Vol. 06, No. 4 (2024) 1593-1602, doi: 10.24874/PES06.04.018

Proceedings on Engineering Sciences

www.pesjournal.net

PATH PLANNING OF HUMANOID ROBOTS FOR STABLE MOTION USING REINFORCEMENT LEARNING BASED FUZZY LOGIC CONTROLLER

Rajeev Kumar
Laxman Singh[1]
Rajdev Tiwari

Received 06.02.2024.
Received in revised form 25.06.2024.
Accepted 08.07.2024.
UDC – 004.85

Keywords:

Humanoid Locomotion, Reinforcement learning, Fuzzy logic, Webots

A B S T R A C T

The locomotion and route planning of humanoid robots has become one of the utmost promising areas of research as humanoids are used more frequently in various fields of industrial automation and manufacturing. In this study, an effective solution for the navigation of humanoid robots is promised by fuzzy logic controllers. The fuzzy rule-base built and tuned by a human operator must be kept accurate, consistent, and complete, but this is challenging. One kind of machine learning is the reinforcement learning approach. The field of robotics frequently employs this strategy. When we suppose that the sole information collected is a scalar signal which is a reward or punishment, it seeks to automatically learn the fuzzy rules and to build a control law for a humanoid robot in an unfamiliar environment. The robot navigation in this study makes use of fuzzy controllers and the Q-learning algorithm. The outcomes of the simulation demonstrate appreciable improvements in the robot behaviors and learning rate in compare to latest state of art techniques available in recent literature. The outcomes are evaluated and discussed.

1. INTRODUCTION

A robot is considered humanoid if its general look is inspired by the human body (Kumar et al., 2019). Although certain types of humanoid robots may just represent a portion of the body, such as the upper torso, in general, humanoid robot consists of a torso with two arms, two legs and a head (Panwar & Kumar, 2012). Some humanoids could also be equipped with facial interfaces for their eyes, lips, and face (Kumar et al., 2021). Because it can adjust to changes in its surroundings or in itself, a humanoid robot is autonomous (Lindner et al., 2023). Robots designed to look like humans are made to perform some of the mental and physical tasks that people do every day. Engineers, cognitive scientists, and linguists are just a few of the experts who work together to develop robots that are as human-like as feasible (Marchetti et al., 2022). Their designers wanted the robot to recognize human aptitude, intelligence and behave similar to humans. If they are successful, humanoids may ultimately collaborate with us (Song & Kim, 2022).

There are several challenges to overcome in order to create a humanoid robot. The most difficult part is keeping the robot balanced while it executes its activity. Because gravity impacts how much we weigh, much like humans, robots are affected by gravitational force.

[1] Corresponding author: Laxman Singh
Email: laxman.mehlawat02@gmail.com

As a result, research in this area is still inspiring and challenging (Liu et al., 2023). Humanoids robots have gained widespread acceptance in society with the advancement of science and technology (Siciliano & Khatib, 2018). In a broad sense, humanoids are seen as entertainment robots, and in a limited sense, as robots that can aid humans. To replicate human deftness in a synthetic humanoid robot navigation system is a difficult task for researchers. The development of humanoid robots for environmental exploration and other humanoid robots to enhance maneuverability in a congested environment are both being made possible by the advancement of technology (Kumar, 2013). The biomechanics of human locomotion can be incorporated using this platform. As a result, robotics research places a high value on the application of locomotion and route planning to humanoid robots (Kumar et al., 2023).

The problem of robotic agents' locomotion has been studied by many researchers. A multiple robot's path planning with few source and target points including various obstacles is visualized in Fig. 1 to show the need of controller program. In 1965, Zadeh (Zadeh, 1999) invented the fuzzy notion for the first time. In the humanoid robot navigation system, Parhi (2005) adopted a fuzzy logic-based control paradigm. Samant et al. (2016) have suggested a technique for humanoid robot interaction in a congested environment. In order to validate the methodology, they use fuzzy logic in their experimental setup. A gait control strategy has been put forth by Wang et al. (2011) To address the issue of a humanoid robot's excessive energy consumption, they used fuzzy logic and an iterative process. An artificially intelligent humanoid robot has been proposed by Dadios et al. (2012) they demonstrated the humanoid robot's capacity to maintain balance, walk and circumvent obstructions. Mohanty and Parhi created a number of smart techniques for humanoid robot navigation that were inspired by nature (Singh et al., 2009).

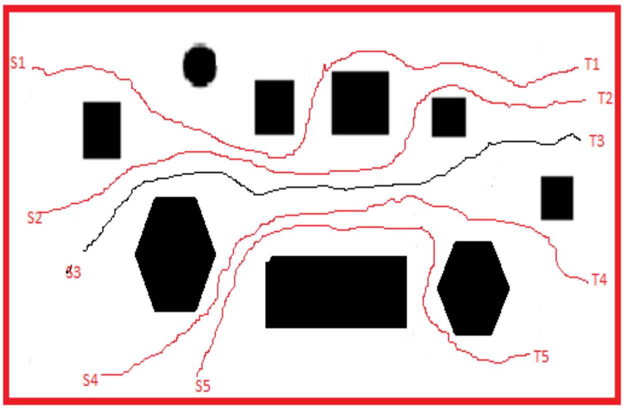


Figure 1. Multiple robots path planning with Source and Target Points

Fuzzy logic was employed by Pandey et al. as a possible locomotion strategy to help humanoid robotic agents avoid obstacles in complex environments (Pandey & Sonkar et al., 2014, Pandey & Parhi, 2014). Through numerous simulations and trials, they confirmed their method. Fuzzy logic has also been employed by Lei and Qiang (2010) to enhance the robot's ability to quickly and accurately detect the ball. An uneven terrain control method for humanoid robots has been presented by Zhong and Chen (2016) For the creation of neural networks and fuzzy logic controllers, a specific swarm optimization approach was employed. A flexible humanoid robot designed to assist elderly people has been introduced by Mohamed and Capi (2012), they talked about kinematics, mechatronics, and other robot specifics. The kinematics of a humanoid robot has been established by Flaherty et al. (2013). For the development of humanoid robots, Pierezan et al. approved the modified self-adaptive differential evolution (MSaDE) method, to confirm the effectiveness of MSaDE (Pierezan et al., 2017); they have conducted a number of trials. Children with disabilities like cerebral palsy and orthotics have been discussed by Wang et al. (2016). For the purpose of improving their treatment plans and reducing discomfort, they have deployed NAO humanoid robots. Robots with human-like features are trained to work like humans. A number of scholars (Pothal & Parhi, 2015; Eliot et al., 2012; Parhi & Mohanta, 2011; Kundu & Parhi, 2016) have made an effort to build the control architecture for humanoid robot route planning in complicated situations and have tested the effectiveness using appropriate simulation and experimental platforms. A humanoid robot and a person have imitated one other's poses, according to Lei et al. (2015). A pose similarity metric-based study was used to assess the imitation investigation between humans and humanoid robots.

The majority of researchers have attempted humanoid robot navigation and path planning in challenging contexts, as can be seen from the thorough review of the literature. On the navigation of humanoid robots, however, very few studies have been documented. Only certain environmental circumstances are suitable for the development of navigational algorithms. A reliable control method that can guide humanoid robots through difficult terrain regardless of the weather is required (Kofinas et al., 2013).

A petri-net control scheme is also employed to avoid the collision during multiple humanoid locomotion in the simulation environment. This scheme consists of few steps and provides priority to the humanoids based on the position of target, if the target is near to a particular humanoid than this humanoid has highest priority to move in the target direction and other humanoids exist on the workspace treated as static obstacles for a specified amount of time. (Muni et al., 2020).

The major contribution of the authors in this research work includes

1. Proposed a novel hybrid reinforcement learning based fuzzy logic controller for

humanoid robot locomotion in the workspace contains static and dynamic obstacles.

2. Focuses on the collision free locomotion of the single and multiple humanoid robots in the defined workspace.
3. Compare and contrast the proposed algorithm with state of art algorithms available in the literature and shows the superiority of the proposed technique in relation of locomotion speed, path smoothness and time taken to travel from source to destination.

In this study, the fuzzy rule base is employed to ascertain the necessary limits and velocities for navigating around obstacles in the specified workspace and reaching the intended target position without any harm. The controller takes into account sensory data related to the lengths and angles of obstacles towards the target as inputs.

For the implementation purpose of our idea, we use the Webot simulation software for simulations, and use an AMD RYZEN 5000 series processor machine equipped with 16 GB of RAM.

The paper is organized as follows, second section formulate the problem, the third section describes about the fuzzy logic and navigational system for humanoid robots, fourth section explains the meaning of reinforcement learning and Q-learning with the development of reinforcement learning based fuzzy logic controller algorithm, in section-5 we implement and discuss the obtained results to show the superiority of proposed algorithm over others. Section-6 concludes the outcome.

2. PROBLEM FORMULATION FOR OPTIMIZED LOCOMOTION

Take into consideration a humanoid robot moving through a terrain with a start point (X_{sp}, Y_{sp}), a target point (X_{tp}, Y_{tp}), and an obstacle at (X_{ob}, Y_{ob}). The terrain is made up of several robots that interact with one another to form dynamic barriers. Making the humanoid robot intelligent is the main goal in order to enable it to avoid both dynamic and static obstacles and arrive at the destination with the least amount of travel time. The robot should also ensure that the path is smooth and that it is the safest possible. In this section, the objective function for navigation is designed with following criteria in mind. The NAO humanoid robot is used to perform simulation experiments. The specific characteristics and configuration of NAO may be checked from Aldebaran robotics website.

In mathematical terms the objectives should be formulated to circumvent impediments and facilitate seamless movements along the shortest route towards the goal. The following two tactics will be implemented to achieve the tasks with minimal computational expenses.

1. **Goal seeking strategy**: The shortest and easiest route drawing is the basic task of the robot's route planning strategy for which the concept of Euclidean distance is utilized. For each point the robot distance from the target is updated to obtain the shortest distance between target and the robot. It should be expressed mathematically as

$$f_1(p,q) = d[p_r(i), q_r(i), (p_i, q_i)] \quad (1)$$

Where p_r (i), q_r (i) is the coordinate of the robotic agent at i^{th} location. There should be n points between initial and goal point, and the resultant will be minimum.

2. **Obstacle avoidance strategy**: This is also one of the fundamental requirements to avoid obstacle exist in the workspace for safe locomotion. It depends on the positions of obstacle and robot. Functionally we can express this problem as

$$f_2(p,q) = \left[p_{ob}(j), q_{ob}(j), \left(p_r(i), q_r(i)\right)\right] \quad (2)$$

There must be a safe distance between the robotic agent at i^{th} position and j^{th} obstacle.

3. FUZZY LOGIC AND FUZZY NAVIGATIONAL SYSTEM FOR HUMANOID ROBOTS

3.1 Fuzzy logic

Humanoid locomotion can be categorized into two distinct types: leg motion and trunk motion. The calculation of leg movement can be determined by the prevailing environmental variables. For example, if there is an obstacle during the leg's swinging phase, the foot may be able to surpass the obstacle by moving higher, while keeping the trunk stable. To optimize locomotion stability, the trunk should progress while the humanoid robot ascends a hill.

Fuzzy logic is a highly reliable control mechanism. Fuzzy logic is easily comprehensible, constructible, and applicable. Fuzzy logic can be applied to tackle engineering challenges using simple IF-THEN or IF-ELSE statements. Fuzzy logic was initially conceived as a means of efficiently handling large volumes of data by assigning values to different variables. Rather than serving as a control mechanism, it is a form of mathematical reasoning aimed at solving problems. Instead of simply determining true or false, it measures the degree of truth. The four fundamental components of a fuzzy logic controller include fuzzification of input variables, knowledge base, fuzzy reasoning, and defuzzification. The work utilized fuzzy logic to devise a trajectory for a humanoid robot navigating through a crowded area with randomly placed obstacles. The subsequent part of the discussion elucidates the fundamental principles of fuzzy logic. The objective of the study is to design a fuzzy controller that can be used for path planning of a humanoid robot inside a specified workspace.

3.2 Fuzzy navigational system for humanoid robots

The primary goal of any robotic control strategy is to ensure the robot maintains the greatest possible distance from obstacles while also minimizing the distance to the intended target. The robot employs a fuzzy logic system component to govern its artificial intelligence. This module ensures the equilibrium and steadiness of the robot during tasks such as walking and kicking. The software of the microcontroller, which is capable of being adjusted and altered, integrates fuzzy logic. As the implementation is done through software, this procedure takes place within the microcontroller. The tilt sensor provides the input values, while the output values determine the correct positions of the servo motors. The basic model of fuzzy logic controller is shown in Fig. 2.

Fuzzy logic is a methodology used in control systems to emulate the way humans make decisions by considering imprecise, ambiguous, erroneous, noisy, or incomplete input data. Fuzzy logic, as a general principle, transforms precise sensor inputs, represented as crisp values into membership values ranging from 0 to 1. After obtaining the membership values and establishing the set of rules, fuzzy reasoning is employed. The system employs a fuzzy set created from the preceding stage to govern the servo motors. Fuzzy logic systems may effectively interpret imprecise data and produce practical outcomes. Moreover, the robot can be operated without the need for excessively complex mathematical computations. Furthermore, due to the fuzzy logic system's ability to rectify these faults, the physical structure of the robot does not necessarily have to be very exact and intricate. Since fuzzy logic is implemented using software, making adjustments to the system is easier, more cost-effective, and does not necessitate more physical space, which would only increase the weight of the robot.

The navigational characteristics must be carefully taken into account for humanoid robot navigation. Maintaining a minimal distance from the desired target and a maximum distance from the barrier is the fundamental goal of the control algorithm. Here, the controller's inputs for obstacles are the Very close (VC), Close (CL), distant (DT), Very distant (VD) and Bearing Angle (BA) to the target. Move left (MLT), Move ahead left (MAL), Move ahead (MAH), Move ahead right (MAR), Move right (MRT) are obtained as the intended outputs once the controller has been processed. Following is a description of how fuzzy rules and the controller operate.

MLT: Move left
MAL: Move ahead left
MAH: Move ahead
MAR: Move ahead right
MRT: Move right

3.3 Fuzzy Rule base

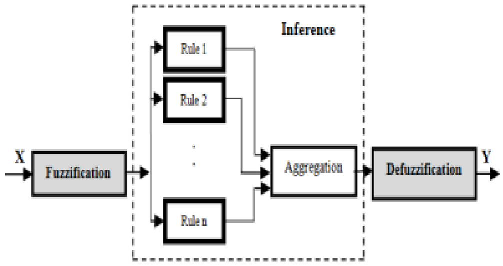


Figure 2. Basic model of fuzzy logic controller

Fuzzy linguistic variables for input and output

	LT	AL	AH	AR	RT
VC	MAH	MAR	MAL	MAL	MAH
CL	MAH	MAH	MRT	MAH	MAH
DT	MAH	MAH	MAR	MAH	MAH
VD	MAH	MAH	MAH	MAH	MAH

VC: Very close
CL: Close
DT: Distant
VD: Very distant
LT: Left
AL: Ahead left
AH: Ahead
AR: Ahead right
RT: Right

3.4 Fuzzy membership functions

1. If obstacle is very close (VC) and in left (LT) then move ahead (MAH)
2. If obstacle is very close (VC) and in ahead left (AL) then move ahead right (MAR)
3. If obstacle is very close (VC) and in ahead (AH) then move ahead left (MAL)
4. If obstacle is very close (VC) and in ahead right (AR) then move ahead left (MAL)
5. If obstacle is very close (VC) and in right (RT) then move ahead right (MAH)
6. If obstacle is close (CL) and in left (LT) then move ahead (MAH)
7. If obstacle is close (CL) and in ahead left (AL) then move ahead (MAH)
8. If obstacle is close (CL) and ahead (AH) then move right (MRT)
9. If obstacle is close (CL) and ahead right (AR) then move ahead (MAH)
10. If obstacle is close (CL) and in right (RT) then move ahead (MAH)

11. If obstacle is distant (DT) and in left (LT) then move ahead (MAH)
12. If obstacle is distant (DT) and in ahead left (AL) then move ahead (MAH)
13. If obstacle is distant (DT) and ahead (AH) then move ahead right (MAR)
14. If obstacle is distant (DT) and ahead right (AR) then move ahead (MAH)
15. If obstacle is distant (DT) and in right (RT) then move ahead (MAH)
16. If obstacle is very distant (VD) and in left (LT) then move ahead (MAH)
17. If obstacle is very distant (VD) and in ahead left (AL) then move ahead (MAH)
18. If obstacle is very distant (VD) and ahead (AH) then move ahead (MAH)
19. If obstacle is very distant (VD) and ahead right (AR) then move ahead (MAH)
20. If obstacle is very distant (VD) and in right (RT) then move ahead (MAH)

4. REINFORCEMENT LEARNING

Reinforcement learning (RL) is a machine learning approach that allows for the solution of a problem within a constrained timeframe by leveraging experimentally obtained information. An agent acquires the ability to optimize its interaction with a changing environment by employing reinforcement learning, which involves a process of experimentation and adjustment. The agent receives a scalar value as a reward for each action it takes. The agent aims to devise a strategy for decision-making that will optimize the anticipated total of discounted rewards. In the conventional framework of reinforcement learning, an agent engages with its environment through actions and perceptions. The agent assesses the state of the environment at each time step t and selects an action. The agent responds by receiving the scalar reinforcement signal r_t from the environment, which then transitions into state s_{t+1}. The agent should make judgments that aim to maximize the cumulative value of the reinforcement signal over the long term. Through a systematic process of trial and error, aided by a diverse set of methodologies, it can gradually acquire proficiency in this task. The membership function plots are shown in Fig. 3.

4.1 Q-learning

Q-Learning is an iterative dynamic programming that is used to address multistage decision issues. Among temporal difference algorithms, it is the most popular. The method consists of three primary components: an evaluation function, a reinforcement function, and an updating function. The objective of Q-learning is to construct a Q-function that maps the current state S_t and action a_t to a utility value $Q(s_t, a_t)$, that predicts the total future discounted rewards obtained from the current

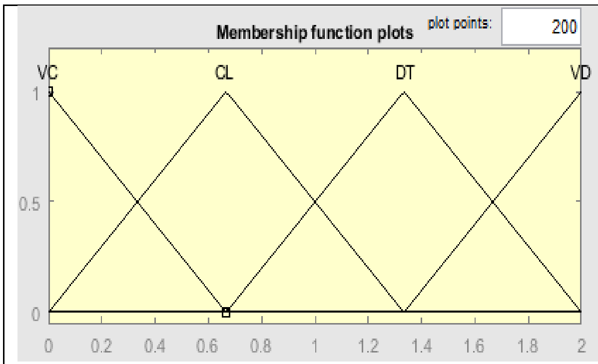


Figure 3a. Input variable "Distance (D)"

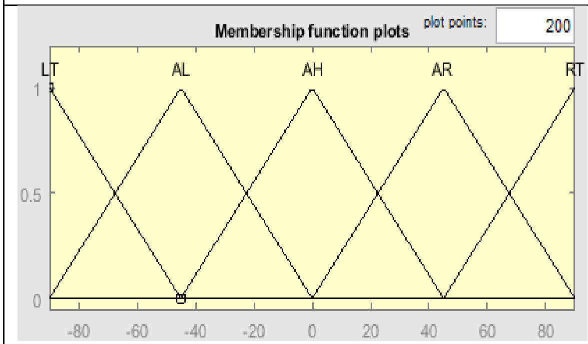


Figure 3b. Input Variable "Angular direction (ϕ)"

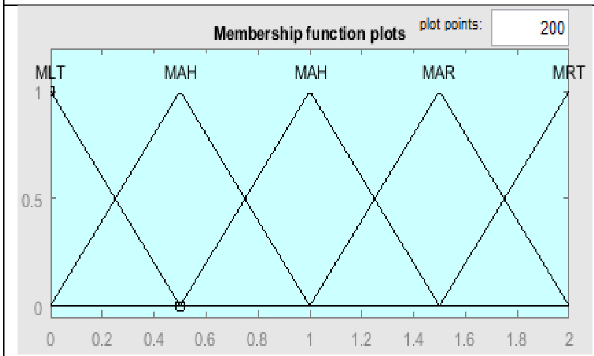


Figure 3c. Output variable "Deviation (δ)"

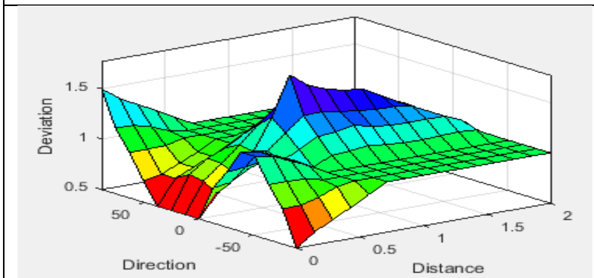


Figure 3d. Output control surface shows the relationship between input and output variables

Figure 3. Membership function plots of input and output variables of proposed fuzzy controller

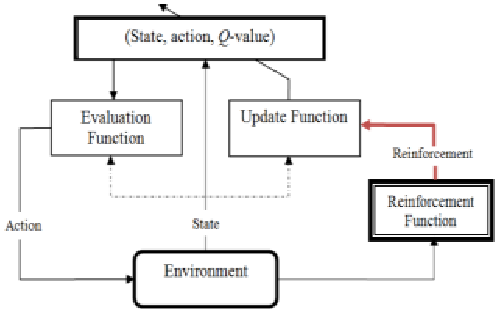


Figure 4. Model of Q-learning technique

action a_t. In that it learns the optimal policy function incrementally as it interacts with the environment after each transition (s_t, a_t, r_t, s_{t+1}). The model of Q-learning process is shown in Fig 4. Fig. 5 and Fig. 6 demonstrates the pseudo code of Q-learning technique and fuzzy Q-learning technique respectively.

1. Choose the FIS structure.
2. Initialize randomly $q[i,j]$, $i=1,\ldots m$ (m: rule number).
   $j=1,\ldots N$ (N: Number of proposed conclusions).
3. $t=0$, observe the state s_t
4. For each rule i, compute $w_i(s_t)$
5. For each rule i, choose a conclusion with the EEP.
6. Compute the action $A(s_t)$ and correspondence quality $Q(s_t, A(s_t))$
7. Apply the action $A(s_t)$. Observe the new state s'_t.
8. Receive the reinforcement r_t.
9. For each rule i, compute $w_i(s'_t)$.
10. Compute a new evaluation of the state value.
11. Update parameters $q[i,j]$ using this evaluation.
12. $t \leftarrow t+1$, Go to 5.

Figure 5. Pseudo code of Q-Learning technique

1. Initialize $Q(s,a)$ (with zeros or random values), $\forall(s,a)\in(S,A)$
2. Repeat (for each episode)
3. Initialize S_0
4. Repeat (for each step episode)
5. Choose a_t from S_t using a policy derived from Q.
6. Take action a_t, observe new state S_{t+1} and the reward r_t.
7. Update the quality function $Q(s,a)$.
8. $S_t = S_{t+1}$

Until S_t is terminal

Figure 6. Pseudo code for fuzzy Q-learning algorithm

4.2 Optimization of fuzzy systems using Q-learning algorithm

The most promising approaches to illustrate quality functions with continuous spaces of states and actions are fuzzy inference systems (FIS). The objective is to roughly determine the related Q-value for each state using the following equation:

$$s \rightarrow y = \widehat{Q} = FIS(s) \tag{3}$$

The concept behind this optimization is to suggest multiple outcomes for each rule and to link each outcome with a quality function that will be assessed over time. The training process allows for the acquisition of optimal rules that optimize future rewards. The Q-learning algorithm with fuzzy logic is referred to as the RL based fuzzy logic algorithm. Fig. 7 shows the flow chart of proposed reinforcement learning (RL) based fuzzy logic controller for humanoid robots.

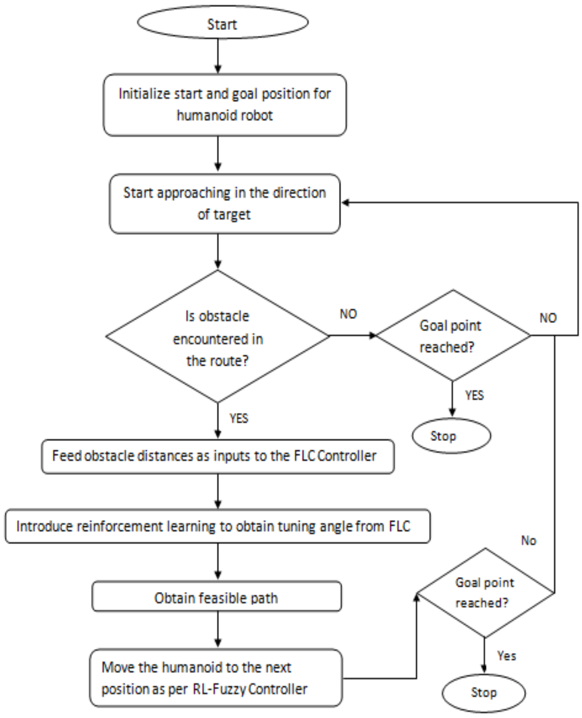


Figure 7. Flowchart of proposed RL based fuzzy logic controller for humanoid robots

5. IMPLEMENTATION AND DISCUSSION OF THE PROPOSED RL BASED FUZZY LOGIC CONTROLLER IN HUMANOID ROBOT PATH PLANNING

We use Webots simulator to implement all considered techniques including proposed controller. To perform experiments, we consider a humanoid NAO available in the Webots environment under proto nodes (Webots projects) to robots' tab followed by Softbank to NAO.

5.1 Locomotion of single robot using RA-ISMO, GWOC, PA-FLC, Simple fuzzy controller, and RL based fuzzy controller

The locomotion of humanoid robots is simulated using a simulated NAO in Webots software. The simulation takes place on a terrain with dimensions of 200*240. The terrain comprises a single humanoid NAO robot, a single source, a single target, and randomly positioned static obstacles. This setup is designed to assess the resilience of the proposed controller. Simulations are conducted for all controllers under consideration to compare their performance in terms of path length and time required to travel from the source to the target. Table 1 indicates that proposed technique shows an improvement of 5.31%, 1.25%, 4.75%, and 3.43% in path length with respect to RA-ISMO, GWOC, PA-FLC and Simple fuzzy controller. The values in bold indicates the average path length for a given technique. Table 2 shows the time spent to cover the same path from a specified source to the designated target and we have noticed an improvement of 45.02%, 8.16%, 13.17% and 9.56% in the proposed technique with respect to RA-ISMO, GWOC, PA-FLC and simple fuzzy controller. The values in bold indicates the average time taken by a particular algorithm.

Table 1. Relationship between simulated length (cm.) for RA-ISMO, GWOC, PA-FLC, simple fuzzy controller and RL based fuzzy controller

SI.NO.	RA-ISMO	GWOC	PA-FLC	Simple FC	RL based FC
1	320.1	304.56	315.58	311.76	301.66
2	313.5	305.78	316.78	312.22	301.12
3	316.9	306.54	315.91	312.00	303.21
4	320.2	305.91	316.44	311.44	301.78
5	315.9	304.86	315.50	312.19	302.34
6	316.4	305.71	316.04	310.98	301.11
7	319.6	304.89	315.22	312.11	301.13
8	320.4	306.56	316.34	313.67	301.88
9	315.2	304.98	315.89	312.32	301.33
10	319.2	305.23	316.66	311.88	301.69
Avg.	**317.74**	**305.50**	**316.04**	**312.06**	**301.72**

Table 2. Relationship among simulated time spent (S) for RA-ISMO, GWOC, PA-FLC, simple fuzzy controller and RL based fuzzy controller

SI.NO.	RA-ISMO	GWOC	PA-FLC	Simple FC	RL based FC
1.	57.19	41.46	43.53	42.98	38.22
2.	52.44	41.72	43.68	41.77	39.71
3.	54.26	41.35	43.41	42.66	38.66
4.	51.11	41.89	43.79	42.87	38.01
5.	57.00	41.46	43.92	41.55	39.44
6.	57.66	41.58	43.67	42.11	38.99
7.	57.75	41.33	42.98	41.99	38.67
8.	58.58	41.98	43.65	42.01	37.22
9.	56.25	41.89	43.55	42.58	37.89
10.	55.92	41.67	43.43	41.22	38.12
Avg.	**55.82**	**41.63**	**43.56**	**42.17**	**38.49**

5.2 Locomotion of multiple robots using RA-ISMO, GWOC, PA-FLC, Simple fuzzy controller, and RL based fuzzy controller

To validate the efficacy of the proposed hybrid controller for multiple humanoid locomotion, two numbers of NAOs are considered for simulation of path planning in complex terrain. The results are demonstrated in the table 3, table 4 and figure 8a and 8b. Table 3 indicates that proposed technique shows an improvement of 7.94 %, 3.39 %, 4.93% and 5.04 % in path length with respect to RA-ISMO, GWOC, PA-FLC and Simple fuzzy controller. The values in bold indicates the average path length for a given technique. Table 4 shows the time spent to cover the same path from a specified source to the designated target and we have noticed an improvement of 45.71%, 9.71%, 15.54% and 10.01% in the proposed technique with respect to RA-ISMO, GWOC, PA-FLC and simple fuzzy controller. The values in bold indicates the average time taken by a particular algorithm.

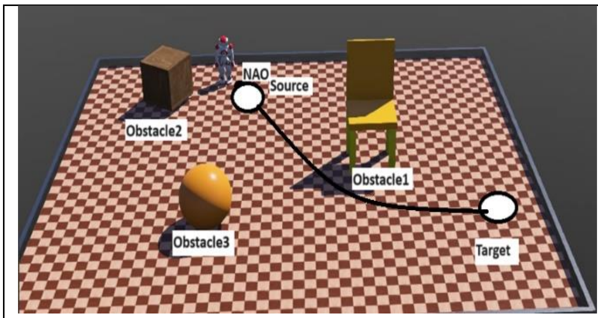
Table 3. Relationship between simulated length (cm.) for RA-ISMO, GWOC, PA-FLC, simple fuzzy controller and RL based fuzzy controller for two humanoids

SI.NO.	RA-ISMO	GWOC	PA-FLC	Simple FC	RL based FC
1.	322.11	312.44	315.00	318.09	302.65
2.	321.45	311.56	317.56	317.77	301.89
3.	322.67	312.87	315.88	318.45	302.67
4.	321.89	312.12	316.04	317.98	298.99
5.	322.77	311.99	315.11	318.39	301.43
6.	333.73	312.02	316.76	318.77	301.66
7.	331.32	311.86	317.88	317.88	302.00
8.	334.11	312.12	318.62	317.43	303.12
9.	335.66	311.62	318.00	316.76	302.36
10.	335.76	312.31	316.23	317.54	301.66
Avg.	**328.15**	**312.09**	**316.71**	**317.91**	**301.84**

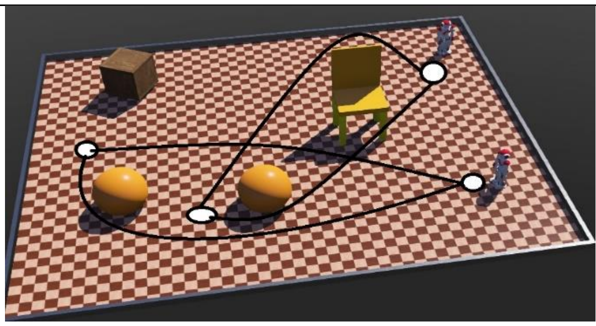
Table 4. Relationship among simulated time spent (S) for RA-ISMO, GWOC, PA-FLC, simple fuzzy controller and RL based fuzzy controller for two humanoids

SI.NO.	RA-ISMO	GWOC	PA-FLC	Simple FC	RL based FC
1.	55.32	43.76	45.54	42.63	38.71
2.	61.45	43.00	46.32	43.28	41.22
3.	58.33	42.69	45.97	42.33	40.56
4.	59.89	43.43	46.00	42.59	39.74
5.	57.62	44.21	46.48	44.01	38.55
6.	58.11	43.32	45.99	44.40	39.37
7.	60.78	44.08	46.71	45.21	40.12
8.	60.19	43.54	44.27	43.89	41.01
9.	54.00	43.99	45.58	44.98	38.69
10.	50.23	44.12	46.47	43.94	39.58
Avg.	**57.92**	**43.61**	**45.93**	**43.73**	**39.75**

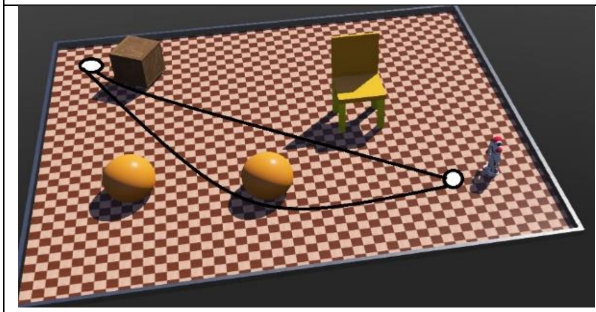
Figure 8 shows the Webots environment consisting of few static obstacles and paths from source to target. The environment consists of one humanoid NAO in figure 8a and 8b with three and four static obstacles respectively whereas figure 8c and 8d consists of two humanoid NAO placed at random positions and environment has 4 static obstacles. The humanoid also treated as a dynamic obstacle for one another and moves towards the target with obstacle avoidance. We use petri-net scheme to decide the priority of humanoid to move toward the target direction.



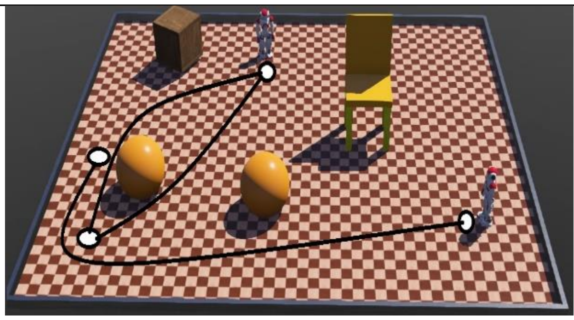
a. Single humanoid NAO with three obstacles



c. Two humanoid NAO with 4 obstacles one source and destination



b. Single humanoid NAO with four obstacles



d. Two humanoid NAO with 4 obstacles one source and destination with position change

Figure 8. Single and multiple humanoid path planning with static obstacles in Webots

6. Conclusion

This work presents a novel approach to collision-free movement planning for single and multiple NAOs in dynamic complicated terrain. The approach combines a Reinforcement Learning based fuzzy controller with a petri-net model. The model of the controller has been built using the conventional Q-learning process. The proposed controller undergoes testing in a simulated environment, and the obtained results are compared to cutting-edge techniques in terms of path length and travel time. The gathered findings demonstrate that the suggested method produces superior results in terms of travel time and path length while travelling from a specified source to a targeted destination.

References:

Dadios, E. P., Biliran, J. J. C., Garcia, R. R. G., Johnson, D., & Valencia, A. R. B. (2012). Humanoid robot: Design and fuzzy logic control technique for its intelligent behaviors. *Fuzzy Logic Controls, Concepts, Theories and Applications*. doi:10.5772/37188

Eliot, E., BBVL, D., & Parhi, D. R. (2012). Design & kinematic analysis of an articulated robotic manipulator. *International Journal of Mechanical and Industrial Engineering*, *1*(1). doi:10.47893/IJMIE.2014.1177

Flaherty, R. O., Vieira, P., & Grey, M. (2013). Kinematics and inverse kinematics for the humanoid robot HUBO2+. Georgia Institute of Technology.

Kofinas, N., Orfanoudakis, E., & Lagoudakis, M. G. (2013). Complete analytical inverse kinematics for NAO. *13th International Conference on Autonomous Robot Systems (Robotica)*, 1-6. doi:10.1109/Robotica.2013.6623524

Kumar, R. (2013). A new approach to design of CPG model for stable humanoid locomotion using neural oscillators. *International Journal of System and Software Engineering, 1*(1).

Kumar, R., Singh, L., & Tiwari, R. (2019). CPG design in bipedal locomotion by machine learning techniques: A review. *IJCA, 182*(50), 1-8. doi:10.5120/ijca2019918682

Kumar, R., Singh, L., & Tiwari, R. (2021). Path planning for the autonomous robots using modified grey wolf optimization approach. *Journal of Intelligent & Fuzzy Systems, 40*(2021), 9453-9470. doi:10.3233/JIFS-201926

Kumar, R., Singh, L., & Tiwari, R. (2023). Novel reinforcement learning guided enhanced variable weight grey wolf optimization (RLV-GWO) algorithm for multi-UAV path planning. *Wireless Personal Communications, 131*(1). doi:10.1007/s11277-023-10534-w

Kundu, S., & Parhi, D. R. (2016). Navigation of underwater robot based on dynamically adaptive harmony search algorithm. *Memetic Computing, 8*(2), 125-146. doi:10.1007/s12293-016-0179-0

Lei, S. H., & Qiang, W. A. (2010). Research on humanoid robot soccer system based on fuzzy logic. *International Journal of Computer Networks, 2*(1), 38. doi:10.5815/ijcnis.2010.01.06

Lei, J., Song, M., Li, Z. N., & Chen, C. (2015). Whole-body humanoid robot imitation with pose similarity evaluation. *Signal Processing, 108,* 136-146. doi:10.1016/j.sigpro.2014.08.030

Lindner, T., Wyrwal, D., & Milecki, A. (2023). An autonomous humanoid robot designed to assist a human with a gesture recognition system. *MDPI Electronics, 12*(12). https://doi.org/10.3390/electronics12122652

Liu, J., Misaros, M., Stan, O. P., Donca, I. C., & Miclea, L. C. (2023). Autonomous robots for services—State of the art, challenges, and research areas. *Sensors, 23*(10), 4962. doi: 10.3390/s23104962

Marchetti, A., Dio, C. D., Manzi, F., & Massaro, D. (2022). Robotics in clinical and developmental psychology. In *Comprehensive Clinical Psychology* (2nd ed., Vol. 10, pp. 121-140). doi: 10.1016/B978-0-12-818697-8.00005-4

Mohamed, Z., & Capi, G. (2012). Development of a new mobile humanoid robot for assisting elderly people. *Procedia Engineering, 4,* 345-351. doi:10.1016/j.proeng.2012.07.183

Muni, M. K., Parhi, D. R., & Kumar, P. B. (2020). Implementation of grey wolf optimization controller for multiple humanoid navigation. *Computer Animation and Virtual Worlds, 31*(8). doi:10.1002/cav.1919

Pandey, A., Sonkar, R. K., Pandey, K. K., & Parhi, D. R. (2014). Path planning navigation of mobile robot with obstacle avoidance using fuzzy logic controller. *IEEE 8th International Conference on Intelligent Systems and Control (ISCO),* 39-41. doi:10.1109/ISCO.2014.7103914

Pandey, A., & Parhi, D. R. (2014). MATLAB simulation for mobile robot navigation with hurdles in a cluttered environment using minimum rule-based fuzzy logic controller. *Procedia Technology, 14*, 28-34. doi:10.1016/j.protcy.2014.08.005

Panwar, V., & Kumar, R. (2012). Stable biped locomotion using central pattern generators based on Matsuoka neural oscillators. *IET International Conference on Information Science and Control Engineering, 3*, December 7-9, Shenzhen, China. doi:10.1049/cp.2012.2453

Parhi, D. R. (2005). Navigation of mobile robots using a fuzzy logic controller. *Journal of Intelligent Robotic Systems, 42*(3), 253-273. doi:10.1007/s10846-004-7195-x

Parhi, D. R., & Mohanta, J. C. (2011). Navigational control of several mobile robotic agents using Petri-potential-fuzzy hybrid controller. *Applied Soft Computing, 11*(4), 3546-3557. doi:10.1016/j.asoc.2011.01.027

Pierezan, J., Freire, R. Z., & Weihmann, L. (2017). Static force capability optimization of humanoid robots based on modified self-adaptive differential evolution. *Computer Operations Research, 84*, 205-215. doi:10.1016/j.cor.2016.10.011

Pothal, J. K., & Parhi, D. R. (2015). Navigation of multiple mobile robots in highly cluttered terrains using adaptive neuro-fuzzy inference system. *Robotica and Autonomous Systems, 72*, 48-58. doi:10.1016/j.robot.2015.04.007

Samant, R., Nair, S., & Kazi, F. (2016). Development of autonomous humanoid robot control for competitive environment using fuzzy logic and heuristic search. *IFAC Journal, 49*(1). doi:10.1016/j.ifacol.2016.03.082

Siciliano, B., & Khatib, O. (2018). Humanoid robots: Historical perspective, overview, and scope. In *Humanoid Robotics: A Reference* (pp. 3-8). doi:10.1007/978-94-007-6046-2_64

Singh, M. K., Parhi, D. R., & Pothal, J. K. (2009). ANFIS approach for navigation of mobile robots. *IEEE International Conference on Advances in Recent Technologies in Communication and Computing*, 722-731.

Song, C. S., & Kim, Y. K. (2022). The role of human-robot interaction in consumers' acceptance of humanoid retail service robots. *Journal of Business Research, 146*, 489-500. doi:10.1016/j.jbusres.2022.03.087

Wang, L. Y., Liu, Z., Zeng, X. J., & Zhang, Y. (2011). Gait control of humanoid robots via fuzzy logic and iterative optimization. *Chinese Control Conference*.

Wang, Z., Peyrodie, L., Cao, H., Agnani, O., Watelain, E., & Wang, H. (2016). Slow walking model for children with multiple disabilities via an application of humanoid robot. *Mechanical Systems and Signal Processing, 68*, 608-619. doi:10.1016/j.ymssp.2015.07.010

Zadeh, L. A. (1999). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Set Systems, 100*, 9-34. doi:/10.1016/S0165-0114(99)80004-9

Zhong, Q. B., & Chen, F. (2016). Trajectory planning for biped robot walking on uneven terrain taking stepping as an example. *CAAI Transactions on Intelligence Technology, 1*(3), 197-209. doi:10.1016/j.trit.2016.10.009

Rajeev Kumar
Dr. APJ ABDUL KALAM Technical University, Lucknow, India
crajiv1978@gmail.com
ORCID 0000-0001-6805-2032

Laxman Singh
Krishna Institute of Engineering & Technology, Ghaziabad
India
Laxman.mehlawat02@gmail.com
ORCID 0000-0002-0956-6989

Rajdev Tiwari
Ceekh Edunix Pvt. Ltd.
Noida-58
India
rajdevtiwari@yahoo.com
ORCID 0000-0002-0950-8690