THE BASE OF AUTOMOTIVE WORLD - REINFORCEMENT LEARNING AND ITS APPLICATIONS

Dr Shilpa Gite shilpa.gite@sitpune.edu.in
Urjit Tembhurnikar tembhurnikar.urjit@sitpune.edu.in
Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India.

ABSTRACT

In this era of the modern automotive world, day by day things is getting more and more automatic. Humans are inventing new things to reduce their workload. Nowadays the technology has been so advanced that there are robots everywhere like the robotic arms in the industry to lift highly weighted objects, the well-known human interacting bots like Alexa, Siri, Cortana which take our voice commands as an input and do the tasks which we order them to do. In short, this is an automotive world and what is needed to make things automatic to train them to learn on their own, do the work on their own. This, in turn, gives birth to the concept of Reinforcement Learning. Training such a robot or model which learns by its own to survive or to make decisions in an environment and gives the most optimal solution as an output. This article mainly highlights what is Reinforcement learning and how it can be beneficial for us in our day to day life in this advanced world. This article also contains the applications of reinforcement learning.

Keywords: Reinforcement Learning, Neurodynamic Programming, RL applications, Self-learning, etc.

1. INTRODUCTION:

Reinforcement learning is one of the fascinating branches of Artificial Intelligence [11]. In the era of this automotive world, Sutton is considered as the father of the modern computational AI that is Reinforcement Learning technique. Reinforcement learning is the future of AI that is Artificial General intelligence to be particular. This Reinforcement Learning is considered to be the advance study in deep learning which is the beautiful combination of dynamic programming and Neural networks lead to a system of learning by trial and error method. This term Reinforcement Learning was introduced to the world in 1970 for the first time. There are four general types of reinforcement learning which are positive, negative, punishment, extinction. They are three general steps to develop a reinforcement environment and that is

1) Creating a simulation environment
2) To denote simulation, we have to add a state vector in the model

3) Now at the last, we have to add a reward/punishment system in the model to make it work in the simulation.

1.1 Reinforcement Learning

AI says that a type of dynamic programming that trains models using a system containing reward and punishment is called Reinforcement Learning [5]. In an environment, an RL model or agent learns to survive or to make decisions on its own. There are three types of machine learning which are Supervised Learning, Unsupervised Learning and Reinforcement Learning.

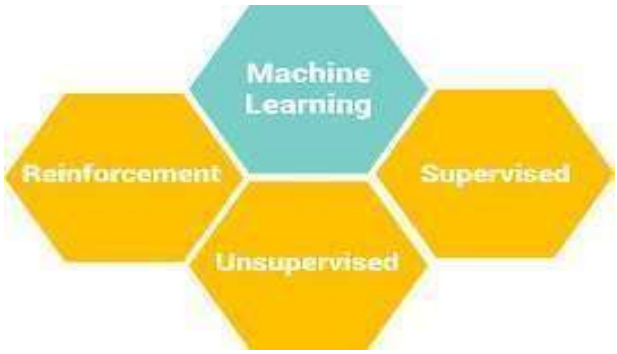

Fig 1: Basic types of machine learning

If the attempt of the model gets successful, then it gets a reward and vice versa (gets a penalty if it loses or takes wrong decisions).

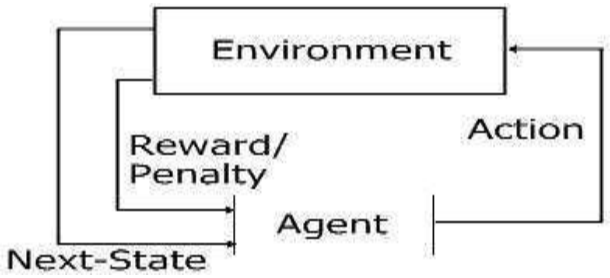

Fig 2: Reinforcement Learning mechanism in terms of environment agent and action taken by the agent.

In this real world, the best example where Reinforcement learning is applied is none other than The Driverless Cars. Let us have a brief idea of RL. Reinforcement learning is also known as Neuro-Dynamic Programming. As the name describes itself, it mainly consists of Neural networks as well as Dynamic programming algorithms (NN + DP) [1]. What it does is, it learns how to make better decisions by inspecting their actions taken earlier. In less anthropomorphic DP terms used in a book called RL by Richard S. Sutton and Andrew G. Barton is “observing their behavior”. It is related to restoration and to upgrading their agents working in the environment through the mechanism of reinforcement [1]. With the help of DP, we can achieve terms in classical DP like value and policy relation. It relates to schemes to improve the quality of the optimal-cost-to-go or

Q factors or optimal policy now it is clear that using NDP is much more advisable to get better performance. As it proves that DP has many limitations which hardly result in better behaviourof the agent [1]. Success is often obtained using methods whose properties are not fully understood. It is proved by the observation that, if once the task of estimation and computing is obtained offline then it will be easier to the agent to work accordingly as well as it turns in better performance of the agent (it gets help in making accurate and fast decisions), and which means it gives us better results in real-time[8].

The scientists research on the designs which can be able to solve learning problems in scientific as well as economic interest, and then they will evaluate the algorithms through the mathematical approach. And this approach is nothing but Reinforcement learning[3]. The approach should be goal-directed and of interactive nature. The approach we explore is called reinforcement learning. And we do interactions with the environment through this approach.

The invention of a model of RL which is beneficial in the solving-learning process in the scientific and economic field in which the computational, as well as the mathematical approach, is used.

With the help of a camera, it is easier to collect the real-time information offered by sensors in it. It is a better way to control traffic. This information can be useful for transportation planning and analyzing user behaviour and much more[10]. For example, in the parking area, we can get to know which parking slot is vacant as well as nearer to our location. And this, in turn, it helps in maximization of the resource and minimization of the load in traffic. But there is uncertainty in the future development of such information (i.e. non-deterministic)[10]. However, algorithms which are used in minimal pathfinding problems should consider the new paths which are generated in the learning process.

Fig 3: Finding the best optimal place to park [11]

Fig 4: finding an optimal-go-to path to the destination [11].

There are two terms in RL that are, Exploration and Exploitation. Exploration is nothing but improving knowledge for long term benefit and on the other hand, exploitation is gaining knowledge for short term benefit. To achieve more optimization the model has to follow the static route in its reward-penalty system but when we do so then it may reduce the performance because the environment is changing continuously. To get better performance we must plan the agent's trajectory and it is advisable to calculate its action policies instead of keeping a record of its static routes. It is suggested to follow the trajectory of the agent working in a particular environment which is passed through all the adverse conditions. It is practicable to predict human behaviour based on actions taken by him/her previously in the same type of situation or environment[11]. And to compute these all things Neurodynamic approach offers many tools for such perdition as well as surveys for better results.

2.3 Types of Reinforcement: There are two types of Reinforcement:

A. Positive RL

Positive reinforcement leads to save rewarding steps and act accordingly.

Advantages of reinforcement learning are:

- Gains better Performance
- Saves the rewards for a long period.

Disadvantages of reinforcement learning:

- If one applies a lot of reinforcement then it may occur overloading of levels, which in turn may result in failure or it may vanish results

B. Negative RL

Negative RL leads to saving the penalties gained by the model and it tries not to take that step again which in turn results in better strength of the model.

Advantages:

- Behaviour shoot up
- It keeps resistance in performing minimum standards.

Disadvantages:

- It may result in slow and minimal performance.

Meta-RL algorithms suffer from poor sample efficiency using on-policy data yet training meta-RL models on off-policy data introduces challenges such as a mismatch between meta-training time and meta-test time. To address these challenges, the researchers introduce **Probabilistic Embeddings for Actor-critic RL(PEARL: Probabilistic Embeddings for Actor-critic RL)**, which combines existing off-policy algorithms with the online inference of probabilistic context variables:

 # Achievement Of RL

The experimental evaluation on six continuous control meta-learning environments demonstrates that **PEARL** outperforms the previous state-of-the-art approaches in terms of:

- **Sample efficiency** by using 20-100× fewer samples during meta-training.
- **asymptotic performance** with the results improved by 50-100%in five out of six domains.

3.APPLICATIONS OF RL:

3.1 Real-life applications of Reinforcement Learning –

RL offers a great service in the industries which are moving towards automation. As there is a robotic arm for lifting heavy loads, as well as that are tremendously used in the manufacturing industry.

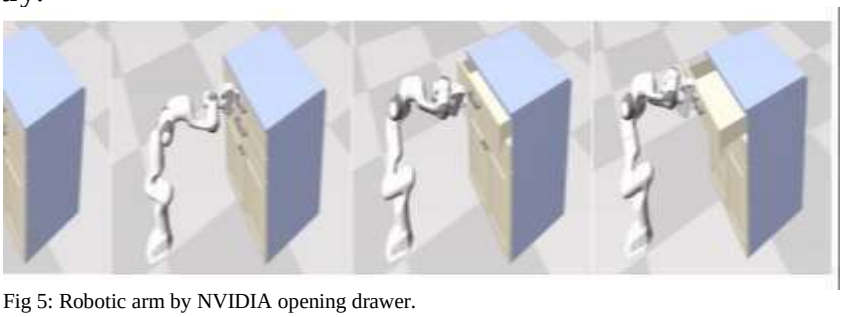

Fig 5: Robotic arm by NVIDIA opening drawer.

RL has implemented in most of the machine learning areas as well as data processing techniques. It can be used to learn the history of the stock market as well as stock prices of various shares, and this machine learning application can be used as a future market price predicting model.

There is a teaching-learning system invented for students based on RL which keeps track on the performance of a student and encourages him/her to get the best output through them.

One of the best examples we can put in our consideration is that A research was going on AI that plays chess games. And after training the model in every possible way in 1997 that robot won against the world chess champion [7]. It was impossible without RL.

In this modern era, Google has created the self-driving(driverless) car in 2012 which uses the RL model to learn in various environments with the help of Computer vision [7].

3.2 WHAT ARE POSSIBLE BUSINESS APPLICATIONS?[10]

1. RUBIK'S CUBE (COMBINED UNCERTAINTY AND BATHYMETRY ESTIMATION)

The suggested approach to training an algorithm in a simulated environment might be used to train robotic hands for further applications in manufacturing and warehouse operations.[6]

2. Adoption of RL in Gaming agent [8],[10]

RL technology can be used to create a non-living player or agent in a gaming area where the agent robot will challenge the human player [2]. This is one of the biggest steps in the Gaming industry. Some of the classic examples are the Football FIFA game, some of the fighting games, PUBG(The bots). [10] The chess game is one of the best areas where the RL model is used. The robotic agent learned by attempting tons of tries and beaten The World champion in 1997.

3.3 WHAT'S THE KEY ACHIEVEMENT RL TECHNOLOGY?

- The robot hand can successfully manipulate the Rubik's cube to a solved state:20% of the time for maximally difficult starting blocks requiring 26 face switches [5].60% of the time for simpler scrambles that require 15 rotations. The robot hand is robust enough to deal with perturbations during the manipulation, such as tying fingers together or putting a blanket over the cube.

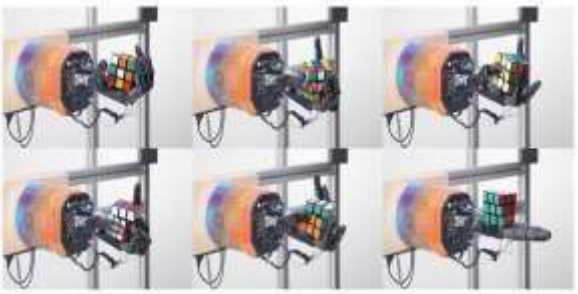

Figure 1: A five-fingered humanoid hand trained with reinforcement learning and automatic domain randomization solving a Rubik's cube.

Fig 6: There is a Robotic arm which is fully trained to solve the Rubik's cube. Whatever is the condition it will solve it in a few seconds.

The formula used [3]::

$$At = \text{Argmax} [Qt(a) + c. \text{Sqrt}(Lt / Nt(a))]$$

Exploration *Exploitation*

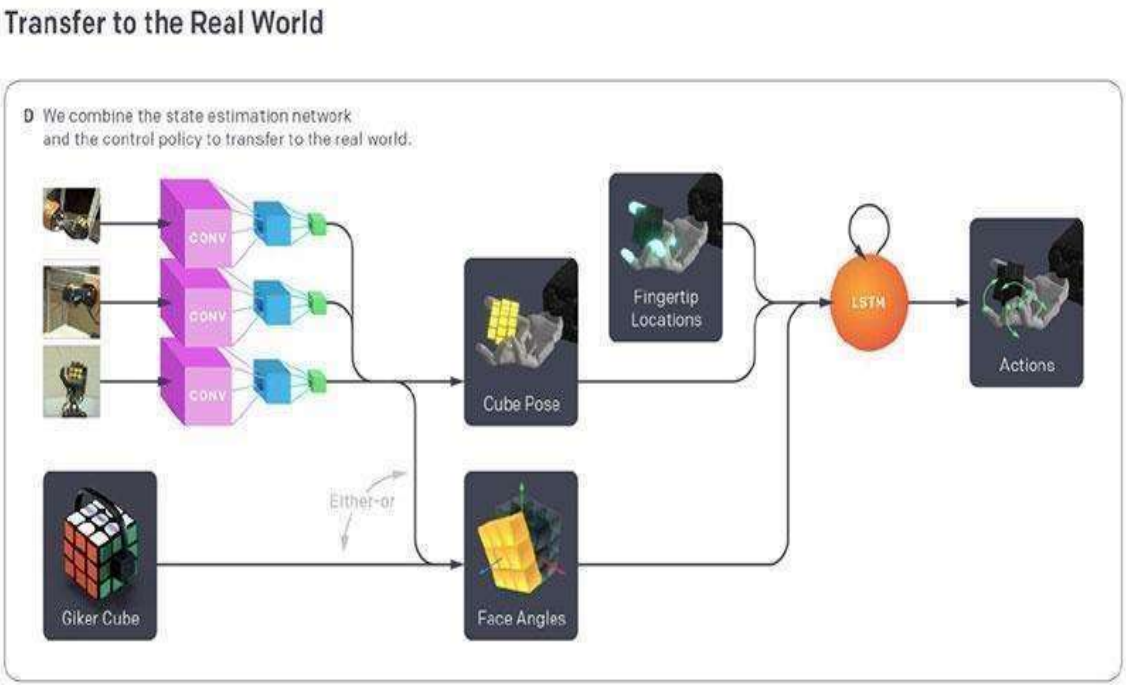

Fig 7: The fig shows

4. CONCLUSION AND FUTURE DIRECTIONS:

This paper focuses on Reinforcement Learning and its various applications in the real world. In our day to day life RL may play a very important role by acquiring and adopting it properly. This paper also states the real-life applications of RL as well as Business Applications along with automotive industries. No matter how far this world has progressed, still, the machine has not been able to defeat the entire thinking capacity of the human brain. Extending to general-purpose systems that can quickly adapt to the changing environment. Improving adaptation to real-world dynamics during the first few moves of the manipulation task can be done by RL. Therefore, there is a thrust area in the AI domain such as RL which works on a feedback system just like a human being and has tremendous potential. The future scope is the right implementation of RL and its incorporation into real-life applications.

5. REFERENCES:

1. Sutton, R. S., Barto, A. G., & Williams, R. J. (1992). Reinforcement learning is direct adaptive optimal control. *IEEE Control Systems Magazine*, *12*(2), 19
2. Abbeel, P., Ganapathi, V., & Ng, A. Y. (2006). Learning vehicular dynamics, with application to modelling helicopters. In *Advances in Neural Information Processing Systems* (pp. 1-8). Wikipedia search
3. Ahtiainen, J. Exploration and Other Applications of Reinforcement Learning in Robotics.
4. Littman, M. L., & Szepesvári, C. (1996, July). A generalized reinforcement-learning model: Convergence and applications. In *ICML* (Vol. 96, pp. 310-318).
5. Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of deep learning and reinforcement learning to biological data. *IEEE transactions on neural networks and learning systems*, *29*(6), 2063-2079.
6. Mahmud, M., Kaiser, M. S., Hussain, A., & Vassanelli, S. (2018). Applications of deep learning and reinforcement learning to biological data. *IEEE transactions on neural networks and learning systems*, *29*(6), 2063-2079.
7. Kormushev, P., Calinon, S., & Caldwell, D. G. (2013). Reinforcement learning in robotics: Applications and real-world challenges. *Robotics*, *2*(3), 122-148.
8. Reinforcement Learning course on **Coursera**. Platform.
9. Artificial Intelligence course on **Coursera** platform.