International Journal of Integrated Engineering, Vol. 7 No. 2 (2015) p. 20-27

Reinforcement Learning Adaptive PID Controller for an Under-actuated Robot Arm

Adel Akbarimajd[1*]

[1]Faculty of Electrical Engineering, University of Mohaghegh Ardabili, Ardabil, Iran.

Abstract: An adaptive PID controller is used to control of a two degrees of freedom under actuated manipulator. An actor-critic based reinforcement learning is employed for tuning of parameters of the adaptive PID controller. Reinforcement learning is an unsupervised scheme wherein no reference exists to which convergence of algorithm is anticipated. Thus, it is appropriate for real time applications. Controller structure and learning equations as well as update rules are provided. Simulations are performed in SIMULINK and performance of the controller is compared with NARMA-L2 controller. The results verified good performance of the controller in tracking and disturbance rejection tests.

Keywords: Robot manipulator, Under-actuated mechanism, Adaptive PID controller, Reinforcement learning

1. Introduction

Under-actuated robot manipulator is a kinematic chain wherein total degree of the freedom of the mechanism is more than actuators. Under actuated manipulators are advantageous from the minimalism viewpoint in robotics where a task is performed with less energy consuming actuations. In fact dynamic of the mechanism is exploited instead of fighting it [1][2]. Moreover, studies on under-actuated can be beneficial in building of fault tolerant mechanisms, as when some joints of a fully actuated manipulator fail, the task can be continued before need for repairing them.

Control of under-actuated manipulators is a challenging issue because of their nonlinear characteristics and the lake of global controllability. Fortunately, It was proven that these manipulators have small-time locally controllability on an open subset of their zero velocity section, which allow them to follow any path in this subset [3]. This fact makes adaptive controllers as suitable choice for under-actuated manipulators.

Among different controllers, PID controllers are the most popular ones due to their simple implementation and high reliability. Moreover, in most cases model-free methods are available for tuning of PID parameters. As a result PID controllers have been extensively used in industries. Nevertheless, in time variant systems where the controller parameters should be adjusted according to variations in system dynamics, achieving good control performance is difficult.

Designing good performance adaptive PID controllers have been a challenging issue in recent years. In adaptive PID controllers, controller parameters should be tuned according to changes in system dynamics. Different structures have been introduced for adaptive PID controller. Three main categories of such structures include conventional adaptive PID controllers [4]-[6], fuzzy adaptive PID controllers [7]-[9] and evolutionary based adaptive PID controllers [10]-[12]. Conventional adaptive controllers exhibit low performance behavior, fuzzy adaptive PID controllers require prior knowledge of the system to be adequately tuned and evolutionary based adaptive PID-controllers are not appropriate for fast dynamic systems because of their required training time.

Neural network based adaptive controllers those employ supervised learning methods (ex. [10]) can be categorized in evolutionary adaptive controllers. As mentioned earlier, training process of these controllers need a period of time to be converged, which makes these controllers unsuitable for online instant applications. Unlike supervised learning approaches, there is no reference pattern in reinforcement learning methods. Reinforcement learning, which has origin in behaviorist psychology, adopts a test and verification method where the learning agent interacts with its environment and learns from the consequence of its actions [13][14]. As there is no reference to which convergence of algorithm is anticipated, the results of reinforcement learning can be instantly utilized, hence this learning approach can be employed in online and real time applications.

*Corresponding author: akbarimajd@uma.ac.ir

penerbit.uthm.edu.my/ojs/index.php/ijie

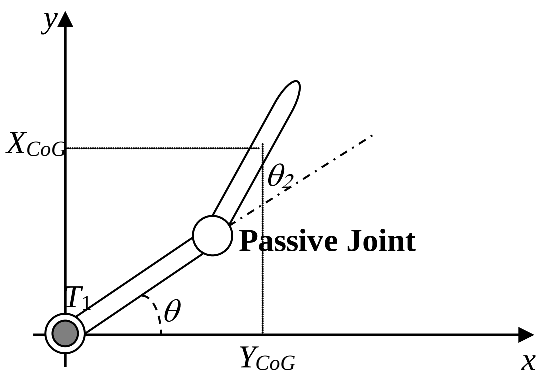

Fig. 1. Two-DoF under-actuated manipulator. First joint is active and the second one is passive

Actor-critic learning is one of the variants of reinforcement learning which provides a systematic method for simultaneously obtaining optimal action and expected value in real time [15][16]. Actor critic approach has been successfully used in different tasks related to robotics [17]-[20]. In this paper we employ an actor critic learning neural network based adaptive PID controller for motion control of a 2R under actuated robot whose second joint is passive (See Fig. 1). Parameters of the controller are tuned on-line and the controller adapts itself with variations of system dynamics.

The proposed method is a model-free intelligent approach, hence, the performance of the controller is compared with one of the recently published works that has the same approach of using a model-free intelligent technique[38].

2. Relate works to Control of Under-actuated Manipulators

There are lots of works reported in the framework of control of under-actuated manipulators. Arai and his colleagues [21][22] presented a control strategy based on holding of brakes on the passive joint in a 3-DoF under-actuated manipulator. They afterwards proposed a methodology to stabilize the trajectories of passive joint [23]. In [24] a measure of the dynamic coupling between the active and the passive joints was exploited as a cost function of an optimal control strategy that was applied to control of under-actuated manipulator. For a gravity-assisted under-actuated manipulator, a nonlinear closed loop control law that is guaranteed to be stable in positioning one unactuated joint at a time was presented in [25] where a Lyapunov function is introduced to prove the convergence of that control scheme. A robust adaptive control scheme for an under-actuated free-flying space robot is devised in [26]. In [27] a feedback linearization decoupling dynamic control scheme for one-passive joint under-actuated manipulators is proposed. Fuzzy sliding mode control was employed in [28] to control a 3R under-actuated manipulator. A motion planning method for a 3R under-actuated manipulator was presented by Lynch et al [29] and subsequently kinematical controllability of under-actuated systems was illustrated [30]. A controller was designed for a class of under-actuated manipulators to render the closed-loop equilibrium at the origin globally attractive [31]. De Luca and his colleagues used nilpotent approximations to control an under-actuated 2R manipulator [32][33] and performed a simple test to show that a planar 2R manipulator does not satisfy STLC conditions; therefore it will show spinning motion while steering from a certain configuration [34]. A point-to-point control method for a 2R planar under-actuated manipulator was introduced in [35] where passive link is firstly moved into its desired position, then the second link is moved into its desired position keeping the passive link at rest. A fuzzy controller for an under-actuated manipulator was developed in [36] whose member functions are optimized using genetic algorithm. Xin et. al. devised an energy based swing-up controller that uses a new Lyapunov function based on that transformation for *n*-link revolute planar robot with any one of the joint being a passive joint [37]. Akbarimajd and Kia have proposed neural network based nonlinear autoregressive moving average controller to stabilize a 2-DoF passive manipulator [38].

As review of above literature shows, different approaches have been proposed in order to control passive manipulators. However, as our proposed method is a neural network based intelligent and model-free method, to evaluate performance of our method, it would be fair to compare it with a similar one. Among recent works in this area [38] is also neural network based intelligent and model-free method. Therefore we will compare the results of the controller with that work. Details will be provided in next sections.

3. Reinforcement learning Adaptive PID controller

3.1 Control structure

Fig. 2 shows schematic diagram of employed adaptive PID controller in which learning approach is based on reinforcement learning actor-critic idea. The controller is incremental PID controller with coefficients $K(t)=[k_I(t),k_P(t),k_D(t)]$. In Fig. 1, bold thick lines show vector signals and thin lines show scalar signals. $y_d(t)$ and $y(t)$ are desired and real outputs of the system. Error signal *e(t)* in converted to state vector $\mathbf{x}(t)$ via a state convertor block. Actor critic learning process in encircled inside dashed line. Illustrated learning process includes three basic components of an actor-critic learning: an actor, a critic and Stochastic Action Mediator (SAM). Actor is used for estimation of policy function and realizes a mapping from current state to prior PID parameters $K'(t)=[k'_I(t),k'_P(t),k'_D(t)]$ those will not directly contribute in control process.

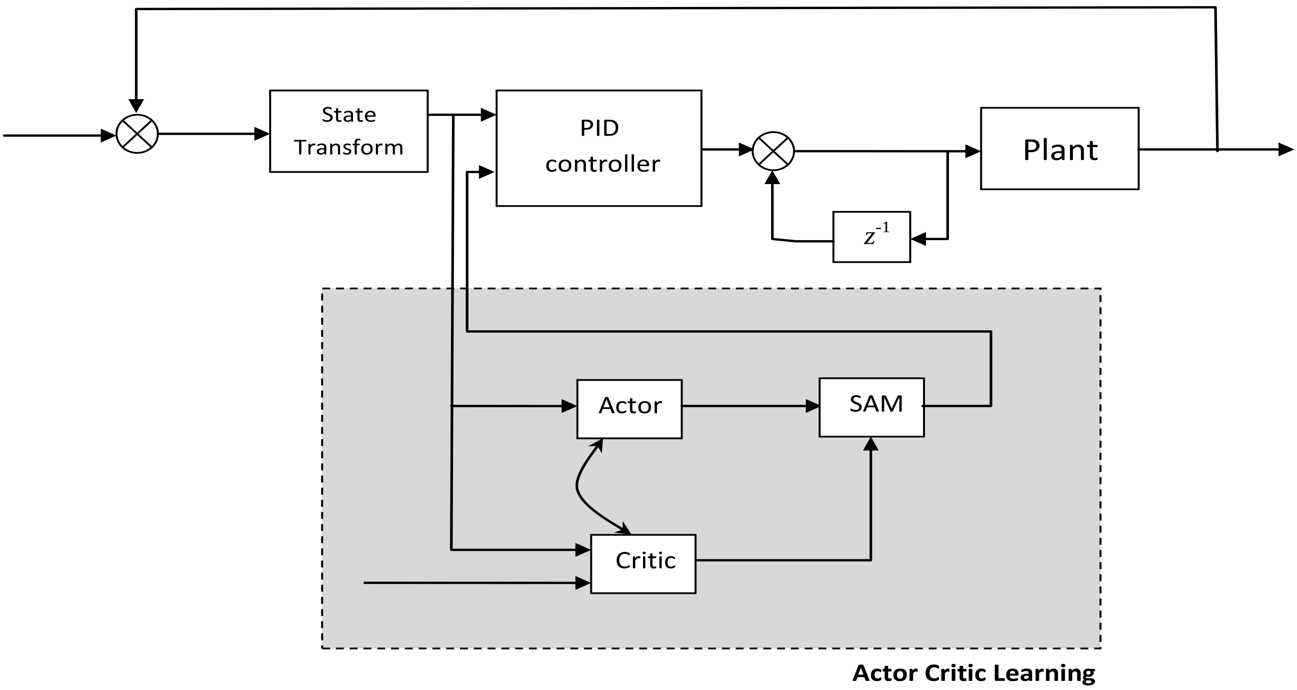

Fig. 2. Structure of actor-critic learning adaptive PID controller

SAM is used to generate real PID coefficients $K(t)$ based on prior PID parameters $K'(t)$ suggested by the actor and estimated signal $V(t)$. The critic receives state and immediate external reinforcement signal (say immediate reward) from the environment and produces error signal $\delta_{td}(t)$ and estimation value function $V(t)$. $\delta_{td}(t)$ is directly prepared for the actor and the critic and it is behaved as a basis of updating parameters of the actor and the critic. $V(t)$ is sent to SAM and is employed to modify output of the vector.

Effect of the error signal and its variations on performance of control should be simultaneously considered in designing of external reinforcement signal $r(t)$. Thus, $r(t)$ is defined as:

$$r(t) = \alpha r_e(t) + \beta r_{ec}(t) \tag{1}$$

where α and β are weighting factors and:

$$r_e(t) = \begin{cases} 0 & |e(t)| < \varepsilon \\ -0.5 & otherwise \end{cases} \tag{2}$$

$$r_{ec}(t) = \begin{cases} 0 & |e(t)| < |e(t-1)| \\ -0.5 & otherwise \end{cases} \tag{3}$$

and ε determines tolerable error band.

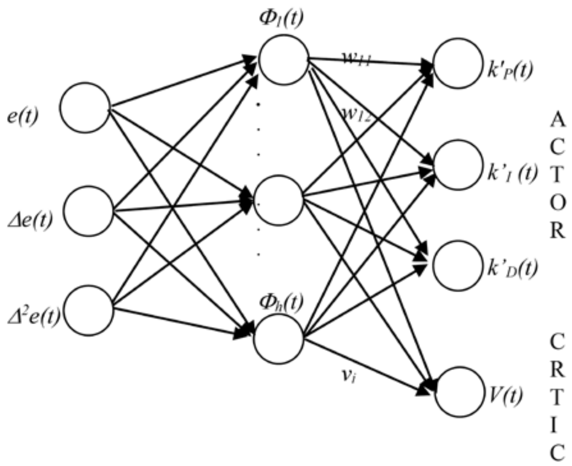

Fig. 3. RBF network employed for learning of actor-critic

3.2 Actor critic learning based on RBF

RBF is a multilayered feed forward neural network. Structure of RBF is shown in Fig. 2. A RBF network is employed for implementation of learning processes of value function of the critic and the policy function of the actor. Layeres of the network and their role is described as the sequel. In input layer each neuron is a system state variable x_i and state vector $\mathbf{x}(t) \in R^3$ is directly supplied to the next layer known as hidden layer. In hidden layer, the kernel function of each neuron is selected to be Gaussian function where output of j^{th} neuron is given as:

$$\Phi_j(t) = \exp\left(-\frac{\|\mathbf{x}(t) - \boldsymbol{\mu}_j(t)\|^2}{2\sigma_j^2(t)}\right), \quad j = 1,\ldots,h \qquad (4)$$

where $\boldsymbol{\mu}_j$ is center vector and σ_j is width scalar of j^{th} neuron and h is the size of the hidden layer. Finally output layer is composed of two parts including actor part and critic part. m^{th} output of the actor part can be calculated as:

$$\mathbf{K}'_m(t) = \sum_{j=1}^{h} w_{mj}(t)\Phi_j(t), \quad m = 1,2,3 \qquad (5)$$

Then real PID parameters are obtained as:

$$\mathbf{K}(t) = \mathbf{K}'(t) + n_k(0, \sigma_V(t)) \qquad (6)$$

where n_k is Gaussian noise with $\sigma_V(t) = \frac{1}{1+e^{2V(t)}}$.

When $V(t)$ is large n_k is small and vice versa and this provides a good tradeoff between exploration and exploitation.

In actor-critic learning the actor learns the policy function and the critic learns the value function using TD-error $\delta_{TD}(t)$ which by itself is calculated as:

$$\delta_{TD}(t) = r(t) + \gamma V(t) - V(t-1) \qquad (7)$$

where $0 < \gamma < 1$ is discount factor. Performance index of the learning system is defined as:

$$E(t) = \frac{1}{2}\delta_{TD}^2(t) \qquad (8)$$

Weights of network $w_{mj}(t)$, $v_j(t)$ and $\mu_{ij}(t)$ are updated to minimize above index and through a gradient descent method and chain rule (for details see [39]).

3.3 Controller design

The Based on discussions of previous sections, stages of designing of adaptive PID controller can be illustrated as the following:

Step 1. Arbitrarily set initial values for parameters of the learning system including $\alpha, \beta, \varepsilon, \gamma, \alpha_A$ and network parameters including $a_c, \eta_c, \eta_\mu, \sigma_j(0), w_{mj}(0), v_j(0), \mu_{ij}(0)$

Step 2. Read real output $y(t)$ and calculate $e(t)$, $\Delta e(t)$, $\Delta^2 e(t)$

Step 3. Receive immediate reward $r(t)$

Step 4. Obtain outcome of actor $\mathbf{K}'(t)$ and value function of critic $V(t)$.

Step 5. Calculate real parameters of PID controller $\mathbf{K}(t)$ and accordingly find control signal $u(t)$

Step 6. Apply $u(t)$ to the system and get output $y(t+1)$ and $r(t+1)$ for next time step.

Step 7. Calculate $\mathbf{K}'(t+1)$ and $V(t+1)$

Step 8. Calculate TD error $\delta_{TD}(t)$

Step 9. Update weights of the actor and the critic

Step 10. Update weights of RBF network.

Step 11. If the final time is not achieved, go to Step 2.

4. NARMA-L2 Controller

In this section we briefly introduce NARMA-L2 controller that have been used in [38] to control passive manipulator of Fig. 1. NARMA-L2 is one of the most appropriate architectures for prediction and control of time variant nonlinear systems. This control technique is based on input output linearization. Block diagram of NARMA-L2 controller is shown in Fig. 4. There are two basic steps in NARMA-L2 including identification step and controller design step.

In identification step, the following nonlinear autoregressive moving average model is adopted for the system:

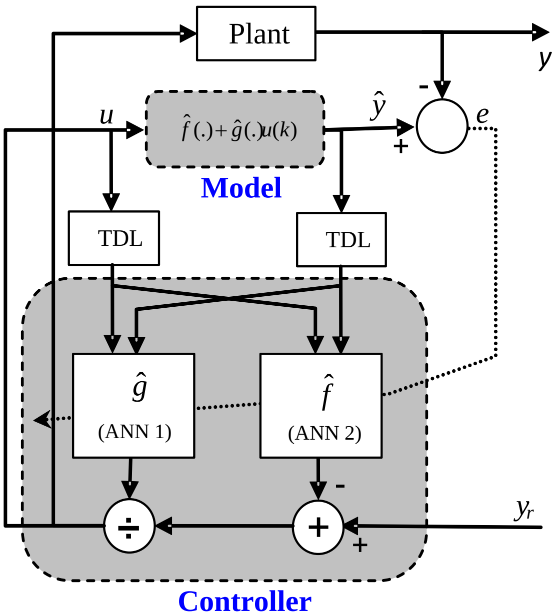

Fig. 4. NARMA-L2 Controller

$$\hat{y}(k+d) = \hat{f}\big(y(k), y(k-1),..., y(k-n+1),$$
$$u(k-1),...,u(k-m+1)$$
$$+\hat{g}\big(y(k), y(k-1),..., y(k-n+1),$$
$$.u(k-1),...,u(k-m+1)\big)u(k) \quad (9)$$

where $u(k)$ and $y(k)$ are the system input and output respectively and d is the relative degree. The positive integers m and n are respectively the number of measured delayed values of inputs and outputs. $\hat{f}$ and $\hat{g}$ are approximated by two MLP neural networks (see Fig. 4).

In controller design step, using (9) the control rule is given by:

$$u(k) = \frac{A(k)}{B(k)} \quad (10)$$

where

$$A'(k) = y_r(k+d) - \hat{f}\big(y(k),..., y(k-n+1),$$
$$u(k),...,u(k-m+1)\big)$$

and

$$B'(k) = \hat{g}\big(y(k),..., y(k-n+1),$$
$$u(k),...,u(k-m+1)\big)$$

The control rule (10) is not realizable since input computation of $u(k)$ requires the output signal $y(k)$. A more practical form can be represented as the following

$$u(k+1) = \frac{A'(k)}{B'(k)} \quad (11)$$

where

$$A(k) = y_r(k+d) - \hat{f}\big(y(k),..., y(k-n+1),$$
$$u(k-1),...,u(k-m+1)\big)$$

and

$$B(k) = \hat{g}\big(y(k),..., y(k-n+1),$$
$$u(k-1),...,u(k-m+1)\big)$$

Rule (11) is realizable for $d>1$. For more details about this controller see for example [38].

5. Simulation Results and Discussion

In order to design proposed adaptive PID controller for under-actuated manipulator of Fig. 1, control signal u and system output y should be determined. For this manipulator, control signal would the torque applied to the base joint. Output signal should be selected according to control goal. Namely the output signal could be Y_{CoG} or X_{CoG} of the second link or a function of these coordinates. Design of a regulator for an appropriate selection of output function, renders CoG of the passive link stay in a specific area.

A model of 2R planar under-actuated manipulator was constructed using SimMechanics in SIMULINK. Links are similar with the parameters masses $m_1=m_2=2$kg, moments of inertia $I_1=I_2=0.1$N.m. and length $l_1=l_2=0.2$m. The output is selected as y coordinate of the CoG of the second link, i.e. $y=Y_{CoG}$. Input is torque applied to the active joint i.e. $u=T_1$. We also assumed that the joints are frictionless.

Using abovementioned input-output pair of signals both the proposed controller and NARMA-L2 controller are designed. The results are devised and compared at the sequel.

5.1 Tracking test

In tracking simulation, the manipulator is initialized at fully extended configuration which means joint angles are $\theta_1=0$ and $\theta_2=0$. In this configuration $Y_{CoG}=0.0$. We applied a square wave reference trajectory for Y_{CoG} as it is shown in Fig. 4. The reference signal is 0.5m for first 1 second then it switches down to 0 at t=1sec. The system is simulated with both the proposed controller and NARMA-L2 controller of [38]. Fig 5 shows snapshots of simulations with the prposed controller. Fig. 6 shows output of the system (Y_{CoG}) with the proposed controller and NARMA-L2 of [38]. With both controllers the system can track the reference. However, the performance of the proposed controller is better that the controller of [38] in terms of transient response. The evidence is that transient time corresponding to the proposed controller and controller of [38] are t_s=2.7sec and t_s=5sec respectively. Moreover, the response of the proposed controller has no overshoot where the overshoot of NARMA-L2 is about 11%. It is noteworthy that these values are calculated after t=1sec after learning time of the controller. In the first second the controller learns the plant and its performance is not very good (however it is still better that NARMA-L2) which is usual event in adaptive controllers.

5.2 Disturbance rejection test

In this set of simulations, we also assumed that the robot is in the fully extended configuration. Simulation starts and the robot remains in this state until t=1sec. At time t=1sec a pulse disturbance with 0.15sec in time duration and 0.25N.m in amplitude – as shown in Fig. 7 – is added to the control signal. In fact, a disturbance torque is inserted to the active joint.

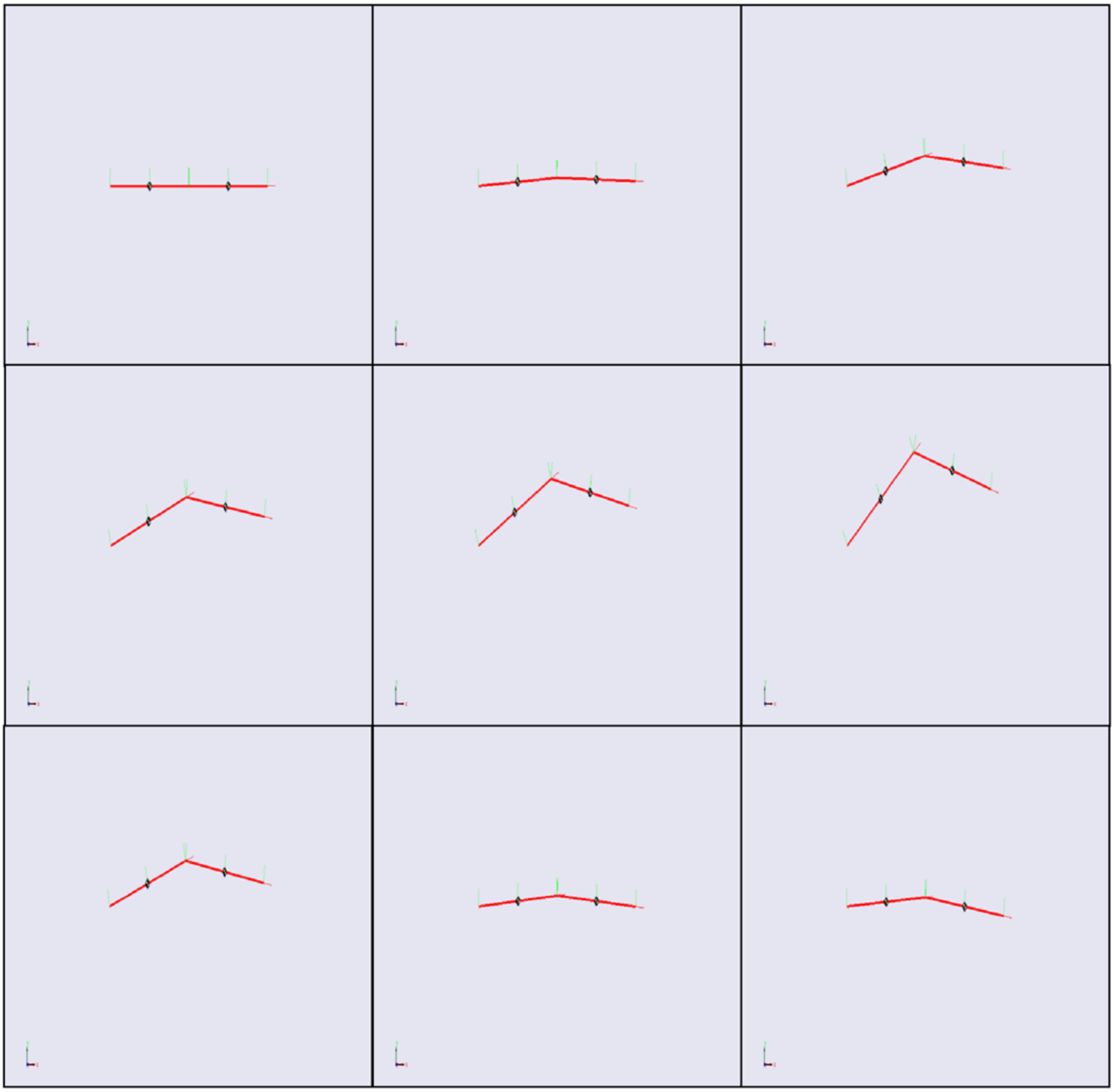

Fig. 5. Snapshots of tracking simulation

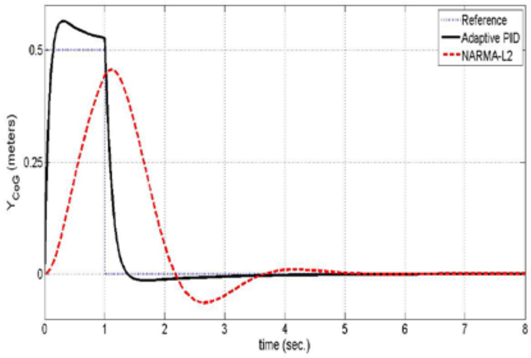

Fig. 6. Tracking simulation of the proposed Adaptive-PID controller and NARMA-L2 controller

As a result the passive joint deviates from its initial position but the proposed controller and controller of [38] both restore Y_{CoG} back to zero. Over again, time response of the proposed controller is better than NARMA-L2. In next simulation we increased the magnitude of the disturbance to 0.35N.m. From Fig. 8 it is evident that the system with NARMA-L2 controller has become unstable. The proposed controller can reject disturbances with magnitudes less 0.35N.m while this limit for controller of [38] is 0.3N.m. Actually, the proposed controller is better disturbance rejection performance than that of [38] in terms of both response time and disturbance tolerable limit.

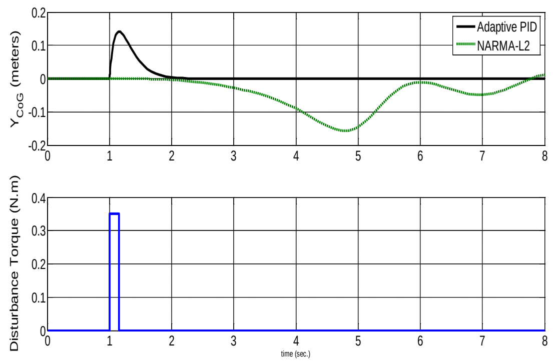

Fig. 7. Disturbance simulation of the proposed Adaptive-PID controller and NARMA-L2 controller (top) by a disturbance with magnitude 0.25N.m (bottom). Both systems can restore the response.

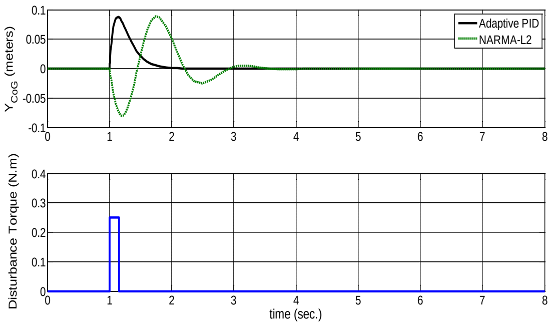

Fig. 8. Disturbance simulation of the proposed Adaptive-PID controller and NARMA-L2 controller (top) by a disturbance with magnitude 0.35N.m (bottom). NARMA-L2 is unstable.

Above simulation results verify that the proposed adaptive PID controller has stabilized the mechanism and it is robust to external disturbances and its performance is better than NARMA-L2 controller.

6. Summary

An adaptive PID controller tuned by an actor-critic reinforcement learning was successfully employed in a under actuated manipulator. It was illustrated that the controller has good performance in one coordinate tracking. The controller could resist against pulse disturbances. Simulations in MATLAB/SIMULNK and comparisons with NARMA-L2 controller verified these arguments.

By extending the idea for multivariable controllers, the controller will be able to perform full coordinate tracking.

References

[1] K. M. Lynch, “Toppling Manipulation”, in IEEE International Conference on Robotics and Automation, Michigan, May 1999

[2] Control of Underactuated Manipulators", Website of Advanced Industrial Science and Technology of Japan (AIST): http://www.aist.go.jp/MEL/soshiki/robot/biorobo/arai/uam_e.html

[3] K. M. Lynch , N. Shiroma, H. Arai, K. Tanie, "Motion Planning for a 3-DOF Robot with a Passive Joint", in Proceeding of 1998 IEEE International Conference on Robotics and Automatation, pp 927 – 932, Leuven, May1998

[4] A. Pomerleau, A. Desbiens, & D. Hodouin, “Development and evaluation of an auto-tuning and adaptive PID controller”. *Automatica*, *32*(1), 71-82., 1996

[5] H. B. Siahaan, H. Jin, and M. G. Safonov, “An adaptive PID switching controller for pressure regulation in drilling” In *Proceedings of the IFAC workshop on automatic control in offshore oil and gas production (ACOOG 2012)* pp. 90-94, May 2012.

[6] I. Mizumoto, I., D. Ikeda, T. Hirahata, & Z. Iwai, “Design of discrete time adaptive PID control systems with parallel feedforward compensator”. *Control Engineering Practice*, *18*(2), 168-176., 2010

[7] I. Z. Zhang, W. J. YANG, and A. X. Zhang, “Research on Fuzzy Self-adaptive PID Control and Its Emulation [J]”. Computer Simulation, 9, 132-135, 2009

[8] F. Zhang, Y. H. Lu, F. Qiao, and C. C. Bai, “Self-Tuning Fuzzy Adaptive PID Pitch Control of Wind Power Systems” *Applied Mechanics and Materials*, *394*, 404-409, 2013

[9] Z. Wang, S. C. Wang, X. Chen, Y. Li, and Y. Y. Jiang, “The Slip Frequency Control Based on Fuzzy Adaptive PID Algorithm” *Advanced Materials Research*, *614*, 1595-1599, 2013

[10] B. Guo, H. Liu, Z. Luo and F. Wang, “Adaptive PID controller based on BP neural network”. In *Artificial Intelligence, 2009. JCAI'09. IEEE International Joint Conference*, pp. 148-150, April 2009

[11] L. dos Santos Coelho, “Tuning of PID controller for an automatic regulator voltage system using chaotic optimization approach”. *Chaos, Solitons & Fractals*, *39*(4), 1504-1514, 2009

[12] B. Yang, G. A. Liang, J. H. Peng, S. H. Guo, W. Li, S. M. Zhang, and S. Bai, “Self-adaptive PID controller of microwave drying rotary device tuning on-line by genetic algorithms”. *Journal of Central South University*, *20*, 2685-2692, 2013

[13] L. P. Kaelbling, M. L. Littman, A. W. Moore “Reinforcement learning: A survey”. *arXiv preprint cs/9605103*., 1996

[14] A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 1998

[15] H. Kimura and S. Kobayashi, “An Analysis of Actor/Critic Algorithms Using Eligibility Traces: Reinforcement Learning with Imperfect Value Function”, In *ICML*, pp. 278-286, July 1998

[16] V. R. Konda, and J. N. Tsitsiklis, “Actor-Critic Algorithms” In *NIPS* Vol. 13, pp. 1008-1014, November 1999

[17] B. Kim, J. Park, S. Park, and S. Kang, “Impedance learning for robotic contact tasks using natural actor-critic algorithm”. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, *40*(2), 433-443, 2010
[18] E. A. Pohlmeyer, B. Mahmoudi, S. Geng, N. Prins, and J. C. Sanchez, “Brain-machine interface control of a robot arm using actor-critic rainforcement learning” In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* pp. 4108-4111, 2012
[19] C. Li, R. Lowe, and T. Ziemke, “Crawling Posture Learning in Humanoid Robots using a Natural-Actor-Critic CPG Architecture” In *Advances in Artificial Life, ECAL*, vol. 12, pp. 1182-1190, 2013
[20] K. Wang, Y. Jia, J. Du, and F. Yu, “Online Actor-Critic Learning for Motion Control of Non-holonomic Mobile Robot”, In *Electrical, Information Engineering [21]and Mechatronics 2011*, pp. 1763-1770, Springer London, 2012
[21] H. Arai, K. Tanie, S. Tachi, Dynamic control of a manipulator with passive joints in operational space, IEEE Transactions on Robotics & Automation 9, 85–93, February. 1993
[22] H. Arai, S. Tachi, "Position control of a manipulator with passive joints using dynamic coupling", IEEE Transactions on Robotics & Automation 528–534., August 1991
[23] H. Arai, K. Tanie, N. Shiroma, "Nonholonomic control of a three-dof planar underactuated manipulator", IEEE Transactions on Robotics & Automation 14 681–695, October 1998.
[24] M. Bergerman, Y. Xu, "Optimal control sequence for underactuated manipulators", in: Proceedings of the IEEE International Conference on Robotics and Automation, Minneapolis, USA, pp. 3714–3719., April 1996
[25] K. Kobayashi, T. Yoshikawa, "Controllability of under-actuated planar manipulators with one unactuated joint", In the Proceeding of International Conference on Intelligent Robots and Systems, (IROS 2000)., vol.1 pp 133 – 138, Takamatsu, 2000
[26] A. Benallegae, " Adaptive control for flexible-joint robots using a passive systems approach", Journal of Control Engineering Practice, Volume 3, Number 10, pp. 1393-1400, October 1995
[27] J.-H. Shin, J.-J. Lee, "Dynamic control of underactuated manipulators with free-swinging passive joints in Cartesian space", in: Proceedings of the IEEE International Conference on Robotics and Automation, , pp. 3294–3299., Albuquerque, NM, 1997
[28] K. Won; K. Min-Soeng; S. Jin-Ho; L., Ju-Jang, "Fuzzy Sliding Mode Control for a Robot Manipulator with Passive Joints ", Proceedings of Control, Automation and Robotics, pp. 324-327, 1998
[29] K.M. Lynch, N. Shiroma, H. Arai, K. Tanie, "Collision-free trajectory planning for a 3-DOF robot with a passive joint", International Journal of Robot Research 19 (12) 1171–1184, 2000
[30] F. Bullo, K.M. Lynch, "Kinematic controllability for decoupled trajectory planning in underactuated mechanical systems", IEEE Transactions on Robotics & Automation 17, 402–412., August 2001
[31] M. Reyhanoglu, S. Cho, N.H. McClamroch, "Discontinuous feedback control of a special class of underactuated mechanical systems", International Journal of Robust and Nonlinear Control 24, 265–281., July 2000
[32] A.D. Luca, R. Mattone, G. Oriolo, "Stabilization of an underactuated planar 2R manipulator", International Journal of Robust and Nonlinear Control 24 181–198., July 2000
[33] A.D. Luca, G. Oriolo, "Trajectory planning and control for planar robots with passive last joint", International Journal of Robotics Research 21 (5–6) 575–590, 2002
[34] A.D. Luca, S. Iannitti, "A simple STLC test for mechanical systems underactuated by one control", in: Proceedings of the IEEE International Conference on Robotics and Automation, Washington, DC, pp. 1735– 1740. May 2002
[35] A. Mahindrakar, S.Rao, R.N. Banavar, "Point-to-point control of a 2R planar horizontal underactuated manipulator", Journal of Mechanism and Machine Theory vol. 41 pp, 838–844, 2006
[36] Q. Liu, Y. Yu, Q. Xia, L. Su, "A new fuzzy method for the motion control of underactuated robots based on genetic algorithm", In the Proceedings of IEEE International Conference on Fuzzy Systems, pp 999 – 1003, Hong Kong, June 2008
[37] X. Xin, J.H. She, T. Yamasaki, "Swing-up control based on virtual composite links for n-link underactuated robot with passive first joint", In the Proceeding of 47th IEEE Conference on Decision and Control, 2008. CDC, Cancun, December 2008.
[38] A. Akbarimajd, S. Kia, “NARMA-L2 Controller for 2-DoF Underactuated Planar Manipulator”, 2010 11th. Int. Conf. Control, Automation, Robotics and Vision, Singapore, 5-8th December 2010
[39] X. Wang, Y. H. Cheng, Wei, Sun, “A Proposal of Adaptive PID Controller Based on Reinforcement Learning”, J China Univ Mining and Technol , 17(1): 0040–0044, 2007.