コネクショニストモデルによる日本語複文解析モジュールの構築
Constructing a Japanese Complex Sentence Parsing Module with Connectionist Models

Minoru MOTOKI, Katsuhiko TSUCHIMOCHI and Yoshio SHIMAZU

Dept. of Electric Eng., Kyushu Sangyo University
2-3-1 Matsukadai, Higashi-ku, Fukuoka-shi 813-8503, Japan
E-mail:motoki,shimazu@te.kyusan-u.ac.jp

1. Introduction

We have been trying to construct a Japanese understanding systems by using artificial neural networks. This paper realizes a complex sentence parsing system that is one of the principal subsystem. The basic function of the parsing system is conversions from syntactic structures to case structures. The system deals with complex sentences with embedded clauses and compound sentences that express time relations and causal relations.

When we construct a general system, we divide the system into several subsystems and adopt applicable neural network models for those subsystems. And we realize the desired system by integrating those subsystems. We call this constructing method, **the integrated connectionist model**. Concerning to natural language processing, many previous works by symbolic rule-based models have been proposed and already proved the way of dividing into subsystems. Thus, natural language processing is a good application of the integrated connectionist model. This research intends to break through some problems of language processing technologies and we expect that this integrated connectionist model will be proved to be a cognitive science model.

2. Philosophy of Our Research

2.1. Parallelism of Language Columns

When we started to construct a natural language understanding system with connectionist models, we introduced a working hypothesis as follows: Our brain has the whole world image consisting of many small parallel worlds. Moreover, we made a hypothesis the existence of **language columns** to express the small worlds[1]. The language column includes allot the all subsystems necessary to language processing such as the sentence parsing system and the sentence generating system. We realize the subsystem in a modular approach[2]. Furthermore, we will construct a parallel system of the language columns that express those of the small worlds to realize human language (figure 1).

The modules consisting of the sentence parsing system and the sentence generating system are realized by neural networks. The difference of the connectionist approach, that is, the method using neural networks from the symbolic rule-based approach is that the task of the constructing module is realized by implicit learning of the neural network instead of by formulating of explicit rules.

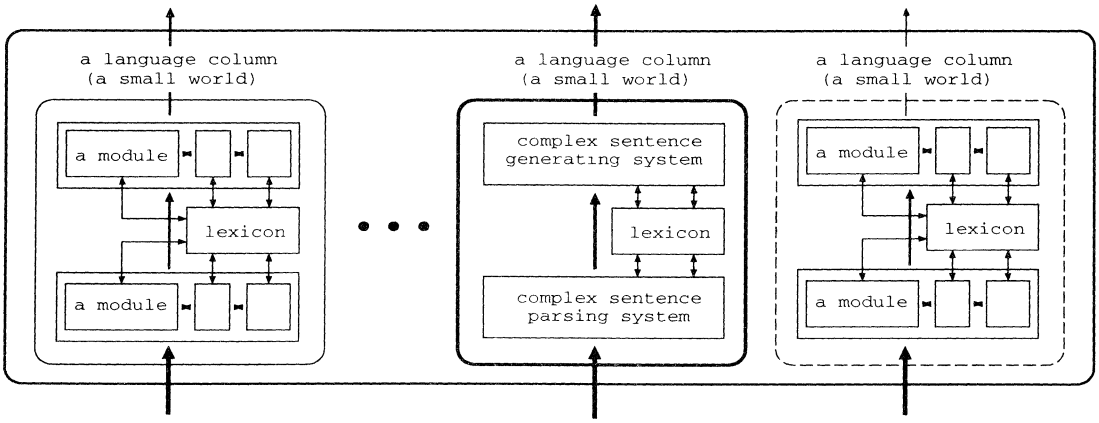

Figure 1: Parallelism of Small Worlds

2.2. Subsymbolic Approach

The other aspect of connectionism different from the conventional symbolism is the way of expressing words. In the connectionism, the localist representation and the distributed representation are proposed and in the latter case words are expressed with multiple neuron units. Naturally words are expressed with real valued vectors. Also, these real valued vectors are able to be expressed as position vectors in a multi-dimensional space. Moreover, a sentence gathering words is able to be expressed as position vectors in a multi-dimensional space. Therefore, we can express the relations among words and among sentences as the very high dimensional analogue relations. Hence word manipulations are able to be realized in not only a surface level but also subsymbolic level.

3. Corpus and Word Representations

3.1. Corpus

The corpus should be discussed in detail, because the small world of the language column is formed by the column. Therefore, we chose words used computational experiments, along with a policy as follows: (1) The corpus constructed by the words should be able to express small worlds divided from the whole real world, and (2) the corpus is all of the produced sentence except semantically unacceptable sentences. In this paper, we chose the words in the theme "I am(was) observing our garden in a winter day.". Chosen words and simple sentence are shown in appendix. As the result, we can make simple sentences and complex sentences like these:

- 妻が花壇に雪割草を植えた。
  (Wife planted a hepatica in the flower bed.)
- 木立に登る子供をやさしく妻が叱る。
  (Wife scolds child who climbs the tree tenderly.)

3.2. Concept-Vectors

We adopt distributed representation in our integrated connectionist model in accordance with subsymbolic approach. We express words with real valued vectors whose elements are [0.0,1.0] and composed of 12 dimensions. And we call these vectors, **the Concept-Vectors**. Two types of concept-vector are proposed. One is the fixed representation which vectors are random and fixed, the other is the adaptive representation acquired with the FGREP module that is proposed by Miikkulainen[3]. The adaptive representation shows the distribution that reflects the usage of each word in the corpus.

4. Network Architecture

4.1. Complex-Sentence-Parser

The schema of the complex sentence parser is shown in figure 2. The model can be divided into five modules: Simple-Sentence-Parser, Stack, Segmenter, Compressor and Wanderer Compressor. The modules are trained in their respective tasks separately and in parallel. On executing, the modules form a network of networks, each feeding its output to the input of another modules.

4.2. Simple-Sentence-Parser

One of the main modules of the complex sentence parser is the simple sentence parser. Figure 3 shows its network architecture. The task of the network is the assignment of the case-roles to the independent words of input sentences. We adopt the seven cases which are based on the case grammar proposed by Fillmore. There are AGENT, OBJECT, GOAL, LOCATION, INSTRUMENT for nouns, and ACTION, ACTION(PAST) for verbs.

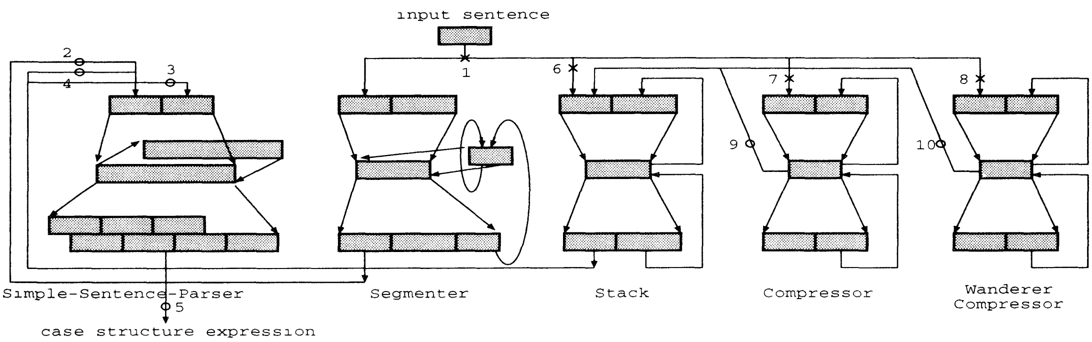

Figure 2: Complex-Sentece-Parser

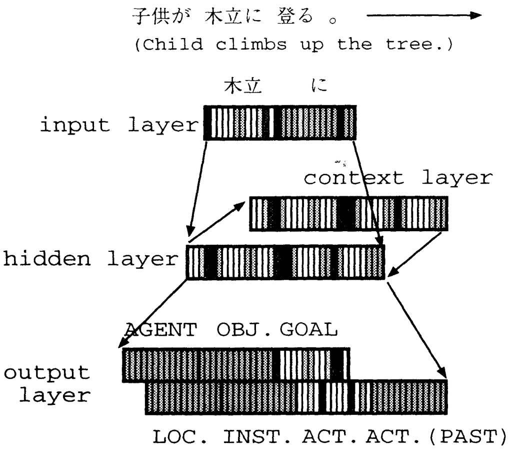

Figure 3: Simple-Sentece-Parser

For dealing with the flexible component length of the sentence, the network is constructed by use of the Elman type neural network. The input layer has 24 neural units (12 units × 2 slots) for two Concept-Vectores which correspond to two words of an input phrase. This phrase means a pair of one independent word and one dependent word. While the output layer has 84 units (12 units × 7 slots). If the network has successfully learned the task, each Concept-Vectors of the input independent words fill correct case slots in the output layer. In Japanese, verb phrases that restrict sentence meanings locate at the last of the sentences. Naturally verb should be shown at an early stage. Therefore, we adopted the reverse input method, in which sentence constituents are input in the reverse order.

4.3. Stack

The Stack is constructed by using RAAM(Recursive Auto Associative Memory) network. The Stack can realize the behavior like a stack of the computer algorithm. The task of the Stack against the Complex-Sentence-Parser is as follows:

1. Realizing reverse order input of a sentence intending to input vectors into the Simple-Sentence-Parser.
2. Inputting the activated pattern of the Stack hidden layer into the input of Segmenter, so that the Segmenter an act considering the preseding whole words of a complex sentence.

4.4. Segmenter

The Segmenter is constructed by use of the Jordan network which can learn time series. The Segmenter works as a governor of the whole of networks. The Concept-Vector of the input word and the activated pattern of Stack's hidden layer are input into the input layer of the Segmenter. The output layer of the Segmenter includes the units, the number of which is the same number as the input layer units, and 10 units for the 'control signal'. This control signal means the signal that control information transfer among the networks. The Segmenter is trained so that the output layer can output the same pattern of input activation, and control signals in order that the whole of networks executes its task in appropriate time sequence. The main task of the Segmenter is detecting verbs and outputting the control signals which makes the Stack executes popping action through a whole simple sentence. Hence the Segmenter segments simple sentences from complex sentences.

The Segmenter governs the whole of networks by control signals which are composed of 6 excitatory signals and 4 inhibitory signal. The excitatory control signal opens the gate when the value is 1. On the contrary, the inhibitory control signal opens the gate when the value is 0. The excitatory signals are shown with '○', while the inhibitory control signals are shown with '×' in figure 2.

4.5. Compressors

There are two kinds of the Compressors. One is the Compressor which compresses such expressions of more than one word as a adjective and a noun. The other is the Wanderer Compressor which compress the expressions of separated words like a adverb and a verb.

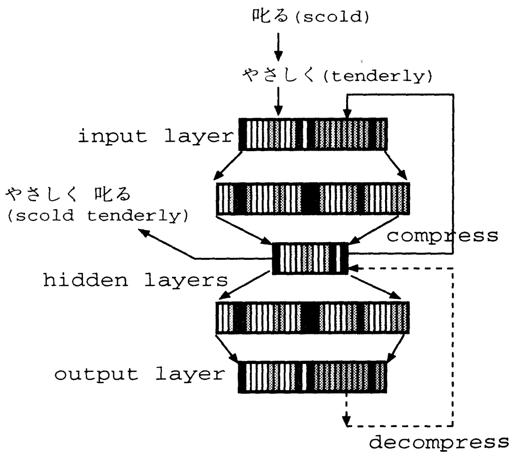

Figure 4: Wanderer Compressor

They are constructed by RAAM network (figure 4). They have two assemblies in the input layer and output layer, and trained with the auto-associative learning against adjectives and nouns, adverbs and verbs. Then compressed expression of those phrases (adj+noun, etc.) are aqcuired from the hidden layer.

5. Summary

We proposed a Japanese Complex-Sentence-Parser based on philosophies that are the parallelism of the language columns and the subsymbolic approach. The Complex-Sentence-Parser produce the case struture piles that are used by the complex sentence generating system, too. We are realizng the propose model on the computational simulations.

References

[1] A.F.C.Sharkey: " On Combining Artificial Neural Nets," *Connection Science*, vol.8, Nos 3 & 4, pp.99-313 (1996).

[2] A.F.C.Sharkey: " Modulatiry, Combining and Artificial Neural Nets," *Connection Science*, vol.9, No.1, pp.3-10 (1997).

[3] Miikkulainen,R.: " Subsymbolic Parsing of Embedded Structures," *Computional Architecutures Integrating Neural and Symbolic Processes.edited by R.Sun and L.A.Bookman*, pp.153-186 (1996).

Appendix

The followings shows words and simple sentence templates that are used in this research. Complex sentences will be produced to concatinate the simple sentences.

1.Words

noun(12)	妻 (wife) 子供 (child) 犬 (dog) 猫 (cat) 雪 (snow) 雨 (rain) 雪割草 (hepatica) 木立 (tree) 塀 (fence) 庭 (garden) 花壇 (flower bed) 犬小屋 (kennel)
verb(8)	登る (climb) 入る (enter) 走る (run) 遊ぶ (play) 降る (fall) 植える (plant) 吠える (bark) 叱る (scold)
adjective(4)	黄色い (yellow) 白い (white) 冷たい (cold) 小さい (small)
adverb(4)	さんさんと (-) やさしく (tenderly)
postpositional particle(4)	が (*ga*) を (*wo*) で (*de*) に (*ni*)
auxiliary verb(2)	た (*ta*:past) だ (*da*:past)
conjuctive particle(2)	して (*shite*:and) から (*kara*:cause)

2.Simple Sentence Template

Sentence Templates	Case-roles
子供 が 塀 に 登る (the child climbs up the fence)	agt goal
子供 が 木立 に 登る (the child climbs up the tree)	agt goal
猫 が 塀 に 登る (the cat climbs up the fence)	agt goal
猫 が 木立 に 登る (the cat climbs up the tree)	agt goal
犬 が 犬小屋 に 入る (the dog enters the kennel)	agt goal
猫 が 犬小屋 に 入る (the cat enters the kennel)	agt goal
子供 が 庭 を 走る (the child runs in the garden)	agt loc
犬 が 庭 を 走る (the dog runs in the garden)	agt loc
猫 が 庭 を 走る (the cat runs in the garden)	agt loc
子供 が 庭 で 遊ぶ (the child plays in the garden)	agt loc
子供 が 雪 で 遊ぶ (the child plays with snow)	agt inst
雪 が 庭 に 降る (snow falls in the garden)	agt goal
雨 が 庭 に 降る (rain falls in the garden)	agt goal
雪 が 木立 に 降る (snow falls in the garden)	agt goal
雪 が 犬小屋 に 降る (snow falls in the garden)	agt goal
妻 が 雪割草 を 花壇 に 植える (my wife plants hepatica in the flower bed)	agt obj goal
子供 が 雪割草 を 花壇 に 植える (the child plants hepatica in the flower bed)	agt obj goal
犬 が 子供 に 吠える (the dog barks the child)	agt obj
犬 が 猫 に 吠える (the dog barks the cat)	agt obj
妻 が 子供 を 叱る (my wife scolds the child)	agt obj
妻 が 犬 を 叱る (my wife scolds the dog)	agt obj