A learning automaton is one type of machine learning algorithm studied since 1970s. Learning automata select their current action based on past experiences from the environment. It will fall into the range of reinforcement learning if the environment is stochastic and a Markov decision process (MDP) is used.

History

Research in learning automata can be traced back to the work of Michael Lvovitch Tsetlin in the early 1960s in the Soviet Union. Together with some colleagues, he published a collection of papers on how to use matrices to describe automata functions. Additionally, Tsetlin worked on reasonable and collective automata behaviour , and on automata games . Learning automata were also investigated by researches in the United States in the 1960s. However, the term learning automaton was not used until Narendra and Thathachar introduced it in a survey paper in 1974.

Definition

A learning automaton is an adaptive decision-making unit situated in a random environment that learns the optimal action through repeated interactions with its environment. The actions are chosen according to a specific probability distribution which is updated based on the environment response the automaton obtains by performing a particular action.

With respect to the field of reinforcement learning , learning automata are characterized as policy iterators . In contrast to other reinforcement learners, policy iterators directly manipulate the policy π. Another example for policy iterators are evolutionary algorithms .

Formally, Narendra and Thathachar define a stochastic automaton to consist of:

a set X of possible inputs,

a set $\Phi = \{ \Phi_1 , ..., \Phi_s \}$ of possible internal states,

a set $\alpha = \{ \alpha_1 , ..., \alpha_r \}$ of possible outputs, or actions, with $r \leq s$,

an initial state probability vector $p(0) = \blacksquare p_1 (0), ..., p_s (0) \blacksquare$,

a computable function A which after each time step t generates p (t +1) from p (t), the current input, and the current state, and

a function $G : \Phi \rightarrow \alpha$ which generates the output at each time step.

In their paper, they investigate only stochastic automata with r = s and G being bijective , allowing them to confuse actions and states.

The states of such an automaton correspond to the states of a "discrete-state discrete-parameter Markov process ". [1] At each time step t =0,1,2,3,..., the automaton reads an input from its environment, updates p (t) to p (t +1) by A , randomly chooses a successor state according to the probabilities p (t +1) and outputs the corresponding action. The automaton's environment, in turn, reads the action and sends the next input to the automaton.

Frequently, the input set X = { 0,1 } is used, with 0 and 1 corresponding to a nonpenalty and a penalty response of the environment, respectively; in this case, the automaton should learn to minimize the number of penalty responses, and the feedback loop of automaton and environment is called a "P-model". More generally, a "Q-model" allows an arbitrary finite input set X , and an "S-model" uses the interval [0,1] of real numbers as X . [2]

A visualised demo [3] [4] / Art Work of a single Learning Automaton had been developed by μSystems (microSystems) Research Group at Newcastle University.

Finite action-set learning automata

Finite action-set learning automata (FALA) are a class of learning automata for which the number of possible actions is finite or, in more mathematical terms, for which the size of the action-set is finite. [5]

See also

Reinforcement learning

Game theory

Automata theory

Literature

Philip Aranzulla and John Mellor (Home page): Mellor J and Aranzulla P (2000): "Using an S-Model Response Environment with Learnng [sic] Automata Based Routing Schemes for IP Networks ", Proc. Eighth IFIP Workshop on Performance Modelling and Evaluation of ATM and IP Networks, pp 56/1-56/12, Ilkley, UK. Aranzulla P and Mellor J (1997): "Comparing two routing algorithms requiring reduced signalling when applied to ATM networks", Proc. Fourteenth UK Teletraffic Symposium on Performance Engineering in Information Systems, pp 20/1-20/4, UMIST, Manchester, UK.

Mellor J and Aranzulla P (2000): "Using an S-Model Response Environment with Learnng [sic] Automata Based Routing Schemes for IP Networks ", Proc. Eighth IFIP Workshop on Performance Modelling and Evaluation of ATM and IP Networks, pp 56/1-56/12, Ilkley, UK.

Aranzulla P and Mellor J (1997): "Comparing two routing algorithms requiring reduced signalling when applied to ATM networks", Proc. Fourteenth UK Teletraffic Symposium on Performance Engineering in Information Systems, pp 20/1-20/4, UMIST, Manchester, UK.

Narendra K.; Thathachar M.A.L. (July 1974). "Learning automata – a survey" (PDF) . IEEE Transactions on Systems, Man, and Cybernetics . SMC-4 (4): 323– 334. CiteSeerX 10.1.1.295.2280 . doi : 10.1109/tsmc.1974.5408453 .

Tsetlin M.L. Automation theory and modeling of biological systems. Academic Press; 1973. [permanent dead link]

References