Learning Automata: A Comparative Analysis of Estimator Algorithms

Seyyed Amir Hadi Minoofam
PhD Candidate
Faculty of Computer and Information Technology Engineering
Qazvin Branch, Islamic Azad University
Qazvin, Iran, 34199-5195
Email: minoofam@qiau.ac.ir

Azam Bastanfard
Assistant Professor
Faculty of Media Engineering
Islamic Republic of Iran Broadcast University
Tehran, Iran, 19395-1746
Email: bastanfard@iribu.ac.ir

Abstract— Learning automata (LAs) play a crucial role as a reinforcement scheme to solve engineering problems even in non-stationary environments. However, their low rate of convergence has considered as the main drawback in the works of literature. Estimator algorithms have suggested as a successful attempt to alleviate the inconvenience. The aim of this paper is to discuss various types of estimator algorithms and categorize them based on their intrinsic properties. Also, an overall comparison between existing algorithms based on the conventional measures is presented. The associated analysis provides a foundation for identifying strengths and weaknesses of estimator algorithms in the field of LAs, as well as general guidelines for future improvements and innovations.

Keywords—Bayesian estimator; estimator algorithms; learning automata; maximum likelihood estimator.

I. INTRODUCTION

Learning automata is a reinforcement learning approach which is an important area of research in artificial intelligence. LAs were introduced as a tool for adaptive learning [1]. A stochastic automaton embarks its operation without having any knowledge about the random environment. Then, by choosing actions based on a probability vector interacts with the environment. After that, the automaton observes the environment’s response as a reward or penalty and updates its action probabilities to learn the optimal action gradually and finally converges to it.

The low rate of convergence was considered as a main drawback of LAs in the literature [2]. Discretizing the probability space was suggested by Thathachar and Oommen in 1979 to improve the rate of convergence [3]. In a discretized automaton, the probabilities are restricted as a finite set of values in the interval $[0, 1]$. Another improvement was introduced by Thathachar and Sastry in 1985 in the form of proposing estimator algorithms [4]. In this novel approach, the history of LAs is used to maintain an estimate of probabilities as an estimator vector. Therefore, updating the probability vector is took place not only based on the probability vector but also based on the estimator vector. Discretized version of estimator algorithms was introduced as well to enhance both speed and accuracy [5].

In systems with incomplete knowledge about the environment, LAs have found numerous applications in a variety of fields. They have been used in game playing [6-12], pattern recognition [13], parameter optimization [14-16], object partitioning [17, 18], channel selection in cognitive networks [19], multi-objective analysis [20], assigning capabilities in prioritized networks [21], telephony routing [22, 23], solving knapsack problems [24], priority assignment in a queuing system [25], optimally allocation of limited resources [26, 27], statistical decision making [28], service selection [29], distribution approximation [30], numerical optimization [31], natural language processing [32], web crawling [33], string taxonomy [34], microassembly path planning [35], graph partitioning [36], multiagent learning [37], distributed scheduling [38], batch sequencing in manufacturing systems [39], network conflict avoidance [40], chain inventory control [41], photonic LANs [42], dynamic channel allocation [43], map learning [44], adaptive signal processing [45], star networks [46], PID controllers [47], digital filter design [48], vehicle path control [49], vehicle suspension systems [50], broadcast communication [51], controlling client-server systems [52], and control of power systems [53] to mention a few.

This paper aims to introduce an analytical comparison between estimator algorithms while considering state of the art. Also, a categorization based on the internal techniques in algorithms is presented.

The structure of the following sections is as follows. An overview of learning automata and their properties is given in Section II. Section III consists of considering various kinds of estimator algorithms as well as an overall categorization. The analytical comparison of the estimators and some discussions are provided in Section IV. Finally, the paper is concluded in Section V with some significant remarks about future works.

II. LEARNING AUTOMATA

Learning automaton is a finite state machine that attempts to learn the optimal action by interacting with a random environment. The formal definition of LAs and environment are described in Definition 1 and 2 respectively [7].
Definition 1. A stochastic automaton is defined as a quintuple \( \langle \phi, \alpha, \beta, F, G \rangle \) where:

\( \phi = \{ \phi_1, \phi_2, \ldots, \phi_s \} \) is the internal state set (\( 2 \leq s \leq \infty \)),
\( \alpha = \{ \alpha_1, \alpha_2, \ldots, \alpha_r \} \) is the actions of the environment,
\( \beta \) is the input set,
\( F: \phi \times \beta \rightarrow \phi \) is the state transition mapping,
\( G: \phi \rightarrow \alpha \) is the output mapping.

Definition 2. An environment is a triple \( \{ \alpha, \beta, c \} \) where:

\( \alpha = \{ \alpha_1, \alpha_2, \ldots, \alpha_r \} \) is the input set,
\( \beta \in \{0, 1\} \) is the output or response, in which 0 and 1 correspond to reward and penalty, respectively
\( c = \{ c_1, c_2, \ldots, c_r \} \) is a vector of penalty probabilities.

The automaton selects an action according to a probability vector. The chosen action triggers the environment to respond according to the penalty probabilities of an action. This response can be a reward or a penalty. The automaton modifies its state according to the response, selects another action, and the mentioned procedure is continued until the convergence of the automaton. It means that the automaton ultimately decides the optimal action more frequently than the others. Figure 2 shows the schematic of the interaction between automata and environment.

III. ESTIMATOR ALGORITHMS

This section provides an overview of estimator algorithms from beginning to state of the art. To the best of our knowledge, nearly all estimator algorithms could be categorized as two main classes which are as follows.

A. Maximum likelihood estimators

These traditional approaches estimate the action probability vector is utilizing famous maximum likelihood estimator (MLE). Different kinds of this class were introduced in the literature which is described as below.

- Pursuit Algorithm (PA)

Thathachar and Sastry provided pursuit algorithm in 1986 [54]. The pursuit algorithms are characterized by the fact that they pursue an action that is currently reckoned to be the optimal action. The algorithm involves three steps. The first step is to pick an action based on the probability distribution. In the second step, the probability vector is updated based on the response from the environment. Finally, the third phase is to update the running estimates of the action that is rewarded.

The second step organizes as two different ways. In the reward-penalty (RP) scheme, if the reckoned optimal action is rewarded, its probability is increased while the other’s decrease, and if it gets a penalty, the procedure is operated in a reverse manner. However, reward-inaction (RI) scheme, functions as the same in the reward case, but takes no probability updating in the penalty case.

- Discretized Pursuit Algorithm (DPA)

Discretized version of pursuit algorithms was introduced by Oommen and his colleagues in 1992 [55]. In this class, the probability vector is allowed to take finite discrete values in the interval \([0, 1]\). The algorithm substitutes the multiplication operations in the PAs by addition or subtraction to keep the
sum of probabilities to be unity. Oommen et al. demonstrated that DPA is about 60% faster than PA [55].

- **Generalized Pursuit Algorithm (GPA)**

  A generic version of the pursuit algorithm introduced by Agache and Oommen in 2002 [56, 57]. In GPA, the probability of pursuing a wrong action is minimized by pursuing all the actions with higher reward estimates than the selected action.

- **Discretized Generalized Pursuit Algorithm (DGPA)**

  Agache and Oommen introduced DGPA in 2002 [57]. The DGPA pursues all the actions that have higher estimates than the currently selected action in different discrete steps. Simulation results provide superiority of the DGPA over other estimators.

- **Tathachar Sastry Estimator Algorithm (TSE)**

  Thathachar and Sastry have introduced a more sophisticated algorithm in 1985 [58]. The probability updating is done based on both the action probability and the reward estimate vectors. The algorithm increases the probabilities of the action that have higher estimated than currently optimal action decrease the probabilities of others.

- **Discrete TSE Algorithm (DTSE)**

  Discretized Lanctôt and Oommen suggested TSE in 1991 [55]. This method is a sophisticated algorithm that has two necessary and sufficient properties of being \( \varepsilon \)-optimal which are moderation and monotonicity properties [55].

- **Generalized TSE Algorithm (GTSE)**

  Various generalization of TSE algorithm could be introduced, but Agache and Oommen were suggested a generalization in 2004 [59]. The algorithm provides a different weight matrix which also keeps the probability vector as valid probabilities in the interval \([0, 1]\).

**B. Bayesian estimators**

These novel approaches utilize Bayesian theorem in a way that its parameters are computed by Thompson sampling [60] using \( \beta \) distribution. Thus the computation intractability of Bayesian approach is relaxed. Varieties types of Bayesian estimator algorithms are briefly described below.

- **Bayesian Learning Automaton (BLA)**

  Granmo has suggested a Bayesian learning automaton in 2008 [61, 60]. In this algorithm, the estimation procedure is provided by Bayesian theorem. It has an infinite discrete state with the absence of probability vector.

- **Bayesian Pursuit Algorithm (BPA)**

  Zhang and her co-workers presented BPA with the use of Bayesian estimator in 2011 [62, 63]. This algorithm is based on utilizing multiple values of the posterior and can update probabilities in one of the RI or RP schemes.

- **Discretized Bayesian Pursuit Algorithm (DBPA)**

  Zhang and her colleagues introduced DBPA in 2012 [63, 64] with the utilization of both discretization and Bayesian techniques. The simulations demonstrated that DBPA is about 20% faster than BPA in convergence speed.

- **Generalized Bayesian Stochastic Estimator (GBSE)**

  GBSE was introduced by Jiang et al. in 2016 [65]. This algorithm is based on a stochastic estimator and not only improve the initialization phase of the algorithm but also increase the probabilities with higher stochastic estimates.

From the above description, the estimator algorithms can be categorized as Figure 2. This categorization demonstrates that any estimator algorithm can be used MLE or Bayesian estimators whereas each continuous algorithm could have a discrete version by using discretization approach.

![Figure 2. Categorization of estimator algorithms](image-url)

The comparison of the algorithms and their rational discussion is considered in the next section.

**IV. COMPARISON AND DISCUSSION**

This section provides a consideration of characteristics of estimator LAs. Although ML-based estimators have been used as a standard approach for estimation, they have some disadvantages [65]. Thus Bayesian estimators have been proposed to overcome these obstacles. Simulation results show the superiority of Bayesian Estimators over ML-based estimators [63, 65].

Table 1 demonstrates the main characteristics of ML-based estimators. As it is evident from Table 1, continuous algorithms like PA, TSE, GPA and GTSE have only one parameter for learning to be tuned. But discretized ones have
more parameters which make them more involved. Among the discretized algorithms in Table 1, DTSE is the most complex one because it needs three parameters to be tuned. All the mentioned algorithms could be used in stationary environments.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA</td>
<td>Parameters: $0&lt;\alpha&lt;1$</td>
</tr>
<tr>
<td>TSE</td>
<td>Parameters: $0&lt;\alpha&lt;1$</td>
</tr>
<tr>
<td>GPA</td>
<td>Parameters: $0&lt;\beta&lt;1$</td>
</tr>
<tr>
<td>GTSE</td>
<td>Parameters: $0&lt;\lambda&lt;1$</td>
</tr>
<tr>
<td>DPA</td>
<td>Parameters: $n&gt;0, \Delta=1/(m)$</td>
</tr>
<tr>
<td>DTSE</td>
<td>Parameters: $n&gt;0, \beta&gt;0, \Delta=1/(m0)$</td>
</tr>
<tr>
<td>DGPA</td>
<td>Parameters: $n&gt;0, \Delta=1/(m)$</td>
</tr>
</tbody>
</table>

Table 2 demonstrates the comparison between Bayesian estimators. It can be seen that BLA has the least parameters among them and therefore it can be tuned easier than the others. Also, it could be utilized in both stationary and non-stationary environments.

As it can be extracted from Table 1 and Table 2, the more parameters of Bayesian estimators in comparison to ML-based estimators is an expense paid by the designers to develop more accurate algorithms with an increase in convergence speed. However, in this case, the additional complexity of the environment leads to the reduction in accuracy. The main reason lies in the intricate design of Bayesian algorithms which requires a precise tradeoff between tunable values of parameters. Hence, the investigation of new estimators to overcome this drawback is an open problem for more investigation.

V. CONCLUSION AND FUTURE WORK

The low performance of classic learning automata makes them unacceptable when operating in environments providing a high penalty. Some modifications in LAs are taken to alleviate the low-performance problem, e.g. discretization and estimation.

In this paper estimator LAs are analyzed which are the fastest family among various types of LAs approaches [57, 63, 65]. Estimator LAs are divided according to their estimation strategies which are mainly MLE or Bayesian. All of them could also be presented in the continuous or discretized form. The overall categorization with an analytical comparison of their characteristics is also considered in this paper.

Developing estimator algorithms in modified versions could be taken into account as a future work. Moreover, the application of the state of the art estimator algorithms is remaining as an open area of research.

REFERENCES

25th Iranian Conference on Electrical Engineering (ICEE2017)

[18] Oommen BJ, Ma DC. Stochastic automata solutions to the object partitioning problem. Carleton University, School of Computer Science; 1986.


