

Title	再帰型回路網による文法の獲得
Author(s)	原田, 哲治
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/737
Rights	
Description	Supervisor: 櫻井 彰人, 知識科学研究科, 修士

Learning grammars with recurrent neural networks

Tetsuji HARADA

School of knowledge Science,
Japan Advanced Institute of Science and Technology
March 2001

Keywords: connectionist symbol processing, simple recurrent network (SRN), recursive auto-associative memory (RAAM), context-free grammar (CFG), parsing, grammatical inference, embedded relative clauses , holistic computation

1 Introduction

Most of language acquisition models that have been constructed so far are based on traditional AI approaches. On the other hand, artificial neural networks (ANNs), contrasting with traditional AI approaches, have many great abilities such as ability to learn, generalization capability and robustness. But they are poor in representing compositional structures and manipulating them, and are considered to be unable to perform symbol processing, specifically natural language processing. However, various connectionist models to represent compositional structures were proposed around 1990. Since then, there have been many researches on connectionist symbol processing. The primary purpose of this work is to construct a recurrent neural network (RNN) architecture that learns grammars.

One of primary issues addressed in connectionist symbol processing researches is how to represent and manipulate compositional structure, such as lists and trees, with ANNs. Although these structured data have variable length, ANNs store and process data as fixed-width patterns. The RAAM network proposed by Pollack is a promising technique for structured data processing. In this work, RAAM is used to represent parse trees.

Formal language theories say that some important properties of natural language grammars can not be represented by regular grammar (RG). One example is the self embedding of noun phrases. It is known that context-free grammars (CFGs) can represent them. Although Many researches have clarified that RNN have abilities to learn RG from sample sentences, they also have shown that learning CFG by RNNs is quite difficult. Parsing context-free languages (CFLs) needs stacks; Therefore, RNN is required to learn stack representation and its manipulation. There are some models that learn to have stack capability with limited depth, there is no network that succeed in learning recursive rules.

The networks proposed here learn to parse sentences generated from context-free grammars in a way similar to natural languages. If learned RNN can correctly parse sentences with deeper embedded clauses than learned, the RNN may be said it has learned stack representations and its manipulations.

2 Model and experiments

The model proposed here consists of RAAM network and SRN. In these experiments, grammars that networks tried to learn were English-flavored CFG, Japanese-flavored one, and a simple CFG that generates language $\{a^n b^n\}$. Successfully learned RAAM can construct distributed representations of structured parse trees, called “RAAM representation of parse tree” in this paper, by embedding subtrees recursively in real number vectors formed by fixed number of units. SRN learns to output the RAAM representation of the largest subtree that includes the current terminal symbol (hereafter, called “word”), but not unseen word that are presented in order. RAAM learns first, then SRN are fed with RAAM representations of parses tree as target in training. It is also fed with lookahead words when the network learns natural language-flavored grammars.

Experiments for the model was performed as simulations on digital computers. Some learned networks performed to have generalization capability of parsing $\{a^n b^n\}$ sentences twice longer than used in training. It is not clear whether the network recognizes the end of sentence because the network might output correct parse trees by only counting the number of “b” inputs.

The networks can learn parsing natural language-flavored sentences of up to 10 word long, of up to depth four of embedded clauses. Namely we obtained the networks that output correct parse trees and subtrees for all words of sentences of 10 word long, that were used in training. Therefore we conclude that the proposed model has capability for parsing natural language-flavored sentences. There are a few networks that learned to have generalization capability to parse natural language-flavored sentences longer than used in training. However no SRN succeeded in having generalization capability

to parse natural language-flavored sentences with deeper embedded clauses. A few RAAM networks performed to represent correct parse trees for some sentences of four embedded depth by trained with natural language-flavored sentences of less than three embedded depth.

The novel aspects of this work are as follows. i) The model proposed here is novel in researches of grammatical inference with connectionist model. This work investigates generalization capabilities of the model for longer and deeper sentences in detail. It is one of interest aspects of this work to try to learn a grammar that has right recursive rules, and also to learn another grammar that has left recursive rules with the same network architecture. ii) This work is the first to show that RAAM has the generalization capability for the embedded clauses of language $\{a^n b^n\}$.