Title	確率的学習アルゴリズムを用いた有限状態オートマトンの抽出に関する研究
Author(s)	近藤, 雅之
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/361
Rights	
Description	Supervisor:林 幸雄, 知識科学研究科, 修士

Study on Stochastic Learning of Finite-State Automata by Neural Networks

Masayuki Kondou

School of Knowledge Science,
Japan Advanced Institute of Science and Technology

February 13, 2002

Keywords: learning algorithm, simple recurrent network(SRN), finite-state automata(FSA), regular language.

There have been a lot of researches on symbol processing by neural network models. Among them,there are researches on learning grammar. In such researches a recurrent neural network (RNN),which can treat time series data,have been used to learn grammar.

Schreiber et.al[1988] made a simple recurrent network(SRN) learn sentenses of which were generated by finite-state automata(FSA).SRN is the most fundamental model of RNN.In the research,he claimed that the cluster of the activation pattern in learned context units expresses each state of the FSA.Moreover,Giles et.al[1992] reported that neural networks several times learned the perfect FSA.Futhermore,Omlin has reported that "Of particular concern — and an open issue — is fidelity of the extraction process,i.e. how accurately the extracted knowledge corresponds to the knowledge stored in a neural network".Therefore,in the study of acuisition of grammar,how to extract the FSA from internal states of neurons in neural network models is one of the most important issues.

In converntional researches on the FSA extracion,a sigmoid function is applied as the output function of an element which constitues a network. However a sigmoid function outputs a continuous value.Thus,it is difficult to extract the state transition of the FSA, which network will learn,from activation pattern of hidden units.On the other hand,if the elements a linear-threshold element,the FSA can be extracted easily.Because the linear-threshold element outputs a discrete value.However,the learning algorithms which were developed for multilayer perceptron(MLP) can be applied only for one linear-threshold element.

Resently, Sakurai proposed a stochastic learning algorithm for MLP(S-MLP)[2001]. This algorithm can be applied to the multilayer perceptron which consists of a linear-threshold element. Sakurai claims that the algorithm has a global-convergency and rapidity-conbergency. In this work,we focus on the feature that this algorithm can be applied to a network in which the output of each element is discrete.

The purpose of this work is to apply S-MLP to SRN and to investigate wheter the SRN learn FSA using S-MLP. I try to make SRN learn sentences which an generated by a regular grammar using S-MLP.And I also confirm that the FSA can be extracted from hidden units of a learned network.I also disucussed what kind of FSA has been extracted as a result.

In experiments,we make the network learn Tomita grammar, which is a one of the simplest regular grammar. The Tomita grammar has grammar from the 1st to 7th as fallows.

1 1^*.

2 $(10)^*$.

3 any string without an odd number of consecutive 0's after an odd number of consecutive 1's.

4 any string without more than 2 consecutive 0's.

5 aniy string of even length which,making pairs,an odd nomber of (01) or (10).

6 any string such that the difference between the numbers of 1's and 0's is 3n.

7 $0^*1^*0^*1^*$.

In this work, we used the 3rd,4th,6th,and 7th grammar. As a teaching signal, wheser the current state is accepted or not at that time is given.Thus the learning can considered to be full learning about the grammar (though it is actually a limited learn for a finite set of samples).

Initial values for connecting weights were chosen randomly from integers -2 $\sim$ 2, and the output of context units were to set 1.In each grammar, I tried 100 times and counted the number of successful trials.At this time,the length of string and the number of data being fixed ,the input data were generated at random. As a result,all of the trials were successful. Next,I tested the data which cover all bit patterns. In this test, the length of a string was set up with 3-5. A FSA was tried to be extracted from hidden units in the learned network. Consequently, in the 3rd,4th,and 7th grammar,a minimal FSA was extracted. Moreover,FSA with comparatively few states were able to be extracted in the 6th grammar,too.

From these results, the S-MLP can be applied to a SRN. Minimal FSA was extracted from learned SRN with S-MLP. Therefore, it is shown that the SRN using S-MLP can learn FSA.