ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 7| July 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

FLAPPY BIRD AUTOMATION USING REINFORCEMENT ALGORITHM

O S Sumukh[1], Prinson Fernandes[2], Merin Meleet [3]

[1,2,3]*Department of Information Science and Engineering, R V College of Engineering*
Mysore Road, Bengaluru - 560059

ABSTRACT

One of the most popular subjects being investigated in AI nowadays is game learning. Addressing such issues require proper domain specific knowledge.So one such game was developed ie., flappy bird where the agent learns itself on how to avoid the obstacles and also tries to maximize the score based on the rewards and punishments it receives. No prior knowledge was given to the agent regarding the environment. Instead of utilizing raw pixels, the agent was trained using domain-specific features such as the bird's speed, the distance between pipes, and the height of the pipes, which significantly simplifies the feature space and avoids the need for deeper models to automatically extract underlying data. The agent was trained using the NeuroEvolution of Augmenting Topologies(NEAT) algorithm and is talked about in this paper.

KEYWORDS-*NEAT,feature extraction,pygame,reward,fitness,reinforcement learning*

I. INTRODUCTION

With the introduction of video games during the past ten years, this field of study has seen a significant extension and enrichment, allowing us to examine a wider variety of issues of significant commercial, social, economic, and scientific significance..With the increase of usage of phones there has also increase in the number of people playing mobile games which include the flappy bird game.Dong Nguyen, a Vietnamese video game programmer and designer, created the addictive game Flappy Bird through his company dotGears.[2] .The mobile game was introduced in May 2013, and many people praised it for its high level of difficulty while also criticizing it. The game appears to be extremely straightforward, but because of the rapid game dynamics, significant environment diversity, and enormous search space, the scores are typically low[1]. This paper discusses how to automate such a game using reinforcement learning. When we need an agent to carry out a task but there isn't a single "right" method to do it, reinforcement learning can be helpful[4]. It is about taking suitable action to maximize reward in a particular situation.NEAT algorithm is one such reinforcement learning algorithm,which we have used to train the bird and will be discussing in this paper.

II. METHODOLOGY

In the first step, the flappy bird game was built using the pygame library available on Python. A slight modification to the game was done. The speed of the pipes coming towards the bird increases as the bird progresses through the game. This modification was done to make sure that the bird learns to adapt to varying speeds of the game and be able to encounter new examples much often.

NEAT reinforcement learning algorithm was used to enable the Reinforcement Learning(RL) agent to play the game. The NEAT algorithm is a genetic algorithm which is a heuristic method for solving a wide range of optimization issues by simulating the process of natural selection. It uses the concepts of selection, crossover and mutation to get the best possible aspects from the previous generation and use it in the current generation. Each genome in NEAT is encoded directly using a list of node genes and a list of connection genes. Input, hidden, and output nodes are all indicated by node genes, while connection genes hold the details of connections as well as an invention number for historical markers. By doing so, NEAT can line up the corresponding genes quickly during the crossover process.NEAT grows and evolves the topology by structural mutation. A new node can be added to an existing connection or a new connection can be made between two unconnected nodes. NEAT can thereby increase genomic diversity, investigate the solution space, and steer clear of local minima.

The NEAT algorithm is applied to the flappy bird game as follows. Learning is done through generations where each generation consists of ten birds. Each bird has a neural network associated with it. Initially, each neural network has three input nodes that are fully connected to one output node and the weights associated with each connection are random. The nodes and connections can be added or deleted in future generations. A generation ends when every bird in that generation dies, either by colliding with one of the poles or flying out of the screen. At

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 7| July 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

the end of each generation, the two fittest birds are selected(selection) using which the next generation of birds is created(crossover). There is also a possibility of mutation, where the weights are changed randomly in order for the bird to explore new examples in the search space.

The three input nodes, in the neural networks of the birds in the first generation, depend on the vertical position of the bird, the vertical position of the top pipe and the vertical position of the bottom pipe. The existing methods extracted raw pixels from the frame and trained the agent to avoid the obstacles, but here game specific features have been used to train the agent. This feature engineering approach reduces the search space drastically and eliminates the need of deeper models to automatically extract underlying features. This results in faster training speed where the time required for training the agent to do well falls from a matter of hours to minutes.

III. MODELING AND ANALYSIS

Reward Policy: The reward policy required to assign fitness to the birds in each generation is as follows

Table 1 - Reward Policy

Description	Reward
Bird is alive for a frame	0.1
Bird passes through a pipe	+10
Bird collides with a pipe	-10
Bird flies off screen	-20

Population: The first parameter to be tested is the population(number of birds) in each generation. The tests were run for upto 15 generations in each case, where the threshold fitness to be reached by the birds was 300.

Table 2 - Table for population in each generation

Population	Best Fitness	Best fitness generation	Achieved threshold fitness?
10	277.8	13	No
50	301.2	8	Yes
100	328.6	2	Yes

The above results clearly show that increasing the population of birds would ultimately result in achieving the threshold fitness faster, using up less number of generations. This result is expected as an increase in the number of birds would imply an increase in the number of neural networks with random weights and would therefore result in covering much of the search space faster. However, for the purposes of showing how fitness changes as generations increase and to make the task of learning more challenging to the RL agent, we selected a population of 10 birds for the Reinforcement Learning task.

Number of hidden nodes: The second parameter to be tested is the number of hidden nodes in the neural network used in the NEAT algorithm. The tests were run for upto 15 generations in each case, where the threshold fitness to be reached by the birds was 300. The population of birds in each generation was kept constant(10 birds).

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 7| July 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

Table 3 -Table for number of hidden layers

Number of hidden nodes	Best Fitness	Best fitness generation	Achieved threshold fitness?
0	353.9	10	Yes
3	-2	4	No
5	-2.1	11	No

The above results clearly show that increasing the number of hidden nodes decreases the best fitness drastically for a small number of generations (15 here). This result is expected as increasing the complexity of the neural network would mean that it would try to find more of the underlying features of the environment and therefore take more time to learn. Therefore, we chose 0 hidden nodes as this would take a much lesser time to get to the best fitness and therefore would take less time for the RL agent to learn to play the game.

Threshold for jumping: In this project, we use a tanh activation function as the output function which gives an output value between -1 and 1 based on the inputs to the neural network, weights and the bias. The bird jumps only if the output from the neural network is greater than the threshold. The tests were run for upto 15 generations in each case, where the threshold fitness to be reached by the birds was 300. The population of birds in each generation was kept constant(10 birds).

Table 4-Table for jump Threshold

Jump threshold	Best Fitness	Best fitness generation	Achieved threshold fitness?
-0.5	91.6	12	No
0	328.4	7	Yes
0.5	353.9	12	Yes

IV.RESULTS AND DISCUSSION: The bird reached a score of 40 after training the agent for 50 generations,which is shown in figure below.

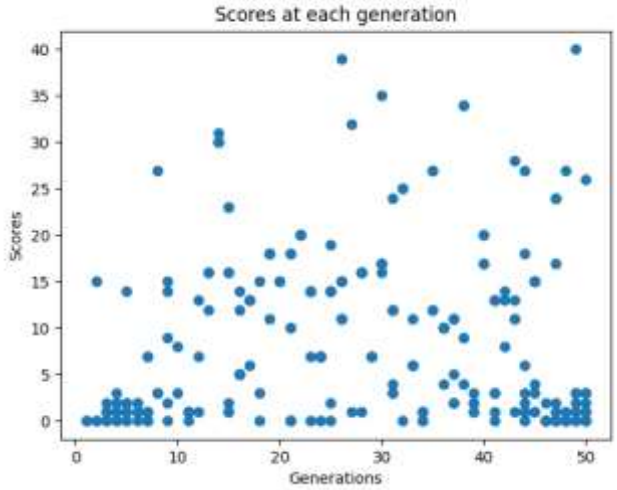

Fig :Generations vs scores

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 8| Issue: 7| July 2022|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2022: 8.205 || ISI Value: 1.188

For the initial generation the scores are comparatively less as the agent has not yet learnt but as it learns(generations increase) it is observed that the scores increase. There may also be cases where the scores drop as the generations increase. This is because the agent encounters a situation that it has not seen before and is not able to adapt from what it has already learnt. The agent then does exploration and learns something new to eventually increase the scores in future generations. There can also be cases where the scores are high in earlier generations. This is because the flappy bird game is quite simple and the random weights assigned to a neural network might be very close to the best weights.

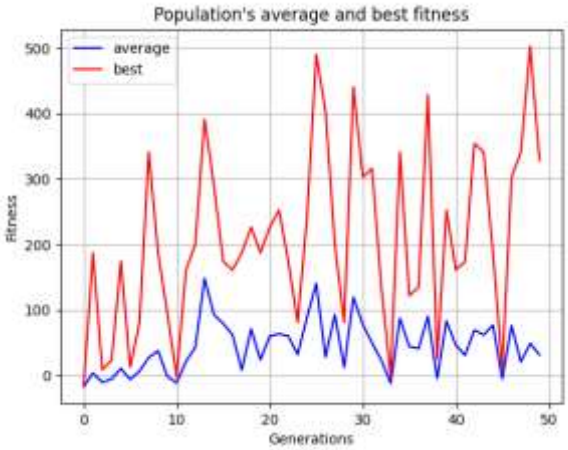

Fig :Generations vs Fitness

The above results imply that the fitness keeps increasing as the number of generations increase but we also encounter some pitfalls when the bird does come across an arrangement of pipes it has not experienced before. This happens because the speed of the game increases as the bird progresses through the game. Due to this the bird might not have learnt to adapt to this increase in speed. However, the next generation of birds learn from this experience and continuously try to improve their performance by also taking into account the newly encountered arrangement.

V.CONCLUSION

The flappy bird agent learns how to play the game without any human inputs by using the NEAT reinforcement learning algorithm, within a finite amount of iterations. Using features specific to the game instead of creating a CNN of each frame of the game increased the speed of learning by a great extent.The results shows that the bird learns as the number of generations increase and better scores are achieved.The best fitness in each generation increases as the generations increase

VI.REFERENCES

1. *AndreBrandao,Pedro Pires,Petia Georgieva, "Reinforcement Learning and Neuroevolution in Flappy Bird Game", Springer Nature Switzerland AG,2019*
2. *Isabelle Savoye, Catherine M Olsen, David C Whiteman, Anne Bijon, Lucien Wald, LaureenDartois, Françoise Clavel-Chapelon,"Performance Analysis of Flappy Bird Playing AgentUsing Neural Network and Genetic Algorithm",Springer Nature Switzerland AG,2018*
3. *Naveen Appiah,Sagar Vare, "Playing FlappyBird with Deep Reinforcement Learning",cs231n.stanford.edu, 2016*
4. *Kevin Chen, "Deep Reinforcement Learning for Flappy Bird ", cs229.stanford.edu, 2015*
5. *Naud Ghebre, Ethan Jen, Matthew McBrien, Arihan Shah, Nahom Solomon, "Flappy Bird: Neuroevolution vs NEAT", mmcbrien.com, 2017*
6. *Pilcer, Louis-Samuel & Hoorelbeke, Antoine & D'andigne, Antoine, "Playing Flappy Birdwith Deep Reinforcement Learning",Playing Flappy Bird with Deep Reinforcement Learning. 10.13140/RG.2.2.13159.96165, 2018*
7. *Cedrick Argueta, Austin Chow, Cristian Lomeli , "Deep Reinforcement Learning and Transfer Learning with FlappyBird", cs221.stanford.edu, 2018*
8. *Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G. Bellemare and Joelle Pineau , "An Introduction to Deep Reinforcement Learning",Foundations and Trends in Machine Learning: Vol. 11, No. 3-4. DOI: 10.1561/2200000071, 2018*
9. *Lucas, S.M.; Kendall, G,"Evolutionary computation and games", IEEE Computational Intelligence Magazine,2006*
10. *Nie, Y, How does a Flappy Bird Learn to Fly Itself How does a Flappy Bird Learn to Fly Itself, Fall Conference of KISM ,2016*
11. *Sahin, A, Atici, E., Kumbasar, Fuzzified flappy bird control system,IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2016,*
12. *Chen, K, Deep Reinforcement Learning for Flappy Bird. Cs229.Stanford. Edu, 6 (2015)*
13. *V.Mnih, K.Kavukcuoglu X-L. "Human-level Control through Deep Reinforcement Learning", Nature. Vol.10, February 2015.*
14. *C.Clark, A.Storkey. "Teaching Deep Convolutional Neural Networks to Play Go"arXiv.December 2014.*