# THESIS INFORMATION

| | |
|---|---|
| Title: | Research on reinforcement learning-based robust adaptive control algorithms for nonlinear systems |
| Major: | Automation |
| Code: | 62.52.60.01 |
| Ph.D student: | Nguyen Tan Luy |
| Advisors: | Dr. Nguyen Thien Thanh |
| | Dr. Hoang Minh Tri |
| University: | Ho Chi Minh City University of Technology, Vietnam National University-Ho Chi Minh City |


**Abstract**

The optimal control problem for nonlinear systems is constrained directly by the solution of Hamilton-Jacobi-Bellman (HJB) equation and the robust optimal control problem is constrained directly by the solution of Hamilton-Jacobi-Isaacs (HJI) equation. These are nonlinear partial differential equations that have been proven to be impossible to solve analytically. Since then, the problems for approximating off-line or online HJB and HJI solutions are devoted. The reinforcement learning (RL) method, at first, derived from the dynamic programming (DP) theory, and then, developed into adaptive dynamic programming (ADP) method, becomes one of the most effective online methods to approximate HJB and HJI solutions. This thesis propose reinforcement learning-based robust adaptive control algorithms for nonlinear systems, in which Online Adaptive Dynamic Programming (OADP) and Online Robust Adaptive Dynamic Programming (ORADP) are two main analyzed and designed algorithms. OADP algorithm is used to approximate a HJB solution for the nonlinear system with known dynamics, and then extended to ORADP algorithm to approximate HJI solution for the nonlinear system without absolutely knowing knowledge of internal dynamics.


**Novel contributions**

**A. Theory:**

1. The standard ADP structures with two approximators are used to transform into nonlinear control structures with only single approximator for OADP to avoid the complex computation and waste of resources in order to accelerate the speed of update processes.

2. Novel update laws for the approximator's parameters and the novel algorithms are designed. In the algorithm OADP, parameter update laws are synchronized in one iterative step to increase the speed of convergence.

3. Any stability control law to initialize algorithm OADP is not needed; Therefore, design procedures become more flexible.

4. The algorithm OADP guarantee that cost functions are minimized, parameters of approximators and control laws converge to suboptimal values while all closed-system states and the approximate errors are bounded by UUB (Uniform Ultimate Bounded) standard.

5. The standard ADP structures with two approximators are used to transform into nonlinear control structures with only single approximator for ORADP to avoid the complex computation and waste of resources in order to accelerate the speed of update processes. The algorithm is applied for nonlinear system with unknown knowledge of system internal dynamics.

6. Novel update laws for the approximator's parameters and the novel algorithms are designed. In the algorithm ORADP, parameter update laws are synchronized in one iterative step to increase the speed of convergence.

7. The algorithm ORADP does not use knowledge of system internal dynamics, so procedures of system identification are avoided.

8. Any stability control law to initialize algorithm ORADP is not needed; Therefore, design procedures become more flexible.

9. The algorithm guarantee that cost functions is minimized, parameters of approximators and control laws converge to suboptimal values while all closed-system states and the approximate errors are bounded by UUB (Uniform Ultimate Bounded)

standard without using knowledge of system internal dynamics.

**B. Practice:**

To verify the application ability of ORADP algorithm, simulation and experiment for WMR (Wheeled Mobile Robot) are conducted. It is shown that when ORADP algorithm is applied to control WMR, some novel advantages compared with other adaptive control algorithms have been gained.

1. The separation of kinematic and dynamic controllers that commonly used in adaptive control for WMR is unnecessary. By doing that, we can avoid depending on the designer's experience in choosing the parameters for the kinematic controller.

2. Identifying directly or indirectly uncertainty, unstructured and unmodeled dynamics in the robot models is not required.

3. Using ORADP algorithm, the performance index function related to both kinematic, dynamic tracking errors and control energy is minimized.

The ORADP algorithm is continuously designed extendedly for the cooperative control problem of multiple MIMO nonlinear systems without using the knowledge of system internal dynamics. Initially, graph theory is used to establish distributed communication configures for multiple cooperative nonlinear systems. Then, ORADP algorithm is expanded to become the robust adaptive cooperative control algorithm. Simulation results of synchronous control for the swarm robot system show the effectiveness of the extended ORADP algorithm.

**Development trend**

1. Extending ORADP to design algorithms without using input matrices.

2. Extending ORADP to control nonlinear systems more general, such as the nonaffine nonlinear systems with unknown structure.

3. Extending ORADP to develop in nonlinear control theory with output feedback.

4. Extending ORADP in hierarchical reinforcement learning to speed up convergence.

5. Applying ORADP to do experiment on swarm robots or other biological systems.

6. Integrating biological technology into ORADP to enhance practical abilities.

| **Advisors** | | **Ph.D student** |
|---|---|---|
| **Dr. Nguyen Thien Thanh** | **Dr. Hoang Minh Tri** | **Nguyen Tan Luy** |