International Journal of Electrical and Computer Engineering
ISSN 0974-2190, Volume 17, Number 1 (2025), pp. 1-9

http://www.irphouse.com

DC Motor Control Using Actor-Critic Reinforcement Learning

Sweety G. Chabukswar[1],
[1]M. Tech. Electrical Machine and Drive, Government College of Engineering, Chh. Sambhaji Nagar
E-mail: [1]sweetychabukswar321@gmail.com

N. J. Phadkule[2]
[2]Associate Professor, Department of Electrical Engineering, Government College of Engineering, Chh. Sambhaji Nagar
E-mail: [2] phadkule@gmail.com

ABSTRACT

This paper investigates the application of the Actor-Critic reinforcement learning method for DC motor control. Unlike conventional approaches such as PID controllers, which rely on fixed parameters and accurate modelling, the Actor-Critic framework enables autonomous policy learning through direct interaction with the environment. In this setup, the actor network proposes actions (motor voltage) based on the current system state (motor speed and position), while the critic network evaluates these actions via a value function. Over successive iterations, both networks refine their outputs to optimize performance. Altogether, these results position the Actor-Critic approach as not only robust but also highly scalable, offering a promising pathway toward smarter, adaptive motor control,
Keywords: DC motor, reinforcement learning, Actor-Critic, motor control.

1. Introduction

Reinforcement Learning (RL) presents a promising machine learning technique applicable to the control of dynamic systems, such as the DC motor under consideration.[3] The primary goal of DC motor speed control is to achieve the desired speed within a specified reference range as quickly as possible while rejecting environmental disturbances such as load variations and changes in operating conditions.[3,12,13] The most commonly used DC motor speed control technique is based onthe Proportional Integral Derivative (PID) controller family which can consider the error signal, the change in the error signal, and the sum of the error signal in order to produce a control signal. [2,3,8,11]

Precise speed control of direct current (DC) motors presents a difficult problem, particularly when external disruptions and system irregularities are present. [1,2] Traditional regulation techniques, such as Proportional-Integral (PI) and Proportional-Integral-Derivative (PID) regulators, are sometimes insufficient in managing these matters. Reinforcement Learning

(RL), conversely, is a potent technique for addressing intricate and dynamic systems given its capacity to independently acquire optimal behavior in an evolving setting byutilizing a process of experimentation. Fundamentally, it entails an entity that learns to refine its actions based on the input it obtains from its surroundings. [1,2,6,8,9] In this scenario, the RL algorithm learns an ideal regulation strategy that lessens the disparity between the observed and desired motor speed. Due to the growing interest from the scientific community in research activities related to artificial intelligence, there are various works related to this topic [1,2,7,10,]

The control system is based on reinforcement learning of the critic – actor type. The critic is represented by a neural network that evaluates the efficiency of the actions generated by the actor (which is similar to the controller in conventional control systems).[2,4,5,6,] Critic tuning (neural network training) is done online using the technique known as Temporal Difference Learning. Temporal Difference (TD) learning is a reinforcement learning method that updates value estimates based on the difference between consecutive predictions, effectively combining aspects of dynamic programming and Monte s Carlo methods. [2,12] In control systems, TD learning can be utilized to enhance decision-making and policy optimization. The method is based on solving on-line (at some sampling moments) an Actor Critic type optimization problem [6][2]. The objective function is written in integral form - which justifies the name of the method. This objective function is approximated by a simple neural network that is trained online and uses as activation functions some polynomials that depend on the state of the system.[2,3,8,9,10] The main advantage of the method is that it is not necessary to determine the dynamics of the system, which is the reason why we can place this approach in the category of adaptive-optimal methods. However, it is necessary to know (or estimate) all the states of the system[2,6,7]. Electric motors are integral to various industries such as manufacturing, transportation, aerospace, and robotics. Ensuring precise motor control is crucial for optimizing performance, enhancing energy efficiency, and maintaining system reliability. Reinforcement learning (RL), which mimics intelligent decision-making through trial and error, is increasingly regarded as an effective method in machine learning [3,11,13]. Unlike supervised or unsupervised learning, RL learns directly from its interactions with the environment, making it suitable for dynamic and uncertain scenarios [2,6,12,].

The central challenge in reinforcement learning is balancing exploration (discovering new strategies) with exploitation (optimizing known strategies). This becomes more complex when working with continuous state and action spaces, which are typical in industrial systems. By focusing on integral reinforcement learning, we can address these challenges and enhance the control of DC motors.

2. Controller Design

2.1 Reinforcement Learning

Reinforcement learning is a form of machine learning where an agent learns to make decisions by interacting with its environment. In this context, RL is used for controlling a DC motor, where the agent's task is to determine the optimal control inputs (motor voltage) based on the motor’s current state (speed, position, etc.).

The RL framework used here follows the Actor-Critic methodology, where:
Actor: Responsible for selecting the control input based on learned parameters. The actor is responsible for selecting the control input (applied motor voltage, u). It learns a policy that

maps the current state (motor speed and position) to an optimal action (voltage adjustment). The actor updates its policy based on feedback from the critic

Critic: Evaluates the quality of the action taken by the actor and provides feedback to improve future decisions. The critic evaluates the action taken by the actor by estimating a value function. It measures how good the selected action was based on the reward signal. The critic learns via **Temporal Difference (TD) learning**, refining its estimate of the value function over time.

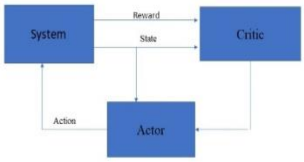

Fig 1 -Actor Critic

2.2 DC Motor Dynamics

The DC motor is modelled using the following differential equations:

● **Electrical Equation (Kirchhoff's voltage law):**

$$L\frac{di_a}{dt} + Ri_a + K\omega = u$$

Where:
L: Armature inductance
R: Armature resistance
i_a : Armature current
K: Motor constant (back-emf constant and torque constant)
ω: Angular velocity
u: Applied voltage (control input)
Kω is the back electromotive force (emf) induced by the motor's rotation.
The term Ldi_a/dt captures the inductive effect, and Ria represents resistive losses
This equation represents thearmature circuit, where the applied voltage u drives the armature current. The control objective is to adjust voltage so that speed reaches reference speed smoothly.

Mechanical Equation (Newton's second law):

$$J\frac{d\omega}{dt} + b\omega = Ki$$

This equation describes the motor's rotational dynamics.
Where
J: Moment of inertia

$J\frac{d\omega}{dt}$ represents the torque required to accelerate the rotor.
b: Friction coefficient
bω: accounts for frictional losses.

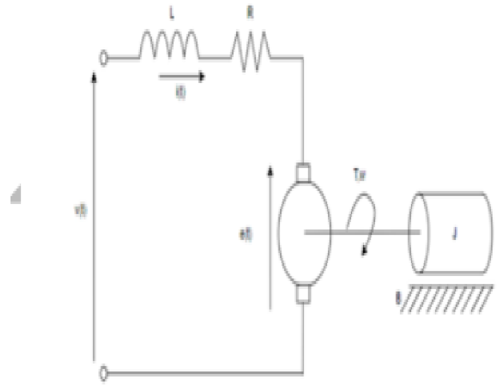

Fig 2- Circuit Diagram of DC Motor

3 Actor-Critic Algorithm.
The Actor-Critic (AC) algorithm is a reinforcement learning (RL) approach that combines:

3.1 Actor (Policy Improvement)
The actor is responsible for adjusting the control input (V/I) using a Controlled weight *w*. It updates the policy based on the error signal. The update rule for actor weight is

$$w = w + \alpha(error \times \omega - w) \times dt \qquad (3)$$

α = learning rate for the actor
ω = motor speed (used to update control actions)
In this Set up, the actor learns the control strategy using policy gradient methods and executes actions during training.

3.2 Control Law (Policy Output
$$\boldsymbol{u = w \times e(t) + Ki \int_0^t e(t)dt} \ (4)$$

Where:
w = dynamically adjusts the control input based on learning.
$\boldsymbol{\int_0^t e(t)dt}$ = Integral Error
Ki=0.1

3.3 Critic (Value Function Estimation)
The critic evaluates the quality of actions taken by the actor. Critic evaluates theresult performance by comparing the motors speed to desired set point.

$$\theta = \theta + \beta(error - \theta) \times dt$$

β = learning rate for the critic
θ = theta adjusted to improve future action evaluations.

3.4 Reward Function (Negative of Error Squared)

The Critic evaluates the current state through a value function influenced by the actor's action learning the value function via temporal difference error.

$$reward = -(error^2 + 0.1 \times Integral\ error^2)$$

This function Penalizes large tracking errors and Encourages learning control actions that reduce the error.

4. Results and Discussion

4.1 Simulation Setup: The DC motor model is simulated in MATLAB using the Actor-Critic RL algorithm. The state space includes motor speed and position, while the action space corresponds to the applied voltage. The reward function penalizes deviations from the desired motor speed, encouraging minimal oscillations and faster convergence.

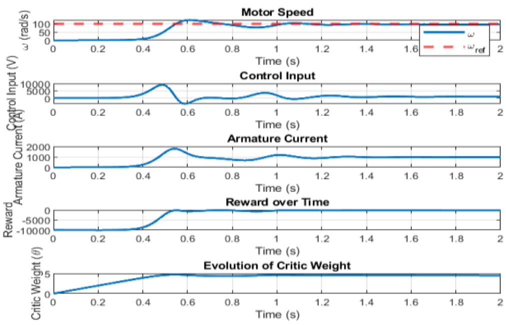

Fig 3– Speed, Critic Weight, Armature Current VS Time

4.2 Result Interpretation: The generated graphs illustrate the performance of the actor-critic reinforcement learning (RL) algorithm in controlling the speed of a DC motor. The graph shows the motor speed over time, where the actual speed (blue line) starts from zero and gradually approaches the desiredreference speed, red dashed line). The effectiveness of the RL controller is reflected in how quickly and smoothly the motor reaches this target without excessive oscillations or delays.

The control input (u), which is the voltage applied to the motor. Initially, a higher voltage is supplied to accelerate the motor, after which the control input stabilizes as the desired speed is maintained. Any excessive fluctuations in this graph might indicate instability in the learning process. The graph depicts the armature current, which is responsible for generating torque. The current is initially high to overcome inertia but decreases as the motor reaches steady-state operation.

The reward function, which is based on the negative squared error between the actual and desired speed. A steadily increasing reward (less negative) indicates that the RL algorithm is learning effectively and improving performance over time. However, unstable or fluctuating rewards may suggest improper learning rate parameters.

Overall, these plots collectively demonstrate the working of the actor-critic RL algorithm, where the actor adjusts the control input, and the critic evaluates the performance based on the reward signal. Fine-tuning parameters like learning rates (α, β\alpha, \beta) is essential to ensure smooth and efficient motor control.

The Evolution of Critic Weight graph illustrates how the critic weight changes over time as the reinforcement learning (RL) algorithm adapts to control the DC motor. In the actor-critic framework, the critic evaluates the effectiveness of the control actions taken by the actor by estimating the value function, which represents the expected long-term reward.

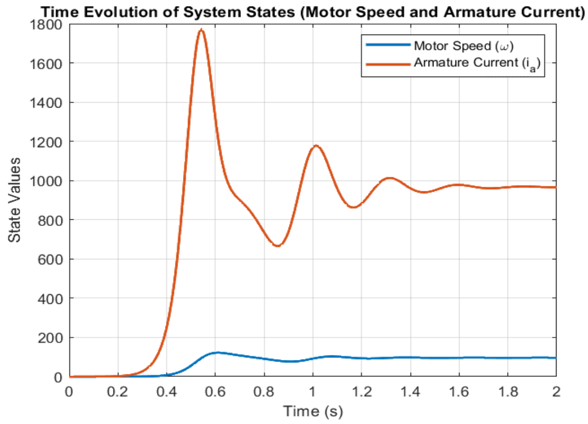

Fig 4- Graph of Time Evolution of System States

The Time Evolution of System States graph represents how the motor speed and armature current evolve over time as the reinforcement learning (RL)-based controller regulates the DC motor.

The motor speed curve shows how the system responds to the control input, aiming to reach the desired reference speed. Ideally, the speed should rise smoothly towards the target value with minimal oscillations or overshoot.

The armature current curve represents the electrical current flowing through the motor, which influences torque generation. Initially, a high current may be required to accelerate the motor, but it should stabilize as the system reaches steady state.

This plot provides insights into the dynamic behavior of the system under the actor-critic RL control strategy. If the motor speed stabilizes quickly and the armature current remains within safe limits, the controller is effectively learning to optimize motor performance. However, if large oscillations or instability appear in either variable, it may indicate issues with learning rates, reward functions, or system parameters that need further tuning.

4.1 Performance Metrics

The performance of the Actor-Critic controller is evaluated using the following metrics.
Settling Time: The time required for the motor to stabilize at the target speed. In this case Settling Time is 0.55 Sec. At 0.55 Sec DC motor reaches Reference Speed 100 rpm. Settling Time is 0.55 Sec. Another settling time is 1.04 Sec. At 1.04 Sec DC motor reach at Reference Speed 100 rpm.

Overshoot: The peak deviation from the desired motor speed. In this case Overshoot at 0.61 Sec. At 0.61 Sec DC motor reaches its Maximum peak overshoot Speed 122.17 rpm. At 0.91 Sec DC motor reaches its Minimum peak overshoot Speed 76.99 rpm.

Stability: The ability of the system to maintain constant speed without oscillations. In this case Settling Time is 1.56 Sec. At 1.56 Sec DC motor reach at Speed 96.0397 rpm. After 1.56 Sec motor runs at Constant Speed.

4.2 Comparison with Conventional Methods

Traditional control methods such as Proportional-Integral-Derivative (PID) controllers have been widely adopted for DC motor speed regulation due to their simplicity and ease of implementation. However, these methods rely heavily on accurate mathematical modelling of the system and are sensitive to parameter variations, disturbances, and nonlinearities
In contrast, the Actor-Critic reinforcement learning (RL) controller offers several key advantages.

Criteria	PID Controller	Actor-Critic RL Controller
Adaptability	Fixed gain values; manual tuning needed	Learns optimal control policy dynamically
System Modelling Requirement	Requires accurate system model	Model-free; learns directly from environment
Response to Disturbances	Limited adaptability	Adjusts to disturbances in real time
Nonlinearity Handling	Poor performance in nonlinear systems	Capable of handling complex, nonlinear dynamics
Training Time	None; parameters set manually	Requires training phase (offline or online)
Performance	Higher	Lower overshoot and faster convergence

Metrics	overshoot and longer settling	

In our MATLAB simulations, the PID controller exhibited larger overshoot and longer settling times when subjected to parameter variation or external disturbances. The RL-based controller, by contrast, adapted its policy during training and consistently reduced the error over time, resulting in smoother, more stable responses.

This comparison highlights the strength of reinforcement learning, especially in applications where the system behaviour is complex, time-varying, or poorly modelled. While RL methods involve a higher initial computational cost during training, the long-term benefits in robustness and performance make them a viable alternative for intelligent motor control in modern automation environments.

5. Conclusion

This study demonstrates the efficacy of the Actor-Critic reinforcement learning framework for real-time control of DC motors. The RL-based controller dynamically adjusts motor voltage based on continuous feedback from the environment, enabling adaptive, data-driven policy learning without requiring an explicit model of the motor dynamics. Simulation results in MATLAB confirm that the Actor-Critic method successfully achieves rapid convergence to the desired speed, reduces settling time, minimizes overshoot, and ensures stable long-term performance under varying operating conditions.

The critic's value function and the actor's policy are continually refined through temporal difference learning, making the system responsive to disturbances and nonlinearities. Performance metrics such as a settling time of 0.55 seconds and minimal oscillations illustrate the controller's superiority over conventional PID-based techniques. These outcomes validate the Actor-Critic approach as a robust, scalable solution for intelligent motor control in dynamic environments, with significant potential for deployment in industrial automation and robotics applications.

Future work may include extending the approach to multi-motor systems, implementing hardware-in-the-loop testing, and integrating safety constraints for real-world deployment.

6. References

1. Rossi, F., Grosso, G., &Gajani, G. S. (2023). A Reinforcement Learning-Based Controller for Optimal Speed Control of a DC Motor Using Deep Q-Network Algorithm. In *IEEE EUROCON 2023 - 20th International Conference on Smart Technologies*, 181-186.
2. Jeen Ann Abraham and Sanjeev Shrivastava. Dc motor speed control using machine learning algorithm. International journal of engineering research and technology, 7, 2018.
3. Konda, V. R., &Tsitsiklis, J. N. (2000). Actor-Critic Algorithms.*Neural Information Processing Systems (NeurIPS)*.
4. Li, M., & Yang, Z. (s2022). Application of Actor-Critic Methods for Motor Speed Control. *IEEE Transactions on Industrial Electronics*, *69*(5), 1525-1535.
5. Zhang, Y., Liu, D., & Chen, X. (2021). Reinforcement Learning for Robotic Arm Control Using Actor-Critic Methods. *IEEE Robotics and Automation Letters*, *6*(3), 1445-1452.
6. Guan, Q., Yao, X., Lin, Z., Wang, J., Iu, H., Fernando, T., & Zhang, X. (2024). A Robust

Control Scheme for PMSM Based on Integral Reinforcement Learning. *IEEE Transactions on Transportation Electrification*.

7. Staessens, T., Lefebvre, T., & Crevecoeur, G. (2022). Adaptive Control of a Mechatronic System Using Constrained Residual Reinforcement Learning. *IEEE Transactions on Industrial Electronics, 69*(10), 10447-10456.
8. Schenke, M., Kirchgässner, W., &Wallscheid, O. (2020). Controller Design for Electrical Drives by Deep Reinforcement Learning: A Proof of Concept. *IEEE Transactions on Industrial Informatics, 16*(7), 4650-4658.
9. Deng, Z., Huo, X., Du, Q., & Liu, Q. (2024). Reinforcement Learning-Based Data-Driven Control Design for Motion Control Systems. In *2024 36th Chinese Control and Decision Conference (CCDC)*, 5745-5749.
10. Lu, P., Huang, W., & Xiao, J. (2021). Speed Tracking of Brushless DC Motor Based on Deep Reinforcement Learning and PID. In *2021 7th International Conference on Condition Monitoring of Machinery in Non-Stationary Operations (CMMNO)*, 130-134.
11. Lee, J. Y., Park, J., & Choi, Y. (2012). Integral Reinforcement Learning with Explorations for Continuous-Time Nonlinear Systems. *International Joint Conference on Neural Networks (IJCNN)*, 1042-1047.
12. Puwei Lu, Wenkai Huang, and Junlong Xiao. Speed tracking of brush less dc motor based on deep reinforcement learning and pid. In 2021 7th International Conference on Condition Monitoring of Machinery in Non-Stationary Operations (CMMNO), pages 130–134, 2021.
13. SevilayTufenkci, BarisBaykanAlagoz, GurkanKavuran, CelaleddinYeroglu, Norbert Herencsar, and ShibenduMahata. A theoretical demonstration for reinforcement learning of pi control dynamics for optimal speed control of dc motors by using a twin delay deep deterministic policy gradient algorithm. Expert Systems with Applications, 213:119192, 2023.
14. Zhe Song, Jun Yang, Xuesong Mei, Tao, and Muxun Xu. Deep reinforcement learning for permanent magnet synchronous motor speed control systems. Neural Computing and Applications, 33:1–10, 05 2021.

Author's biography

First Author Name: Miss. Sweety Gorakhnath Chabukswar is currently pursuing M. Tech in Electrical Machine and Drive at Government College of Engineering, Aurangabad. She completed her B. Tech in Electrical Engineering. Her areas of interest in Machine learning.

Second Author Name: Dr. Nitin Jaykumar Phadkule is currently serving as the Head of the Department of Electrical Engineering at Government College of Engineering, Aurangabad. He holds a Ph.D. in Electrical Engineering and postgraduate degrees in both Electrical Power Systems and Instrumentation. He has also recently completed a PG Certificate in Machine Learning from the Indian Institute of Science (IISc), Bengaluru. His areas of interest include power system analysis, renewable energy integration, and deep learning for signal processing. He has developed several MATLAB apps for teaching and learning, which are widely used by students and educators alike. He is a member of the Institution of Engineers (India) and the Indian Society for Technical Education (ISTE).