早稲田大学大学院情報生産システム研究科

博士論文概要

論文題目

Study on Robustness and Adaptability of Genetic Network Programming with Reinforcement Learning for Mobile Robot

申請者

Siti SENDARI

情報生産システム工学専攻
ニューロコンピューティング研究

2012 年 12 月

In the real implementation of autonomous mobile robots, the environments change dynamically with unknown environments and inexperienced situations, which make agents do tasks inappropriately. In order to behave appropriately, the agents should be robust to the unknown environments and also have adaptability mechanisms to change the structures/parameters of the controllers adaptively when inexperienced situations occur.

In classical methods, designing the controllers of the agents using differential equations or conventional network structures have problems, where those methods are difficult to represent all the actual agent behaviors in the dynamic environments. Generally speaking, although the experiences of an expert are needed to design the controllers, the evolutionary algorithms, such as Genetic Algorithm, Genetic Programming and Genetic Network Programming (GNP) could generate the optimized programs for robots.

GNP is one of evolutionary algorithms to automatic program generation, which generates intelligent rules for agent behaviors. The structures of GNP are constructed and optimized in evolution, where the experiences of an expert are not needed. GNP has some advantages which make GNP suitable for implementing robots, that is, (1) re-usability of the nodes which make the structures more compact and use small memory; and (2) applicability to Partially Observable Markov Decision Process (POMDP). Furthermore, GNP was enhanced to GNP with Reinforcement Learning (GNP-RL) by integrating Reinforcement Learning, which has the ability to learn the node transitions and change the functions adaptively by selecting the appropriate sub nodes when troubles occur.

GNP-RL was successfully implemented to navigate the mobile robots, where GNP-RL can determine the alternative function adaptively using the learning algorithm. The adaptability of GNP-RL using Sub node Selection (SS method) has been analyzed; however, the adaptability mechanism of GNP-RL is not enough when severe troubles occur. In addition, the robustness of GNP-RL in the noisy environments has not been analyzed yet. Thus, the robustness and adaptability of GNP-RL to the changing environments are studied in this research.

In Chapter 1, the research backgrounds, motivation, objectives and outline of the thesis are described. The objectives of this research are to improve the robustness of Genetic Network Programming with Reinforcement Learning (GNP-RL) in uncertain environments and adaptability when inexperienced changes of the environments occur. The improvement of the robustness is studied in chapter 2, while the improvement of the adaptability is studied in Chapter 3 to 5.

In Chapter 2, Fuzzy Genetic Network Programming with Reinforcement Learning (Fuzzy GNP-RL) has been proposed. The structures of Fuzzy GNP-RL are based on GNP-RL, where Fuzzy logic is integrated to the judgment nodes, then the node transitions of Fuzzy GNP-RL can be determined probabilistically. The proposed method is simulated using Webots simulator and evaluated for the wall following behaviors of a Khepera robot. The robustness of the proposed method is studied by using different training and implementation environments. As a result, it is clarified that Fuzzy GNP-RL improves the robustness in unknown environment. Furthermore, the robustness of Fuzzy GNP-RL in the dynamic environments is also studied by introducing Gaussian noises during the training phase. The results show that the robustness of Fuzzy GNP-RL is improved more. Based on the results, Fuzzy judgment nodes are used in Chapter 3, 4 and 5.

In Chapter 3, Fuzzy Genetic Network Programming with Two-Stage Reinforcement Learning using Branch connection Selection method, i.e., Fuzzy GNP-TSRL (BS) has been proposed. Fuzzy GNP-TSRL (BS) combines the Sub node Selection method (SS method) like Fuzzy GNP-RL and Branch connection Selection method (BS method). The SS method and BS method are learned in the first stage and in the second stage of RL, respectively. In Fuzzy GNP-TSRL (BS), the actions are divided into two groups using two kinds of Q tables, that is, Q_{SS} table and Q_{BS} table. Fuzzy GNP-TSRL (BS) enhances the adaptability mechanism of Fuzzy GNP-RL, which can adaptively change the programs by selecting not only the appropriate functions but also the appropriate next node connections. The adaptability of Fuzzy GNP-TSRL (BS) is evaluated when inexperienced troubles occur due to some sensors' break in the implementation. The adaptability of Fuzzy GNP-TSRL (BS) which uses Two-Stage Reinforcement Learning is compared with that of Fuzzy GNP-RL which uses One-Stage Reinforcement Learning. The results show that the adaptability mechanism of Fuzzy GNP-TSRL (BS) works effectively and efficiently, therefore the performance of Fuzzy GNP-TSRL (BS) becomes better than Fuzzy GNP-RL.

In Chapter 4, Changing ε-greedy policy and learning rate α are studied for improving the performance of Fuzzy GNP-TSRL (BS). ε-greedy and learning rate α are set at larger values in early generations and decreased gradually in latter generations. Thus, larger exploration is carried out in early generations, however, larger exploitation is carried out in later generations. The changing ε-greedy and learning rate α are also studied in the implementation phase, that is, when sudden changes occur, the exploration is increased and gradually decreased until the end of life time. The performance of Fuzzy GNP-TSRL (BS) with changing parameters of ε-greedy policy and learning rate α is studied by comparing with that of fixed parameters. As a result, the changing ε-greedy and learning rate α can improve the performance of Fuzzy GNP-TSRL (BS), which means that the balance of the exploration and exploitation can be controlled efficiently and effectively by the proposed method.

In Chapter 5, another method of Two-Stage Reinforcement Learning based on GNP is studied, that is, Fuzzy Genetic Network Programming with Two-Stage using Credit branch Selection method (Fuzzy GNP-TSRL (CS)). The proposed method combines the Sub node Selection method (SS method) of Fuzzy GNP-RL and Credit branch Selection method (CS method). While Fuzzy GNP-TSRL (BS) provides selecting alternative functions and alternative connections, Fuzzy GNP-TSRL (CS) provides selecting alternative functions like Fuzzy GNP-RL, but this method can skip the nodes when they are considered as harmful nodes. The results show that Fuzzy GNP-TSRL (CS) is superior than Fuzzy GNP-TSRL (BS) and Fuzzy GNP-RL, because skipping harmful nodes may recover the troubles more quickly which cannot be done by Fuzzy GNP-TSRL (BS) and Fuzzy GNP-RL. Thus, Fuzzy GNP-TSRL (CS) determines the node transitions more efficiently and effectively, as a result, the adaptability can be improved.

Chapter 6 concludes the thesis by describing the achievements of the proposed methods.