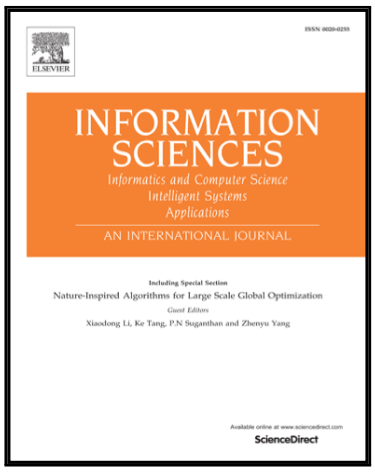

An adaptive Decision-making Method with Fuzzy Bayesian Reinforcement Learning for Robot Soccer

Haobin Shi , Zhiqiang Lin , Shuge Zhang , Xuesi Li , Kao-Shing Hwang

PII: S0020-0255(18)30043-4
DOI: 10.1016/j.ins.2018.01.032
Reference: INS 13386

To appear in: *Information Sciences*

Received date: 24 June 2016
Revised date: 10 January 2018
Accepted date: 14 January 2018

Please cite this article as: Haobin Shi , Zhiqiang Lin , Shuge Zhang , Xuesi Li , Kao-Shing Hwang , An adaptive Decision-making Method with Fuzzy Bayesian Reinforcement Learning for Robot Soccer, *Information Sciences* (2018), doi: 10.1016/j.ins.2018.01.032

An adaptive Decision-making Method with Fuzzy Bayesian Reinforcement Learning for Robot Soccer

Haobin Shi[a*], Zhiqiang Lin[a], Shuge Zhang[a], Xuesi Li[a], Kao-Shing Hwang[b]

[a] *School of Computer Science, Northwestern Polytechnical University, Xi'an, Shaanxi Province 710072, China*

[b] *Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan*

* *Corresponding author at: School of Computer Science, Northwestern Polytechnical University, Xi'an, Shaanxi Province 710072, China*

E-mail address: shihaobin@nwpu.edu.cn

Abstract

A robot soccer system is a typical complex time-sequence decision-making system. Problems of uncertain knowledge representation and complex models always exist in robot soccer games. To achieve an adaptive decision-making mechanism, a method with fuzzy Bayesian reinforcement learning (RL) is proposed in this paper. To extract the features utilized in the proposed learning method, a fuzzy comprehensive evaluation method (FCEM) is developed. This method classifies the situations in robot soccer games into a set of features. With the fuzzy analytical hierarchy process (FAHP), the FCEM can calculate the weights according to defined factors for these features, which comprise the dimensionality of the state space. The weight imposed on each feature determines the range of each dimension. Through a Bayesian network, the comprehensively evaluated features are transformed into decision bases. An RL method for strategy selection over time is implemented. The fuzzy mechanism can skillfully adapt experiences to the learning system and provide flexibility in state aggregation, thus improving learning efficiency. The experimental results demonstrate that the proposed method has better knowledge representation and strategy selection than other competing methods.

Keywords: Situation Evaluation; Robot Soccer; Fuzzy Bayesian; Reinforcement learning

1. Introduction

Robot soccer game is an example of a multi-agent cooperative confrontation platform where multiple agents need to complete complex tasks within a dynamic and uncertain environment [11][2]. A complete decision-making system for robot soccer games comprises two parts: situation evaluation and decision-making [37]. Situation evaluation makes assessment according to the uncertain environmental information and multi-robot situations. Decision-making involves selecting the most appropriate strategy on the basis of the assessment made of the situation in order to achieve a favorable outcome. These two parts are the key technologies of a robot soccer system and have attracted much research interest.

Multi-attribute expert decision-making methods have been frequently used for decision-making [37][32]. The experience knowledge can be effectively expressed by a multi-attribute expert decision-maker. However, the decision-maker is susceptible to a lack of experience and may break down when the available experience does not cover well the experience domain. Hence, this method may perform poorly in dynamic scene evaluation such as robot soccer games [15]. To resolve this problem, another method called the knowledge representation and rule-based reasoning method with fuzzy logic is applied to dynamic decision-making [33][25][39]. This approach can effectively quantify the expert's experience and make the agent's decision-making process more in line with human thinking. However, many effective training samples are necessary and the results may be difficult to quantify for decision inference. The use of only a few samples may lead to a slightly lower efficiency, hence, this method may be not appropriate for evaluating a dynamic scene. Knowledge inference methods, such as the blackboard model, logic template matching and Bayesian network inference technology, are employed for decision-making in some studies [19][28]. Knowledge inferred using a Bayesian network has the ability to deal objectively with the uncertainty problem [19][3][9]. It can accurately explain the causal relationship and degree of correlation between the variables, contributing to predict the occurrence of events through directed edges and conditional probability distribution. The information of each node for the Bayesian network has a significant impact on the inference results. Since its representation ability for the input data is weak, node information has a strong influence on the results of inference. Thus when the input information is insufficient for quantitative analysis, this method may be less appropriate and it is limited in its application to the evaluation of robot soccer under such a complex dynamic environment.

For strategy selection, a Bayesian network is usually applied [19][3]. It is sensitive to *a priori* knowledge, which is normally difficult to be acquired, and hence cannot work well in a dynamic environment. Therefore, the single Bayesian network-based strategy selection method is somewhat limited in its applications. A neural network is an alternative for conducting strategy selection [12]. Neural networks have good abilities of linear and nonlinear mappings and self-adaptation, so that they can be applied to strategy selection. However, the method of neural network-based strategy selection needs a large amount of offline training and has high complexity, making it a shortcoming for dynamic strategy selection. Reinforcement learning (RL) systems have been applied to selecting the most appropriate decision scheme in