formed. For instance, as depicted in Fig. 1, a hybrid valuable strategy is depicted as the composition of novel states, represented as purple circles, into valuable strategies, represented as gray paths. The incorporated novel states will influence the optimization performed on the valuable part of the strategies. As a consequence, a creative strategy will emerge from a novelty spreading process. It indicates how much of the novel states are influencing the value of the strategy being optimized. When the influence of the novel states reach a threshold, the consolidated strategies are finally considered to be creative.

4. Experimental setup

In the digital collectible card games universe, *HearthStone* is one of the most played games in the last year [Blizzard 2018, Góes et al. 2016]. The game mechanics, fully described by [Silva 2017] and [Góes et al. 2016], are summarized in alternating turn matches between two players, where each player tries to destroy a controllable avatar called hero from the opposite player by playing cards that represent minions and special abilities. According to our estimations, HearthStone posses around $\sum_{p=0}^{m}(2.85 \times 10^{51})^p$ possible paths, where p is the size of the path and m the maximum allowed path size. This turns the game hard to solve, what motived us in using it to evaluate our proposal, since it represents a number greater than the estimation for the GO game [Silver et al. 2016].

Our experiments are divided into three parts. The first one focus in evaluating the ability of the proposed symbolic model of the HT, deployed in an agent called HoningStone, when generating creative card combos for HearthStone. Next, a non-creative version of the HearthBots was evaluated, in the Metastone HearthStone simulator, when using only the Proximity ANN. This HearthBot in special was deployed in a tesla GPU, that handles up to 4 million neurons, in order to verify its ability to learn and play, against known and unknown enemies based on the MCTS, without temporal learning methods. For the last part, we developed eight HearthBot versions, where four of them learn from scratch and other four used a behavioral action model. For the models that learn from scratch, one of them is a baseline that incorporates the basic reactive and Q-Learning temporal learning methods found in [Teng and Tan 2015]. Two of them incorporate only the Expectation ART or the UAM-ANN, and the last one incorporate the full HARP. For the behavioral version, those configurations were replicated. They were primarily evaluated in terms of win rate performance, also in Metastone, when playing against known and unknown enemies based on the MCTS and BC-Greedy, learning convergence, search space exploration capabilities, and storage space used by the UAM-ANN's neuron clusters. Moreover, we also conducted complementary experiments to evaluate their performance, in terms of execution time, Bayesian surprise behavior, and convergence.

The second part of our experiments was conducted in a tesla K20 equipped computer. On the other hand, the first and last parts were conducted in a personal laptop and do not rely on GPU computing. To obtain our results, for all three parts, we performed an exhaustive number of simulations to avoid statistical fluctuations. For HoningStone, 1440000 simulations were performed, 324000 for the GPU-based HearthBot, and more than 626300 for all other HearthBots.

5. Results

Our results show that HoningStone can generate combos that are more creative than a greedy randomized algorithm. Although the combos generated by HoningStone are certainly new and playable from an expert point of view, their efficiency has yet to be proven in real games. Finally, the behavior generated from HoningStone indicates that the adapted GRASP can still guide the objective function towards a better solution, being local or global optima, after its adaptation to implement the described processes by the HT.

With respect to the GPU based HearthBot, our results show that it was able to overcome the MCTS by obtaining an average win rate performance above 80% for some cases using only 4% of the tesla's capacity. When playing against unobserved enemies, it obtained an average win rate performance above 60% for most of the cases. On the other hand, when playing against observed enemies, the full HARP version of the Creative HearthBots that learns from scratch obtained an average win rate performance above 90%, and its behavioral version obtained an average win rate above 95% for most of the cases. When playing against unknown enemies learning from scratch, it obtained an average win rate above 90%, and its behavioral version obtained an average above 95%. If compared to our baseline, the full HARP Creative HearthBot used 3000 fewer neurons to reach better solutions and it explored better the search space by a factor of 2.

6. Conclusions

This paper summarizes the Master's Thesis developed by [Silva 2017], in which the main contribution proposes the HARP, the fusion of the HT (Honing Theory) and the ART (Adaptive Resonance Theory), that empowers agents to output creative decisions or artifacts through ANNs (Adaptive Neural Networks). Our proposals were evaluated in agents called HoningStone and HearthBot that are able to generate creative card combos, learn and play the digital collectible card game HearthStone, by forming creative strategies, respectively.

Other domain, besides HearthStone, could be explored in order to verify the behavior of the proposals. For example, the proposed HARP could assist astronauts in generating novel solutions when confronting unpredictable situations. It can help doctors in finding novel and useful solutions to deal with the unforeseen, help physicists to develop and advance new theories about the fabric of existence, and could care for novel ways to deal with climate change. Those applications are future research topics of our interest focusing on the improvement of the quality of life for the human kind.

References

Amorim, A., Góes, L. F. W., Silva, A. R., and Junior, C. R. F. (2017). Creative flavor pairing: Using rdc metric to generate and assess ingredients combinations. In *International Conference on Computational Creativity*, pages 33 – 40.

Blizzard, E. (2018). Available at: https://us.battle.net/hearthstone/en/.

Gabora, L. (2010). Revenge of the "Neurds": Characterizing Creative Thought in Terms of the Structure and Dynamics of Memory. *Creativity Research Journal*, 22(1):1–13.

Góes, L. F. W., Silva, A. R., Rezende, J., Amorim, A., Franca, C., Zaidan, T., Olimpio, B., Ranieri, L., Morais, H., Luana, S., and Martins, C. A. P. S. (2016). Honingstone: Building creative combos with honing theory for a digital card game. *IEEE Transactions on Computational Intelligence and AI in Games*, PP(99):1–1.

Junior, C. R. F., Goés, L. F. W., Amorim, A., Rocha, R., and Silva, A. R. (2016). Dependent creativity: A domain independent metric for the assessment of creative artifacts. In *International Conference on Computational Creativity*, pages 68 – 75.

Ramos, S. L., Goés, L. F. W., Ponciano, L., França, C., and Morais, H. (2016). Avaliação da percepção de jogadores sobre a criatividade de combos do jogo digital de cartas hearthstone. In *Brazilian Computing Society conference on Games (SBGames)*, pages 190 – 193.

Silva, A. R. (2017). Creative agent reasoning through adaptive neural networks. Master's thesis, Pontifícia Universidade Católica de Minas Gerais.

Silva, A. R. and Góes, L. F. W. (2017). Hearthbot: An autonomous agent based on fuzzy art adaptive neural networks for the digital collectible card game hearthstone. *IEEE Transactions on Computational Intelligence and AI in Games*, PP(99):1–1.

Silva, A. R. and Machado, A. M. C. (2016). Control of mobile robots with amorphic architecture. *IEEE Latin America Transactions*, 14(7):3093–3101.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489.

Teng, T. H. and Tan, A. H. (2015). Fast reinforcement learning under uncertainties with self-organizing neural networks. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, volume 2, pages 51–58.