Kubilay Agi
UID: 304784519
8 December 2017

Assignment 10 Report: *Clever Machines Learn How to be Curious*

There are two major driving forces that contribute to the learning of a topic: extrinsic motivation and intrinsic motivation. Extrinsic motivation is a reward system that drives an entity toward its goal by rewarding it for successes and positive behavior, but penalizes it for negative behavior. Since the twentieth-century, research has been conducted on how to employ the learning reinforcement systems present in humans to create intelligent machines. Algorithms for machine learning and artificial intelligence use predominately extrinsic motivation reward systems. There are supervised and unsupervised types of machine learning that employ extrinsic motivation. More recently, intrinsically motivated machines have been created. Intrinsically motivated entities are capable of adapting to their environment and extrinsically motivated machines are effective and efficient in completing assigned tasks. Both have advantages and disadvantages over the other in different environments.

Extrinsically motivated machines allow the user to precisely define what they want to be done. For example, Stanford University created AlphaGo, which was a program reinforced with machine learning that was able to defeat 18-time world champion Go player (Go is a complex board game).[1] However, although it is helpful in pushing the machine to do exactly what the user wants it to do, there are drawbacks to the extrinsic reward system. The first issue comes with supervised machine learning: the training of the machine by the supervisor is time consuming. Additionally, the machine may develop unexpected and undesired behaviors if it is unsupervised. Since the reward system must be created specifically to address certain scenarios or tasks, the

[1] “Self-play Code Excites Machine-Learning World”, Katyanna Quach, The Register.

code for this is not portable to other machine learning agents. The programmers writing the code for the machine learning agent are forced to account for as many cases as possible. This way, they can avoid a situation in which the machine learning agent has no way to assess whether it is succeeding or failing when performing a certain action.

The solution to the issues with extrinsic motivation systems seems to be intrinsic motivation. Pulkit Agrawal and Deepak Pathak are doing research at UC Berkeley on employing intrinsic motivation in machine learning. They have built a machine learning agent to play the Nintendo video game *Super Mario Brothers* where the player's objective is to move through a world filled with several obstacles and dangers. This approach to machine learning was inspired by psychological studies about humans, and more specifically about toddlers. Studies show that toddlers were more interested in toys that surprised them the most.[2] The algorithm for the machine learning agent is based off of favoring surprise over familiarity. The process for the machine to explore the world is to develop a prediction for the screen after a few frames. It then makes its moves and the more incorrect the model is, the more surprised it is, and therefore the machine receives a larger incentive to continue its current actions. In this fashion, the machine is able to explore the world around it. Perhaps the greatest benefit to this algorithm for machine learning is that the machine does not get stuck often because if it gets stuck then the screen remains the same for a long period of time, and therefore the surprise it receives will be minimal. The machine will try new things to remove itself from its current position to gain more reward. The algorithm for curiosity-driven machine learning is essentially a greedy algorithm that seeks the maximum reward (in this case, maximum surprise).[3]

---

[2] "The origins of Inquiry", Laura Schulz, MIT.
[3] "Intrinsically Motivated Reinforcement Learning", Satinder Singh, University of Michigan.

This method of teaching the machine how to behave solves several problems of the extrinsic reward system. First of all, it allows the programmers to develop an algorithm that is more portable for different environments because the machine is not looking for anything in particular; it is only making a prediction model and checking its results. Additionally, there is much less time spent looking for grey areas that programmers may have missed when writing an extrinsic motivation system for their machine learning agent. Intrinsic motivation also requires far fewer resources to progress in its learning. Large data sets are not needed for the machine to develop its knowledge This is not the case with extrinsically motivated machine learning agents. However, there are issues with this form of motivation as well. First of all, the machine may get distracted by a piece of its environment. The example that the article presented was a television screen with static on it; the screen is constantly changing, so the machine is continually surprised by this and gets stuck. However, the researchers at UC Berkeley have developed a solution to this problem by developing general features of the environment instead of focusing on each individual pixel. Another issue comes with the environment that it was tested in. The machine learning agent had success in the *Super Mario Brothers* video game, but this game is very linear and prevents the player from moving away from the final destination. The article did not acknowledge the complications that would arise if this type of algorithm was used in an open environment where there are no clear paths guiding it to the finishing position.

Even though the use of curiosity in machine learning has been researched since 1991, the curiosity-based machine learning is rarely used on large-scale applications that physically interact with humans. Intrinsic curiosity holds potential in improving the speed at which machines learn and the quality of the learning by “focus[ing] the attention of the machine and guid[ing] exploration".[4] However, beyond video games and contributing to other types of machine learning, it is difficult to find applications that require solely curiosity-based machine learning. The article suggests that curiosity would be "machine-learning overkill" in the context of autonomous vehicles. The article points out the limitations of intrinsic motivation systems, but it fails to suggest fields in which it would be useful.

Instead of focusing on creating solely curiosity driven machines, perhaps researchers should base their research in creating an algorithm that will be supplemental in machine learning. Pairing the intrinsic motivation algorithm with an extrinsic motivation could be useful because it would allow the machine to do most of the learning on its own, but it would allow the programmers to steer it in the right direction if the machine goes astray. Under the right circumstances, curiosity-based learning can be powerful. However, the time where intrinsically-motivated systems can contribute beneficially to the world on their own has not yet come. Progress in the field of psychology in better understanding human curiosity would also lead to improvements in intrinsically-motivated machines. But for now, resources would be better spent focus on improving our current machine learning agents.

---

[4] "Clever Machines Learn How to be Curious", John Pavlus, QuantaMagazine

Works Cited

1. Pavlus, John. *Clever Machines Learn How to Be Curious*. 19 Sept 2017. Quanta Magazine. https://www.quantamagazine.org/clever-machines-learn-how-to-be-curious-20170919/

2. Barto, Andrew. *Intrinsically Motivated Learning of Hierarchical Collections of Skills.* http://www.lira.dist.unige.it/teaching/SINA_08-09/SINA_PREV/library/ICDL2004/pdfs/24.pdf

3. Baldassarre, Gianluca. *Intrinsically Motivated Learning Systems: An Overview.* https://link.springer.com/chapter/10.1007/978-3-642-32375-1_1#Sec3

4. Singh, Satinder. *Intrinsically Motivated Reinforcement Learning*. University of Michigan. http://web.eecs.umich.edu/~baveja/Papers/FinalNIPSIMRL.pdf

5. Schulz, Laura. *The origins of inquiry: inductive inference and exploration in early childhood.* MIT. http://eccl.mit.edu/papers/TICS.pdf

6. Hacking Note. *How to Choose a Machine Learning Algorithm*. https://www.hackingnote.com/en/data-science-in-theory/algorithm-pros-and-cons/

7. Quach, Katyanna. *How DeepMind's AlphaGo Zero Learned All by Itself to Trash World Champ AI AlphaGo.* 18 October 2017. https://www.theregister.co.uk/2017/10/18/deepminds_latest_alphago_software_doesnt_need_human_data_to_win/