An Integrated Design Method using Reinforcement Learning Model

Yingjun Hao[1,a],Shaosong Wan[1,a], Jian Cao[2,b] ,Cong Yan[2,c]

[1]XiJing University, Xi'an, Shaanxi Province 710123, P.R. China

[2]Air Force Engineering University, Xi'an, Shaanxi Province 710051, P.R. China

[a]12156555@qq.com, [b]cao_jian1972@163.com, [c]yancong@qq.com

Keywords: reinforcement learning (RL), design, method, application

Abstract. This is followed by a summary of future research directions, and possibilities for technology transition that are currently underway. RL can autonomously get optional policy with the knowledge obtained by trial-and-error and continuously interacting with dynamic environment. Firstly, the model and theory of reinforcement learning is given. Then, a description of the controller architecture and associated stability analysis is given, followed by a more in-depth look at its application to a tiltrotor aircraft.

Introduction

Fig.1 Screenshot of integrated design of tiltrotor configurations and applications using RL

Reinforcement learning (RL) technology develops from some subjects such as statistics, control theory and psychology and so on, and has a very long history, but it is not until the late 80s and early 90s that reinforcement learning technology obtains the wide research and application in some fields such as artificial intelligence, machine learning, automatic control and so on [1] . Reinforcement learning is an important machine learning method [2], its learning technology is divided into three types: non-supervised learning, supervised learning and reinforcement learning. Reinforcement learning is an online learning technology which is different from supervised learning and non-supervised learning. The reinforcement signals provided by the environment in reinforcement Learning is to make a kind of appraisal to the action quality of intelligent Agent, but not tell intelligent Agent how to generate the correct action. Because the external environment provides a little information, intelligent Agent must depend on its own experience to learn, by which intelligent Agent obtains the appropriate appraisal value of the environment state and revises own action strategy to adapt to the environment. Above is followed as Fig.1. In this paper, we first survey the model and theory of reinforcement learning. Then, we roundly present the main reinforcement learning algorithms, including Sarsa, temporal difference, Q-learning and function approximation. Finally, we briefly introduce some applications of reinforcement learning and point out some future research directions of reinforcement learning.

© 2015. The authors - Published by Atlantis Press

RL Model

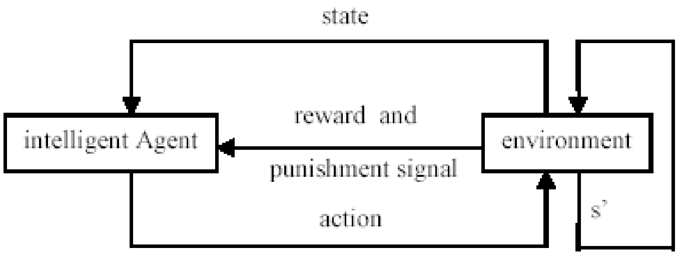

Fig.2 The basic model of RL

The basic model of reinforcement learning is shown in Fig.2. Intelligent Agent can perceive the environment and choose an action to obtain the biggest reward value by continuously interacting with the environment. The interactive interface of intelligent Agent and environment includes action, reward and state.

When each time reinforcement learning system interacts with the environment, the system first accepts the input of environment states, and then the output of action acts the environment according to the internal inference mechanism. Finally, the environment changes to new states' after accepting the action. The system accepts the input of the new state s' and obtains the rewards and punishment signal r of environment for the system. Reinforcement learning system's goal is to learn an action strategy Π: $S \rightarrow A$, the strategy enables the action of the system choice to obtain the largest cumulative reward value of environment [3]. The basic theory of reinforcement learning technology is: If a certain system's action causes the positive reward of the environment, the system generating this action lately will strengthen the trend, this is a positive feedback process; otherwise, the system generating this action will diminish this trend. It is followed as Fig.3.

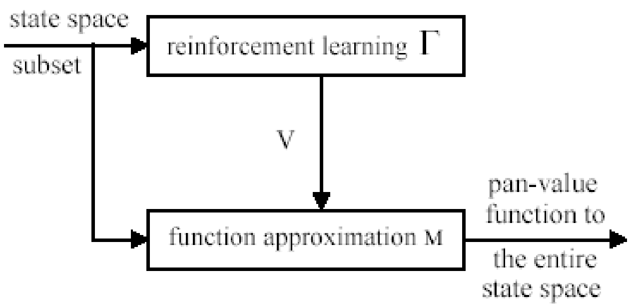

Fig.3 Function approximation RL structure

At present function approximation reinforcement learning method usually uses the supervised learning method, such as state cluster [4], decision tree [5], function interpolation and artificial neural networks and so on; the artificial neural network is the hot spot.

RL Model Application

Reinforcement learning is mainly used in process control, dispatch management, robot control, game competition and information retrieval etc. The widest application is the field of intelligent control and intelligent robot. In the process control field, the typical application example is the inverted pendulum control system and ARCHON system; In the dispatch management, the most successful application is Crites and Barton's elevator scheduling problem, they apply a step reinforcement learning algorithm to the operation scheduling including 4 lifts and 10 floors; In the game competition, the most successful application is Backgammon's chess game invented by Gerry through using instantaneous difference algorithm, it uses the TD error to train three BP neural network and achieves a very high level; In the robot field, HeePakBeen uses the fuzzy logic and reinforcement learning to achieve landbased mobile robot navigation system, Wilfriedllg uses reinforcement learning to make hexapod insect robot coordinate their six-legged actions, Christopher uses reinforcement learning to control the robot arm action, Mataria uses an improved Q-learning algorithm to make 4 robots perform the foraging task; The information retrieval is mainly used in the Internet-related information collection and information filtering, namely, the user can inquire most cared information from the mass of Web information.

Integrated Design using RL

The desire for enhanced agility and functionality demands that the aircraft perform over an increased range of operating conditions characterized by dramatic variations in dynamic pressure and nonlinear aerodynamic phenomena. Furthermore, the use of nonlinear actuation systems increases the complexity of the control design. Therefore, there is presently a strong interest in the development of real-time adaptive control methods that are applicable to flight control problems where the aircraft characteristics are poorly understood or are rapidly changing. In high angle-of-attack (AoA) flight, the aerodynamics are poorly understood and expensive to model. Alternately, variation in dynamic response may occur due to battle damage or component failure, requiring rapid on-line reconfiguration of the control system to maintain stable flight and reasonable levels of handling qualities.

Traditional flight control designs involve linearizing the vehicle dynamics about several operating conditions throughout the flight envelope, designing linear controllers for each condition, and blending these point designs with an interpolation scheme. This 'gain-scheduling' approach, which tends to be rather tedious, may produce a control law that does not globally possess the desirable properties exhibited locally by its constituent point designs. Although gainscheduling has historically proven successful in a variety of applications, future designs will benefit from more advanced methods which explicitly account for the intrinsic nonlinearities of the system.

Nonlinear control techniques, such as feedback linearization, rely heavily upon accurate knowledge of the plant dynamics. However, some aerodynamic effects are very difficult to model. An example is the asymmetric nose vortex shedding that causes 'phantom yaw' in high-AoA flight. The direction in which the vortex is shed is unpredictable, as it depends greatly on the imperfections in the surface of the vehicle. Accounting for such uncertain effects using robust control is, in general, a conservative approach and may sacrifice achievable performance. In contrast, a control system that adapts to the nonlinear dynamics of various flight regimes as they occur has the potential to achieve superior performance throughout the full envelope. To date, most adaptive flight control designs have addressed the issue of the uncertain aerodynamic effects within the context of linear control. Aircraft of the future will benefit from an adaptive control system based on the full nonlinear dynamics of the vehicle while avoiding prohibitively complex gain scheduling.

RL Augmented Model Inversion Architecture

This section details the architecture of the RL augmented model inversion as applied to the tiltrotor aircraft. Fig.4 contains a diagram of the architecture used for implementation of ACAH control in the pitch channel. The command filter serves both to limit the input rate, and as a model for desired response. This allows for straightforward implementation of ADS-33 handling qualities specifications. The presented architecture provides for excellent results in the longitudinal application, as is shown in following sections. Identical construction applies to the lateral channels. Results in the lateral channels show similar performance.

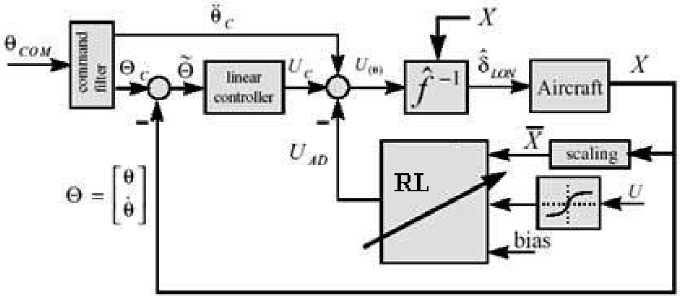

Fig.4 RL augmented model inversion architecture in the longitudinal channel configured for ACAH

Summary

The effectiveness of a controller architecture, which combines adaptive feedforward RL with feedback linearization, has been demonstrated on a variety of flight vehicles. The boundedness of tracking error and control signals is guaranteed. The architecture can accommodate both linear-in-the-parameters networks, as well as single-hidden-layer perceptron RL model. Both theoretical and experimental research is planned to expand and improve the applicability of the approach, and to demonstrate practical utility in the areas of cost reduction and improved flight safety.

References

[1] P. K. Menon, E. J. Ohlmeyer. Integrated Design of Agile Missile Guidance and Autopilot Systems. Control Engineering Practice, vol. 9, pp. 1095-1106. (2012)

[2] N. F. Palumbo, B. E. Reardon, R. A. Blauwkamp. In-tegrated Guidance and Control for Homing Missiles. Johns Hopkins Application Technical Design, vol. 25, no. 2, pp. 121-139. (2011)

[3] M. Xin, S. N. Balakrishnan, E. J. Ohlmeyer. Integrated Guidance and Control of Missiles with θ-D Method. IEEE Transactions on Control Systems Technology, vol. 14, pp. 6, pp. 981-992. (2010)

[4] Singh S Agents and reinforcement learning [M] San Matco CA USA Miller freeman publish Inc 2012.

[5] Bush R R & Mosteller F. Stochastic Models for Learning [M] . New York :Wiley ,2005.