iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IV **Month of publication:** April 2021

DOI: https://doi.org/10.22214/ijraset.2021.33821

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

Reinforcement Learning Paradigm & Decision Making

Reena Hooda
Assistant Professor, CSE, Indira Gandhi University Meerpur (Rewari)- Haryana -122502, India.

Abstract: Objectives- In an era of robotics, artificial intelligence playing a significant role in training machines acts like a human. Present paper targeted towards reinforcement learning in AI, its functioning and challenges.
Methods- The Agent has to enter in a new state or not is based on the score matrix of rewards and penalties. Agent path is described by network model in Google Docs comprises nodes, links, alternatives to destination. If total rewards for a path i > total number of penalties for path i, that path is selected, it changed the state, otherwise, the highest number of penalties bars Agent and state is the same.
Findings- Neural networks are the best example of mimicking the human brain for decision making and exactly, deep learning got its space in adding to artificial intelligence with enhanced accuracy in decisions with the layered architecture of learning from examples and experiences. This opens up a vast scope of research for improvements. Reinforcement learning is the paradigm that focused on experiences and decision-making for unknown and unpredicted circumstances even if no experience or training data for such cases is available. Reinforcement learning is based on the summing of rewards and penalties of a path with the assumption that the length of the path in terms of time taken in traversing is optimal. Otherwise, a path with longer distances merely to increase the rewards has an adverse effect on the long-term performance of the agent.
Novelty- The present paper targeted towards the meticulous investigation & understanding of reinforcement learning through detailed graphical representation of its functioning with computations in tabular form enriching the existing literature and theories. The real-world environment is represented as a network model containing different states as its nodes and moving ahead by rewards-based computations and to reach the target optimally along with depiction of challenges.
Keywords: States, Rewards, Penalties, Deep Learning, Machine Learning.

I. INTRODUCTION

Robotics is no more science fiction, it is the reality and future of a technological world where every field of society & economy employs robots whether it is a military, health, engineering, home care, production, more advanced analytics, and many more. Reinforcement learning helps robots to act more intelligently even in unknown scenarios, unobserved conditions. So, every path towards intelligent machine learning goes through reinforcement learning. Therefore, the paper interacts with the reinforcement learning model, its applications, and working along with the limitations.

II. LITERATURE SURVEY

Reinforcement learning is a machine learning paradigm aimed to make a way between unknown facts and Current Knowledge at hand. This helps in decision-making with partially supervised and personally unsupervised behaviour. [1] This learning helps in automated decisions based on the outcome of current knowledge plus observations and computations for acquiring the potentiality of outcomes in an aggregate manner. [1] The method becomes more useful as it pretends to act in an environment where nothing is found exactly suitable to a decision and not any current model is 100% suitable. [1]

Reinforcement learning does not provide help in decision making nor does the designer give any directions. It lets the agent learn from its own mistakes and actions. Search the environment for particular conditions and make its own decisions based on various hits and trials. It continuously upgraded & modified its algorithms, learned until satisfied the designer with minimum penalties in actions and observations. The model took random actions and chose the best stream of actions near to the goal like an automated car without a driver can run without accidents or human hazards by leaning from given route information to routes of unknown areas or lands. [1] [2]. In this era of highest unpredictability, artificial intelligence inferences based on if-then-else fails. Reinforcement learning plays a role to train the model to make its own decisions on rewards & penalties-based learning, though to program such learning requires highly skilled programmers. [2]

A. *Applications of Reinforcement Learning*

The applications of reinforcement learning include statistics, simulation-based Optimization, operational analysis, control models, Game Theory, robotics even Economics for equilibrium models. Reinforcement learning is dynamic programming that learns from the environment and feedbacks the new conditions and States which helps in growing up its intelligence to work in a dynamic environment more suited to the neurology of human beings optimally Not only satisfactory, however, much better than the cognition behaviour of human. [1] The basic applications of reinforcement learning are:

1) Training the agent to control an autonomous car in a live environment. [1]
2) Building a Model to act as an automated pilot specially used in the military.
3) In gaming like chess. [1] [3]
4) Robotics [1] [2] like motion control. [8]
5) Developing prosthetic legs that connect with the nervous system or learn to walk by observing and recognizing people's walking patterns smoothly. [2]
6) Data processing [3] and strategic business planning. [8]
7) Training system [3] for example Training system for the students based on their custom requirements. [8]
8) Dialogue Agent is now popular on the internet nowadays. [4]
9) Health care for the better treatment of the patient. [4]

III. THE FUNCTIONALITY OF REINFORCEMENT LEARNING

Reinforcement learning is lies between supervised and unsupervised methods work based on the interactions with the environment, actions taken & state changed because of rewards and penalties cumulatively to achieve a target. Then fixation of policy and storage for the future reference, the decision is not taken as a unit, however, it is the outcome of a sequence of decisions, labelling of dependent decisions as a whole to get the desired target.

The functioning of the reinforcement learning model is defined in figure 1 that is created in Google Docs under the drawing option by freely available clip arts [6] [10] used in the figure.

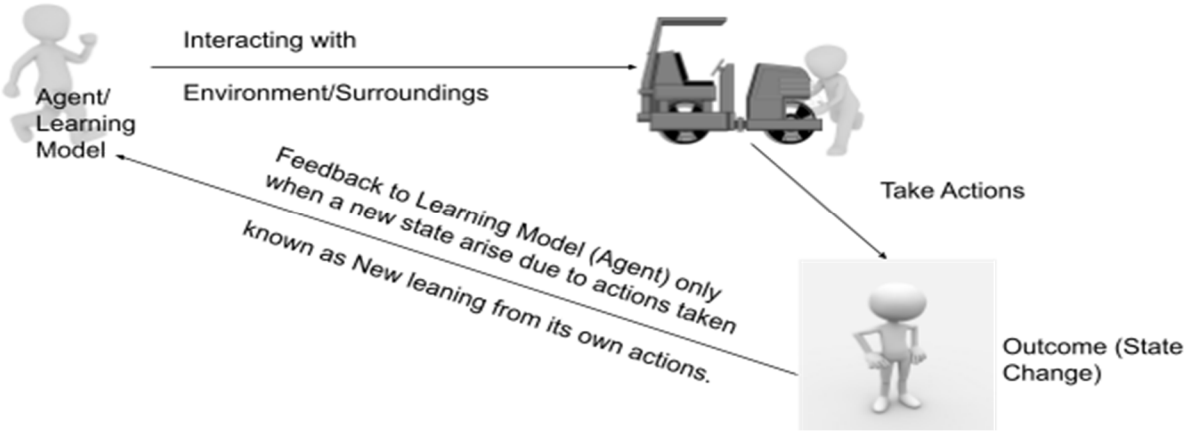


Figure 1: Functioning of Reinforcement Model

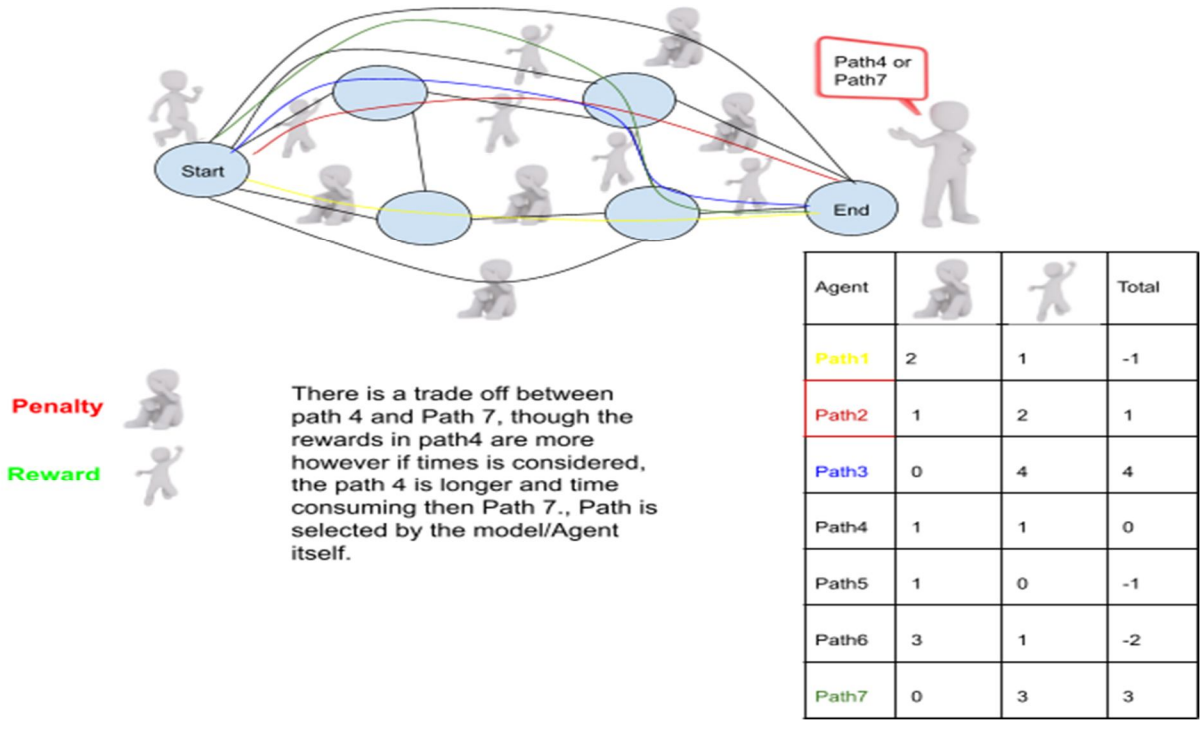
The observations from a certain environment can be fully or partially. If the agent directly observed the various states, it is called fully observed, and if the states contain noise or the agent cannot access all the states, the observations can be described as partially. [1] Reinforcement learning is the training of an agent in an uncertain environment to make optimal decisions. [2] It is one of the machine learning methods that helps the model or agent to decide a sequence of actions based on environmental observations and learning from its past actions [2] as shown in figure 1. It provides a continuous learning process through interacting with unseen conditions and deciding the actions being taken. [2] It is just like a programmer does in writing its programs; change the codes to get the desired outcome. Similar manner based on rewards or penalties this field of artificial intelligence learns and upgrades itself to maximize its rewards in the long term and minimize penalties for the best solutions for its designer. With reinforcement learning, the solution with maximum rewards can be computed also help in deciding when to take or not to take any action to minimize the penalties and enhance rewards and which situation needs an action with precaution not to overload with the states just to minimize the rewards so optimality of the principal is always a requisite. [8]

A. *Principles of Optimality*

To make reasonable decisions and better future actions, the agent has to check not only the current outcomes of its actions but the future consequences of the actions taken by the agent. [1] This the agent has to make a trade-off between long-term occurrences or the optimality and short-term rewards or penalties, this is reinforcement learning. [1] Therefore, this type of learning can be gained only with the interactions with the environment and best suited with games based on the challenges like chess or target hits, robotics, simulations, etc. [1]

B. *Calculation of Rewards and Penalties for Decision Making while Traversing A Path*

The reinforcement learning model takes an initial state and moves towards the ultimate solution (desired outcome). In between, many hurdles and rewards are found as in a target game, the model has to take the initial start, observe the things, and try to achieve the maximum rewards while avoiding the penalties. Each step or action is taken will convert the agent into the next state with the result as a reward or penalty for this selected state. The procedure is repeated until the final state is achieved. After reaching the final state, the path with the maximum rewards is finalized and stored, this new storage or the information is known as the learning of the model. [3] Penalties and rewards can be assigned by the programmer in advance for the various steps with this reason reinforcement learning can be considered partially supervised as the user has some control over it. But the model takes its step without any help from the user to get into the next state and therefore also can be considered as an unsupervised method, this way, reinforcement learning is called partially supervised and partially unsupervised learning method of machine learning.[3] The main difference between the reinforcement learning method and supervised learning method is that agent in reinforcement learning act from their own learning experiences whereas in supervised learning the instructions are provided to the model like training data-set and the labelling of the different actions to be taken based on the external inputs and given conditions. [5] The working of Agent is graphically represented by network model of States & actions based on rewards & penalties computed in tabular form in figure 2 that is designed and created in Google docs under the drawing tool by freely available clip arts. [9][10] [11]



Agent			Total
Path1	2	1	-1
Path2	1	2	1
Path3	0	4	4
Path4	1	1	0
Path5	1	0	-1
Path6	3	1	-2
Path7	0	3	3

Figure 2: Traversing of Agent to an unknown path to reach the target and calculating total Rewards and Penalties.

The penalties and rewards can be considered as negative or positive behaviour; their positive behaviour helps in moving ahead and negative hinder the agent to avoid penalties or hazards within or motto of increased efficiency and achieve the standard performance.

IV. RESULTS AND DISCUSSIONS

The Reinforcement learning model act as an agent that interacts with the environment reacts, takes actions based on previous experiences, gets a penalty if there is a wrong action to avoid wrong behaviour, rewarded for their right actions toward the targeted behaviour. Every new state is an action and feedback to an agent to proceed further until the goal is achieved. [4] The stored sequence or overall positive behaviour increases the speed of the actions where the negative behaviour protects from barriers to maximize the accuracy of the model to get a better solution in less time in the future. To implement the reinforcement learning model and building an environment Markov Decision Process is used to describe certain states and exchanges and set of rewards. [1] [4] The major reinforcement learning methods are Q- learning is also known as the "off-policy method", SARSA is described as state-> action-> reward-> state-> action also known as "on- policy method". [4] [7] Other popular methods of reinforcement learning are Deep Q Networks [7], Deep Deterministic Policy Gradient to work in unknown states. [4] Idea of reinforcement learning is based on a newborn child that learns from the interactions with the environment. In the same way, the model works on cause and affects factors, actions taken to achieve your target, and their consequences. [5] Taking actions by own computations, getting penalties or rewards, known closed-loop problems as the model or the learner is not provided with the instructions. The agent starts with an initial state, sensing the environment, taking the actions based on the Goal, and then transferring from the initial state to a new state based on the rewards and penalties. [5]

A. Challenges in Reinforcement Learning

Implementing a reinforcement learning model is not an easy task, first issue is how much penalties or rewards should be finalized, moreover choosing an optimal method is another issue as there is always a tradeoff between the time taken and the rewards got, the other issues are given below.

1) Simulating the real-world environment is the toughest task and becomes more hectic to face the real challenge. [11]
2) The only way to train the model is penalties and rewards', thus interacting and training the neural network model to behave in a more realistic way like a human being is another challenge. The only way of learning and improvements to face the environment. Then what actions are better is another issue in reinforcement learning. Another task to change or update memory and erase or keeping the old data to act optimally is the real challenge of this paradigm. [11]
3) Trade-offs between the exploitation (using experiences to choose an action) and Exploration (if new actions/ interactions arise) to choose the best from the list of paths (states) to maximize rewards.
4) Policymaking [12] is another challenge to elect the action sequences to gain maximum rewards and minimize penalties for better outcomes.

V. CONCLUSIONS

Overloading of the states to gain the highest score is a big challenge in reinforcement learning thus there are lots of research scopes in this area to improve the optimality. It is a paradigm that is in a growing stage and has a vast scope because of usability of machine learning in day-to-day operations and developments in the "internet of things" and robotics adds importance to this field to enhance the accuracy and perform reasonably more like human beings with increased speed & performance.

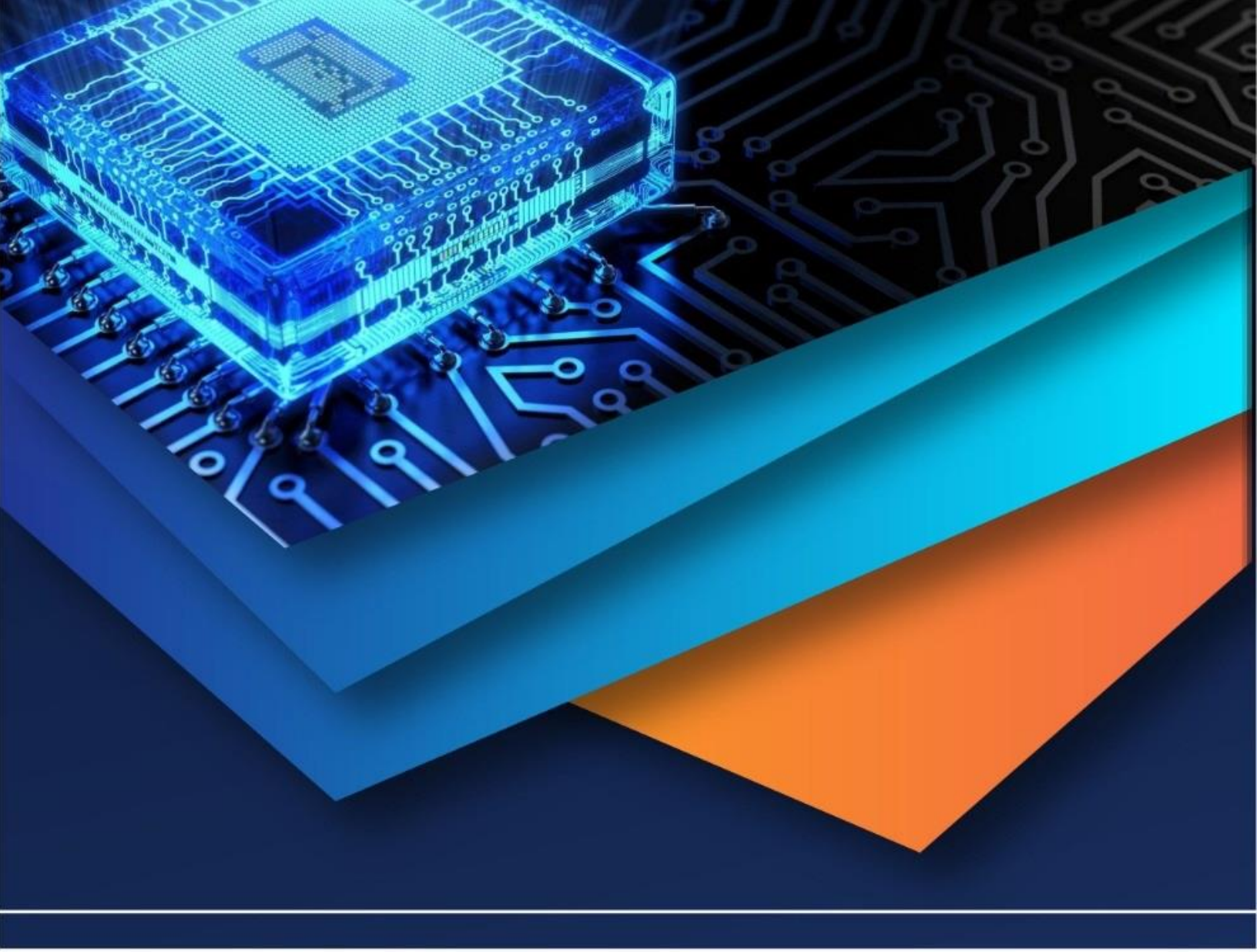
A. Limitations and Recommendations

The proposed network model in this paper elaborates the best path selected by the agent based on rewards and penalties as well as considering the time factor. However, the maximum and the minimum number of rewards & penalties is based on the designer's choice thus provide scope for enhancing the method to select an optimal range of rewards & penalties. The other factors like obstacles on the way to ignore the time factor or the distance still have scope for enhancements. The study & analysis of various defined models like Q- learning, SARSA, Deep Q Networks, Deep Deterministic Policy Gradient, and other methods can be the future scope of work.

REFERENCES

[1] Wikipedia contributors. Reinforcement learning. Wikipedia, The Free Encyclopedia. 2021 Mar, 02:59 UTC. Available at: https://en.wikipedia.org/w/index.php?title=Reinforcement_learning&oldid=1014088722. Accessed January 01, 2021.
[2] Błażej Osiński and Konrad Budek. What is reinforcement learning? The complete guide". 2018 Jul. Retrieved from: https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/ Date accessed: 01/01/2021
[3] Prateek Bajaj. Reinforcement learning. Geeksforgeeks. 2020 May, Retrieved from https://www.geeksforgeeks.org/what-is-reinforcement-learning/#:~:text=Reinforcement%20learning%20is%20an%20area,take%20in%20a%20specific%20situation. Date accessed: 2020

[4] Shweta Bhatt, Youplus. Tutorials on 5 Things You Need to Know about Reinforcement Learning. 2018 Mar. Retrieved from: https://www.kdnuggets.com/2018/03/5-things-reinforcement-learning.html
[5] Richard S. Sutton and Andrew G. Barto. Reinforcement Learning: An Introduction". A Bradford Book the MIT Press Cambridge, Massachusetts London, England. 2014-2015. Retrieved from: https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf
[6] https://freesvg.org/roller-vector-image
[7] Robert Moni. Reinforcement Learning algorithms — an intuitive overview. 2019. Retrieved from: https://medium.com/@SmartLabAI/reinforcement-learning-algorithms-an-intuitive-overview-904e2dff5bbc
[8] Guru99. A tutorial on Reinforcement learning-tutorial. 202. Retrieved from: https://www.guru99.com/reinforcement-learning-tutorial.html. Date accessed: 01/02/2021
[9] https://www.pxfuel.com/en/search?q=3d+figures
[10] https://www.wallpaperflare.com/search?wallpaper=3d+man&page=2



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY