# กรอบงานวิเคราะห์ความหมายของคำกริยาเรียงสองตัวแบบติดกันในภาษาไทย

นางสาวจุฑาพักตร์ ปักษี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551

# A SEMANTIC ANALYSIS FRAMEWORK OF TWO CONSECUTIVE THAI SERIAL VERBS

Miss Jutapuck Pugsee

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Computer Engineering
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic year 2008

| | |
|---|---|
| Thesis Title | A SEMANTIC ANALYSIS ALGORITHM OF TWO CONSECUTIVE THAI SERIAL VERBS |
| By | Miss Jutapuck Pugsee |
| Field of Study | Computer Engineering |
| Advisor | Associate Professor Wanchai Rivepiboon, Ph.D. |

---

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

.................................................. Dean of the Faculty of Engineering

(Associate Professor Boonsom Lerdhirunwong, Dr.Ing.)

THESIS COMMITTEE

.................................................. Chairman

(Professor Boonserm Kijsirikul, Ph.D.)

.................................................. Advisor

(Associate Professor Wanchai Rivepiboon, Ph.D.)

.................................................. Examiner

(Assistant Professor Pattarasinee Bhattarakosol, Ph.D.)

.................................................. Examiner

(Assistant Professor Wirote Aroonmanakun, Ph.D.)

.................................................. External Examiner

(Assistant Professor Arnon Rungsawang, Ph.D.)

จุฑาพักตร์ ปักษี : กรอบงานวิเคราะห์ความหมายของคำกริยาเรียงสองตัวแบบติดกันในภาษาไทย. (A SEMANTIC ANALYSIS FRAMEWORK OF TWO CONSECUTIVE THAI SERIAL VERBS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร.วันชัย ริ้วไพบูลย์, 161 หน้า.

การวิเคราะห์คำกริยาเรียงภาษาไทยเป็นปัญหาที่ยุ่งยากเนื่องจากคำกริยาเรียงมีหลายประเภทและปรากฏในโครงสร้างวากยสัมพันธ์ที่หลากหลาย เช่น กริยา-กรรม กริยา-กริยาวิเศษณ์ กริยา-บุพบท เป็นต้น การที่จะเข้าใจความหมายของประโยคที่มีกริยาเรียงได้ จำเป็นต้องวิเคราะห์โครงสร้างวากยสัมพันธ์และความหมายของกริยาเรียงให้ได้ก่อน ดังนั้นงานวิจัยนี้จึงนำเสนอกรอบงานวิเคราะห์ความหมายของคำกริยาเรียงสองตัวแบบติดกันในภาษาไทย โดยวิเคราะห์รูปแบบวากยสัมพันธ์และคำที่มีความสัมพันธ์กับคำกริยาเรียงนั้น เพื่อที่จะหาข้อมูลที่เป็นประโยชน์สำหรับการวิเคราะห์ความหมาย ขั้นตอนหลักของกรอบงานจะแบ่งออกเป็น 2 ส่วนที่สำคัญ คือ การวิเคราะห์วากยสัมพันธ์โดยใช้ไวยากรณ์แบบโครงสร้างวลีภาษาไทย และการวิเคราะห์ความหมายโดยใช้ทฤษฎีวิเคราะห์ความสัมพันธ์

ผลลัพธ์ที่ได้แสดงให้เห็นว่า เทคนิคนี้สามารถวิเคราะห์รูปแบบความสัมพันธ์ด้านความหมายของคำกริยาเรียง โดยการพิจารณารูปแบบวากยสัมพันธ์ และความสัมพันธ์ของคำที่เกี่ยวข้องกับคำกริยาเรียงนั้น นอกเหนือจากนั้นการตีความหมายของคำกริยาเรียงสามารถตัดสินได้จาก การพิจารณารูปแบบวากยสัมพันธ์ของคำกริยาเรียง คำที่มีความสัมพันธ์กับคำกริยาเรียง และคำกริยาตัวแรกในกริยาเรียงนั้น

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต
สาขาวิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่ออ.ที่ปรึกษาวิทยานิพนธ์หลัก
ปีการศึกษา 2551

# # 4671808121 : MAJOR COMPUTER ENGINEERING

KEYWORDS : SERIAL VERB CONSTRUCTIONS / SEMANTIC ANALYSIS / NATURAL LANGUAGE UNDERSTANDING

JUTAPUCK PUGSEE : A SEMANTIC ANALYSIS FRAMEWORK OF TWO CONSECUTIVE THAI SERIAL VERBS. ADVISOR : ASSOC.PROF. WANCHAI RIVEPIBOON, Ph.D., 161 pp.

Analysis of Thai serial verbs is a difficult problem because there are many types of serial verb constructions (SVCs) and they can occur in various syntactic patterns, e.g. verb-object, verb-adverb, verb-preposition. To understand the meaning of a sentence containing SVCs, the syntax and semantic of SVCs must be resolved first. Therefore, the research proposes the semantic analysis framework concerning two consecutive Thai serial verbs. This framework analyzes the syntactic patterns and related words of SVCs to find useful information for semantic analysis. The main processes of this framework are separated into two parts. The first part is syntactic analysis using Thai phrase structure grammars, and the second part is semantic analysis using the relation analysis theorems.

The result shows that this technique can analyze the semantic binding patterns of SVCs by considering syntactic patterns and relations of related words. Furthermore, the semantic interpretations of SVCs can be determined by syntactic patterns, related words and the first verb of consecutive SVCs.

Department : Computer Engineering Student's Signature ..............................

Field of Study : Computer Engineering Advisor's Signature ..............................

Academic Year : 2008

# ACKNOWLEDGEMENTS

I am especially deeply grateful to my thesis advisor, Associate Professor Dr. Wanchai Rivepiboon, for his continuous support in the Ph.D program. He provided me useful guidance and suggestions throughout this research period. I also wish to express my special thanks to Assistant Professor Dr. Pattarasinee Bhattarakosol who provided me helpful instructions of doing research. She taught me how to be a good researcher. Moreover, I would like to thank Professor Dr. Martha W. Evens who looked after me when I did my research oversea. Without all of them, my thesis would have never been accomplished.

I would like to give my special thanks to other my thesis committees, Professor Dr. Boonserm Kijsirikul Assistant Professor Dr. Wirote Aroonmanakun and Assistant Professor Dr. Arnon Rungsawang, for giving valuable comments, reading and reviewing my research.

I would like to thank Thailand's Commission on Higher Education, Ministry of Education for funding. This fund supported for studying in Ph.D. program and gave me a chance to have research collaboration in oversea.

My thanks also go to all my past and present lecturers for their lectures and instructions. Moreover, I wish to thank all my colleagues and friends at the Computer Engineering Department for their everlasting friendships.

Finally, I would like to express my sincere gratitude to my parents and family for their supports and encouragements throughout my life.

.

CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER I
# INTRODUCTION

## 1.1 Motivation

To make a computer understand natural languages, it is necessary to make them understand the syntax and semantics of a sentence. Therefore, developing software to interpret or capture information from a document require us to develop software for syntactic and semantic interpreters. Much research in natural language understanding emphasizes semantic analysis. Semantic analysis must be performed to ascertain the meaning of a word, or a series of words. So, the benefit of this process leads to valuable information for further problem solving, such as automatic life saving information or an automatic life control mechanism.

Considering forms of sentences, most consist of a subject, a verb, and perhaps an object, but some sentences have other extended verbs called serial verbs. The serial verb construction, also known as verb serialization, is a syntactic phenomenon common to many languages, such as Thai, and Chinese languages. Contrary to subordination, where one clause is embedded in another, verb serialization strings two verbs together in a sequence in which neither verb is subordinated to the other (Tallerman, 1998). Therefore, sentences with serial verbs are difficult to analyze because the meaning of a serial verb is a composition of the meanings of its components. Moreover, there are various semantic interpretations. For example, "เลือกใช้/lueak-chai/," which means "select to use," is composed of two verbs: "เลือก/lueak/" and "ใช้/chai/". The meaning of this example with two consecutive Thai serial verbs can be discovered by combining the meaning of both verbs and recognizing that the second verb also emphasizes the purpose of the first verb. Understanding this example requires that we choose one of many types of semantic interpretation, the one that applies to purposive serial verbs.

Since serial verbs strongly influence the meaning of a sentence, the serial verb analysis is needed before semantic understanding can be completed. Therefore, this research will design a framework for semantic analysis of two consecutive Thai serial verbs using some machine leaning techniques to discover the patterns of two consecutive Thai serial verbs. Methods in the framework will create a relationship between syntactic information and semantic information using the grammatical knowledge and knowledge from the patterns of serial verbs. Furthermore, a lexicon of some two consecutive Thai serial verbs will be generated to provide information about serial verbs for other research.

## 1.2 Objectives

1.2.1 To propose a framework for analyzing the semantics of two consecutive Thai serial verbs.

1.2.2 To analyze the patterns of two consecutive Thai serial verbs using machine learning techniques for semantic analysis.

1.2.3 To create a lexicon for two consecutive Thai serial verbs.

## 1.3 Scope and Assumptions

1.3.1 Research on Thai serial verbs found in the ORCHID corpus obtained from the National Electronics and Computer Technology Center of Thailand (NECTEC) (ORCHID, 2003).

1.3.2 We will analyze the patterns of only those two consecutive serial verbs with a first element is among the top ten most frequent serial verbs in the ORCHID corpus.

## 1.4 Research Methodology

1.4.1 To collect basic information for the research.

1.4.2 To design methods for the syntactic analysis of sentences.

1.4.3 To design methods for the semantic analysis of two consecutive Thai serial verbs.

1.4.4 To design a method for analyzing the patterns of two consecutive Thai serial verbs using machine learning.

1.4.5 To design and create a lexical database of two consecutive Thai serial verbs.

1.4.6 To evaluate the performance of our proposed semantic analysis framework and analyze the results.

1.4.7 To make conclusions and produce documentation of our work.

## 1.5 Contribution

1.5.1 A framework for semantic analysis of two consecutive Thai serial verbs.

1.5.2 The patterns of two consecutive Thai serial verbs.

1.5.3 A lexical database of two consecutive Thai serial verbs.

## 1.6 Organization of the Thesis

Chapter II describes the theoretical background that is used in this thesis and reviews research works that are related to this research.

Chapter III describes the primary resources used to create our lexical database, the ORCHID Corpus (ORCHID, 2003), the Thai Computational Linguistics Laboratory's Computational Lexicon (TCLLEX, 2005) and Lexitron (LEXITRON, 2005).

Chapter IV presents our proposed framework to analyze syntactic information of sentences and semantic information of two consecutive Thai serial verbs.

Chapter V presents our experiments and discusses the results.

Chapter VI draws the conclusions of what have been achieved in this thesis. Future research directions related to this thesis are suggested.

# CHAPTER II
# BACKGROUND AND LITERATURE REVIEW

In this chapter, we describe some important theoretical background on the characteristics of Thai language, and research on Thai serial verbs constructions. In addition, we briefly describe semantic analysis, and some machine learning techniques. In the last part of this chapter, we discuss some ontologies and lexicons.

## 2.1 Characteristics of Thai language

### 2.1.1 Thai Sentence Structures

Thai sentences are composed of a sequence of words with unmarked word boundaries. Thai sentences consist of different phrases that appear in various positions according to their grammatical roles (Pankhuenkhat, 1998). Thai is essentially an SVO Language in which sentences are made up of a noun phrase subject followed by a verb phrase predicate. The basic forms of Thai sentence structure are Subject + Verb and Subject + Verb + Object. However, Thai sentence structure has a number of variations, such as Subject + Verb + Complement, Subject + Verb + Adverbial, Subject + Verb + Object + Adverbial or Subject + Verb + Object + Complement (Pankhuenkhat, 1998). In this study, Thai sentences with the structure Verb + Object are considered as an incomplete sentence. They will be excluded from this study, although they can occur in Thai speech and Thai written sentences. In Thai sentences it is difficult to identify where one sentence ends and a new sentence begins because a space can be put between two sentences, or between two clauses, or between two phrases. Moreover, sometimes two consecutive sentences are written without a space between them.

### 2.1.2 Thai Phrases

In (Pankhuenkhat, 1998), Thai phrases are classified into four different categories: Noun phrases, Verb phrases, Adverb phrases and Vocative phrases. We adopted this classification to classify Thai phrases that occur in our phrase structure grammar. Our categories of Thai phrases are Noun Phrases, Verb Phrases, Adverb Phrases, and Prepositional Phrases, which are described below:

1. Noun Phrase (NP): is composed of a noun, pronoun or nominal noun and their modifiers. The modifier can be a determiner, verb, adjective, or classifier. For example, "ปากกาสีฟ้า/pak-ka-si-fa/" (ปากกา/pak-ka/(pen) สีฟ้า/si-fa/(blue)) means "a blue pen") is composed of the basic noun "ปากกา/pak-ka/(pen)" and the modifier "สีฟ้า/si-fa/(blue)".

2. Verb Phrase (VP): is composed of a single verb as a main verb and its modifiers. The VP can be a verb combined with a negator as NEG + VERB to express the negation as "ไม่ไป/mai-pai/" – ไม่/mai/(not) ไป/pai/(go) means "not go". Moreover, the VP can be a verb combined with the pre-verb auxiliary as "ควรไป/khuan-pai/" – ควร/khuan/(should) ไป/pai/(go) means "should go".

3. Adverb Phrase (ADVP): is composed of an adverb and its modifiers. The ADVP can be just a single adverb, e.g., "มาก/mak/(much)" or an adverb with its modifiers. The ADVP can be combined with a negator (NEG) as "ไม่เร็ว/mai-reo/" – ไม่/mai/(not) เร็ว/reo/(fast) meaning "not fast," "ไม่ปกติ/mai-pa-ka-ti/" – ไม่/mai/(not) ปกติ/pa-ka-ti/(normal) means "not normal".

4. Prepositional Phrase (PP): is composed of a preposition and its object. The object is always a noun or a noun phrase, but it can contain a determiner, verb, or classifier. For example, "ในปีต่อไป/nai-pee-tor-pai/" – ใน/nai/(in) ปี/pee/(year) ต่อไป/tor-pai/(next) meaning "in the next year," "ที่โรงเรียนนี้/thi-rong-rian-ni/" – ที่/thi/(at) โรงเรียน/rong-rian/(school) นี้/ni/(this) means "at this school".

### 2.2 Research on Thai Serial Verb Constructions (SVCs)

We will now discuss some major previous research on serial verb constructions in Thai: (Thepkanjana, 1986), (Chuwicha, 1992), (Wilawan, 1993), (Muansuwan, 2002) and (Sudmuk, 2005). Some definitions and classifications of serial verbs from these researches are basic information for proposed semantics analysis framework.

### 2.2.1 Serial Verb Construction in Thai (Thepkanjana, 1986)

This work broadly defines SVCs as a surface form in which more than one verb phrase is strung together. SVCs are considered as one kind of multi-verb construction. The other kinds of multi-verb constructions are compound verbs and consecutive constructions. The functions of serial verbs in Thai are described as follows:

- complementing the initial verb, which results in the semantic implications of causative, passive, and resultative
- indicating direction and aspect
- acting as grammatical markers and/or case markers for which she uses the specialized term ‘coverb’
- indicating purposive and simultaneous actions

Moreover, this research classifies serial verbs in Thai into seven types: causatives, complements of modality verbs, resultatives, passives, directional serial verbs, aspectual serial verbs, and simultaneous serial verbs. Thepkanjana gives the following examples:

The examples of the first verb in causative SVCs are “ทำ/tham/(make)” and “ให้/hai/(give)”. The examples of the first verb of SVCs in complements of modality verbs are “อยาก/yak/(want)” and “พยายาม/pha-ya-yam/(try)”. The example of resultative SVCs “กินอิ่ม/kin-im/” illustrates the fact that this construction conveys the meaning that an event happens to the subject, and that event is followed by another event also occurring to the same person. The second event is the result of the first event. The example of passive SVCs “ถูกกัด/thuk-kat/(be bitten)” illustrates that the first verb of SVCs of this type “ถูก/thuk/” is a full verb expressing the inherent concept of passivity. The examples of the first verb of SVCs in the category of directional serial verbs illustrate that the initial verb in this group must be a motion verbs following the geometric shape of the path, such as “ตรง/trong/(straight)”. The other constraints are the direction with respect to the previous path, such as “ย้อน/yon/(reverse)”, the direction with respect to the outside world, such as “เลย/loei/(pass)” and the direction with respect to speech act participants, such as “ไป/pai/(go)” and “มา/ma/(come)”. For the aspectual serial verbs, Thepkanjana's research analyzes strings of verbs and aspectual verbs as complex verbs. The aspectual verbs will signal meanings only if they are serialized and occurred with certain types of verbs. The simultaneous serial verbs of this research are classified into two kinds that are purposive simultaneous serial verbs, e.g. “มาคุย/ma-khui/(come to talk)”, and simultaneous action serial verbs, e.g. “ยืนคุย/yuen-khui/(stand and talk)”.

#### 2.2.2 Clausehood in serial verb constructions in Thai (Chuwicha, 1992)

This study aims at investigating the syntactic and semantic properties of serial verb constructions (SVC) in Thai and determining the clausehood of the constructions. It is found that SVCs can be classified into two groups:

- the basic SVC, consisting of only two verbs
- the complex SVC, consisting of more than two verbs

Four syntactic and semantic properties are found to indicate compactness within the SVCs. They are sharing the same subject, having the same time reference, having the same aspect and not allowing a negative word to occur in between. The SVCs that have all the four properties are considered to have the highest degree of compactness or unity. The semantic properties of SVCs are analyzed in term of the kinds of event expressed by each verb in the SVC and the temporal relations between the events in the SVC. It is found that there are systematic relationships between the clausal structure and the semantic properties of SVCs. The temporal relations proposed in the mono-clausal SVCs are different from those proposed in the non-mono clausal SVCs.

#### 2.2.3 Reanalysis of so-called serial verb constructions in Thai (Wilawan, 1993)

Wilawan studied SVCs in four languages in four different language families: Thai, Khmer, Mandarin Chinese, and Yoruba. This research classifies SVCs in Thai into two major groups according to the syntactic properties of the first verb in the series, which is called the “regent verb”. The first group is SVCs with intransitive regents and the second group is SVCs with transitive regents. Each major group is classified by intransitive dependents and transitive dependents. The detail of these SVCs is described as follows.

- Manner intransitive verb and intransitive verb
- Manner intransitive verb and transitive verb
- Transitive verb and transitive verb with a shared object NP

Intransitive regent and intransitive dependents: SVCs in this group consist of two verbs with a shared NP subject. The regent verb is interpreted as referring to the manner in which the action of the dependent verb is performed. The first verb can only be a manner intransitive verb and the second verb is an intentional intransitive verb, such as “เดินร้องไห้/doen-rong-hai/(walk and cry)”.

Intransitive regents and transitive dependents: In this construction, the first verb must be a manner intransitive verb and the second verb must be an intentional transitive verb, such as “นั่งทำ/nang-tham/(sit and do something)”. This research analyzes that the second verb is a complement of the first verb because the first verb only occurs with particular verbs, and the part of the sentence headed by the second verb is analyzed as a subordinate construction.

Transitive regents and intransitive dependents: This so-called serial verb construction contains a manner transitive verb “ใช้/chai/(use)” and an infinitival intransitive verb complement, such as “ใช้หายใจ/chai-hai-chai/(use something to breath)”.

Transitive regent and transitive dependents: This construction consists of a purpose transitive regent verb with its transitive dependent, so it has only one overt NP object between two verbs. This research analyzes the second verb in this construction as a subordinate dependent of the first verb because the shared NP object must appear after the first verb.

### 2.2.4 Verb Complex in Thai (Muansuwan, 2002)

Muansuwan studies three SVCs in Thai: directional SVCs, the aspectual construction in SVCs, and the adjoining construction in sequential SVCs. These SVCs are described bellow.

The first verb in directional SVCs denotes a manner of motion and the noninitial verbs encode a directed motion, which includes information about the paths, directions, and/or deictic center of the motion events. Each verb in a directional SVC describes the same as the first verb. Thus, directional SVC can be interpreted as a single event obtaining from different perspectives. Thai Directional serial verb constructions include two kinds of syntactic structures: a recursive co-head structure and a complementation structure. All verbs in co-head structures have equal status, meaning that the occurrence of one verb is not determined by another. Verbs in complementation structures do not have equal status (one verb is the complement of the others). For aspectual SVCs, this research describes the position of aspect morphemes in Thai in term of their positions in the clause. Furthermore, this research defines the sequential SVCs as a construction that consists of a sequence of verbs, which share at least one argument and encode eventualities that occur in sequence. Adjoining constructions (AJCs) are a subset of the sequential SVCs. Muansuwan argues further that the resulting verb in an Adjoining Construction behaves the same way as the complements or inherent arguments of the matrix verb in that the resulting verb cannot be separated from the matrix verb and its arguments by an adverb. This research uses Head-Driven Phrase Structure Grammar (HPSG) as the theoretical basis for analysis.

### 2.2.5 The Syntax and Semantics of Serial Verb Constructions in Thai (Sudmuk, 2005)

Sudmuk examines both the syntax and the semantics of the Serial Verb Constructions (SVCs) in Thai in a Lexical Functional Grammar (LFG) framework. It is proposed that there are eight types of the SVCs in Thai with respect to the restriction of the limited sets of verbs that occurs in the series, which are the Motion SVCs, Posture

SVCs, Take-SVCs, Use-SVCs, Open Class SVCs, Give-SVCs, Causative SVCs, and Resultative SVCs. Each type of SVCs is described below:

Motion SVCs: There are two types of the motion SVCs depending on the set of verbs that occur in the sequence: motion – deictic SVCs and motion – directional SVCs. The set of the first verb in the motion- deictic SVCs is limited to the manner-of motion verbs such as “เดิน/doen/(walk)”, “วิ่ง/wing/(run)” and the set of its second verb is limited to the deictic verbs that are “มา/ma/(come)” and “ไป/pai/(go)”. The set of the first verb in the motion – directional serialization is limited to the manner-of-motion verbs.

Posture SVCs: The set of the first verb in the posture SVCs is the postural verb, and the second verb in series can be any verbs, that is, this verb class is open (called the open class for this research). Serial verbs in this class have only one semantic interpretation, the simultaneous interpretation. The postural verb indicates the action of the body in a particular position while the agent carries out an action, e.g. “นั่งร้องเพลง/nang-rong-phleng/(sing while sitting)”.

Take-SVCs: The first verb in the series of the Take-SVCs is the verb “เอา/ao/(take)” and the second verb belongs to the open class verb; this second verb is usually an activity verb. They have two semantic interpretations, the sequential interpretation or the purposive interpretation, e.g. “เอามีดตัดเนื้อ/ao-mit-tat-nuea/” has two meanings “take the knife and cut the meat”, or “take the knife to cut the meat”.

Use-SVCs: The first verb in series of the Use-SVCs is the verb “ใช้/chai/(use)” and the second verb belongs to the open class verb; again it is usually an activity verb. When the Use-SVCs are added more verb phrase, which is the verb that denotes the intentional action. They have either the sequential or purposive interpretation.

Open Class SVCs: There are two verbs in series in the open class SVCs. Both of them can be any verb in the open class. They have two semantic interpretations, which are the sequential or purposive interpretation.

Give-SVCs: There are two verbs in series in the Give-SVCs. The first verb is limited to a verb “ให้/hai/(give),” and the second verb belongs to the open class verb. They have two semantic interpretations, which are the sequential or purposive interpretation (the second verb denotes the purpose of the first verb).

Causative SVCs: There are two verbs in series in the Causative SVCs. The first verb is restricted to the verb “ทำ/tham/(make)”. For the second it may be any intransitive verb in the open class, but it may not be a transitive verb.

Resultative SVCs.: The Resultative SVCs are composed of two verbs in series. Both verbs belong to the open class verbs; the second verb is also serves as the result verb. They have only one semantic interpretation, which is the result interpretation. The second verb shows the result state of the agents or themes after the action of the first verb.

All eight types of the SVCs in Thai are shown in Table 2.1, and their semantic interpretations are shown in Table 2.2.

Table 2.1: Types of SVCs in Thai (Sudmuk, 2005)

| Types of Serialization | Class of Verb |
|---|---|
| Motion SVCs | Manner-of-motion verb + deictic/directional verb |
| Posture SVCs | Postural verb + open class verb |
| Take-SVCs | Take + open class verb<br>(usually an activity verb) |
| Use-SVCs | Use + open class verb |
| Open class SVCs | Open class verb + open class verb |
| Give-SVCs | Give + open class verb |
| Causative SVCs | Make + open class verb |
| Resultative SVCs | Open class verb + open class verb (result) |

Table 2.2: The semantic interpretation of SVCs in Thai (Sudmuk, 2005)

| Types of Serialization | The Semantic Interpretation |
|---|---|
| Motion SVCs | Overlapping, Purposive |
| Posture SVCs | Simultaneous |
| Take-SVCs | Sequential, Purposive |
| Use-SVCs | Sequential, Purposive |
| Open class SVCs | Sequential, Purposive |
| Give-SVCs | Sequential, Purposive |
| Causative SVCs | Causative |
| Resultative SVCs | Resultative |

## 2.3 Semantic Analysis

Semantics is the study of relationships between signs or symbols and what they represent. The goal of computational semantics is to find techniques for automatically constructing semantic representations. Lexical semantics explain the systematic meaning-related connections among words and the internal meaning-related structure of each word, and stores this information in a semantic lexicon. Compositional semantics combine the meanings of the component words to produce a logic form that represents the meaning of the sentence using the syntactic analysis of the sentence.

To analyze semantics, several approaches have been proposed. In (Sarikaya et al., 2003), semantic analysis involves finding the semantic units that span words or word groups and the relationship among all units in a sentence. Syntactic features, such as tags on words and the roles of words in sentences are used as semantic features. It is concluded that the combination of a number of features may improve performance over the best single feature.

Research by Bejan (Bejan et al., 2004) produced a method for automatically assigning semantic roles to the contents of English sentences. Semantic roles include agents, targets, items and goals. To identity semantic roles, their researches use many features of FrameNet (e.g., Phrase type, Position, Voice, Head word, Part of speech, Named entity class, Support_verbs, Coreness, Content word, Parse tree path, etc.). Some researchers use semantic grammars to analyze semantics (Minker, 2001).

### 2.4 Machine Learning Techniques

Much research on natural language understanding emphasizes on semantic analysis (Li et al., 2003). Solving this problem requires improvement in the representation of language information and in semantic analysis methods. (M`arquez, 2000) describes some work combining fields of machine learning and natural language processing. More recently, (Murata et al., 2002) developed new tagging methods using three different machine learning methods. Another research project, (Sebastiani, 2002) surveys classification of texts using some machine learning techniques. Although none of these studies have the same focus as our research, all of them show that machine learning techniques help to enhance the performance of natural language processing. We describe some details of machine learning techniques briefly in the following sections.

#### 2.4.1 Neural Networks

(Sebastiani, 2002) defines a neural network classifier as a network of units. The input units represent features, the output units represent the category or categories of interest (in our research these are meaning classes), and the weights on the edges connecting units represent dependence relations. For classifying meaning, its features and feature weights are loaded into the input units; the activation of these units is propagated forward through the network, and the value of the output units determines the choice of the meaning class. A typical way of training neural networks is back propagation, whereby the term weights of a training document are loaded into the input units; and if a misclassification occurs, the error is "back propagated" so as to change the parameters of the network and eliminate or minimize the error. The simplest type of neural network classifier is the perceptron, which is a linear classifier. Other types of linear neural network classifiers implement a form of logistic regression. Figure 2.1 shows a basic neural network.

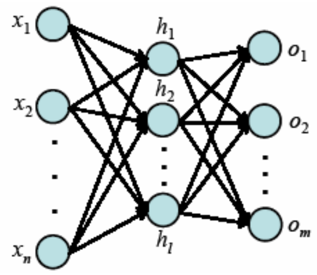



Figure 2.1: A Neural Network.

2.4.1.1 Probabilistic Neural Networks (PNNs)

The PNN takes its basic concept from the Bayesian statistical classifier which is the optimal statistical classifier. Specht proposed a method to formulate the weighted-neighbor method described above in the form of a neural network (Specht, 1990) so PNN is conceptually similar to K-Nearest Neighbor (k-NN) models. A Diagram of PNN appears in Figure 2.2.

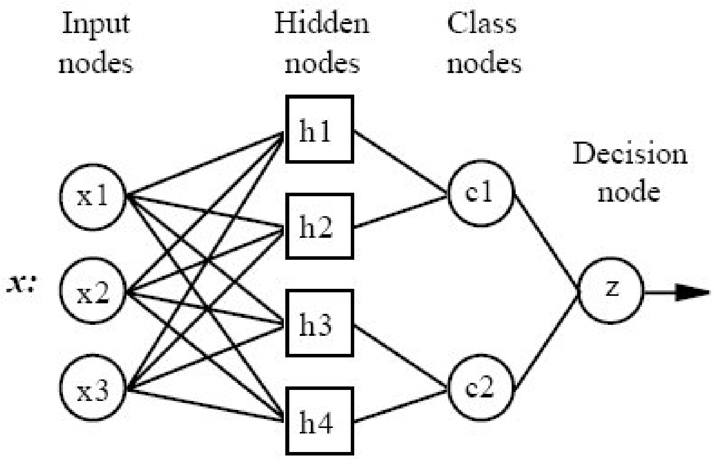



Figure 2.2: A Probabilistic Neural Network.

The basic idea is that a predicted target value of an item is likely to be about the same as other items that have close values of the predictor variables. A probabilistic neural network builds on this foundation and generalizes it to consider all of the other points (Wasserman, 1993). The distance is computed from the point being evaluated to each of other points, and a radial basis function (RBF) (also called a kernel function) is applied to the distance to compute the weight (influence) for each point. The radial basis function is so named because the radius distance is the argument to the function.

$$\text{Weight} = \text{RBF}(\text{distance})$$

Four layers of a PNN are listed as follows.

- Input layer - There is one neuron in the input layer for each predictor variable. The input neurons then feed the values to each of the neurons in the hidden layer.
- Hidden layer - This layer has one neuron for each case in the training data set. The neuron stores the values of the predictor variables for the case along with the target value. When presented with the x vector of input values from the input layer, a hidden neuron computes the Euclidean distance of the test case from the neuron center point and then applies the RBF kernel function using the sigma value(s). The resulting value is passed to the neurons in the pattern layer.
- Pattern layer / Summation layer - There is one pattern neuron for each category of the target variable. The actual target category of each training case is stored with each hidden neuron; the weighted value coming out of a hidden neuron is fed only to the pattern neuron that corresponds to the hidden neuron category. The pattern neurons add the values for the class they represent.
- Decision layer - The decision layer compares the weighted votes for each target category accumulated in the pattern layer and uses the largest vote to predict the target category.

### 2.4.2 Decision Trees

(M`arquez, 2000) Decision trees are the basis of a method of supervised learning from examples that represents one of the most popular approaches within the field of Artificial Intelligence for dealing with classification problems. In concept learning decision trees are sometimes translated into rules and eventually pruned to represent the concept. The most used system is C4.5-Rule. Decision trees are a way to represent rules underlying training data with hierarchical sequential structures that recursively partition the data. In all these fields of application, decision trees have been used for data exploration with some of the following purposes: Description, Classification and Generalization. A decision tree classifier is a tree where internal nodes are labeled by features, branches departing from them are labeled by the weight of each feature from the training data, and leaves are labeled with the names of the meaning classes.

### 2.4.3 Naive Bayes

(M`arquez, 2000) says that Naive Bayes classification can be viewed as a simple representative of statistical learning methods. This method assumes independence of features. It classifies a new example by assigning the class that maximizes the conditional probability of the class given the observed sequence of features in that example.

Let $\{1,...,k\}$ be the set of classes and $\{x_{i,1},...,x_{i,m}\}$ the set of feature values of $x_i$. The Naive Bayes method tries to find the class that maximizes $P(k \mid x_{i,1},...,x_{i,1})$ :

$$\arg\max_k P(k \mid x_{i,1},...,x_{i,1}) \approx \arg\max_k P(k)\prod_i P(x_{i,j} \mid k)$$

where $P(k)$ and $P(x_{i,j} \mid k)$ are estimated during the training process using relation frequencies.

To avoid the effects of zero counts when estimating the conditional probabilities of the model, a very simple smoothing technique has been used. It consists in replacing zero counts of $P(x_{i,j} \mid k)$ with $P(k)$ / N where N is the number of training examples.

### 2.4.4 Maximum Entropy Principle (ME)

(Murata et al., 2002) The maximum entropy method calculates the distribution of probabilities using the following equations. The category with the maximum probability is judged to be the correct category.

$$P(o \mid h) = \frac{\exp\sum_i \lambda_i f_i(o,h)}{\sum_{o'} \exp\sum_i \lambda_i f_i(o',h)'}$$

where $o$ represents the current word, $h$ represents the set of previous words, and the $f_i$ represents feature indicators that are activated for a certain history.

(Sarikaya et al., 2003) uses the Maximum Entity approach. Maximum Entropy is used to model sentences based on syntactic and higher level semantic information by computing the joint probabilities of a word sequence and a parse tree. Maximizing the entropy makes the distribution more uniform, which is known to provide a very helpful solution to data sparseness problems.

### 2.4.5 K Nearest Neighbors (k NNs)

(M`arquez, 2000) describes the k Nearest Neighbors algorithm, which does not generalize from training examples. Instead, examples are stored in memory and the classification of new examples is based on the classes of the most similar stored examples. After the instance base is built, new instances are classified by matching them to all instances in the instance base, and by calculating for each match the distance between the new instance $x = < x_i, ..., x_m >$ and the stored instance $y = < y_i, ..., y_m >$

The most basic metric for instances with symbolic features is the overlap metric defined as follows:

$$\Delta(x, y) = \sum_{i=1}^{m} w_i \delta(x_i, y_i)$$

where $w$ is the weight for the $i$ th feature and $\delta(x_i, y_i)$ is the distance between two values, which is defined as:

$$\delta(x_i, y_i) = 0 \qquad \text{if } x_i = y_i$$

$$\delta(x_i, y_i) = 1 \qquad \text{otherwise}$$

In the basic implementation, all examples are stored in the memory and classification of a new example is based on a k Nearest Neighbors algorithm using Hamming distance to measure closeness. For k's greater than 1, the resulting sense is the weighted majority sense of the k nearest neighbors where each example votes its sense with strength proportional to its closeness to the test example.

## 2.5 Ontologies and Lexicons

A lexicon is a collection of words, often with some organized information about them. There are different types of lexicons for example, network-based and frame-based lexicons. One example of a network-based hierarchy (taxonomy) is WordNet (Miller et al., 1993). Examples of a frame-based networks are Mikrokosmos (Beale, 1995), the Generative Lexicon (Pustejovsky, 1995), FrameNet (Bejan et al., 2004), and VerbNet (Kipper et al., 2000).

Ontology is a system of concepts relevant for knowledge about actions in the world or a system of symbols representing a system of concepts (Hirst, 2004). A linguistic ontology is a system of symbols representing the concepts (meanings) encoded by natural language (NL) expressions (lexical units, terms, etc.), such as a semantic representation language or an Interlingua. Some linguistic So a linguistic ontology can be used to provide structure to a lexicon.

Research on question-answering, such as (Bejan et al., 2004), produced a method for automatically assigning semantic roles to the contents of English sentences. Semantic roles include agents, targets, items and goals. To identity semantic roles, their researches uses many features of FrameNet (e.g., Phrase type, Position, Voice, Head word, Part of speech, Named entity class, Support_verbs, Coreness, Content word, Parse tree path, etc.).

FrameNet (Fillmore et al., 2003) is a computational lexicography project that extracts information about the linked semantic and syntactic properties of English words from large electronic text corpora. It is based on the theory of Frame Semantics and it is concerned with networks of meanings. The examples of relations in FrameNet are Subframe, Inheritance, See-Also, Using, Inchoative-of and Causative-of. Some researchers use a semantic grammar to analyze semantics (Minker, 2001).

WordNet (Miller et al., 1993) is an on-line lexical database whose design is inspired by current psycholinguistic theories of human lexical memory. English nouns, verbs, and adjectives are organized into synonym sets representing lexical concepts.

WordNet was organized lexical information in terms of word meaning rather than word forms. Semantic relations in WordNet are Synonymy, Antonymy, Hyponomy, Meronymy, and Entailments for verbs. The number of words, synsets and senses in WordNet are shown in Table 2.3. WordNet classifies nouns (24 groups) and verbs (15 groups) into detailed groups. WordNet is a rich resource as far as nouns are concerned, but the semantics of verbs is not well captured.

Table 2.3: WordNet.

| POS | Unique String | Synsets | Word-Sense Pairs |
|---|---|---|---|
| Noun | 117,097 | 81,426 | 145,104 |
| Verb | 11,488 | 13,650 | 24,890 |
| Adjective | 22,141 | 18,877 | 31,302 |
| Adverb | 4,601 | 3,644 | 5,720 |
| Total | 155,327 | 117,597 | 297,916 |

Another interesting lexicon is VerbNet (Kipper et al., 2000), which classifies verbs according to Levin verb classes (Levin, 1993). Each verb class is completely described by the set of its members (senses) and thematic roles for the predicate-argument structure of the members. It includes the syntactic subcategorization information, information about thematic roles, syntax, and basic lexical semantics (Crouch and King 2005). The information about thematic roles are such as patients, instruments, destinations, sources; about syntax information are such as plural, and infinitival. Each frame is composed of a frame description, a syntactic description, semantic predicates, thematic roles, and examples. The thematic roles for "hit" are shown in Figure 2.3.

Agent[+int_control]

Destination[+animate OR [+location -region]]

Destination[+animate]

Source[+location]

Theme[+concrete]

Figure 2.3: An Example of Thematic Roles for "hit"

# CHAPTER III
# RESOURCES

In this chapter, we will describe the ORCHID corpus and some existing Thai lexicons that provide information about words for our lexicon.

## 3.1 The ORCHID Corpus (Thai Part-of-speech Tagged Corpus)

As part of our research we conducted experiments on one of the Thai corpora, ORCHID (Open linguistic Resources CHanelled toward InterDisciplinary research). The scope of this research is to analyze patterns of only those SVCs involving two consecutive serial verbs in which the first element is in the top ten most frequent serial verbs in the ORCHID corpus. ORCHID is an initiative aimed at building linguistic resources to support research in natural language processing. The ORCHID corpus is a collection of annual reports and research proposals produced by different universities and research organizations in Thailand. An overview of the ORCHID corpus is shown in Table 3.1 (ORCHID, 2003; Sornlertlamvanich and Charoenporn, 1998).

**Table 3.1: The ORCHID corpus.**

| | |
|---|---|
| Size of Corpus | 5.8 MB (text file) |
| Number of Sentences | 27,634 sentences |
| Number of Words | 352,113 words |
| Number of Distinct Tags (Word Classes) | 14 categories<br>47 subcategories |

There are 14 word categories and 47 word subcategories in the ORCHID corpus (listed in APPENDIX B). The subcategories are used to help in determining the meanings of words. Thai word classes are similar to English classes, but there is no adjective word class in Thai. Instead, these functions are carried out by a subclass of verbs that always follow the nouns that they modify, for instance, “ปากกาสวย/pak-ka-suai/” – ปากกา/pak-ka/(pen) สวย/suai/(beautiful) means “beautiful pen”. These verbs are called attributive verbs (VATT). Examples in the ORCHID corpus include “อ้วน/uan/(fat)”,

“ดี/di/(good)”. So, the Thai verb class is divided into three subcategories, the action category (VACT), the state category (VSTA) and the attribute category (VATT).

The most important part of a sentence, which describes an action, or an important state or attribute, is the verb. In Thai language, verbs can appear in a series of two or three or sometimes more.

Examples of serial verbs:

“เลือกใช้/lueak-chai/” means “select to use.” – เลือก/lueak/(choose) ใช้/chai/(use).

โครงการวิจัยได้เลือกใช้เครื่องไมโครคอมพิวเตอร์เป็นหลัก (The research project selected mainly to use microcomputers.).

“ใช้ออกแบบ/chai-ok-baep/” means “use to design” – ใช้/chai/(use) ออกแบบ/ok-baep/(design).

ซอฟท์แวร์NECTEC II สามารถใช้ออกแบบวงจรรวมได้จริง (The software can be used to design an integrated circuit practically.).

“ทำหน้าที่กำหนด/tham-na-ti-kam-not/” means “act to define” – ทำหน้าที่/tham-na-ti/(act) กำหนด/ kam-not /(define).

คณะกรรมการนโยบายทำหน้าที่กำหนดนโยบายตลอดจนกำกับการดำเนินงาน (Policy Committees act to define policies and oversee processes.).

Moreover, the form of Thai verbs does not change with tense, so it is difficult to determine the main verb in a verb phrase. The parts of speech in the ORCHID corpus can help solving this problem partially, because the verb is separated from the auxiliary.

**3.2 TCL’s Computational Lexicon (TCLLEX)**

(TCLLEX, 2005) The Thai Computational Linguistics Laboratory (TCL) is a partnership laboratory of the Computational Linguistics Group (Japan), under the Keihanna Human Info-Communication Research Center (KICR) of the National Institute of Information and Communications Technology (NICT). TCL’s computational lexicon is a basic lexical knowledge base for natural language processing research. The authors designed both the terminology and the ontology for structuring the lexicon based on the ideas of computability and reusability. The computational lexicon consists of more than 60,000 Thai-English lexical entries, which are represented in morphological, syntactic, and semantic information. The syntactic information gives the parts of speech of words and the verb patterns in the sentence structures. The semantic information supplies word senses and tries to discover preferences of syntactic arguments of thematic roles. Table 3.2 and Table 3.3 show logical constraints and semantic constraints in TCLLEX. Our lexicon is also based on some semantic information from TCLLEX.

Table 3.2: Logical Constraints.

| Attribute | Value Description |
|---|---|
| Is-a (ISA) | a conceptual class of a given word (X) |
| Equal(EQU) | a word having the same meaning of X |
| Not-equal(NEQ) | a word having the opposite meaning of X |
| Part-of(POF) | a word specifying a part of X |
| Whole-of(WOF) | a word referring to the whole of which X is a part |

Table 3.3: Semantic Constraints.

| Attribute | Value Description |
|---|---|
| Agent(AGT) | an entity initiating the action |
| Object(OBJ) | an entity affected by the action |
| Instrument(INS) | an entity used in the action |
| Location(LOC) | a position or place where an event occurs |
| Time(TIM) | a point or period of time when an event occurs |

The guiding principle of TCLLEX is the idea of structuring a lexicon for computability, so the lexicon of two consecutive Thai SVCs, which is produced by the proposed framework, is designed and structured based on some tag structures of TCLLEX. Moreover, we use the hierarchy of word categories in TCLLEX to classify the categories of words that are related to SVCs.

### 3.3 Lexitron (LEXITRON, 2005)

Lexitron (Lexicon + Electron) is a Thai <--> English corpus-based dictionary created by the National Electronics and Computer Technology Center of Thailand (NECTEC). It contains of 53,000 English-Thai entries with English words translated to Thai words, synonyms, antonyms, and examples of English sentences. It also contains of 35,000 Thai-English entries with Thai words translated to English words, Thai meanings, synonyms, antonyms, and examples of Thai sentences. This dictionary is based on the information from the ORCHID corpus and the online version was developed by a J-builder program and an Apache server. Lexitron classifies words into main categories without subcategories, but it contains synonyms and antonyms of words. Lexitron uses English parts of speech, i.e., noun, verb, adverb, adjective, preposition, conjunction and pronoun, so it lacks some useful information for analyzing Thai sentences. Table 3.4 shows the number of entries in each entry class in Lexitron.

Table 3.4: Lexitron.

| POS | Number of entries |
|---|---|
| Noun | 21,090 |
| Verb | 13,028 |
| Adjective | 2,879 |
| Adverb | 2,876 |
| Conjunction | 218 |
| Pronoun | 133 |
| Others | 629 |

Some methods in the proposed framework use some word information from Lexitron such as collecting the result of word segmentation and finding words surrounding SVCs. In addition, the word category in Lexitron is one of the basic ideas for categories of words related to SVCs in this research.

# CHAPTER IV

# METHODOLOGY

This chapter describes our framework for semantic analysis of two consecutive Thai serial verbs, the creation of a lexicon for these verbs and the detail of some algorithms of some processes in the proposed framework.

## 4.1 An Overview of Framework for Semantic Analysis of Two Consecutive Thai Serial Verbs

Our framework is composed of two main components: one performs the syntactic analysis of sentences and the other performs the semantic analysis of serial verbs. All components of the framework are shown in Figure 4.1.

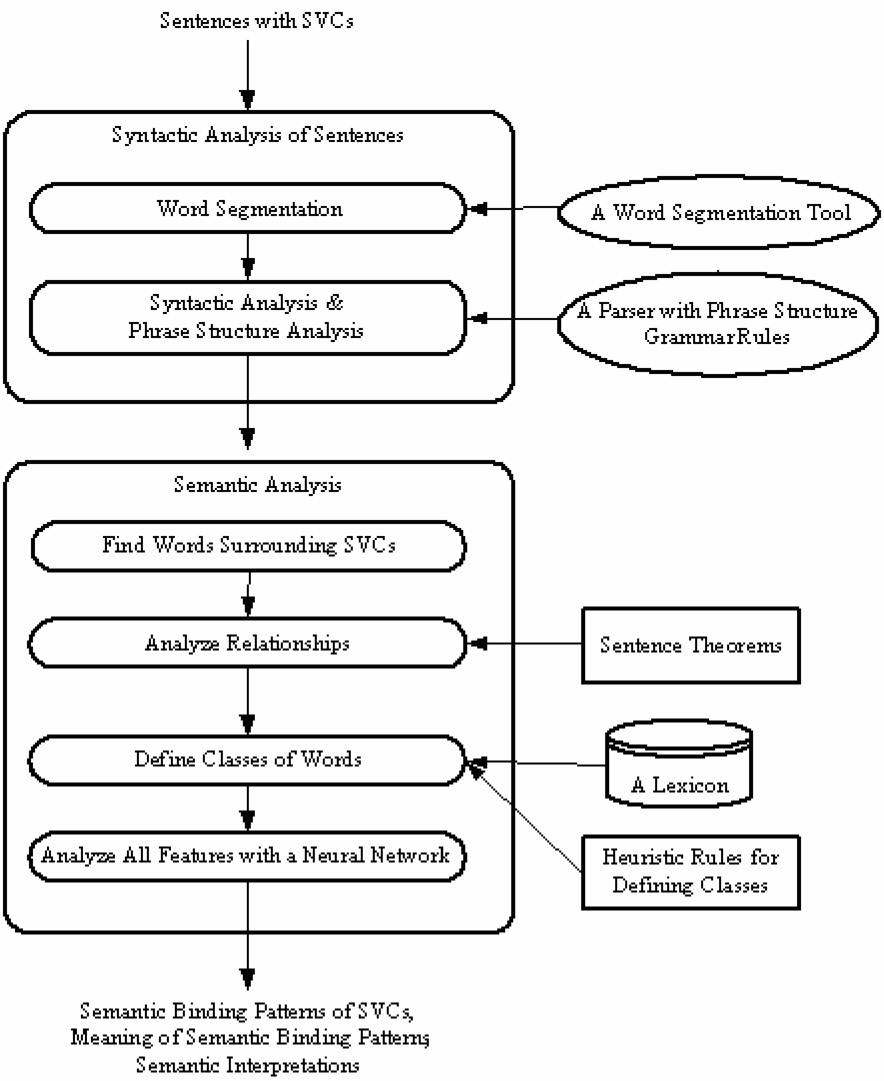



Figure 4.1: Overview of a semantic analysis framework.

As shown in Figure 4.1, the syntactic analysis of sentences is carried out by two modules while the semantic analysis is carried out by four modules. The syntactic analysis of sentences is performed to provide syntactic information about the words in the sentence and about the sentence structures for the semantic analysis part. Machine learning techniques are applied during the semantic analysis process to discover the patterns of serial verbs. Finally, all of the information about serial verbs from the previous analyses is collected to produce a lexicon for two consecutive Thai serial verbs. The details of the syntactic analysis of sentences and the semantic analysis are described in Sections 4.2 and 4.3 respectively

## 4.2 Syntactic Analysis of Thai Sentences

The syntactic analysis component is divided into two parts: word segmentation, followed by syntactic analysis and phrase structure analysis. Figure 4.2 shows an overview of our syntactic analysis of Thai sentences with serial verbs. The details of each step are described below.

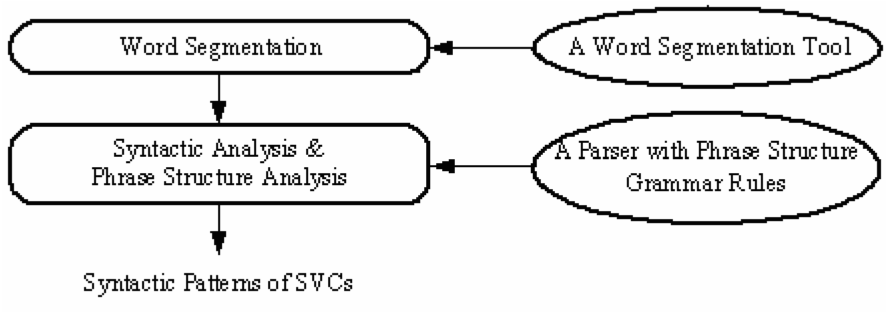



Figure 4.2: Syntactic analysis of sentence processes.

### 4.2.1 Word Segmentation

In the word segmentation part, since the Thai language as written does not represent between words, our research uses a Thai word segmentation tool (SWATH, 2003) to segment individual words and tag their parts of speech (Thai parts of speech are described briefly in Appendix B). This is what happens to the serial verbs of a sentence in the word segmentation module. Each serial verb, which contains more than one verb, will be considered as one individual phrasal verb. On the other hand, other words, such as nouns, auxiliaries, adverbs, adjectives and prepositions will be considered independently. To automatically analyze syntactic structures of sentences by designed grammar rules in the syntactic analysis process, the results of this process are required some editing to handle punctuation and determiners, and to correct compound words. The example of editing is managing punctuation, determiners and correcting compound words. Punctuations and some determiners are not found after word segmentation process. Then, correcting rules for compound words are generated to merge separated words into a compound word. However, these rules may not conform to the linguistic principles. Figures 4.3 and 4.4 state rules to correct compound nouns and adverbs that are segmented into more than one word.

| | | | |
|---|---|---|---|
| FIXN + VACT | = | NOUN | (1) |
| FIXN + VATT | = | NOUN | (2) |
| FIXN + VSTA | = | NOUN | (3) |
| FIXN + XVAM | = | NOUN | (4) |
| NCMN + VACT + NCMN | = | NOUN | (5) |
| PPRS + NPRP + NPRP | = | NOUN | (6) |
| NCMN + VATT | = | NOUN | (7) |
| DETERMINER + NOUN | = | NOUNWORD | (8) |
| NOUN + DETERMINER | = | NOUNWORD | (9) |
| NOUN + DETERMINER + CLASSIFIER | = | NOUNWORD | (10) |

**Figure 4.3: Some rules for nouns and compound nouns.**

Some examples of corrected compound nouns are:

(1) FIXN(“การ/kan/”) + VACT(“ทำงาน/thamngan/”) = NOUN(“การทำงาน/kan-thamngan/”)

(2) FIXN(“ความ/khwam/”) + VATT(“เร็ว/reo /”) = NOUN(“ความเร็ว/khwamreo/”)

(4) FIXN(“ความ/khwam/”) + XVAM(“สามารถ/samat/”) = NOUN(“ความสามารถ /khwam samat/”)

(5) NCMN(“หน่วย/nuai/”) + VACT(“แปลง/plaeng/”) + NCMN(“สัญญาณ/sanyan/”) = NOUN(“หน่วยแปลงสัญญาณ/nuai-plaeng-sanyan/”)

(6) PPRS(“คุณ/khun/”) + NPRP (Name)+ NPRP(Surname) = NOUN(“คุณ/khun/” + Name + Surname)

(7) NCMN(“รูปแบบ/rupbaep/”) + VATT(“สมมาตร/sommat”) = NOUN(“รูปแบบสมมาตร/rupbaep-sommat/”)

(10) NOUN(“ข้อมูล/khomun”-NCMN) + DETERMINER(“8”-DCNM) + CLASSIFIER(“บิต /bit”-CMTR) = NOUNWORD(“ข้อมูล 8 บิต/khomun-8-bit”)

| | | | |
|---|---|---|---|
| FIXV + VATT | = | ADVERB | (1) |
| FIXV + VSTA | = | ADVERB | (2) |

**Figure 4.4: Some rules for adverbs.**

Some examples of corrected adverbs are:

(1) FIXV(“ความ/khwam /”) + VATT(“สำคัญ/samkhan/”)= ADVERB (“ความสำคัญ/khwam samkhan/”)

(1) FIXV(“อย่าง/yang/”) + VATT(“รวดเร็ว/ruatreo/”) = ADVERB (“อย่างรวดเร็ว/yang ruatreo/”)

(2) FIXV(“อย่าง/yang/”) + VSTA(“สูง/sung/”) = ADVERB (“อย่างสูง/yangsung/”)

(2) FIXV(“โดย/doi/”) + VSTA(“ตรง/trong/”) = ADVERB (“โดยตรง/doitrong/”)

### 4.2.2 Syntactic Analysis and Phrase Structure Analysis

The Thai grammar for serial verbs defined in our research is used for syntactic analysis and phrase structure analysis. The syntactic analysis and the phrase structure analysis use a phrase structure grammar (PSG) to analyze parts of sentences. The PSG describing the sentence structures of the language is defined in the form of a set of rules which describe possible modes of combination of words. The basic phrases that can represent the meaning of sentences include noun phrases, verb phrases, adverb phrases and prepositional phrases This research has produced a specific PSG which provides added semantic information of basic sentence structures to analyze syntactic structures of serial verbs in sentences. Therefore, the designed PSG of this framework is valuable than other PSGs. The basic sentence structures are classified into four different categories in terms of their functions: subjects, verbs, objects, and adverbials. The phrase structure trees for the sentences are created one sentence at a time and they are parsed by grammar rules in a parser tool. Some non-terminal symbols in grammar rules are formed to automatically parse sentences by the parser tool.

To solve the problem of ambiguous structure trees, the probability of phrase patterns computed from example sentences in the corpus is applied. The probabilistic values are computed by applying the following formula:

$$prob(rule_i) = \frac{\#\,\text{parsed sentences by rule}_\text{i}}{\#\,\text{parsed sentences by the same left non - terminal symbol rules}}$$

Rules with the highest probability will be chosen to parse sentences. The result of this process is a phrase structure tree of the sentence that represents the subject, the object of the serial verb and some phrases that modify the serial verb, such as adverb phrases, and prepositional phrases. Some parts of the phrase structure grammar (Pugsee and Rivepiboon 2005) are expressed in the form of grammar rules with probabilistic values (p) ordering by the main components to particle components of the sentence as follows. The source codes for the Thai grammar rules in our parser are shown in Figure 4.5.

| | | | | |
|---|---|---|---|---|
| S | --> | SJ + VP | p = 0.003 | (1) |
| S | --> | SJ + VA | p = 0.218 | (2) |
| S | --> | SJ + VO | p = 0.583 | (3) |
| S | --> | SJ + VOA | p = 0.196 | (4) |
| SC | --> | SJCL + VP | p = 0.050 | (5) |
| SC | --> | SJCL + VPA | p = 0.650 | (6) |
| SC | --> | SJCL + VO | p = 0.100 | (7) |
| SC | --> | SJCL + VOA | p = 0.250 | (8) |
| VBA | --> | VP + ADVB | | (9) |
| VO | --> | VP + OJ | | (10) |
| VOA | --> | VO + ADVB | | (11) |
| SJ | --> | NP | | (12) |
| OJ | --> | NP | | (13) |
| ADVB | --> | ADVP | p = 0.310 | (14) |
| ADVB | --> | ADVP+ SUBORDCL | p = 0.028 | (15) |
| ADVB | --> | ADVP+PP | p = 0.113 | (16) |

| | | | | |
|---|---|---|---|---|
| ADVB | --> | COORDCL | p = 0.014 | (17) |
| ADVB | --> | PP | p = 0.261 | (18) |
| ADVB | --> | PP + ADVP | p = 0.021 | (19) |
| ADVB | --> | PP + SUBORDCL | p = 0.035 | (20) |
| ADVB | --> | SUBORDCL | p = 0.218 | (21) |
| NP | --> | noun | | (22) |
| NP | --> | NP + RELCL | p = 0.333 | (23) |
| NP | --> | NP + NP | p = 0.333 | (24) |
| NP | --> | NP + PP | p = 0.333 | (25) |
| VP | --> | auxiliary + VP | | (26) |
| VP | --> | verb | | (27) |
| VP | --> | VP + VP | p = 1.000 | (28) |
| ADVP | --> | auxiliary + ADVP | | (29) |
| ADVP | --> | adverb | | (30) |
| ADVP | --> | ADVP + ADVP | p = 1.000 | (31) |
| PP | --> | preposition + NP | | (32) |
| RELCL | --> | relative pronoun + S | p = 0.006 | (33) |
| RELCL | --> | relative pronoun + VP | p = 0.114 | (34) |
| RELCL | --> | relative pronoun + VBA | p = 0.369 | (35) |
| RELCL | --> | relative pronoun + VO | p = 0.485 | (36) |
| RELCL | --> | relative pronoun + VOA | p = 0.070 | (37) |
| SUBORDL | --> | subordinating conjunction + S | p = 0.080 | (38) |
| SUBORDL | --> | subordinating conjunction + VP | p = 0.040 | (39) |
| SUBORDL | --> | subordinating conjunction + VBA | p = 0.340 | (40) |
| SUBORDL | --> | subordinating conjunction + VO | p = 0.420 | (41) |
| SUBORDL | --> | subordinating conjunction + VOA | p = 0.120 | (42) |
| COORDL | --> | coordinating conjunction + VP | p = 0.000 | (43) |
| COORDL | --> | coordinating conjunction + VBA | p = 0.000 | (44) |
| COORDL | --> | coordinating conjunction + VO | p = 1.000 | (45) |
| COORDL | --> | coordinating conjunction + VOA | p = 0.000 | (46) |
| SJCL | --> | SJ + VB | p =0.040 | (47) |

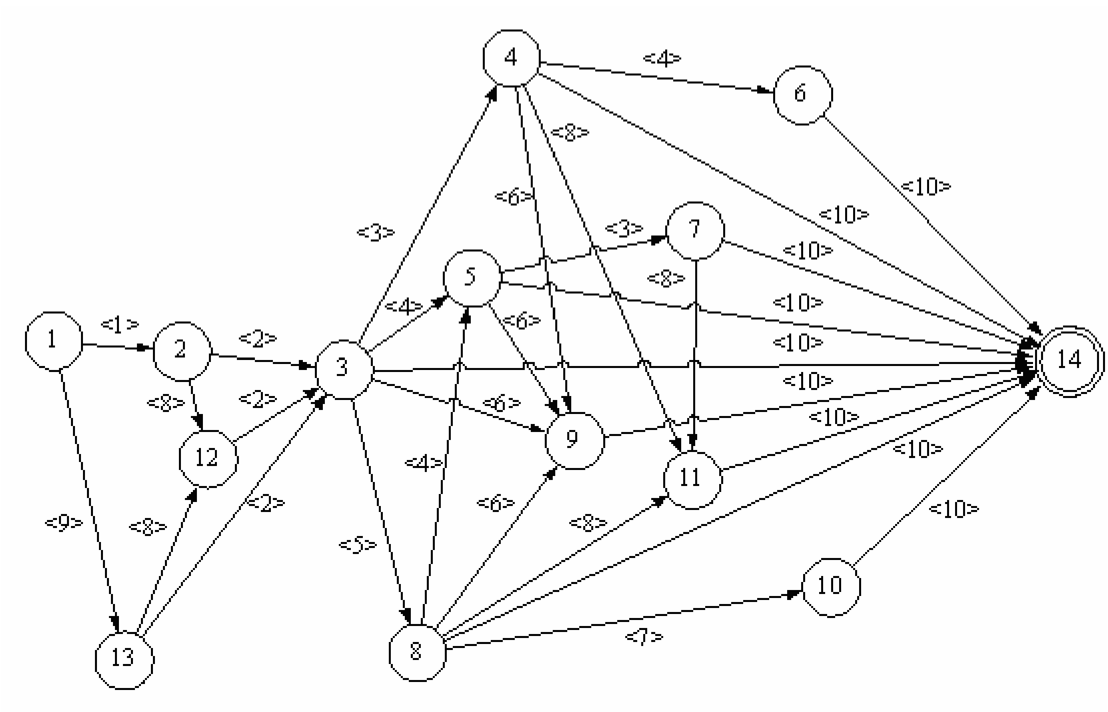
| | | | | |
|---|---|---|---|---|
| SJCL | --> | SJ + VBA | p =0.619 | (48) |
| SJCL | --> | SJ + VO | p =0.095 | (49) |
| SJCL | --> | SJ + VOA | p =0.238 | (50) |

According to the grammar rules described above, some rules have no probabilistic values because the right hand side is the terminal symbol, or there is no alternative path to parse the sentence.

```
grammar Thai // skips whitespace and comments
{
   allstn ::= stn ["SPAC" allstn];
   stn ::= [sj](vp|va|vo|voa)["EITT"]["PNTR"]|[sjcl](vp|va|vo|voa)["EITT"]["PNTR"];
   va ::= vp advb;
   vo ::= vp oj;
   voa ::= vp oj advb;
    sj ::= np;
   oj ::= np;
   sjcl ::= [sj][va|vp|vo|voa];
   advb ::= advp ([pp]|[subordincl])| pp ([advp]|[subordincl]) |subordincl| coordincl;
   np ::= nounword ([{nounword}][pp]| [relativecl]);
   vp ::=  [{helpverb}] verbword [{verbword}];
   advp ::= advword [{advword}][{deter_class}];
   nounword ::=  [deter_class]noun[{noun}][{deter_class}]|pronoun|deter_class;
   verbword ::= verb [{verb}];
   advword ::=  [{helpverb}] adv [{adv}];
   relativecl ::= "PREL" [sj](vp|va|vo|voa);
   subordincl ::= "JSBR" [sj][(vp|va|vo|voa)];
   coordincl ::= "JCRG" (vp|va|vo|voa);
   pp ::= prepword np;
   prepword ::= prep [{prep}];
   pronounword ::= pronoun [{pronoun}];
   deter_class ::= "DDAN"|"DDAC"|"DDBQ"|"DDAQ"|"DIAC"|
"DIBQ"|"DIAQ"|"DCNM"| "DONM"|"CNIT"|"CLTV"|"CMTR"|"CFQC"|"CVBL";
   conj ::= "JCRG"|"JCMP"|"JSBR";
   noun ::= "NCMN" |"NTTL"|"NPRP"|"NLBL" |"NCNM"|"NONM"|"FIXN" (verb|
helpverb)[verb][{noun}];
   verb ::= "VACT"|"VSTA"|"VATT";
   helpverb ::= "XVBM"|"XVAM"|"XVMM"|"XVBB"|"XVAE"|"NEG";
   prep ::= "RPRE";
   adv ::= "ADVN"|"ADVI"|"ADVP"|"ADVS"|"FIXV" verb
[{verb}][{noun}[|("XVAE"[verb][adv]);
   pronoun ::= "PDMN"|"PPRS"|"PNTR";
};
```

Figure 4.5: Source codes for Thai grammar rules

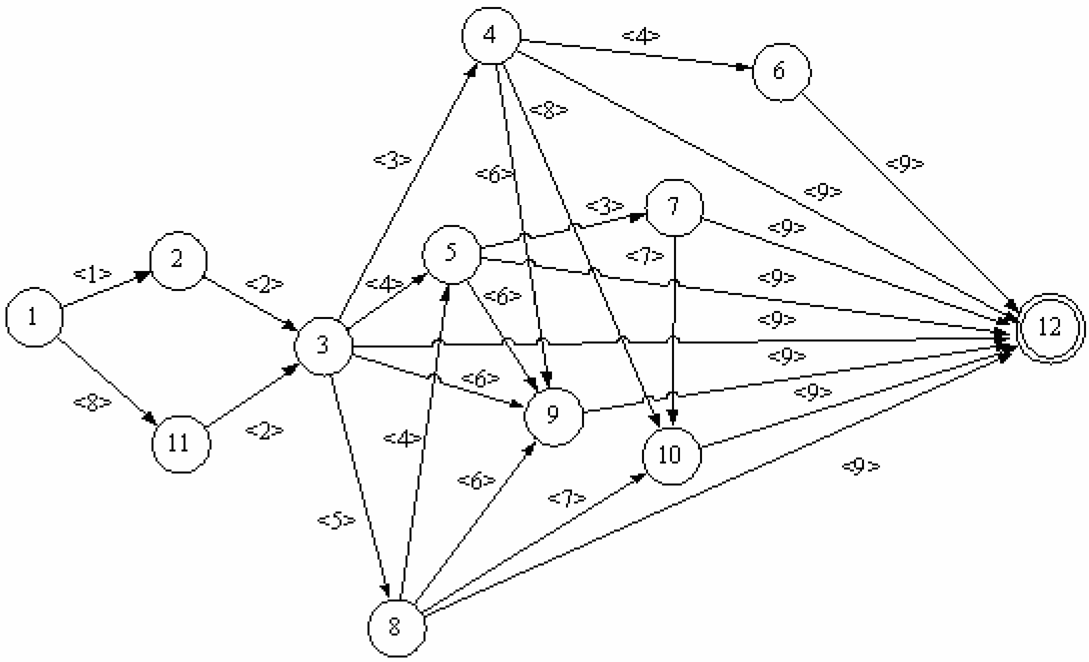
A little semantic information is obtained from the paths through the phrase structure tree, which correspond to the grammar rules selected to parse the sentence. After the syntactic analysis is completed, we get information about the structure of the sentence that is useful for determining the syntactic pattern of the SVCs in the sentence. Based on the structures of Thai sentences with SVCs, Figure 4.6 presents a diagram of all syntactic parts of sentences in which the SVC is the verb in main clause. The diagram in Figure 4.7 expresses all syntactic parts of subordinate clauses in which the SVC occurs as the verb.




<1> refer: noun phrase which is a subject of a sentence
<2> refer: verb phrase which is the verb of main clause in a sentence
<3> refer: prepositional phrase
<4> refer: adverb phrase
<5> refer: noun phrase which is an object of a sentence
<6> refer: subordinating clause
<7> refer: coordinating clause
<8> refer: relative clause
<9> refer: clause which is a subject of a sentence
<10> refer: empty string

Figure 4.6: Diagram of sentences in which SVCs are found as verbs in main clauses.

Referring to Figure 4.6, each circle with a number represents each state during parsing each part of a sentence. State transitions take place, when phrases or clauses are found, that are represented by an edge in the diagram. The paths in the diagram present all possible phrases or clauses in the sentence that appear in sentences containing SVCs in the verb phrase.




<1> refer: relative pronoun
<2> refer: verb phrase which is a verb of a clause
<3> refer: prepositional phrase
<4> refer: adverb phrase
<5> refer: noun phrase which is an object of a clause
<6> refer: subordinating clause
<7> refer: relative clause
<8> refer: subordinating conjunction
<9> refer: empty string of a clause

Figure 4.7: Clause diagram of all SVCs functioning as verbs in subordinate clauses.

Referring to Figure 4.7, each circle with a number represents a state reached during parsing a part of a clause. State transitions take place, when phrases or clauses are found in clauses represented by an edge in the diagram. The paths through the diagram represent all possible phrases or clauses appearing in clauses with an SVC as the verb.

## 4.3 Semantic Analysis

Since the meaning of a word usually depends on its position which is analyzed in the syntactic analysis phase, outputs from the syntactic analysis phase will be inputs for the semantic analysis phase. Thus, the knowledge about serial verbs can be extracted from our proposed framework.

The proposed mechanism for semantic analysis of components is composed of four processes. The first and the second processes are identify the words surrounding the SVCs and analyze the relationships between the words and the SVCs; semantic relations between SVCs and other parts of the sentence structures will be generated. The next step is to define classes of words related to SVCs. The last step is to analyze all features to determine the semantic bindings. Figure 4.8 shows the processes that perform the semantic analysis of sentences with serial verbs.

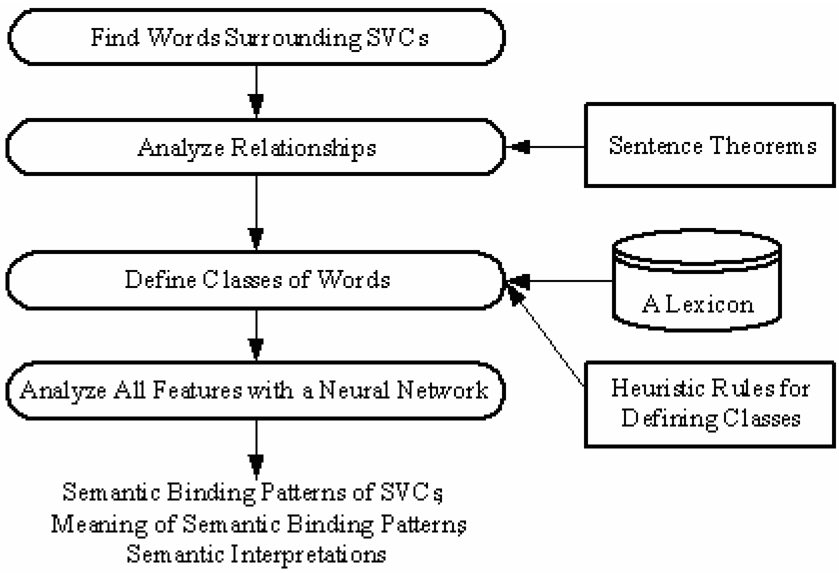



Figure 4.8: Semantic Analysis.

Referring to Figure 4.8, finding words surrounding SVCs and relationships analysis are methods to get context information about two consecutive SVCs. Then, the related words of SVCs are revealed by proposed sentence theorems. Next, the classes of related words are defined to add information about the SVCs. Finally, all relations and classes of related words are analyzed and learned to generate the semantic binding patterns of SVCs and the meaning of those some patterns. Details of these modules are described in the following section.

### 4.3.1 Modules for Finding Words Surrounding SVCs and for the Analysis of the Relationships

After receiving words and syntactic patterns from the syntactic analysis phase, words surrounding SVCs are identified by traversing branches of the phrase structure trees from the nodes that are SVCs in the verb phrases. For example, a noun phrase that has been occurred before SVCs in a sentence is the subject of the SVCs and a noun phrase which has been occurred after the SVCs in the series of word in a sentence is the object of the SVCs.

Once the surrounding words are recognized, the next process is to find the relationships between these words and the serial verbs. The syntactic patterns and definitions with theorems are described in Appendix C. With regard to sentence theorems, the algorithms for relationship analysis are shown in Figures 4.9 and 4.10. Whether SVCs are verbs in main clauses or in subordinate clauses, all syntactic patterns and relations are included in Table 4.1 with results of the relationship analysis SVCs and the other words in the phrases that are related to SVCs.

```
READ first_part
pattern = <NULL>
relation = <NULL>
IF first_part = NP
THEN
  pattern = <NP >
  READ next_part
  IF next_part = VP
  THEN
    pattern = <NP, SV>
    relation = <SUBJECT1>
    READ next_part
    IF next_part = NP
    THEN
      pattern = <NP, SV, NP>
      relation = <SUBJECT1, OBJECT1>
      READ next_part
      IF next_part = NULL
      THEN
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause.

```
    OUTPUT pattern
    OUTPUT relation
  ELSE IF next_part = ADVP
  THEN
    pattern = <NP, SV, NP, ADVP>
    relation = <SUBJECT1, OBJECT1, MODIFYV2>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    ELSE IF next_part = PP
    THEN
      pattern = <NP, SV, NP, ADVP, PP>
      relation = <SUBJECT1, OBJECT1, MODIFYV2, MODIFYV1 >
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = CL_SUB
      THEN
        pattern = <NP, SV, NP, ADVP, PP, CL_SUB>
        relation = <SUBJECT1, OBJECT1, MODIFYV2, MODIFYV1,
COMPLETE2 >
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    ELSE IF next_part = CL_SUB
    THEN
      pattern = <NP, SV, NP, ADVP, CL_SUB>
      relation = <SUBJECT1, OBJECT1, MODIFYV2, COMPLETE2 >
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      END IF
    END IF
  ELSE IF next_part = CL_SUB
  THEN
    pattern = <NP, SV, NP, CL_SUB >
    relation = <SUBJECT1, OBJECT1, COMPLETE2>
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  ELSE IF next_part = CL_REL
  THEN
    pattern = <NP, SV, NP, CL_REL >
    relation = <SUBJECT1, OBJECT1, MODIFYN2>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  ELSE IF next_part = CL_COR
  THEN
    pattern = <NP, SV, NP, CL_COR >
    relation = <SUBJECT1, OBJECT1, JOIN>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  END IF
ELSE IF next_part = PP
THEN
  pattern = <NP, SV, PP>
  relation = <SUBJECT1, MODIFYV1>
  READ next_part
  IF next_part = NULL
  THEN
    OUTPUT pattern
    OUTPUT relation
  ELSE IF next_part = ADVP
  THEN
    pattern = <NP, SV, PP, ADVP>
    relation = <SUBJECT1, MODIFYV1, MODIFYV2>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  ELSE IF next_part = CL_REL
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
      THEN
        pattern = <NP, SV, PP, CL_REL>
        relation = <SUBJECT1, MODIFYV1, MODIFYN3>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      ELSE IF next_part = CL_SUB
      THEN
        pattern = <NP, SV, PP, CL_SUB>
        relation = <SUBJECT1, MODIFYV1, COMPLETE1>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    ELSE IF next_part = ADVP
    THEN
      pattern = <NP, SV, ADVP>
      relation = <SUBJECT1, MODIFYV2>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = PP
      THEN
        pattern = <NP, SV, ADVP, PP>
        relation = <SUBJECT1, MODIFYV2, MODIFYV1>
        READ next_part
        IF next_part = NULL
          THEN
            OUTPUT pattern
            OUTPUT relation
        END IF
      ELSE IF next_part = CL_SUB
      THEN
        pattern = <NP, SV, ADVP, CL_SUB >
        relation = <SUBJECT1, MODIFYV2, COMPLETE1>
        READ next_part
        IF next_part = NULL
        THEN
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
        OUTPUT pattern
        OUTPUT relation
      END IF
    END IF
  ELSE IF next_part = CL_SUB
  THEN
    pattern = <NP, SV, CL_SUB >
    relation = <SUBJECT1, COMPLETE1>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  END IF
ELSE IF next_part = CL_REL
THEN
  pattern = <NP, CL_REL>
  READ next_part
  IF next_part = VP
  THEN
    pattern = <NP, CL_REL, SV>
    relation = <SUBJECT1, MODIFYN1>
    READ next_part
    IF next_part = NP
    THEN
      pattern = <NP, CL_REL, SV, NP>
      relation = <SUBJECT1, MODIFYN1, OBJECT1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = ADVP
      THEN
        pattern = <NP, CL_REL, SV, NP, ADVP>
        relation = <SUBJECT1, MODIFYN1, OBJECT1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        ELSE IF next_part = CL_SUB
        THEN
          pattern = <NP, CL_REL, SV, NP, ADVP, CL_SUB>
          relation = <SUBJECT1, MODIFYN1, OBJECT1, MODIFYV2,
COMPLETE2>
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
            READ next_part
            IF next_part = NULL
            THEN
              OUTPUT pattern
              OUTPUT relation
            END IF
          END IF
        ELSE IF next_part = CL_REL
        THEN
          pattern = <NP, CL_REL, SV, NP, CL_REL>
          relation = <SUBJECT1, MODIFYN1, OBJECT1, MODIFYN2>
          READ next_part
          IF next_part = NULL
          THEN
            OUTPUT pattern
            OUTPUT relation
          END IF
        ELSE IF next_part = CL_SUB
        THEN
          pattern = <NP, CL_REL, SV, NP, CL_SUB>
          relation = <SUBJECT1, MODIFYN1, OBJECT1, COMPLETE2>
          READ next_part
          IF next_part = NULL
          THEN
            OUTPUT pattern
            OUTPUT relation
          END IF
        END IF
      ELSE IF next_part = PP
      THEN
        pattern = <NP, CL_REL, SV, PP>
        relation = <SUBJECT1, MODIFYN1, MODIFYV1>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      ELSE IF next_part = ADVP
      THEN
        pattern = <NP, CL_REL, SV, SDVP>
        relation = <SUBJECT1, MODIFYN1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
        END IF
      ELSE IF next_part = CL_SUB
      THEN
        pattern = <NP, CL_REL, SV, CL_SUB>
        relation = <SUBJECT1, MODIFYN1, COMPLETE1>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    END IF
  END IF
ELSE IF next_part = CL_N
THEN
  pattern = <CL_N>
  READ next_part
  IF next_part = VP
  THEN
    pattern = <CL_N, SV>
    relation = <SUBJECT_CL1>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    ELSE IF next_part = NP
    THEN
      pattern = <CL_N, SV, NP>
      relation = <SUBJECT_CL1, OBJECT1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = ADVP
      THEN
        pattern = <CL_N, SV, NP, ADVP>
        relation = <SUBJECT_CL1, OBJECT1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      ELSE IF next_part = CL_SUB
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
      THEN
        pattern = <CL_N, SV, NP, CL_SUB>
        relation = <SUBJECT_CL1, OBJECT1, COMPLETE2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    ELSE IF next_part = PP
    THEN
      pattern = <CL_N, SV, PP>
      relation = <SUBJECT_CL1, MODIFYV1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      END IF
    ELSE IF next_part = ADVP
    THEN
      pattern = <CL_N, SV, ADVP>
      relation = <SUBJECT_CL1, MODIFYV2>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      END IF
    END IF
  ELSE IF next_part = CL_REL
  THEN
    pattern = <CL_N, CL_REL>
    READ next_part
    IF next_part = VP
    THEN
      pattern = <CL_N, CL_REL, SV>
      relation = <SUBJECT_CL1, MODIFYN_CL1>
      READ next_part
      IF next_part = NP
      THEN
        pattern = <CL_N, CL_REL, SV, NP>
        relation = <SUBJECT_CL1, MODIFYN_CL1, OBJECT1>
        READ next_part
        IF next_part = NULL
        THEN
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
          OUTPUT pattern
          OUTPUT relation
        ELSE IF next_part = CL_SUB
        THEN
          pattern = <CL_N, CL_REL, SV, NP, CL_SUB>
          relation = <SUBJECT_CL1, MODIFYN_CL1, OBJECT1, COMPLETE2>
          READ next_part
          IF next_part = NULL
          THEN
            OUTPUT pattern
            OUTPUT relation
          END IF
        END IF
      ELSE IF next_part = PP
      THEN
        pattern = <CL_N, CL_REL, SV, PP>
        relation = <SUBJECT_CL1, MODIFYN_CL1, MODIFYV1>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    END IF
  END IF
END IF
```

Figure 4.9: Analysis of syntactic patterns and relations in sentences containing with an SVC as verb in main clause (cont.).

```
READ first_part
pattern = <NULL>
relation = <NULL>
IF first_part = PR_REL
THEN
  pattern = <NP_REL, PR_REL >
  READ next_part
  IF next_part = VP
  THEN
    pattern = <NP_REL, PR_REL, SV>
    relation = <SUBJECT_REL1>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    ELSE IF next_part = NP
    THEN
      pattern = <NP_REL, PR_REL, SV, NP>
      relation = <SUBJECT_REL1, OBJECT1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = ADVP
        pattern = <NP_REL, PR_REL, SV, NP, ADVP>
        relation = <SUBJECT_REL1, OBJECT1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        ELSE IF next_part = PP
          pattern = <NP_REL, PR_REL, SV, NP, ADVP, PP>
          relation = <SUBJECT_REL1, OBJECT1, MODIFYV2, MODIFYV1>
          READ next_part
          IF next_part = NULL
          THEN
            OUTPUT pattern
            OUTPUT relation
          END IF
        END IF
      ELSE IF next_part = CL_SUB
        pattern = <NP_REL, PR_REL, SV, NP, CL_SUB>
        relation = <SUBJECT_REL1, OBJECT1, COMPLETE2>
        READ next_part
```

Figure 4.10: Analysis of syntactic patterns and relations in subordinate clauses with an SVC as verb.

```
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  ELSE IF next_part = CL_REL
    pattern = <NP_REL, PR_REL, SV, NP, CL_REL>
    relation = <SUBJECT_REL1, OBJECT1, MODIFYV2>
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  END IF
ELSE IF next_part = ADVP
THEN
    READ next_part
    IF next_part = NULL
    THEN
      OUTPUT pattern
      OUTPUT relation
    END IF
  END IF
ELSE IF next_part = CL_SUB
THEN
  pattern = <NP_REL, PR_REL, SV, CL_SUB>
  relation = <SUBJECT_REL1, COMPLETE1>
  READ next_part
  IF next_part = NULL
  THEN
    OUTPUT pattern
    OUTPUT relation
  END IF
  pattern = <NP_REL, PR_REL, SV, ADVP>
  relation = <SUBJECT_REL1, MODIFYV2>
  READ next_part
  IF next_part = NULL
  THEN
    OUTPUT pattern
    OUTPUT relation
  ELSE IF next_part = PP
  THEN
    pattern = <NP_REL, PR_REL, SV, ADVP, PP>
    relation = <SUBJECT_REL1, MODIFYV2, MODIFYV1>
    READ next_part
```

Figure 4.10: Analysis of syntactic patterns and relations in subordinate clauses with an SVC as verb (cont.).

```
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    ELSE IF next_part = PP
    THEN
      pattern = <NP_REL, PR_REL, SV, PP>
      relation = <SUBJECT_REL1, MODIFYV1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = ADVP
      THEN
        pattern = <NP_REL, PR_REL, SV, PP, ADVP>
        relation = <SUBJECT_REL1, MODIFYV1, MODIFYV2>
    END IF
  END IF
ELSE IF first_part = C_SUB
THEN
  READ next_part
  IF next_part = VP
  THEN
    pattern = <C_SUB, SV>
    READ next_part
    IF next_part = NP
    THEN
      pattern = <C_SUB, SV, NP>
      relation =  <OBJECT1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = ADVP
        pattern = <C_SUB, SV, NP, ADVP>
        relation = <OBJECT1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      ELSE IF next_part = CL_REL
```

Figure 4.10: Analysis of syntactic patterns and relations in subordinate clauses with an SVC as verb (cont.).

```
        pattern = <C_SUB, SV, NP, CL_REL>
        relation = <OBJECT1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    ELSE IF next_part = ADVP
    THEN
      pattern = <C_SUB, SV, ADVP>
      relation = <MODIFYV2>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      END IF
    ELSE IF next_part = PP
    THEN
      pattern = <C_SUB, SV, PP>
      relation = <MODIFYV1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = CL_REL
      THEN
        pattern = <C_SUB, SV, PP, CL_REL >
        relation = <MODIFYV1, MODIFYN3>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    END IF
ELSE IF first_part = NP
THEN
  pattern = <NP >
  READ next_part
  IF next_part = VP
  THEN
```

Figure 4.10: Analysis of syntactic patterns and relations in subordinate clauses with an SVC as verb (cont.).

```
    pattern = <NP, SV>
    relation = <SUBJECT1>
    READ next_part
    IF next_part = NP
    THEN
      pattern = <NP, SV, NP>
      relation = <SUBJECT1, OBJECT1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = ADVP
      THEN
        pattern = <NP, SV, NP, ADVP>
        relation = <SUBJECT1, OBJECT1, MODIFYV2>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    ELSE IF next_part = PP
    THEN
      pattern = <NP, SV, PP>
      relation = <SUBJECT1, MODIFYV1>
      READ next_part
      IF next_part = NULL
      THEN
        OUTPUT pattern
        OUTPUT relation
      ELSE IF next_part = CL_REL
      THEN
        pattern = <NP, SV, PP, CL_REL >
        relation = <SUBJECT1, MODIFYV1, MODIFYN3>
        READ next_part
        IF next_part = NULL
        THEN
          OUTPUT pattern
          OUTPUT relation
        END IF
      END IF
    END IF
  END IF
END IF
```

Figure 4.10: Analysis of syntactic patterns and relations in subordinate clauses with an SVC as verb (cont.).

Table 4.1: Theorems, syntactic patterns and relations.

| Theorems | Syntactic Patterns | SUBJECT1 | OBJECT1 | MODIFYV1 | MODIFYV2 | COMPLETE1 | COMPLETE2 | JOIN1 | SUBJECT_CL | MODIFYN1 | MODIFYN2 | MODIFYN3 | MODIFYN_CL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | S:= <NP, SV, NP> | x | x | | | | | | | | | | |
| 2 | S := <NP, SV, PP> | x | | x | | | | | | | | | |
| 3 | S := <NP, SV, ADVP> | x | | | x | | | | | | | | |
| 4 | S := <NP, SV, ADVP,PP> | x | | x | x | | | | | | | | |
| 5 | S := <NP, SV, PP, ADVP> | x | | x | x | | | | | | | | |
| 6 | S := <NP, SV, NP, ADVP> | x | x | | x | | | | | | | | |
| 7 | S := <NP, SV, NP, ADVP, PP> | x | x | x | x | | | | | | | | |
| 8 | S := <NP, SV, CL_SUB> | x | | | | x | | | | | | | |
| 9 | S := <NP, SV, NP, CL_SUB> | x | x | | | | x | | | | | | |
| 10 | S := <NP, SV, NP, COORDCL> | x | x | | | | | x | | | | | |
| 11 | S := <NP, SV, NP, CL_REL> | x | x | | | | | | | | x | | |
| 12 | S := <NP, SV, PP, CL_SUB> | x | | x | | x | | | | | | | |
| 13 | S := <NP, SV, PP, CL_REL> | x | | x | | | | | | | | x | |
| 14 | S := <NP, SV, ADVP, CL_SUB> | x | | | x | x | | | | | | | |
| 15 | S := <NP, SV, NP, ADVP, CL_SUB> | x | x | | x | | x | | | | | | |
| 16 | S := <NP, SV, NP, ADVP_PP, CL_REL> | x | x | x | x | | | | | | | x | |
| 17 | S := <NP, CL_REL, SV, NP> | x | x | | | | | | | x | | | |
| 18 | S := <NP, CL_REL, SV, PP> | x | | x | | | | | | x | | | |
| 19 | S := <NP, CL_REL, SV, ADVP> | x | | | x | | | | | x | | | |
| 20 | S := <NP, CL_REL, SV, NP, ADVP> | x | x | | x | | | | | x | | | |
| 21 | S := <NP, CL_REL, SV, CL_SUB> | x | | | | x | | | | x | | | |
| 22 | S := <NP, CL_REL, SV, NP, CL_SUB> | x | x | | | | x | | | x | | | |
| 23 | S := <NP, CL_REL, SV, NP, CL_REL> | x | x | | | | | | | x | x | | |
| 24 | S := <NP, CL_REL, SV, NP, ADVP, CL_SUB> | x | x | | x | | x | | | x | | | |
| 25 | S := <CL_N, SV> | | | | | | | | x | | | | |
| 26 | S := <CL_N, SV, NP> | | x | | | | | | x | | | | |
| 27 | S := <CL_N, SV, PP> | | | x | | | | | x | | | | |

Table 4.1: Theorems, syntactic patterns and relations (cont.).

| Theorems | Syntactic Patterns | SUBJE<br>CT1 | OBJE<br>CT1 | MODI<br>FYV1 | MODI<br>FYV2 | COMPL<br>ETE1 | COMPL<br>ETE2 | JOIN1 | SUBJEC<br>T_CL | MODI<br>FYN1 | MODI<br>FYN2 | MODI<br>FYN3 | MODIF<br>YN_CL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 28 | S := <CL_N, SV, ADVP> | | | | x | | | | x | | | | |
| 29 | S := <CL_N, SV, NP, ADVP> | | x | | x | | | | x | | | | |
| 30 | S := <CL_N, SV , NP, CL_SUB> | | x | | | | x | | x | | | | |
| 31 | S := <CL_N, CL_REL, SV ,NP> | | x | | | | | | x | | | | x |
| 32 | S := <CL_N, CL_REL, SV ,PP> | | | x | | | | | x | | | | x |
| 33 | S := <CL_N, CL_REL, SV ,NP, CL_SUB> | | x | | | | x | | x | | | | x |
| 35 | C := <$NP_{REL}$, PR_REL, SV, NP> | x | x | | | | | | | | | | |
| 36 | C := <$NP_{REL}$, PR_REL, SV> | x | | | | | | | | | | | |
| 37 | C := <$NP_{REL}$, PR_REL, SV, PP> | x | | x | | | | | | | | | |
| 38 | C := <$NP_{REL}$, PR_REL, SV, ADVP> | x | | | x | | | | | | | | |
| 39 | C := <$NP_{REL}$, PR_REL, SV, ADVP, PP> | x | | x | x | | | | | | | | |
| 40 | C := <$NP_{REL}$, PR_REL, SV, PP, ADVP> | x | | x | x | | | | | | | | |
| 41 | C := <$NP_{REL}$, PR_REL, SV, NP, ADVP> | x | x | | x | | | | | | | | |
| 42 | C := <$NP_{REL}$, PR_REL, SV, NP, ADVP_PP> | x | x | x | x | | | | | | | | |
| 43 | C := <$NP_{REL}$,REL,SV, CL_SUB, RELCL> | x | | | | x | | | | | | | |
| 44 | C := <$NP_{REL}$, PR_REL, SV, NP, CL_SUB> | x | x | | | | x | | | | | | |
| 45 | C := <$NP_{REL}$, PR_REL, SV, NP, RELCL> | x | x | | | | | | | | x | | |
| 46 | C := <C_SUB, SV, NP> | | x | | | | | | | | | | |
| 47 | C := <C_SUB, SV, PP> | | | x | | | | | | | | | |
| 48 | C := <C_SUB, SV, ADVP_PP> | | | x | x | | | | | | | | |
| 49 | C := <C_SUB, SV, NP, ADVP> | | x | | x | | | | | | | | |
| 50 | C := <C_SUB, SV, NP, RELCL> | | x | | | | | | | | x | | |
| 51 | C := <C_SUB, SV, PP, RELCL> | | | x | | | | | | | | x | |
| 52 | C := <NP, SV, NP> | x | x | | | | | | | | | | |
| 53 | C := <NP, SV, PP> | x | | x | | | | | | | | | |
| 54 | C := <NP, SV, NP, ADVP> | x | x | | x | | | | | | | | |
| 55 | C := <NP, SV, PP, RELCL> | x | | x | | | | | | | | x | |

### 4.3.2 Defining Classes of Words Module

After completion of the process, that finds words surrounding SVCs and analyzes their relationships to the SVC process, the surrounding words and relations related to SVCs have been discovered in a sentence. Then, the class of each surrounding word is determined in defining phases. The first noun of the noun phrase that has been occurred before an SVC in a sentence is labeled as a head noun to define classes. Some word information from a lexicon and some heuristic rules that we designed are used to define classes of head nouns in noun phrases, classes of adverb phrases, and classes of prepositional phrases. Domains or types of nouns are categorized into various classes, such as the person class of nouns referring to persons or organizations (e.g., users, employees, officers), the artifact class referring to artifacts or instruments (e.g., computers, programs, software, instruments), the data class referring to data or information, the activity class referring to activities or processes and the abstract class referring to abstract nouns. The characteristics of each class were chosen and also simplified classes from the semantic hierarchy of noun classes in (Scott, 2003) and noun sets in (TCLLEX, 2005). Figure 4.11 shows the hierarchy of noun classes that we defined for our research. To define classes for adverb and prepositional phrases, we chose beginning words of phrases, (an adverb and a preposition, respectively), used as keywords for classes. For example, “at” and “in” are used to express the location of the action in sentences, while “with” is used to express that the action in a sentence occurs in co-operation with nouns following the preposition.

Concrete Thing
- Living Thing
    - Person
    - Organization
    - Animal
- Non-living Thing
    - Artifact
        - Parts of Artifact
    - Non-artifact

Abstraction
- Activity & Process
- Data & Information (Value, State, Attribute, Characteristics)
- Technology & Knowledge
- Abstract Thing

Figure 4.11: Hierarchy of noun classes.

There are three approaches to define classes of words as shown in Figure 4.12. The first approach cis to to test whether the word satisfies a defined rules. The second approach is to compare words with other words in the database, including matching a substring of words. The final approach is to let user define the class manually.

```
STEP 1: Define by rules
IF There are no rules for defining
THEN   STEP 2: Compare with words in database
       IF   There are no similar words
       THEN STEP 3: Define manually
       END IF
END IF
```

Figure 4.12: The basic steps for defining classes of words

According to the hierarchy of noun classes (Figure 4.11), there are seven different noun classes in noun phrases that are Person or Organization (PER), Artifact & Parts of Artifacts (ART), Non Artifact (NAR), Activity & Process (ACT), Data & Information & Value & Number (DAT), Technology & Knowledge (TEC), and Abstract Thing (ABT). Rules for classification in the noun class are shown in Figure 4.13.

1) PER is a set of nouns that are used to refer to people, groups of people, or organizations. The example of this set is the proper noun or the common noun that means names of persons, groups of people, names of organizations, names of projects and the personal pronoun (PR).

2) ART is a set of nouns that are used to refer to artifacts and the part of them. Examples of this set are common nouns for kinds of equipments, instruments, tools, programs, systems, computers, database systems and their parts.

3) NAR is a set of nouns that are used to refer to concrete things that are not in the artifact classes. Examples of this set are common nouns for rice, flowers and some products.

4) ACT is a set of nouns that are used to refer to activities. Examples of this set are the common nouns that mean activities, processes and work.

5) DAT is a set of nouns that are used to refer to data and information, including values, states and attributes of something. Examples of this set are common

nouns for the age of a person, telephone number, the number of pages in this thesis, information, signals, results, reports, characteristics, features, and qualities.

6) TEC is a set of nouns that are used to refer to technology and knowledge. Examples of this set are the common nouns that mean principles, policy, methods, models, rules, research and fields of knowledge.

7) ABT is a set of nouns that are used to refer to abstract things that are not in the other noun classes. Examples of this set are common nouns like phenomenon, event and relation.

```
word ∈ DAT
IF word ∈ PR OR word = “หน่วยงาน”, “องค์กร”, “โครงการ”, “คณะ”, “เจ้าหน้าที่”, “บริษัท”,
“รัฐ” OR first substring = “ผู้”, “นัก”, “หน่วย”, “ภาค”
THEN word ∈ PER
ElSE IF word = “ซอฟต์แวร์”, “โปรแกรม”, “ระบบ”, “คอมพิวเตอร์”, “อุปกรณ์”, “ฐานข้อมูล”
THEN word ∈ ART
ElSE IF first substring = “การ”, “งาน”
THEN word ∈ ACT
ELSE IF word = “ข้อมูล”, “สัญญาณ”, “รหัส”, “ความถี่”, “ภาพ”
THEN word ∈ DAT
ElSE IF word = ““วิธี”, “หลักการ”, “นโยบาย”, “กฎ”, “โครงสร้าง”
THEN word ∈ TEC
ElSE IF first substring = “ความ” OR word = “ปรากฏการณ์”
THEN word ∈ ABT
ELSE To compare with database
END IF
```

Figure 4.13: Rules for noun classification.

The prepositional phrases are classified into five classes by the beginning words of phrases: Location or Place, Coordination, Purpose, Means, and Source or Destination. Rules for classification of the prepositions are shown in Figure 4.14.

1) Location or Place is the set of prepositions introducing phrases that specify positions or places, i.e., they appear at the beginning of nouns that refer to locations or places. Examples of this set are “ใน/nai/”, “ที่/thi/”, “บน/bon/”.

2) Coordination is the set of prepositions discovering that there is a connection to the following noun. Examples of this set are ”กับ/kap/”, “เข้ากับ/khaokap/”, “ร่วมกับ/ruamkap/”, “เกี่ยวกับ/kiaokap/”, “ร่วมด้วย/ruamduai/”.

3) Purpose is the set of prepositions of purpose. One Example of this set is “สำหรับ/samrap/”.

4) Mean is the set of prepositions of means. Examples of this set are “โดย/doi/”, “ด้วย/duai/”, “ตาม/tam/”, “เฉพาะ/chapho/”.

5) Source or Destination is the set of prepositions of direction that appear at the beginning of nouns that refer to sources or destinations. Examples of this set are “จาก/chak/”, “ยัง/yang/”, “ให้/hai/”, “สู่/su/”.

```
word ∈ Mean
IF word = “ใน”, “ที่”, “บน”
THEN word ∈ Location
ElSE IF word = “จาก”, “ให้”, “สู่”, “ยัง”
THEN word ∈ Source or Destination
ElSE IF word = “สำหรับ”
THEN word ∈ Purpose
ElSE IF substring = “กับ”, “ร่วม”
THEN word ∈ Coordination
ELSE IF word = “โดย”, “ด้วย”, “ตาม”, “เฉพาะ”
THEN word ∈ Means
ELSE To compare with database
END IF
```

**Figure 4.14: Rules for preposition classification.**

The set of adverb phrases is divided into four classes by beginning words of phrases that signify Manner, Cooperating or Separating, Time or Frequency, and Source or Destination. Rules for classification of the adverb are shown in Figure 4.15.

1) Manner is the set of adverbs of manner. Examples of this set are classified into classes by words which are found in our data on SVCs.

– The first word of adverb phrases of manner may be “ได้/dai/”, such as “ได้มาก/daimak/”, “ได้ดีมาก/daidimak/”.

– The first word of adverb phrases of manner may be “ให้/hai/”, such as “ให้ถูกต้อง/haithuktong/”, “ให้ละเอียด/haila-iat/”, “ให้ทันสมัย /thansamai/”.

– The first word of adverb phrases of manner may be “เป็น/ pen/”, such as “เป็นหลัก/penlak/”.

– The first word of adverb phrases of manner may be the adverbial prefix (FIXV) such as “อย่างกว้างขวาง/yangkwangkhwang /” , “อย่างอิสระ/yangitsara/”, “อย่างแท้จริง/yangthaeching/”), “ด้วยดี/duaidi/”, “โดยตรง/doitrong/”.

2) Cooperating or Separating is the set of adverbs that refer to cooperating or separating. Examples of this set are “เข้าด้วยกัน/khaoduaikan/”, “ร่วมกัน/ruamkan/”, “ออก/ok/”, “ออกจากกัน/okchakkan/”.

3) Time or Frequency is a set of adverbs of times and adverbs of frequency. Examples of this set are “ตลอดเวลา/talotwela/”, “อยู่/yu/”, “ในเวลาเดียวกัน/naiweladiaokan/”, “ระยะสั้น/rayasan/”, “อีกครั้ง/ikkrang/”, “อีกทีหนึ่ง/ikthinueng/”, “บ่อย/boi/”, “ครั้งสุดท้าย/khrangsutthai/”.

4) Quantity is a set of adverbs of degree that refer to quantity. Examples of this set are “เร็วขึ้น/reokhuen/”, “ลงได้มาก/longdaimak/”, “มากกว่า 10 เท่า/makkwa 10 thao/”, “ลงมา 10 เท่า/longma 10 thao/”, “ให้มากขึ้น/haimakkhuen/”.

```
word ∈ Manner
IF first substring  = “ได้”, “ให้”, “เป็น” OR first substring  ∈  FIXV
THEN  word ∈ Manner
ELSE IF substring  = “เวลา” , “อยู่” , “บ่อย”  OR first substring  =  “อีก”, “ครั้ง”, “ระยะ”
THEN word  ∈  Time or Frequency
ElSE IF first substring  =  “เข้า”, “ร่วม”, “ออก”
THEN word  ∈  Cooperating or Separating
ElSE IF substring  =  “ขึ้น” ,“ลง” OR  first substring  =  “มาก” “น้อย”
THEN word  ∈  Quantity
ELSE To compare with database
END IF
```

Figure 4.15: Rules for adverb classification.

### 4.3.4 Analysis by a Neural Network

To analyze semantic binding patterns with a machine learning technique, all information from the pervious processes is analyzed by a neural network method. Our research uses the probabilistic neural network function called NEWPNN of MATLAB (MATLAB, 2008). The algorithms of NEWPNN are described in Figure 4.16.

```
%	NEWPNN creates a two layer network. The first layer has RADBAS
%	neurons, and calculates its weighted inputs with DIST, and its net
%	input with NETPROD.  The second layer has COMPET neurons, and
%	calculates its weighted input with DOTPROD and its net inputs with
%	NETSUM. Only the first layer has biases.
%
%	NEWPNN sets the first layer weights to P', and the first
%	layer biases are all set to 0.8326/SPREAD resulting in
%	radial basis functions that cross 0.5 at weighted inputs
%	of +/- SPREAD. The second layer weights W2 are set to T.

load ‘Training data’;
load ‘Target data’;
input = ‘Training data’;
Tc = ‘Target data’;
T = ind2vec(Tc);
spread =1;
net = newpnn(input,T,spread);
Y = sim(net,input);
Yc = vec2ind(Y);
Test = sim(net, ‘Test data’);
‘Answer data’ = vec2ind(Test);
```

Figure 4.16: Algorithms of NEWPNN (MATLAB, 2008).

With regarding to Figure 4.16, there are two layers of the probabilistic neural networks in this research. When an input is presented, the first layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, the complete transfer function on the output of the second layer picks the maximum of these probabilities.

The semantic binding patterns can discover the patterns of two consecutive SVCs. The examples of information are the syntactic features of words in sentences, e.g. part of speech of words, roles in sentences, and structures of sentences, e.g. SVO. Classes of words that are related to SVCs, such as nouns in noun phrases, prepositions in prepositional phrases, and adverbs in adverb phrases are features used to learn semantic interpretation of SVCs. The details of the semantic binding patterns are described in the next section.

### 4.4 Semantic Binding Patterns of SVCs and Semantic Interpretations

Semantic binding patterns in this research are produced by the binding analysis. This analysis is separated into three parts. The first part is to classify semantic binding patterns based on only syntactic structures and relationship analysis. The patterns present each phrase which is connected with SVCs. The second part and the third part are to classify semantic interpretations and some meaning of semantic binding patterns based on all features including syntactic structures, relations from relationship analysis and surrounding word classes. The semantic binding patters are shown in Figure 4.17 and Figure 4.18 that are patterns of SVCs for verbs in main clauses and for verbs in clauses, respectively.

Subject: NP
Verb: SV
Modifyv1: PP

Subject: NP
Verb: SV
Modifyv2: ADVP

Subject: NP
Verb: SV
Modifyv1: PP
Modifyv2: ADVP

Subject: NP
Verb: SV
Modifyv2: ADVP
Modifyv1: PP

Subject: NP
Verb: SV
Complete1: Subordinating Clause

Subject: NP
Verb: SV
Join1: Coordinating Clause

Subject: NP
Verb: SV
Modifyv1: PP
Complete1: Subordinating Clause

Subject: NP
Verb: SV
Modifyv2: ADVP
Complete1: Subordinating Clause

Subject NP
Verb: SV
Object NP

Subject NP
Verb: SV
Object NP
Modifyv2: ADVP

Subject NP
Verb: SV
Object NP
Modifyv2: ADVP
Modifyv1: PP

Subject NP
Verb: SV
Object NP
Complete2: Subordinating Clause

Subject NP
Verb: SV
Object NP
Modifyv2: ADVP
Complete2: Subordinating Clause

Subject_cl: Noun Clause
Verb: SV

Subject_cl: Noun Clause
Verb: SV
Object NP

Subject_cl: Noun Clause
Verb: SV
Modifyv1: PP

Subject_cl: Noun Clause
Verb: SV
Modifyv2: ADVP

Subject_cl: Noun Clause
Verb: SV
Object NP
Modifyv2: ADVP

Subject_cl: Noun Clause
Verb: SV
Object NP
Modifyv2: ADVP
Complete2: Subordinating Clause

Figure 4.17: Semantic binding patterns in main clauses.

Subject: NP
Verb: SV

Subject: NP
Verb: SV
Modifyv1 : PP

Subject: NP
Verb: SV
Modifyv2: ADVP

Subject: NP
Verb: SV
Modifyv2: ADVP
Modifyv1 : PP

Subject: NP
Verb: SV
Modifyv1 : PP
Modifyv2: ADVP

Subject: NP
Verb: SV
Complete1 : Subordinating Clause

Subject NP
Verb: SV
Object NP

Subject NP
Verb: SV
Object NP
Modifyv2 : ADVP

Subject NP
Verb: SV
Object NP
Modifyv2 : ADVP
Modifyv1 : PP

Subject NP
Verb: SV
Object NP
Complete2: Subordinating Clause

Subject -
Verb: SV
Object NP

Subject -
Verb: SV
Modifyv1 : PP

Subject -
Verb: SV
Modifyv2: ADVP
Modifyv1 : PP

Subject -
Verb: SV
Object NP
Modifyv2: ADVP

**Figure 4.18: Semantic binding patterns in clauses.**

According to Figure 4.17, one example of semantic binding patterns is “Subject[NP], Verb[SV], Object[NP]” means that the action is expressed by a SVC and the subject is an agent of this action while the object is a patient of this action. The subject is also an instrument of this action in some sentences. Figure 4.19 shows some meaning of semantic binding patterns. For example, the prepositional phrase and the adverb phrase, that are in the pattern “Subject[NP], Verb[SV], Modifyv1[PP]” and the pattern “Subject[NP], Verb[SV], Modifyv2[ADVP]”, can express some meaning of the action. For example, the Modifyv1 expresses the location of this action and the Modifyv2 express the time or the frequency of this action. All meanings established for semantic binding patterns are shown in Appendix D.

Subject: NP
Verb: SV
Object NP

Agent: subject
Action: SV
Patient: object

Agent: subject
Action: SV
Patient: object
Instrument :subject

Subject: NP
Verb: SV
Modifyv2 : ADVP

Agent subject
Action: SV
Manner modifyv2

Agent subject
Action: SV
Cooperating or Separating modifyv2

Agent subject
Action: SV
Time or Frequency: modifyv2

Agent subject
Action: SV
Quantity modifyv2

Subject: NP
Verb: SV
Modifyv1 : PP

Agent subject
Action: SV
Location modifyv1

Agent subject
Action: SV
Coordination modifyv1

Agent subject
Action: SV
Purpose: modifyv1

Agent subject
Action: SV
Means: modifyv1

Agent subject
Action: SV
Source or Destination: modifyv1

Figure 4.19: Examples of meanings of semantic binding patterns.

The semantic interpretations in this research are classified into two groups: “Purpose”, and “Complement”. This classification is based on functions of serial verbs, types of serial verbs and the semantics from (Thepkanjana, 1986, Muansuwan, 2002 and Sudmuk, 2005). The purpose group means that the second verb in two consecutive SVCs indicates purposive action of the first verb. The complement group means that the second verb of two consecutive SVCs is complement of the first verb. The action of SVCs in the purpose group can be represented by combination of both two action verbs in two consecutive SVCs because two verbs have equal status. On the other hand, the action of SVCs in a complement group can be completed by action of the second verb in two consecutive SVCs.

The examples of Thai two consecutive SVCs in the purpose group are “ใช้ออกแบบ/chai ok-baep/” and “เลือกเล่น/leuaklen/”. The examples of Thai two consecutive SVCs in the complement group are “ต้องการควบคุม/tongkarn khuapkhum/” and “พยายามหา/phayayam ha/”. Some Thai two consecutive SVCs and their interpretations are shown in Appendix E.2.

## 4.5 Creation of a Lexical Database

The information about serial verbs and the information about the argument patterns of serial verbs from syntactic analysis and semantic analysis is very useful for generating a lexicon of serial verbs. The specifications of serial verbs in our lexicon are be composed of the category, the patterns of serial verbs, and the semantic interpretation. Tag definitions of a designed serial verb lexicon are based on information of words from (TCLLEX, 2005) and (LEXITRON, 2005). The patterns of serial verbs and the semantic interpretation will be produced by our method of analysis. An example of an entry in the lexicon that we designed is shown in Figure 4.20.

```
<item id="1">
<wd>
   <th>ใช้ออกแบบ</th>
   <en></en>
</wd>
<syn>
   <cat>v.</cat>
   <scat>vact. vact.</scat>
   <vp1>sub+v+obj+advb</vp1>
</syn>
<sem>
   <intpt>purpose</intpt>
   <rel1>sub+obj+mdfv2
   <agt>sub</agt>
   <pat>dob</pat>
   <ins>sub</ins>
   <manner>mdfv2</manner>
   </rel1>
   <rel2>sub+PREL
   <agt>sub</agt>
   </rel2>
 </sem>
</item>
```

Figure 4.20: An example of an entry in the designed lexicon.

Referring to Figure 4.20, the definitions of each tag in generated lexicon are described in Table 4.2. The detail information of two consecutive Thai serial verbs, which are collected in designed lexicon, is based on the information of SVCs that found in the ORCHID corpus. Sentences with two consecutive SVCs are analyzed by algorithms in the proposed semantic analysis framework to generate valuable information of two consecutive SVCs, such as syntactic verb patterns, semantic binding patterns, and semantic interpretations. All tag definitions are shown in Appendix E.3.

Table 4.2: Tag definitions in designed Thai consecutive SVCs lexicon.

| Tag | Definition |
|---|---|
| <item></item> | No. of words |
| <wd></wd> | Word |
| <syn></syn> | Syntactic pattern |
| <cat></cat> | Word’s category |
| <scat></scat> | Particle word’s category |
| <vp></vp> | Syntactic verb pattern |
| <sem></sem> | Semantic pattern |
| <intpt></intpt> | Semantic Interpretation |
| <rel></rel> | Relation pattern |
| <agt></agt> | Agent of action |
| <pat></pat> | Patient of action |
| <ins></ins> | Instrument of action |
| <manner></manner> | Adverb of manner |

# CHAPTER V
# EXPERIMENTS AND RESULTS

This chapter presents the experimental setting and results. In this research, the experiment was separated into two main parts based on the syntactic analysis and the semantic analysis.

## 5.1 Data Set

Our experiments have been conducted on an important Thai corpus That are ORCHID (ORCHID,2003). All sentences with two consecutive SVCs are gathered from the ORCHID corpus. The goal of the ORCHID initiative is building linguistic resources to support research in natural language processing. The ORCHID corpus is a collection of annual reports and research proposals produced by different universities and research organizations in Thailand. The scope of this research is to analyze only two consecutive serial verbs having as first element one of the top ten most frequent serial verbs in the ORCHID corpus. The list of first verbs contains "ใช้/chai/(use)", "ทำหน้าที่/thamnati/(perform, do duty)", "ต้องการ/tongkan/(want)", "ช่วย/chuy/(help)", "ทดลอง/thot-long/(test,try)", "ไป/pai/(go)", "ดำเนินการ/ dam-noen-kan /(manage, perform)", "พยายาม/pha-ya-yam/(try)", "มา/ma/(come)", "เลือก/lueak/(select)". In ORCHID, there are 262 sentences that have SVCs with two consecutive serial verbs occurring as verbs in the main clause of sentence. Additionally, there are also 124 sentences that SVCs with two consecutive serial verbs occurring as verbs in subordinate clauses of sentences. Therefore, the total number of sentences with SVCs is 386 sentences. This data set is analyzed to discovery the semantic binding patterns and semantic interpretations of two consecutive Thai SVCs. Furthermore, the data set from the ORCHID corpus are separated into training sets and test sets to test the performance of methods in the proposed frameworks.

In order to confirm the semantic binding patterns and semantic interpretations, an unseen data set which are not found in the ORCHID corpus is applied. This Thai corpus is provided by Department of Linguistics, Chulalongkorn University (http://www.arts.chula.ac.th/~ling/ThaiConc/). There are 144 sentences that are used as serial verbs. Within these 114 sentences, there are 110 sentences that have two consecutive SVCs are found as verbs in the main clause of sentences. Additionally, there are also 34 sentences that have two consecutive SVCs are found as verbs in subordinate clauses of sentences.

### 5.2 Experiments in Syntactic Analysis

The segmentation was processed automatically by SWATH (SWATH, 2003) which is a word segmentation program. The phrase structure trees of sentences are formed by parsing with the proposed phrase structure grammar rules. All 386 sentences in the corpus can be segmented by SWATH with designed correcting rules and parsed by proposed Thai grammar rules. Examples of the results of the segmentation and syntactic analysis processes are described below.

Examples of sentences are:

(a) "โครงการวิจัยได้เลือกใช้เครื่องไมโครคอมพิวเตอร์เป็นหลัก /khrongkan–wichai-dai-lueakchai-khrueangmaikhrokhomphiotoe-penlak/ (The research project selected mainly to use microcomputers.)"

(b) "ซอฟท์แวร์NECTEC II สามารถใช้ออกแบบวงจรรวมได้จริง /sopwae-NECTEC II-samat-chaiokbaep-wongchonruam-daiching/ (The software can be used to design an integrated circuit practically.)"

(c) "คณะกรรมการนโยบายทำหน้าที่กำหนดนโยบายตลอดจนกำกับการดำเนินงาน /khana-kammakan-nayobai-thamnathikamnot-nayobai-talotchon-kamkap-kanddamnoenngan/ (Policy Committees act to define policies and oversee processes.)"

Examples of the results of segmentation under correcting rules are shown below in Figure 5.1.

(a) โครงการวิจัย/NCMN, ได้/XVAM, เลือก/VACT, ใช้/VACT, เครื่องไมโครคอมพิวเตอร์/NCMN, เป็น/FIXV, หลัก/VSTA

(b) ซอฟท์แวร์/NCMN, NECTECII/NPRP, สามารถ/XVAM, ใช้/VACT, ออกแบบ/VACT, วงจรรวม/NCMN, ได้/XVAE, จริง/ADVN

(c) คณะ/NCMN, กรรมการนโยบาย/NCMN, ทำหน้าที่/VACT, กำหนด/VACT, นโยบาย/NCMN, ตลอดจน/JCRG, กำกับ/VACT, การ/FIXN, ดำเนินงาน/VACT

โครงการวิจัย@NCMN,ได้@XVAM,เลือก@VACT,ใช้@VACT,เครื่องไมโครคอมพิวเตอร์
@NCMN,เป็น@VSTA,หลัก@VSTA,
ซอฟท์แวร์@NCMN,NECTECII@NCMN,สามารถ@XVAM,ใช้@VACT,ออกแบบ@VACT,
วงจรรวม@NCMN,ได้@XVAE,จริง@ADVN,
คณะ@NCMN,กรรมการนโยบาย@NCMN,ทำหน้าที่@VACT,กำหนด@VACT,นโยบาย
@NCMN,ตลอดจน@JCRG,กำกับ@VACT,การ@FIXN,ดำเนินงาน@VACT,

Figure 5.1: The results of segmentation.

Examples of the results of syntactic analysis and phrase structure analysis are shown in Figures 5.2 (a) and 5.2 (c). The phrase structure trees of those sentences from the parser tool are shown in Figures 5.3 (a) and 5.3 (c) (Parsing analysis is based on the phrase structure grammar proposed in Section 4.2 of Chapter IV).

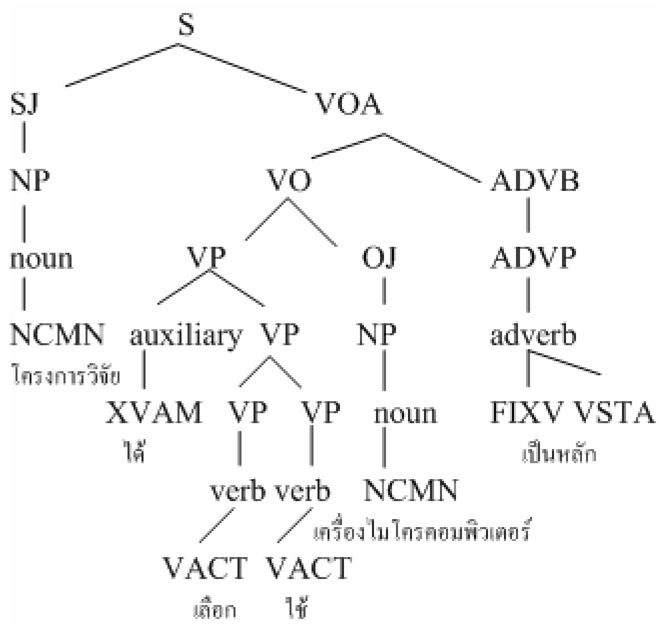



Figure 5.2 (a): The result of syntactic analysis and phrase structure analysis.

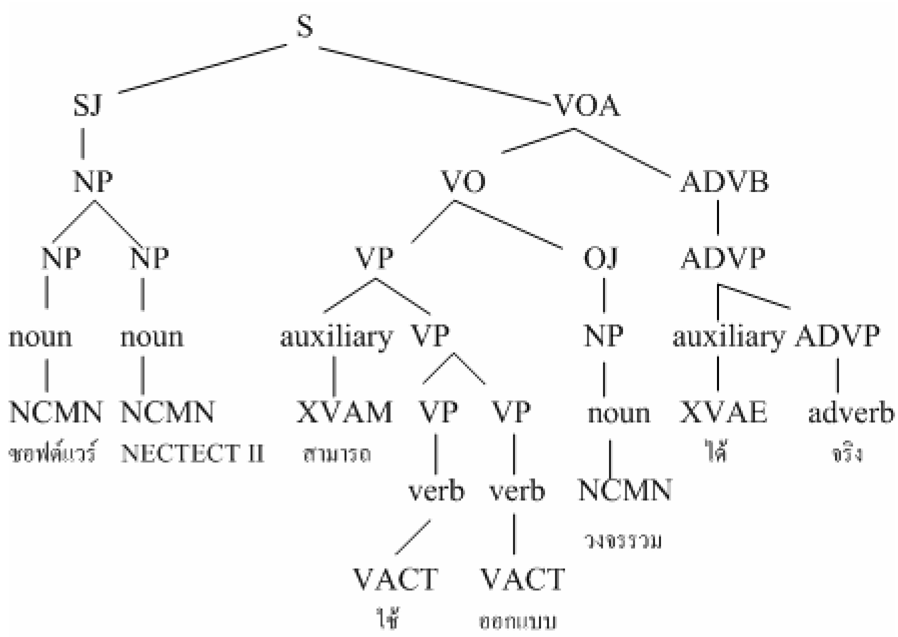



Figure 5.2 (b): The result of syntactic analysis and phrase structure analysis.

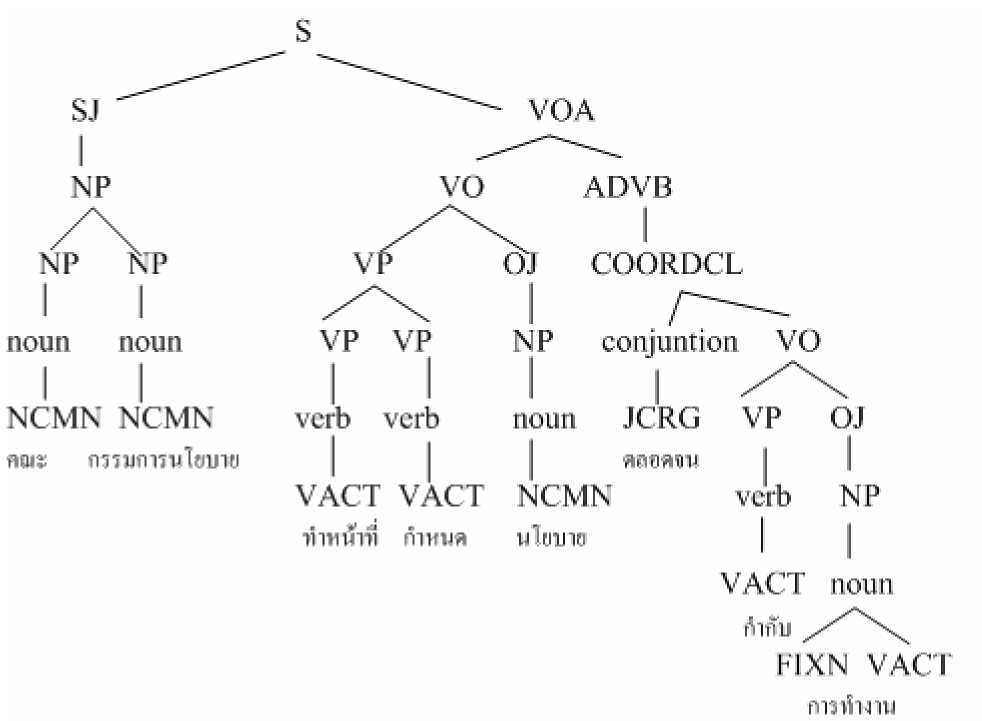



Figure 5.2 (c): The result of syntactic analysis and phrase structure analysis.

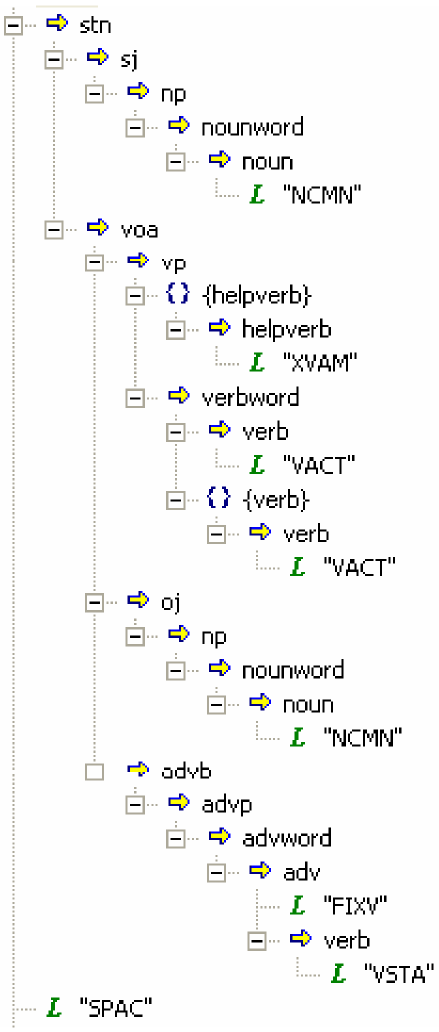



Figure 5.3 (a): The PST.

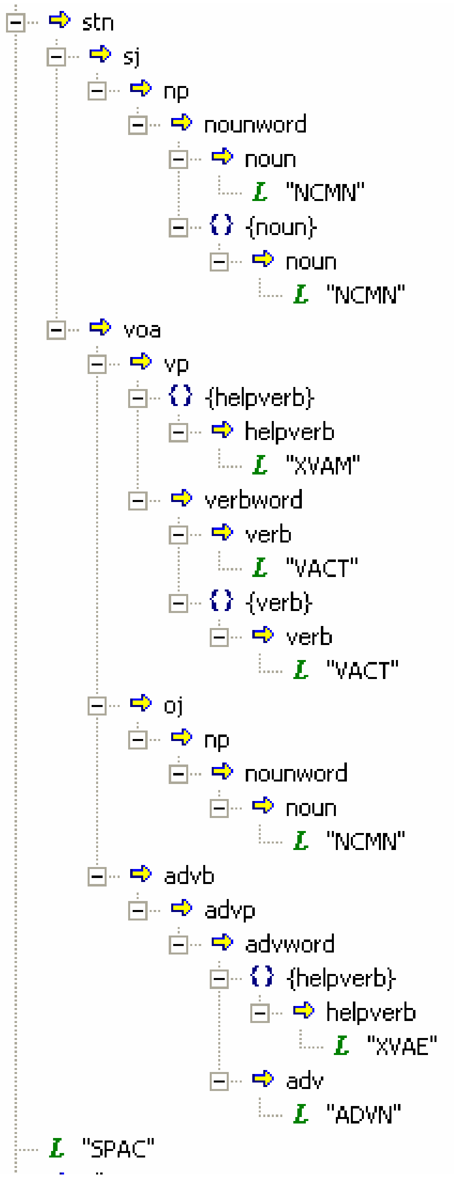



Figure 5.3 (b): The PST.

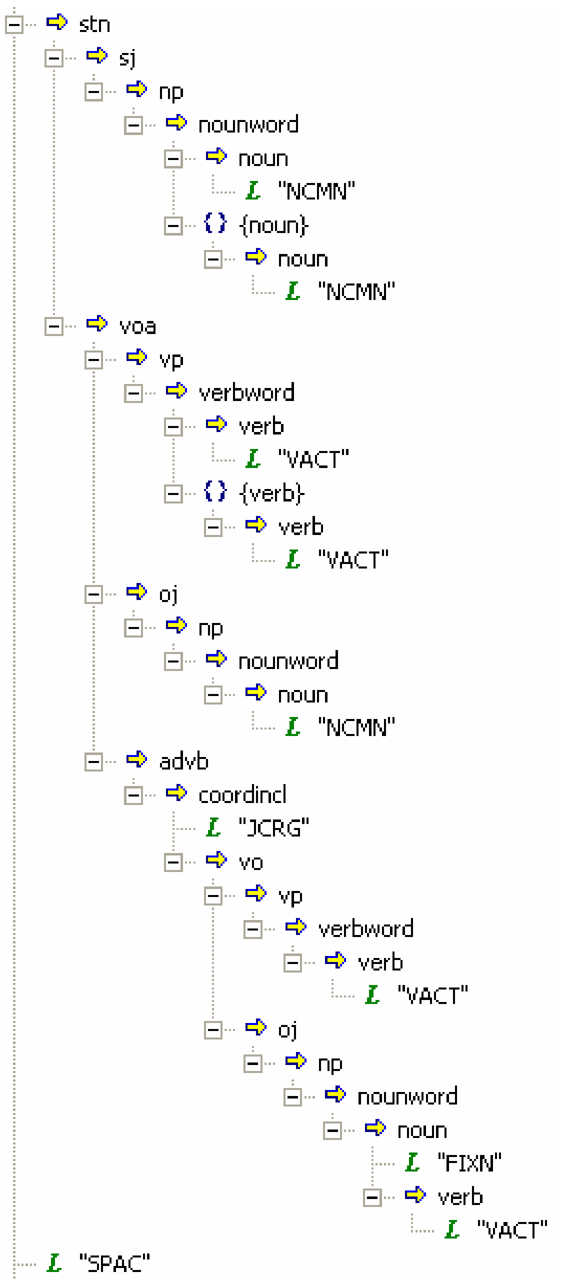



Figure 5.3 (c): The PST.

With regards to Figure 5.3, the phrase structures describe the sentence structures and the order of phrases in the sentences below.

(a) The sentence structure is subject-verb-object-adverbial and the sequential phrase is <NP, VP, NP, ADVP>.

(b) The sentence structure is subject-verb-object-adverbial and the sequential phrase is <NP, VP, NP, ADVP>.

(c) The sentence structure is subject-verb-object-adverbial and the sequential phrase is <NP, VP, NP, CL_ORD>.

## 5.3 Experiments in Semantic Analysis

Sentence structures from phrase structure trees are analyzed to find syntactic patterns of SVCs and relations in the sentence by relation analysis theorems (relationship analysis algorithm in Figure 4.8). Then, the head words or the keywords of phrases that relate to SVCs are defined as classes by rules. All information from previous processes are defined as features of each two consecutive SVCs to learn patterns in the semantic analysis process.

After obtaining syntactic patterns of SVCs, relations and surrounding word classes, this information provides features to analyze semantic binding patterns of SVCs, semantic interpretations, and some meaning of semantic binding patterns. The probabilistic neural networks (MATLAB, 2008) have been used to classify patterns of SVCs automatically based on defined features. The details of the defined features and the classification process are described in Experiment I, Experiment II and Experiment III respectively. The first experiment that we conducted was designed to study on semantic binding patterns of SVCs. The second experiment tests on semantic interpretations. The last experiment tries to learn the meanings of some semantic binding patterns of SVCs.

### 5.3.1 Experiment I

Experiment I is set to classify semantic binding patterns in sentences that are connected to two consecutive SVCs. Each sentence structure and relation in sentences will be used to determine the set of features to learn the binding patterns. The binding patterns of verbs in main clause of sentences are classified into 19 patterns (Figure 4.17); in addition, the binding patterns of verbs in subordinate clauses are classified into 14 patterns (Figure 4.18). All data items are separated into three folds. One fold was used to test the data sets, as well as the remaining two folds were used as the training data sets. Moreover, this experiment is tested for sentences which are not found in the ORCHID corpus. There are 144 sentences in another Thai corpus provided by Department of Linguistics, Chulalongkorn University. These sentences from unseen documents used as the test data sets for the three existing training data sets.

In this experiment, the defined features are sentence structures and relations in sentences. Some examples of features are the relations “Subject” and “Object” (the grammatical roles of noun phrases in the sentences occurring before and after SVCs). Other examples of features are the relations “Modifyv1” and “Modifyv2” which signify a prepositional phrase and an adverb phrase following an SVC. The feature set is generated only from related relations that connect to SVCs, so the examples of unselected relations are “Modifyn1” and “Modifyn2”. Both relations are expressed as relative clause in sentences that occur after noun phrases and the roles of relative clauses which modify nouns or noun phrases that are the subjects and the objects of sentences.

The input of classification is the vector of all features of each sentence and the output is the semantic binding patterns given by the probabilistic neural networks. The performance of the method in Experiment I was measured as accuracy rates by applying the following formula. The results of Experiment I are shown in Table 5.1.

$$Accuracy = \frac{\text{Total number of sentences with correct binding patterns of SVCs}}{\text{Total number of all sentences}} * 100 \rightarrow (1)$$

Table 5.1 (a): Results of Experiment I.

| SVCs' Roles | Number of All Sentences | Data Set Fold | | | Average Accuracy |
|---|---|---|---|---|---|
| | | Test: 1st Training: 2nd + 3rd | Test: 2nd Training: 1st + 3rd | Test: 3rd Training: 1st + 2nd | |
| Verbs in Main Clauses | 262 | 98.85% | 98.85% | 100.00% | 99.23% |
| Verbs in Sub Clauses | 124 | 97.56% | 97.56% | 97.62% | 97.58% |

Table 5.1 (b): Results of Experiment I.

| SVCs' Roles | Number of Unseen Sentences | Data Set Fold | | | Average Accuracy |
|---|---|---|---|---|---|
| | | Test: Unseen Training: 2nd + 3rd | Test: Unseen Training: 1st + 3rd | Test: Unseen Training: 1st + 2nd | |
| Verbs in Main Clauses | 110 | 100.00% | 100.00% | 100.00% | 100.00% |
| Verbs in Sub Clauses | 34 | 100.00% | 100.00% | 100.00% | 100.00% |

According to Table 5.1 (a), the accuracy of corrected binding patterns is nearly 100% because binding patterns of two consecutive SVCs can be classified by formula rules. The error can occur in the case that the pattern was found in the test data only. In Table 5.1 (b), to test for unseen data, the accuracy of corrected binding patterns is 100% because all patterns of unseen sentences were found in ORCHID corpus.

### 5.3.2 Experiment II

Experiment II was designed to classify semantic interpretations. All data items are separated into three folds. One fold was used to be a test data set, as well as the remaining two folds were used to be a training data set. To test for unseen documents, the same sentence set from another Thai corpus in Experiment I is also used as a test data set. The semantic interpretations are classified into two groups that are "Purpose" or "Complement" (described in Section 4.4).

In this experiment, the defined features are sentence structures, relations in sentences and classes of surrounding words. The sentence structures and relations in sentences are the same features described in Experiment I. The adding features are determined by the classes of surrounding words that are related to SVCs. There are different parts of speech in the surrounding words that are the head nouns of noun phrases, the preposition in prepositional phrases and the adverb in adverb phrases. If the relation features are "Subject" and "Object", the examples of added features are noun classes of a subject and an object, such as "Person", "Artifacts" and "Data". If other relations are "Modifyv1" or "Modifyv2", the examples of added features are prepositional classes, such as "Location", "Source or Destination", "Purpose", and "Mean", or adverb classes, such as "Manner" and "Time".

The input of classification is the vector of all features of each sentence; and the output is the semantic interpretation given by the probabilistic neural networks. The experiment was tested on two different feature sets. The first feature set was the sentence structures, relations in sentences and classes of surrounding words. The second feature set, the first verb of two consecutive SVCs was the added feature.

The performance of the method in Experiment II was also divided into two evaluations. The first evaluation was computed as accuracy rates by applying formula(2), and the second evaluation was computed as precision rates by applying formula(3) and formula(4) respectively.

$$Accuracy = \frac{\text{Total number of sentences with correct classification of interpretations}}{\text{Total number of all sentences}} * 100 \rightarrow (2)$$

$$Precision = \frac{\text{Total number of sentences with correct SVCs classified to this interpretation}}{\text{Total number of sentences with SVCs classified to this interpretation}} * 100 \rightarrow (3)$$

$$Recall = \frac{\text{Total number of sentences with correct SVCs classified to this interpretation}}{\text{Total number of sentences with SVCs in this interpretation}} * 100 \rightarrow (4)$$

The results of Experiment II, which was evaluated by the accuracy rates, are shown in Table 5.2. The results of Experiment II, which was evaluated by the precision rates of the purpose group and the complement group, are shown in Tables 5.3 and 5.4 respectively.

Table 5.2 (a): Results of Experiment II in terms of accuracy rates.

| SVCs' Roles | Number of All Sentences | Data Set Fold | | | Average Accuracy |
|---|---|---|---|---|---|
| | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 386 | 61.24% | 63.5% | 71.88% | 65.54% |
| Sentence structures + Relations + Word classes + First verb | 386 | 86.82% | 82.94% | 82.81% | 84.19% |

Table 5.2 (b): Results of Experiment II in terms of accuracy rates.

| SVCs' Roles | Number of Unseen Sentences | Data Set Fold | | | Average Accuracy |
|---|---|---|---|---|---|
| | | Test: Unseen Training: $2^{nd}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 144 | 65.28% | 61.11% | 63.89% | 63.43% |
| Sentence structures + Relations + Word classes + First verb | 144 | 83.33% | 79.17% | 84.72% | 84.41% |

Table 5.3 (a): Results of Experiment II in terms of precision rates of “Purpose”.

| SVCs’ Roles | Number of All Sentences | Data Set Fold | | | Average Precision |
|---|---|---|---|---|---|
| | | Test: 1$^{st}$ Training: 2$^{nd}$ + 3$^{rd}$ | Test: 2$^{nd}$ Training: 1$^{st}$ + 3$^{rd}$ | Test: 3$^{rd}$ Training: 1$^{st}$ + 2$^{nd}$ | |
| Sentence structures + Relations + Word classes | 386 | 59.57% | 61.76% | 73.81% | 65.05% |
| Sentence structures + Relations + Word classes + First verb | 386 | 90.38% | 88.37% | 90.48% | 89.74% |

Table 5.3 (b): Results of Experiment II in terms of precision rates of “Purpose”.

| SVCs’ Roles | Number of Unseen Sentences | Data Set Fold | | | Average Precision |
|---|---|---|---|---|---|
| | | Test: Unseen Training: 2$^{nd}$ + 3$^{rd}$ | Test: Unseen Training: 1$^{st}$ + 3$^{rd}$ | Test: Unseen Training: 1$^{st}$ + 2$^{nd}$ | |
| Sentence structures + Relations + Word classes | 144 | 45.45% | 42.03% | 43.86% | 43.78% |
| Sentence structures + Relations + Word classes + First verb | 144 | 71.43% | 62.50% | 72.73% | 68.89% |

Table 5.4 (a): Results of Experiment II in terms of recall rates of “Purpose”.

| SVCs’ Roles | Number of All Sentences | Data Set Fold | | | Average Recall |
|---|---|---|---|---|---|
| | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 386 | 52.54% | 61.82% | 55.36% | 56.57% |
| Sentence structures + Relations + Word classes + First verb | 386 | 79.66% | 72.73% | 71.43% | 74.60% |

Table 5.4 (b): Results of Experiment II in terms of recall rates of “Purpose”.

| SVCs’ Roles | Number of Unseen Sentences | Data Set Fold | | | Average Recall |
|---|---|---|---|---|---|
| | | Test: Unseen Training: $2^{nd}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 144 | 52.38% | 61.90% | 52.39% | 55.56% |
| Sentence structures + Relations + Word classes + First verb | 144 | 71.43% | 71.43% | 76.19% | 73.02% |

Table 5.5 (a): Results of Experiment II in terms of precision rates of “Complement”.

| SVCs’ Roles | Number of All Sentences | Data Set Fold | | | Average Precision |
|---|---|---|---|---|---|
| | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 386 | 62.20% | 64.21% | 70.93% | 65.78% |
| Sentence structures + Relations + Word classes + First verb | 386 | 84.42% | 80.23% | 79.07% | 81.24% |

Table 5.5 (b): Results of Experiment II in terms of precision rates of “Complement”.

| SVCs’ Roles | Number of Unseen Sentences | Data Set Fold | | | Average Precision |
|---|---|---|---|---|---|
| | | Test: Unseen Training: $2^{nd}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 144 | 78.26% | 79.49% | 77.78% | 78.51% |
| Sentence structures + Relations + Word classes + First verb | 144 | 88.24% | 87.50% | 90.00% | 88.58% |

Table 5.6(a): Results of Experiment II in terms of recall rates of “Complement”.

| SVCs’ Roles | Number of All Sentences | Data Set Fold | | | Average Recall |
|---|---|---|---|---|---|
| | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 386 | 72.86% | 82.43% | 84.72% | 80.00% |
| Sentence structures + Relations + Word classes + First verb | 386 | 92.86% | 93.24% | 94.44% | 93.51% |

Table 5.6(b): Results of Experiment II in terms of recall rates of “Complement”.

| SVCs’ Roles | Number of Unseen Sentences | Data Set Fold | | | Average Recall |
|---|---|---|---|---|---|
| | | Test: Unseen Training: $2^{nd}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $3^{rd}$ | Test: Unseen Training: $1^{st}$ + $2^{nd}$ | |
| Sentence structures + Relations + Word classes | 144 | 70.59% | 60.78% | 68.63% | 66.67% |
| Sentence structures + Relations + Word classes + First verb | 144 | 88.24% | 82.35% | 88.24% | 86.27% |

The accuracy rates in Tables 5.2(a) and 5.2(b) show that the average accuracy of classifying semantic interpretations by the second feature set is about 80%, while the average accuracy of using the first feature set is only 60%. This means that the second feature set can determine semantic interpretations better than that of the first feature set. Like the accuracy rate, the precision rates of classifying both semantic interpretations by the second feature set are more than 80%, as shown in Tables 5.3 and 5.5. Those results show that both semantic interpretations of two consecutive SVCs in this corpus can be determined by the sentence structures, relations, surrounding word classes, and the first verb of two consecutive SVCs. Furthermore, the recall rates of classifying both semantic interpretations are shown in Tables 5.4 and 5.6, respectively.

### 5.3.3 Experiment III

Experiment III was designed to discover the meanings of semantic binding patterns in sentences. Since there are such a large number of different meanings when open class verbs are included, we chose to handle only some important meanings that can be obviously classified using our selected features. These meaning are described as follows:

1) The subject of an SVC can be called as an INSTRUMENT of an action in some patterns. For example, the pattern “Subject[NP], Verb[SV], Object[NP]” can express the semantic roles “Agent[subject], Action[SV], Patient[object], Instrument[subject]”.

2) When the SVC expresses an action and its subject is the agent of that action, the subject is most often a noun in the PERSON class. For example, the pattern “Subject[NP], Verb[SV], Object[NP]” means “Agent[subject(person)], Action[SV], Patient[object]”.

3) The various meanings of constructions Modifyv1[PP] can be classified into two groups of PREPOSITION. In the first group the preposition phrase is “Location or Place”, and “Source or Destination”. In the second group the preposition phrase expresses “Coordination”, “Purpose”, and “Means”.

4) The various meanings of Modifyv2[ADVP] can be classified into two groups of ADVERB. In the first group the adverbial phrase is an adverb of “Manner”, and “Quantity”. In the second group is “Time & Frequency”, and “Cooperating & Separating”

5) The subject of an SVC in a subordinate clause is typically the patient in a relative clause, called as PASSIVE. For example, the pattern “Subject[NP], Verb[SV], Modify2[ADVP]” can mean “Agent[-], Action[SV], Patient[subject], Manner[Modify2]”.

In this experiment, the feature set is composed of sentence structures, relations and surrounding word classes as in Experiment II. All data items are separated into three folds. One fold was used as the test data set, as well as the remaining two folds were used as the training data sets. The performance of the method in Experiment III was evaluated in two different ways. The first evaluation measured the accuracy rate; and the second evaluation measured the precision rate. Different formulas were used to evaluate each meaning as follows:

1) The accuracy rate and the precision rate of the classification of INSTRUMENT.

$$Accuracy = \frac{\text{Total number of sentences with correct identification of instruments}}{\text{Total number of all sentences}} * 100 \rightarrow (5)$$

$$Precision = \frac{\text{Total number of sentences with correct instruments}}{\text{Total number of sentences with identified instruments}} * 100 \rightarrow (6)$$

$$Recall = \frac{\text{Total number of sentences with correct instruments}}{\text{Total number of sentences with instruments}} * 100 \rightarrow (7)$$

2) The accuracy rate and the precision rate of the classification of PERSON.

$$Accuracy = \frac{\text{Total number of sentences with correct identification of persons}}{\text{Total number of all sentences}} * 100 \rightarrow (8)$$

$$Precision = \frac{\text{Total number of sentences with correct persons}}{\text{Total number of sentences with identified persons}} * 100 \rightarrow (9)$$

$$Recall = \frac{\text{Total number of sentences with correct persons}}{\text{Total number of sentences with persons}} * 100 \rightarrow (10)$$

3) The accuracy rate and the precision rate of PREPOSITION classification.

$$Accuracy = \frac{\text{Total number of sentences with correct identification of prepositions}}{\text{Total number of all sentences}} * 100 \rightarrow (11)$$

$$Precision = \frac{\text{Total number of sentences with correct prepositions}}{\text{Total number of sentences with identified prepositions}} * 100 \rightarrow (12)$$

$$Recall = \frac{\text{Total number of sentences with correct prepositions}}{\text{Total number of sentences with prepositions}} * 100 \rightarrow (13)$$

4) The accuracy rate and the precision rate of ADVERB classification.

$$Accuracy = \frac{\text{Total number of sentences with correct identification of adverbs}}{\text{Total number of all sentences}} * 100 \rightarrow (14)$$

$$Precision = \frac{\text{Total number of sentences with correct adverbs}}{\text{Total number of sentences with identified adverbs}} * 100 \rightarrow (15)$$

$$Recall = \frac{\text{Total number of sentences with correct adverbs}}{\text{Total number of sentences with adverbs}} * 100 \rightarrow (16)$$

5) The accuracy rate and the precision rate of PASSIVE classification.

$$Accuracy = \frac{\text{Total number of sentences with correct identification of passive}}{\text{Total number of all sentences}} * 100 \rightarrow (17)$$

$$Precision = \frac{\text{Total number of sentences with correct passive}}{\text{Total number of sentences with identified passive}} * 100 \rightarrow (18)$$

$$Recall = \frac{\text{Total number of sentences with correct passive}}{\text{Total number of sentences with passive}} * 100 \rightarrow (19)$$

The results of Experiment III for the evaluation using accuracy rates are shown in Table 5.7. Additionally, the results of Experiment III, which are evaluated by the precision rates, are shown in Table 5.8 and with the recall rates are shown in Table 5.9.

According to Tables 5.7 and 5.8, it has shown that the average accuracy and the average precision of classifying interesting semantic binding meaning are more than 80% and 70%, respectively. In addition, the average recall rate is 100% as in Table 5.9. These means that this feature set can determine the interesting semantic binding meaning effectively.

Table 5.7: Results of Experiment II in terms of accuracy rates.

| Meaning | Number of All Sentences | Number of Interesting Meaning Sentences | Data Set Fold | | | Average Accuracy |
|---|---|---|---|---|---|---|
| | | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| INSTRUMENT | 386 | 117 | 81.40% | 89.92% | 90.63% | 87.32% |
| PERSON | 386 | 19 | 78.29% | 86.82% | 88.28% | 84.46% |
| PREPOSITION | 386 | 57 | 94.57% | 97.67% | 96.09% | 96.11% |
| ADVERB | 386 | 58 | 95.34% | 96.12% | 95.31% | 95.60% |
| PASSIVE | 124 | 57 | 93.48% | 86.96% | 89.13% | 89.86% |

Table 5.8: Results of Experiment II in terms of precision rates.

| Meaning | Number of All Sentences | Number of Interesting Meaning Sentences | Data Set Fold | | | Average Precision |
|---|---|---|---|---|---|---|
| | | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| INSTRUMENT | 386 | 117 | 62.90% | 75.47% | 76.92% | 71.76% |
| PERSON | 386 | 19 | 87.50% | 83.33% | 87.50% | 86.11% |
| PREPOSITION | 386 | 57 | 72.00% | 87.50% | 78.26% | 79.25% |
| ADVERB | 386 | 58 | 76.00% | 80.00% | 76.00% | 77.30% |
| PASSIVE | 124 | 57 | 86.36% | 76.00% | 89.17% | 80.51% |

Table 5.9: Results of Experiment II in terms of recall rates.

| Meaning | Number of All Sentences | Number of Interesting Meaning Sentences | Data Set Fold | | | Average Recall |
|---|---|---|---|---|---|---|
| | | | Test: $1^{st}$ Training: $2^{nd}$ + $3^{rd}$ | Test: $2^{nd}$ Training: $1^{st}$ + $3^{rd}$ | Test: $3^{rd}$ Training: $1^{st}$ + $2^{nd}$ | |
| INSTRUMENT | 386 | 117 | 100.00% | 100.00% | 100.00% | 100.00% |
| PERSON | 386 | 19 | 100.00% | 100.00% | 100.00% | 100.00% |
| PREPOSITION | 386 | 57 | 100.00% | 100.00% | 100.00% | 100.00% |
| ADVERB | 386 | 58 | 100.00% | 100.00% | 100.00% | 100.00% |
| PASSIVE | 124 | 57 | 100.00% | 100.00% | 100.00% | 100.00% |

## 5.4 Discussion

SVCs are composed of more than one verb and there are various patterns of SVCs. So, to understand the meaning of SVCs is much complicated than the understanding single verbs. Thus, syntactic analysis and semantic analysis of SVCs are actually helpful to find the meaning of SVCs which, in turn, help understanding the natural language statements. The results of this research have shown that finding the syntactic patterns and analyzing related words of SVCs can both extract useful information for the analysis of the semantics of SVCs.

According to the syntactic analysis process, SVCs are the verbs in the main sentences or sometimes they are verbs of the clauses in the sentences. There are various syntactic patterns of SVCs, not only in the main clauses, but also in the subordinate clauses. In the linguistics researches (Thepkanjana, 1986) and (Sudmuk, 2005), SVCs are separated into an independent components, or called as a single particle, before performing semantic and syntactic analysis. However, this method leaves some gaps in the interpretation of words since the meaning also depends on locations of a word and its surroundings. Therefore, this research considers one SVC as one object although this object is composed of many words. Thus, the syntax of SVCs is clearly defined. So, the syntactic patterns can be easily derived from the proposed grammars.

With regards to experiments on semantic analysis, the relationships between related words and SVCs can be found by the relation analysis theorem (Appendix C). Note that these relations have never been proposed by any other researchers as far as we know. Thus, the semantic binding patterns of the SVCs are well analyzed by the syntactic patterns and the relations of related words of SVCs. The relation analysis theorem found that the surrounding words, such as nouns, prepositions or adverbs are related to SVCs and can be used to represent the semantics of SVCs. Experiment I shows that all information from previous analysis can learn automatically for semantic bindings patterns of SVCs. Additionally, it indicates that each syntactic pattern and related words can be transformed into semantic binding patterns using rules.

Therefore, these methods can be applied to other languages in the SVO language family.

Experiment II shows that with only the information about syntactic patterns and related words we cannot determine the correct semantic interpretation as well because there are various patterns in each semantic interpretation. However, the semantic interpretation can be learned automatically after we acquire information about syntactic patterns, related words and the first verb in series.

The results ofExperiment III indicates that some specific meaning of semantic binding patterns can be recognized by syntactic patterns and related words. For the examples, the subject of SVCs “ใช้สอน/chaison/” is the word “ซอฟต์แวร์/sopwae/” that is a noun in the artifact class, so subjects can be agents and instruments of SVCs. Moreover, the most frequent subjects of some SVCs are nouns in the person class, such as “ดำเนินการสนับสนุน/damnoenkan-sanapsanun/”, “พยายามศึกษา/phayayam-sueksa /”.

# CHAPTER VI
# CONCLUSION

This chapter provides a summary of the research and some constraints. Some directions for future work are described at the end of this chapter.

## 6.1 Conclusion

This research presents a semantic analysis framework of two consecutive Thai serial verbs. Our framework is separated into two parts: methods for syntactic analysis of sentences, and methods for semantic analysis of serial verbs. The syntactic analysis output by the syntactic analysis of sentence module produces syntactic patterns of SVCs, while the results of the semantic analysis module are semantic binding patterns of SVCs, the meaning of those semantic binding patterns and the semantic interpretations.

Method for syntactic analysis of sentences analyzes the syntactic information using a Thai segmentation tool and a Thai grammar. The grammar rules and the parser are used to parse sentences and to produce syntactic representations of sentences. The syntactic patterns of SVCs and the word information from syntactic analysis provided essential information for analyzing semantics.

According to methods for semantic analysis of serial verbs, the semantic binding patterns of SVCs, the semantic interpretations, and the meaning of semantic binding patterns are defined by analyzing sentence structures, relations in sentences and surrounding word classes. Moreover, the probabilistic neural network is applied to classify semantics and patterns automatically.

Finally, all information about two consecutive Thai serial verbs, e.g. patterns of serial verbs and the associated semantic interpretation are collected to create our new lexicon. The lexicon provides basic information of serial verbs for other researchers in the future.

## 6.2 Future Work

Our research thus far has studied only those two consecutive serial verbs whose first element is in the top ten most frequent serial verbs in the ORCHID corpus. Since there are various SVCs in real Thai texts, these methods should be tested on other consecutive SVCs in order to increase reliability of the algorithms.

This research, however, can be applied on raw data with no tags of separated words, parts of speech and tags of individual phrases. So, it is much complicated to segment words and to parse the sentence structure. There is also no lexicon including useful information for semantic analysis. Therefore, the semantic analysis cannot be immediately concentrated to analyze actual semantics. In the future work, the semantic analysis should analyze data with functional tags, such as phrase tags, relation tags, and word class tags. When this input is combined with tagged data then the meaning of SVCs can be discovered by a fully automatic system.

# REFERENCES


Beale, S.; Nirenburg, S.; and Mahesh, K. Semantic Analysis in the Mikrokosmos Machine Translation Project. Proceedings of the 2nd Symposium on Natural Language Processing (1995): 297-307.

Bejan, C. A.; Moschitti, A.; Morărescu, P.; Nicolae, G., and Harabagiu, S. Semantic Parsing Based on FrameNet. Proceedings of the 3rd International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (2004): 24-29.

Charoenporn, T.; Kruengkrai, C.; Sornlertlamvanich, V., and Isahara, H. Acquiring Semantic Information in the TCL's Computational Lexicon. Proceedings of the 4th Workshop on Asian Language Resources (2004): 47-53.

Chuwicha, Y. Clausehood in serial veab constructions in Thai. Ph.D. Dissertation, Department of Linguistics, Faculty of Arts, Chulalongkorn University, 1992.

Crouch, R. S.; and King, T. H. Unifying Lexical Resources. Proceedings of Interdisciplinary Workshop on the Identification and Representation of Verb Features and Verb Classes (2005): 32-37.

Fillmore, C. J.; Petruck M. R.L.; Ruppenhofer J.; Wright A. FrameNet in Action. International Journal of Lexicoraphy vol.16 no. 3 (2003): 297-332.

Hirst, G. Ontology and the Lexicon. Handbook on Ontologies in Information Systems (2004): 209-230. Available from: http://ftp.cs.toronto.edu/pub/gh/Hirst-Ontol-2003.pdf [online].

Jindal, N., and Liu, B. Identifying Comparative Sentences in Text Documents. Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (2006): 244-251.

Kipper, K.; Dang, H. T., and Palmer, M. Class-based construction of a verb lexicon. Proceedings of the 17th National Conference on Artificial Intelligence (2000): 691-696.

Levin, B. English Verb Classes and Alternations: A Preliminary Investigation. Chicago, USA: The University of Chicago Press, 1993.

Li, L.; He, Z.; and Yi, Y. Principles and Algorithms of Semantic Analysis. Proceedings of the Second International Conference on Machine Learning and Cybernetics vol. 3 (2003):1613-1618.

M`arquez, L. Machine Learning and Natural Language Processing. Barcelona, Catalonia, Spain: Centre de recerca TALP Departament de Llenguatges i Sistemes Inform`atics, LSI, Universitat Polit`ecnica de Catalunya, 2000. Available from: http://www.lsi.upc.es/~lluism/cursos/docSoria00/soria00.ps.gz [online]

The MathWorks Inc. MATLAB. USA: The MathWorks Inc, 2008. Available from: http://www.mathworks.com [online]

Miller, G. A.; Beckwith, R.; Fellbaum, C.; Gross, D.; Miller, K., and Tengi, R. Five papers on WordNet, Technical report. New Jersey, USA: Cognitive Science Laboratory, Princeton University, 1993. Available from: http://wordnet.princeton.edu/ [online].

Minker, W. Robustness and Portability Issues in Multilingual Speech Processing. Machine Translation Journal vol. 16 no. 2 (2001): 109-126.

Muansuwan, N. Verb Complexes in Thai. Ph.D. Dissertation, Department of Linguistics, University of Buffalo, 2002.

Murata, M.; Ma, Q., and Isahara, H. Comparison of Three Machine Learning Methods for Thai Part-of-Speech Tagging. ACM Transactions on Asian Language Information Processing vol. 1 (2002): 145-158.

National Electronics and Computer Technology Center (NECTEC). Lexitron. Bangkok, Thailand: Information Research and Development Division, National Electronics and Computer Technology Center, 2005. Available from: http://lexitron.nectec.or.th [online]

National Electronics and Computer Technology Center (NECTEC). Smart Word Analysis for THai (SWATH). Bangkok, Thailand: Information Research and Development Division, National Electronics and Computer Technology Center, 2003. Available from: http://www.hlt.nectec.or.th/download/softwares/SWATH_for_Windows.zip [online]

National Electronics and Computer Technology Center (NECTEC), Open linguistic Resources CHanelled toward InterDisciplinary research (ORCHID). Bangkok, Thailand: Information Research and Development Division, National Electronics and Computer Technology Center, 2003. Available from: http://www.links.nectec.or.th/orchid/ [online].

Pankhuenkhat, R. Thai Linguistics. 1st edition. Bangkok: Chulalongkorn University, 1998.

Pugsee, J. and Rivepiboon, W. Meaning Analysis of Sentences Using Phrase Structure Grammar with Sentence Structures. Proceeding of the International Conferences on Foundations of Computer Science (FCS'05), (2005): 111-117.

Sarikaya, R.; Gao, Y., and Picheny, M. Word Level Confidence Measurement Using Semantic Features. Proceedings of Acoustics, Speech and Signal Processing (ICASSP '03 IEEE International Conference) vol. 1 (2003): 604-607.

Sebastiani, F. Machine Learning in Automated Text Categorization. ACM Computing Surveys vol. 34 no.1 (2002): 1-47.

Scott, B. The Logos Model: An Historical Perspective. Machine Translation vol.18 (2003): 1-72.

Sornlertlamvanich, V., and Charoenporn, T. ORCHID: Thai Part-Of-Speech Tagged Corpus. Technical Report: ORCHID Corpus (1997): 5-19.

Specht, D.F. Probabilistic neural networks. Neural Networks vol.3 no.1 (1990): 109-118.

Sudmuk, C. The Syntax and Semantics of Serial Verb Constructions in Thai. Ph.D. Dissertation, Faculty of the Graduate School, The University of Texas at Austin, 2005.

Sukhahuta, R. Information Extraction System for Thai Documents. Ph.D. Dissertation, School of Information Systems, University of East Anglia, 2001.

Tallerman, M. Understanding Syntax. London: Arnold, 1998.

Thai Computational Linguistics Laboratory. TCL's Computational Lexicon. Bangkok, Thailand: Thai Computational Linguistics Laboratory, 2005. Available from: http://www.tcllab.org/tcllex/ [online].

Thepkanjana, K. Serial Verb Construction in Thai. Ph.D. Dissertation, Department of Linguistics, The University of Michigan, 1986.

Wasserman, P.D. Advanced Methods in Neural Computing. New York, USA: Van Nostrand Reinhold, 1993.

Wechsler, S. M. Serial Verbs and Serial Motion. Proceedings of Workshop on Multi-Verb Constructions (2003).

Wilawan, S. A Reanalysis of so-called serial verb constructions in Thai, Khmer, Mandarin, and Yoruba. Ph.D. Dissertation, Department of Linguistics, University of Hawaii at Manoa, 1993.

Yang, Z., Wang, Y., and Kitsuregawa, M. Effective Sequential Pattern Mining Algorithms for Dense Database. Proceeding of Japanese National Data Engineering WorkShop (Mar 2006): 1-7.

# APPENDICES

# APPENDIX A
# PUBLICATIONS

## A.1 International Conferences

1. Pugsee, J. and Rivepiboon, W. Meaning Analysis of Sentences Using Phrase Structure Grammar with Sentence Structures. The International Conferences on Foundations of Computer Science (FCS'05), Las Vegas, USA (2005): 111-117.

2. Pugsee, J.; Evens, M. W. and Rivepiboon, W. Analysis of Thai Sentences with a Serial Verb Using a Semantic Lexicon. The 5th IEEE International Conference on Cognitive Informatics (ICCI 2006), Beijing, China (2006): 484-494.

3. Pugsee, J.; Evens, M. W. and Rivepiboon, W. Sentence Analysis for Discovering Knowledge about Thai Serial Verbs for a Lexical Database. The First International Conference on Knowledge, Information and Creativity Support Systems (KICSS 2006), Ayuttaya, Thailand (2006): 92-99.

## A.2 International Journal

1. (submitted) Pugsee, J.; Bhattarasinee, P. and Rivepiboon, W. Discovering Knowledge About Consecutive Serial Verbs Using Word Information And A Machine Learning Technique. International Journal of Software Engineering and Knowledge Engineering (IJSEKE).

# APPENDIX B
# THAI PARTS OF SPEECH

## B.1 Thai Parts of Speech (ORCHID, 2003)

NOUN: A word that is used to refer to people, animals, places, things or abstractions.

PRONOUN: A word that may replace a noun or noun phrase.

VERB: A word that occurs as part of the predicate of a sentence, or that refers to an action, state or attribute.

AUXILIARY: A word that is used with a verb of a sentence indicating a grammatical function such as aspect, voice, mood or tense. An auxiliary has a rather fixed position relating to the verb of the sentence.

DETERMINER: A word that is used with a noun restricting the meaning of the noun in some way.

ADVERB: A word that describes or modifies the meaning of a verb, another adverb or a sentence.

CLASSIFIER: A word that is used with a noun showing the subclass to which a noun belongs. It was originally used to express the units of a noun in counting. The classifier category is subcategorized into 5 subcategories.

CONJUNCTION: A word that joins words, phrases, or clauses together. The conjunction category is subcategorized into 3 subcategories.

PREPOSITION: A word that is used preceding a noun or a pronoun to link it grammatically to other words or phrases.

INTERJECTION: A word that indicates an emotional state or attitude such as delight, surprise, shock, and disgust, but has no referential meaning.

PREFIX: A word that is added to the beginning of a word changing the meaning or function of the word.

SENTENCE PARTICLE: Ending for affirmative sentence (EAFF):- A word that is added to the end of a sentence to indicate the mood of an utterance, such as affirmation, interrogation, imperative, persuasion, threat and so on. Ending for interrogative sentence particle (EITT):- A word which is added to the end of a sentence to indicate the mood of interrogative, and yes-no question.

NEGATOR: A word that is used to negate a sentence.

PUNCTUATION: A mark or a sign (such as an opening/closing bracket, opening/closing quote, comma, colon, semicolon, dash, exclamation mark, period, and question mark)Our algorithm is composed of two main components, which are syntactic analysis of sentences, semantic analysis of serial verbs. All components of algorithm are shown in Figure 4.1.

Table B.1: Subcategories of Thai Parts of Speech from the ORCHID Corpus.

| No. | POS | Description |
|---|---|---|
| 1 | NPRP | Proper noun |
| 2 | NCNM | Cardinal number |
| 3 | NONM | Ordinal number |
| 4 | NLBL | Label noun |
| 5 | NCMN | Common noun |
| 6 | NTTL | Title noun |
| 7 | PPRS | Personal pronoun |
| 8 | PDMN | Demonstrative pronoun |
| 9 | PNTR | Interrogative pronoun |
| 10 | PREL | Relative pronoun |
| 11 | VACT | Active verb |
| 12 | VSTA | Stative verb |
| 13 | VATT | Attributive verb |
| 14 | XVBM | Pre-verb auxiliary, before negator “ไม่ (not)” |
| 15 | XVAM | Pre-verb auxiliary, after negator “ไม่ (not)” |
| 16 | XVMM | Pre-verb, before or after negator “ไม่ (not)” |
| 17 | XVBB | Pre-verb auxiliary, in imperative mood |

Table B.1: Subcategories of Thai Parts of Speech from the ORCHID Corpus (cont.).

| No. | POS | Description |
|---|---|---|
| 18 | XVAE | Post-verb auxiliary |
| 19 | DDAN | Definite determiner, after noun without classifier in between |
| 20 | DDAC | Definite determiner, allowing classifier in between |
| 21 | DDBQ | Definite determiner, between noun and classifier or preceding quantitative expression |
| 22 | DDAQ | Definite determiner, following quantitative expression |
| 23 | DIAC | Indefinite determiner, following noun; allowing classifier in between |
| 24 | DIBQ | Indefinite determiner, between noun and classifier or preceding quantitative expression |
| 25 | DIAQ | Indefinite determiner, following quantitative expression |
| 26 | DCNM | Determiner, cardinal number expression |
| 27 | DONM | Determiner, ordinal number expression |
| 28 | ADVN | Adverb with normal form |
| 29 | ADVI | Adverb with iterative form |
| 30 | ADVP | Adverb with prefixed form |
| 31 | ADVS | Sentential adverb |
| 32 | CNIT | Unit classifier |
| 33 | CLTV | Collective classifier |
| 34 | CMTR | Measurement classifier |
| 35 | CFQC | Frequency classifier |
| 36 | CVBL | Verbal classifier |
| 37 | JCRG | Coordinating conjunction |
| 38 | JCMP | Comparative conjunction |
| 39 | JSBR | Subordinating conjunction |
| 40 | RPRE | Preposition |
| 41 | INT | Interjection |

Table B.1: Subcategories of Thai Parts of Speech from the ORCHID Corpus (cont.).

| No. | POS | Description |
|---|---|---|
| 42 | FIXN | Nominal prefix |
| 43 | FIXV | Adverbial prefix |
| 44 | EAFF | Ending for affirmative sentence |
| 45 | EITT | Ending for interrogative sentence |
| 46 | NEG | Negator |
| 47 | PUNC | Punctuation |

# APPENDIX C
# SENTENCE THEOREMS

## C.1 Sentence Theorems

**Definition 1** Let D be the domain in the field of Computer and Engineering annual reports and research proposals.

**Definition 2** Let N be a set of noun words.

**Definition 3** Let V be a set of verb words.

**Definition 4** Let $\bullet$ be a set operations to concatenate sets.

$A \bullet B = \{ ab \mid a \in A, b \in B \}$

Let $^*$ and $^+$ be a set operations to define recursively the set.

Let $A = \{a\}$

$$A^* = \bigcup_{i=0}^{\infty} A^i = \{\lambda, a, aa, aaa, aaaa, \ldots\}$$

$$A^+ = \bigcup_{i=1}^{\infty} A^i = \{a, aa, aaa, aaaa, \ldots\}$$

**Definition 5** Let ADJUNCT be a set of words or phrases that adds information to another part of a sentence.

**Definition 6** Let P be a set of prepositions.

$PP = P \bullet N^+$ ; $PP \subset ADJUNCT$

$NP = N^+ \bullet PP^* = N^+ \bullet (P \bullet N^+)^*$

**Definition 7** Let S be a set of sentences in D ; $S \in D$

**Definition 8** Let $S_1$, $S_2$ be a set of sentences in different basic sentence patterns.

Let angular brackets "< >" is to denote ordered sequences.

$S_1 \coloneqq \langle NP, V, NP_2 \rangle$ ; $S_1 \subset S$

$S_2 \coloneqq \langle NP, V \rangle$ ; $S_2 \subset S$

$S_1 \cap S_2 = \varnothing$

**Definition 9** Let SUBJECT be a set of relations from NP to V

SUBJECT : NP $\rightarrow$ V ; NP occur before V in a sequence of the sentence.

Let OBJECT be a set of relations from V to NP

OBJECT : V $\rightarrow$ NP ; NP occur after V in a sequence of the sentence.

**Definition 10** Let SV be set of two consecutive serial verbs.

$SV = V \bullet V = \{ v_1v_2 \mid v_1, v_2 \in V \}$

**Lemma 1** $SV \subset V$

*Proof* $\forall v_1v_2 \ (v_1, v_2 \in V \rightarrow v_1v_2 \in V)$

$\therefore \ SV \subset V$

**Definition 11** Let $S_1{}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 1** $S_1{}' \coloneqq \langle NP, SV, NP_2 \rangle$ ; $S_1{}' \subset S$

Relations of $S_1{}'$ are composed of $SUBJECT_1$ and $OBJECT_1$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

*Proof* $S_1 := \langle NP, V, NP_2 \rangle$ ; $S_1 \subset S$ from Definition 8

$SV \subset V$ from Lemma 1

Therefore $S_1{}' \subset S_1$

$\therefore \ S_1{}' \coloneqq \langle NP, SV, NP_2 \rangle$ ; $S_1{}' \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore \ SUBJECT_1 : NP \rightarrow SV$; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore \ OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

**Definition 12** Let $S_3$ be a set of sentences in a sentence pattern.

**Lemma 2** $S_3 \coloneqq \langle NP, V, ADJUNCT \rangle$ ; $S_3 \subset S$

*Proof* $S_2 := \langle NP, V \rangle$ ; $S_2 \subset S$ from Definition 8

V can be added information by ADJUNCT from Definition 5

Therefore $S_3 \subset S$

$\therefore \ S_3 := \langle NP, V, ADJUNCT \rangle$ ; $S_3 \subset S$

**Definition 13** Let MODIFYV be a set of relations from V to ADJUNCT

MODIFYV : V $\rightarrow$ ADJUNCT ; ADJUNCT occur after V in a sequence of the sentence.

**Definition 14** Let $S_{3.1}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 2** $S_{3.1}' \coloneqq$ <NP, SV, PP>; $S_{3.1}' \subset S$

Relations of $S_{3.1}'$ are composed of $SUBJECT_1$ and $MODIFYV_1$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

*Proof* $S_3 :=$ <NP, V, ADJUNCT> ; $S_3 \subset S$ from Lemma 2

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

Therefore $S_{3.1}' \subset S_3$

$\therefore S_{3.1}' \coloneqq$ <NP, SV, PP> ; $S_{3.1}' \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

**Definition 15** Let ADV be a set of adverbs.

$ADVP = ADV^+$ ; $ADVP \subset ADJUNCT$

**Definition 16** Let $S_{3.2}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 3** $S_{3.2}' \coloneqq$ <NP, SV, ADVP> ; $S_{3.2}' \subset S$

Relations of $S_{3.2}'$ are composed of $SUBJECT_1$ and $MODIFYV_2$

$SUBJECT_1$ : NP → SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_2$ : SV → ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

*Proof* $S_3$ := <NP, V, ADJUNCT> ; $S_3 \subset S$ from Lemma 2

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $S_{3.2'} \subset S_3$

$\therefore$ $S_{3.2'}$ ≔ <NP, SV, ADVP> ; $S_{3.2'} \subset S$

SUBJECT : NP → V from Definition 9

$\therefore$ $SUBJECT_1$ : NP → SV; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

MODIFYV : V → ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV → ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

**Definition 17** Let $S_{3.3'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 4** $S_{3.3'}$ ≔ <NP, SV, ADVP, PP> ; $S_{3.3'} \subset S$

Relations of $S_{3.3'}$ are composed of $SUBJECT_1$, $MODIFYV_2$ and $MODIFYV_1$

$SUBJECT_1$ : NP → SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_2$ : SV → ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$MODIFYV_1$ : SV → PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

*Proof* $S_3$ := <NP, V, ADJUNCT> ; $S_3 \subset S$ from Lemma 2

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

PP $\subset$ ADJUNCT from Definition 6

Therefore $S_{3.3'} \subset S_3$

$\therefore$ $S_{3.3'}$ := <NP, SV, ADV, PP> ; $S_{3.3'} \subset$ S

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset$ MODIFYV

**Definition 18** Let $S_{3.4'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 5** $S_{3.4'}$ := <NP, SV, PP, ADVP> ; $S_{3.4'} \subset$ S

Relations of $S_{3.4'}$ are composed of $SUBJECT_1$, $MODIFYV_1$ and $MODIFYV_2$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset$ MODIFYV

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

*Proof* $S_3$ := <NP, V, ADJUNCT> ; $S_3 \subset$ S from Lemma 2

SV $\subset$ V from Lemma 1

PP $\subset$ ADJUNCT from Definition 6

ADVP $\subset$ ADJUNCT from Definition 15

Therefore $S_{3.4'} \subset S_3$

$\therefore$ $S_{3.4'}$ := <NP, SV, PP, ADVP> ; $S_{3.4'} \subset$ S

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a sentence.

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset$ MODIFYV

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

**Definition 19** Let $S_4$ be a set of sentences in a sentence pattern.

**Lemma 3** $S_4$ := <NP, V, $NP_2$, ADJUNCT> ; $S_4 \subset$ S

*Proof* $S_1$ := <NP, V, $NP_2$> ; $S_1 \subset$ S from Definition 8

V can be added information by ADJUNCT from Definition 5

Therefore $S_4 \subset$ S

$\therefore$ $S_4$ := <NP, V, $NP_2$, ADJUNCT> ; $S_4 \subset$ S

**Definition 20** Let $S_{4.1}{}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 6** $S_{4.1}{}'$ := <NP, SV, $NP_2$, ADVP> ; $S_{4.1}{}' \subset$ S

Relations of $S_{4.1}{}'$ are composed of $SUBJECT_1$, $OBJECT_1$ and $MODIFYV_2$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

*Proof* $S_4$ := <NP, V, $NP_2$, ADJUNCT> ; $S_4 \subset$ S from Lemma 3

SV $\subset$ V from Lemma 1

ADVP $\subset$ ADJUNCT from Definition 15

Therefore $S_{4.1}' \subset S_4$

$\therefore$ $S_{4.1}' \coloneqq$ <NP, SV, $NP_2$, ADVP> ; $S_{4.1}' \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

**Definition 21** Let $S_{4.2}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 7** $S_{4.2}' \coloneqq$ <NP, SV, $NP_2$, ADVP, PP> ; $S_{4.2}' \subset S$

Relations of $S_{4.2}'$ are composed of $SUBJECT_1$, $OBJECT_1$, $MODIFYV_2$ and $MODIFYV_1$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset$ MODIFYV

*Proof* $S_4 \coloneqq$ <NP, V, $NP_2$, ADJUNCT> ; $S_4 \subset S$ from Lemma 3

SV $\subset$ V from Lemma 1

ADVP $\subset$ ADJUNCT from Definition 15

PP $\subset$ ADJUNCT from Definition 6

Therefore $S_{4.2'} \subset S_4$

$\therefore$ $S_{4.2'}$ := <NP, SV, $NP_2$, ADVP_PP> ; $S_{4.2'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset$ MODIFYV

**Definition 22** Let CL be a set of clauses in sentences.

**Definition 23** Let C be a set of conjunctions

Let C_SUB be a set of subordinating conjunctions.

Let C_COR be a set of coordinating conjunctions.

**Lemma 4** C_SUB $\subset$ C

*Proof* $\forall$x (x $\in$ C_SUB $\rightarrow$ x $\in$ C)

$\therefore$ C_SUB $\subset$ C

**Lemma 5** C_COR $\subset$ C

*Proof* $\forall$x (x $\in$ C_COR $\rightarrow$ x $\in$ C)

$\therefore$ C_COR $\subset$ C

**Definition 24** Let CL_SUB be a set of subordinate clauses beginning with a subordinating conjunction and functioning as a adverb.

**Lemma 6** CL_SUB $\subset$ CL

*Proof* $\forall$x (x $\in$ CL_SUB $\rightarrow$ x $\in$ CL)

$\therefore$ CL_SUB $\subset$ CL

**Definition 25** Let CL_COR be a set of coordinate clauses beginning with a coordinating conjunction following another clause.

**Lemma 7** CL_COR $\subset$ CL

*Proof* $\forall x\ (x \in \text{CL\_COR} \rightarrow x \in \text{CL})$

$\therefore$ CL_COR $\subset$ CL

**Definition 26** Let PR be a set of pronouns. PR $\subset$ N

Let PR_REL be a set of relative pronouns. PR_REL $\subset$ PR

**Definition 27** Let CL_REL be a set of relative clauses beginning with a relative pronoun and functioning as a adjective.

**Lemma 8** CL_REL $\subset$ CL

*Proof* $\forall x\ (x \in \text{CL\_REL} \rightarrow x \in \text{CL})$

$\therefore$ CL_REL $\subset$ CL

**Definition 28** Let $S_5$ be a set of sentences in a sentence pattern.

**Lemma 9** $S_5 :=$ <NP, V, CL_SUB> ; $S_5 \subset S$

*Proof* $S_2 :=$ <NP, V> ; $S_2 \subset S$ from Definition 8

$S_2$ can be added information by CL_SUB from Definition 24

Therefore $S_5 \subset S$

$\therefore$ $S_5 :=$ <NP, V, CL_SUB> ; $S_5 \subset S$

**Definition 29** Let COMPLETE be a set of relations from V to CL_SUB

COMPLETE : V $\rightarrow$ CL_SUB ; CL_SUB occur after V in a sequence of the sentence.

**Definition 30** Let $S_{5'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 8** $S_{5'} :=$ <NP, SV, CL_SUB> ; $S_{5'} \subset S$

Relations of $S_{5'}$ are composed of $\text{SUBJECT}_1$ and $\text{COMPLETE}_1$

$\text{SUBJECT}_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $\text{SUBJECT}_1 \subset$ SUBJECT

$\text{COMPLETE}_1$ : SV $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $\text{COMPLETE}_1 \subset$ COMPLETE

*Proof* $S_5 :=$ <NP, V, CL_SUB> ; $S_5 \subset S$ from Lemma 9

$SV \subset V$ from Lemma 1

Therefore $S_{5'} \subset S_5$

$\therefore$ $S_{5'} \coloneqq$ <NP, SV, CL_SUB> ; $S_{5'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

COMPLETE : V $\rightarrow$ CL_SUB from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore$ $COMPLETE_1$ : SV $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

**Definition 31** Let $S_6$ be a set of sentences in a sentence pattern.

**Lemma 10** $S_6 \coloneqq$ <NP, V, $NP_2$, CL> ; $S_6 \subset S$

*Proof* $S_1 :=$ <NP, V, $NP_2$> ; $S_1 \subset S$ from Definition 8

$S_1$ can be added information by CL_SUB or CL_COR

$NP_2$ can be added information by CL_REL

Therefore $S_6 \subset S$

$\therefore$ $S_6 :=$ <NP, V, $NP_2$, CL> ; $S_6 \subset S$

**Definition 32** Let JOIN be a set of relations from V to CL_COR

JOIN : V $\rightarrow$ CL_COR ; CL_COR occur after V in a sequence of the sentence.

Let MODIFYN be a set of relations from NP to CL_REL

MODIFYN : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence.

**Definition 33** Let $S_{6.1'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 9** $S_{6.1'} \coloneqq$ <NP, SV, $NP_2$, CL_SUB> ; $S_{6.1'} \subset S$

Relations of $S_{6.1'}$ are composed of $SUBJECT_1$, $OBJECT_1$ and $COMPLETE_2$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

*Proof* $S_6 := \langle NP, V, NP_2, CL \rangle$ ; $S_6 \subset S$ from Lemma 10

$SV \subset V$ from Lemma 1

$CL\_SUB \subset CL$ from Definition 24

Therefore $S_{6.1}' \subset S_6$

$\therefore S_{6.1}' := \langle NP, SV, NP_2, CL\_SUB \rangle$ ; $S_{6.1}' \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

**Definition 34** Let $S_{6.2}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 10** $S_{6.2}' := \langle NP, SV, NP_2, CL\_COR \rangle$ ; $S_{6.2}' \subset S$

Relations of $S_{6.2}'$ are composed of $SUBJECT_1$, $OBJECT_1$ and $JOIN_1$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$JOIN_1 : SV \rightarrow CL\_SUB$ ; CL_COR occur after SV in a sequence of the sentence and $JOIN_1 \subset JOIN$

*Proof* $S_6 \coloneqq$ <NP, V, $NP_2$, CL> ; $S_6 \subset S$ from Lemma 10

SV $\subset$ V from Lemma 1

CL_COR $\subset$ CL from Lemma 5

Therefore $S_{6.2'} \subset S_6$

$\therefore$ $S_{6.2'} \coloneqq$ <NP, SV, $NP_2$, CL_COR> ; $S_{6.2'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

JOIN : V $\rightarrow$ CL_SUB from Definition 32

CL_COR can occur to join information to sentences.

$\therefore$ $JOIN_1$ : SV $\rightarrow$ CL_SUB ; CL_COR occur after SV in a sequence of the sentence and $JOIN_1 \subset$ JOIN

**Definition 35** Let $S_{6.3'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 11** $S_{6.3'} \coloneqq$ <NP, SV, $NP_2$, CL_REL> ; $S_{6.3'} \subset S$

Relations of $S_{6.3'}$ are composed of $SUBJECT_1$, $OBJECT_1$ and $MODIFYN_2$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$MODIFYN_2$ : $NP_2$ $\rightarrow$ CL_REL ; CL_REL occur after $NP_2$ in a sequence of the sentence and $MODIFYN_2 \subset$ MODIFYN

*Proof* $S_6 \coloneqq$ <NP, V, $NP_2$, CL> ; $S_6 \subset S$ from Lemma 10

SV $\subset$ V from Lemma 1

CL_REL $\subset$ CL from Lemma 6

Therefore $S_{6.3'} \subset S_6$

$\therefore$ $S_{6.3'} \coloneqq$ <NP, SV, $NP_2$, CL_REL> ; $S_{6.3'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_2$ : $NP_2$ $\rightarrow$ CL_REL ; CL_REL occur after $NP_2$ in a sequence of the sentence and $MODIFYN_2 \subset MODIFYN$

**Definition 36** Let $S_7$ be a set of sentences in a sentence pattern.

**Lemma 11** $S_7 \coloneqq$ <NP, V, ADJUNCT, CL> ; $S_7 \subset S$

*Proof* $S_3 :=$ <NP, V, ADJUNCT> ; $S_3 \subset S$ from Lemma 2

$S_3$ can be added information by CL_SUB

ADJUNCT can be PP added information by CL_REL

Therefore $S_7 \subset S$

$\therefore$ $S_7 \coloneqq$ <NP, V, ADJUNCT, CL> ; $S_7 \subset S$

**Definition 37** Let $S_{7.1'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 12** $S_{7.1'} \coloneqq$ <NP, SV, PP, CL_SUB> ; $S_{7.1'} \subset S$

Relations of $S_{7.1'}$ are composed of $SUBJECT_1$, $MODIFYV_1$ and $COMPLETE_1$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

$COMPLETE_1$ : SV $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

*Proof* $S_7 \coloneqq$ <NP, V, ADJUNCT, CL> ; $S_7 \subset S$ from Lemma 11

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

$CL\_SUB \subset CL$ from Lemma 4

Therefore $S_{7.1}{'} \subset S_7$

$\therefore S_{7.1}{'} \coloneqq$ <NP, SV, PP, CL_SUB> ; $S_{7.1}{'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

COMPLETE : V $\rightarrow$ CL_SUB from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore$ $COMPLETE_1$ : SV $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

**Definition 38** Let $S_{7.2}{'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 13** $S_{7.2}{'} \coloneqq$ <NP, SV, PP, CL_REL> ; $S_{7.2}{'} \subset S$

Relations of $S_{7.2}{'}$ are composed of $SUBJECT_1$, $MODIFYV_1$ and $MODIFYN_3$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

$MODIFYN_3$ : $NP_{\text{in PP}}$ $\rightarrow$ CL_REL ; CL_REL occur after PP in a sequence of the sentence and $MODIFYN_3 \subset MODIFYN$

*<u>Proof</u>* $S_7 \coloneqq$ <NP, V, ADJUNCT, CL> ; $S_7 \subset S$ from Lemma 11

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

$CL\_REL \subset CL$ from Lemma 6

Therefore $S_{7.2'} \subset S_7$

$\therefore\ S_{7.2'} \coloneqq \langle NP, SV, PP, CL\_REL \rangle \quad ; \ S_{7.2'} \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore$ $SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

$PP = P \bullet N^+$ and $NP = N^+ \bullet (P \bullet N^+)^*$ from Definition 6

PP can be $P \bullet NP$ when $NP = N^+$

CL_REL can occur to add information to a noun phrase in a prepositional phrase.

$\therefore$ $MODIFYN_3 : NP_{in\ PP} \rightarrow CL\_REL$ ; CL_REL occur after PP in a sequence of the sentence and $MODIFYN_3 \subset MODIFYN$

**Definition 39** Let $S_{7.3'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 14** $S_{7.3'} \coloneqq \langle NP, SV, ADVP, CL\_SUB \rangle \quad ; \ S_{7.3'} \subset S$

Relations of $S_{7.3'}$ are composed of $SUBJECT_1$, $MODIFYV_2$ and $COMPLETE_1$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$COMPLETE_1 : SV \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

*Proof* $S_7 := \langle NP, V, ADJUNCT, CL \rangle$ ; $S_7 \subset S$ from Lemma 11

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

$CL\_SUB \subset CL$ from Lemma 4

Therefore $S_{7.3}' \subset S_7$

$\therefore$ $S_{7.3}' := \langle NP, SV, ADVP, CL\_SUB \rangle$ ; $S_{7.3}' \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

COMPLETE : V $\rightarrow$ CL_SUB from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore$ $COMPLETE_1$ : SV $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

**Definition 40** Let $S_8$ be a set of sentences in a sentence pattern.

**Lemma 12** $S_8 := \langle NP, V, NP_2, ADJUNCT, CL \rangle$ ; $S_8 \subset S$

*Proof* $S_4 := \langle NP, V, NP_2, ADJUNCT \rangle$ ; $S_4 \subset S$ from Lemma 3

$S_4$ can be added information by CL_SUB

ADJUNCT can be PP added information by CL_REL

Therefore $S_8 \subset S$

$\therefore$ $S_8 := \langle NP, V, NP, ADJUNCT, CL \rangle$ ; $S_8 \subset S$

**Definition 41** Let $S_{8.1}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 15** $S_{8.1}' := \langle NP, SV, NP_2, ADVP, CL\_SUB \rangle$ ; $S_{8.1}' \subset S$

Relations of $S_{8.1}'$ are composed of $SUBJECT_1$, $OBJECT_1$, $MODIFYV_2$ and $COMPLETE_2$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

*Proof* $S_8 := \langle NP, V, NP_2, ADJUNCT, CL \rangle$ ; $S_8 \subset S$ from Lemma 12

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

$CL\_SUB \subset CL$ from Lemma 4

Therefore $S_{8.1'} \subset S_8$

$\therefore S_{8.1'} := \langle NP, SV, NP_2, ADVP, CL\_SUB \rangle$ ; $S_{8.1'} \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

**Definition 42** Let $S_{8.2'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 16** $S_{8.2'} :=$ <NP, SV, $NP_2$, ADVP, PP, CL_REL> ; $S_{8.2'} \subset S$

Relations of $S_{8.2'}$ are composed of $SUBJECT_1$, $OBJECT_1$, $MODIFYV_2$, $MODIFYV_1$ and $MODIFYN_3$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

$MODIFYN_3$ : $NP_{\text{in PP}}$ $\rightarrow$ CL_REL ; CL_REL occur after PP in a sequence of the sentence and $MODIFYN_3 \subset$ MODIFYN

*Proof* $S_8 :=$ <NP, V, $NP_2$, ADJUNCT, CL> ; $S_8 \subset S$ from Lemma 12

SV $\subset$ V from Lemma 1

ADVP $\subset$ ADJUNCT from Definition 15

PP $\subset$ ADJUNCT from Definition 6

CL_REL $\subset$ CL from Lemma 6

Therefore $S_{8.2'} \subset S_8$

$\therefore$ $S_{8.2'} :=$ <NP, SV, $NP_2$, ADVP, PP, CL_REL> ; $S_{8.2'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

$PP = P \bullet N^+$ and $NP = N^+ \bullet (P \bullet N^+)^*$ from Definition 6

PP can be $P \bullet NP$ when $NP = N^+$

CL_REL can occur to add information to a noun phrase in a prepositional phrase.

$\therefore$ $MODIFYN_3 : NP_{in\ PP} \rightarrow CL\_REL$ ; CL_REL occur after PP in a sequence of the sentence and $MODIFYN_3 \subset MODIFYN$

**Definition 43** Let $S_9$ be a set of sentences in a sentence pattern.

**Lemma 13** $S_9 \coloneqq \langle NP, CL\_REL, V, NP_2 \rangle$ ; $S_9 \subset S$

*Proof* $S_1 := \langle NP, V, NP_2 \rangle$ ; $S_1 \subset S$ from Definition 8

NP can be added information by CL_REL

Therefore $S_9 \subset S$

$\therefore$ $S_9 \coloneqq \langle NP, CL\_REL, V, NP_2 \rangle$ ; $S_9 \subset S$

**Definition 44** Let $S_{9'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 17** $S_{9'} \coloneqq \langle NP, CL\_REL, SV, NP_2 \rangle$ ; $S_{9'} \subset S$

Relations of $S_{9'}$ are composed of $SUBJECT_1$, $MODIFYN_1$ and $OBJECT_1$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

*Proof* $S_9 \coloneqq$ <NP, CL_REL, V, $NP_2$>; $S_9 \subset S$ from Lemma 13

SV $\subset$ V from Lemma 1

Therefore $S_9{}' \subset S_9$

$\therefore$ $S_9{}' \coloneqq$ <NP, CL_REL , SV, $NP_2$> ; $S_9{}' \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset$ MODIFYN

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

**Definition 45** Let $S_{10}$ be a set of sentences in a sentence pattern.

**Lemma 14** $S_{10} \coloneqq$ <NP, CL_REL, V, ADJUNCT> ; $S_{10} \subset S$

*Proof* $S_3 \coloneqq$ <NP, V, ADJUNCT> ; $S_3 \subset S$ from Lemma 2

NP can be added information by CL_REL

Therefore $S_{10} \subset S$

$\therefore$ $S_{10} \coloneqq$ <NP, CL_REL, V, ADJUNCT> ; $S_{10} \subset S$

**Definition 46** Let $S_{10.1}{}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 18** $S_{10.1}{}' \coloneqq$ <NP, CL_REL , SV, PP> ; $S_{10}{}' \subset S$

Relations of $S_{10.1}{}'$ are composed of $SUBJECT_1$, $MODIFYN_1$ and $MODIFYV_1$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset$ MODIFYN

$MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

*Proof* $S_{10} := \langle NP, CL\_REL, V, ADJUNCT \rangle$; $S_{10} \subset S$ from Lemma 14

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

Therefore $S_{10.1}' \subset S_{10}$

$\therefore S_{10.1}' := \langle NP, CL\_REL , SV, ADJUNCT \rangle$ ; $S_{10.1}' \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

**Definition 47** Let $S_{10.2}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 19** $S_{10.2}' := \langle NP, CL\_REL , SV, ADVP \rangle$ ; $S_{10.2}' \subset S$

Relations of $S_{10}'$ are composed of $SUBJECT_1$, $MODIFYN_1$ and $MODIFYV_2$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

*Proof* $S_{10} := \langle NP, CL\_REL, V, ADJUNCT\rangle$; $S_{10} \subset S$ from Lemma 14

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $S_{10.2'} \subset S_{10}$

$\therefore S_{10'} := \langle NP, CL\_REL, SV, ADJUNCT\rangle$ ; $S_{10'} \subset S$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

**Definition 48** Let $S_{11}$ be a set of sentences in a sentence pattern.

**Lemma 15** $S_{11} := \langle NP, CL\_REL, V, NP_2, ADJUNCT\rangle$ ; $S_{11} \subset S$

*Proof* $S_4 := \langle NP, V, NP_2, ADJUNCT\rangle$ ; $S_4 \subset S$ from Lemma 3

NP can be added information by CL_REL

Therefore $S_{11} \subset S$

$\therefore S_{11} := \langle NP, CL\_REL, V, NP_2, ADJUNCT\rangle$ ; $S_{11} \subset S$

**Definition 49** Let $S_{11'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 20** $S_{11'} := \langle NP, CL\_REL, SV, NP_2, ADVP\rangle$ ; $S_{11'} \subset S$

Relations of $S_{11'}$ are composed of $MODIFYN_1$, $SUBJECT_1$ and $OBJECT_1$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset$ MODIFYN

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

*Proof* $S_{11}$ := <NP, CL_REL, V, $NP_2$, ADJUNCT> ; $S_{11} \subset$ S from Lemma 15

SV $\subset$ V from Lemma 1

ADVP $\subset$ ADJUNCT from Definition 15

Therefore $S_{11}' \subset S_{11}$

$\therefore$ $S_{11}'$ := <NP, CL_REL , SV, $NP_2$, ADVP> ; $S_{11}' \subset$ S

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset$ MODIFYN

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

**Definition 50** Let $S_{12}$ be a set of sentences in a sentence pattern.

**Lemma 16** $S_{12}$ := <NP, CL_REL, V, CL_SUB> ; $S_{12} \subset$ S

*Proof* $S_5$ := <NP, V, CL_SUB> ; $S_5 \subset$ S from Lemma 9

NP can be added information by CL_REL

Therefore $S_{12} \subset S$

$\therefore\ S_{12} := \langle NP, CL\_REL, V, CL\_SUB \rangle$ ; $S_{12} \subset S$

**Definition 51** Let $S_{12}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 21** $S_{12}' := \langle NP, CL\_REL, SV, CL\_SUB \rangle$ ; $S_{12}' \subset S$

Relations of $S_{12}'$ are composed of $SUBJECT_1$, $MODIFYN_1$ and $MODIFYV_2$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$COMPLETE_1 : SV \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

*Proof* $S_{12} := \langle NP, CL\_REL, V, CL\_SUB \rangle$; $S_{12} \subset S$ from Lemma 16

$SV \subset V$ from Lemma 1

Therefore $S_{12}' \subset S_{12}$

$\therefore\ S_{12}' := \langle NP, CL\_REL, SV, CL\_SUB \rangle$ ; $S_{12}' \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore\ SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore\ MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore\ COMPLETE_1 : SV \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

**Definition 52** Let $S_{13}$ be a set of sentences in a sentence pattern.

**Lemma 17** $S_{13} := \langle NP, CL\_REL, V, NP_2, CL \rangle$ ; $S_{13} \subset S$

*Proof* $S_6 := \langle NP, V, NP_2, CL \rangle$ ; $S_6 \subset S$ from Lemma 10

NP can be added information by CL_REL

Therefore $S_{13} \subset S$

$\therefore S_{13} := \langle NP, CL\_REL, V, NP_2, CL \rangle$ ; $S_{13} \subset S$

**Definition 53** Let $S_{13.1}{'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 22** $S_{13.1}{'} := \langle NP, CL\_REL, SV, NP_2, CL\_SUB \rangle$; $S_{13.1}{'} \subset S$

Relations of $S_{13.1}{'}$ are composed of $SUBJECT_1$, $MODIFYN_1$, $OBJECT_1$ and $COMPLETE_2$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

*Proof* $S_{13} := \langle NP, CL\_REL, V, NP_2, CL \rangle$ ; $S_{13} \subset S$ from Lemma 17

$SV \subset V$ from Lemma 1

$CL\_SUB \subset CL$ from Lemma 4

Therefore $S_{13.1}{'} \subset S_{13}$

$\therefore S_{13.1}{'} := \langle NP, CL\_REL, SV, NP_2, CL\_SUB \rangle$ ; $S_{13.1}{'} \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore$ $OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore$ $COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

**Definition 54** Let $S_{13.2}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 23** $S_{13.2}' := \langle NP, CL\_REL , SV, NP_2, CL\_REL \rangle$ ; $S_{13.2}' \subset S$

Relations of $S_{13.2}'$ are composed of $SUBJECT_1$, $MODIFYN_1$, $OBJECT_1$ and $MODIFYN_2$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$MODIFYN_2 : NP_2 \rightarrow CL\_REL$ ; CL_REL occur after $NP_2$ in a sequence of the sentence and $MODIFYN_2 \subset MODIFYN$

*Proof* $S_{13} := \langle NP, CL\_REL, V, NP_2, CL \rangle$ ; $S_{13} \subset S$ from Lemma 17

$SV \subset V$ from Lemma 1

$CL\_REL \subset CL$ from Lemma 6

Therefore $S_{13.2}' \subset S_{13}$

$\therefore$ $S_{13.2}' := \langle NP, CL\_REL , SV, NP_2, CL\_REL \rangle$ ; $S_{13.2}' \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore$ $SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset$ MODIFYN

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_2$ : $NP_2$ $\rightarrow$ CL_REL ; CL_REL occur after $NP_2$ in a sequence of the sentence and $MODIFYN_2 \subset$ MODIFYN

**Definition 55** Let $S_{14}$ be a set of sentences in a sentence pattern.

**Lemma 18** $S_{14}$ := <NP, CL_REL, V, $NP_2$, ADJUNCT, CL> ; $S_{14} \subset S$

*Proof* $S_8$ := <NP, V, $NP_2$, ADJUNCT, CL> ; $S_8 \subset S$ from Lemma 12

NP can be added information by CL_REL

Therefore $S_{14} \subset S$

$\therefore$ $S_{14}$ := <NP, CL_REL, V, $NP_2$, ADJUNCT, CL>; $S_{14} \subset S$

**Definition 56** Let $S_{14'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 24** $S_{14'}$ := <NP, CL_REL , SV, $NP_2$, ADVP, CL_SUB>; $S_{14'} \subset S$

Relations of $S_{14'}$ are composed of $SUBJECT_1$, $MODIFYN_1$, $OBJECT_1$, $MODIFYV_2$ and $COMPLETE_2$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset$ SUBJECT

$MODIFYN_1$ : NP $\rightarrow$ CL_REL ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset$ MODIFYN

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset$ MODIFYV

$COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

*Proof* $S_{14} := \langle NP, CL\_REL, V, NP_2, ADJUNCT, CL \rangle$; $S_{14} \subset S$ from Lemma 18

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

$CL\_SUB \subset CL$ from Lemma 4

Therefore $S_{14}' \subset S_{14}$

$\therefore S_{14}' := \langle NP, CL\_REL, SV, NP_2, ADVP, CL\_SUB \rangle$; $S_{14}' \subset S$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore MODIFYN_1 : NP \rightarrow CL\_REL$ ; CL_REL occur after NP in a sequence of the sentence and $MODIFYN_1 \subset MODIFYN$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

**Definition 57** Let CL_N be a set of noun clauses beginning with a noun and functioning as a noun.

**Lemma 19** $CL\_N \subset CL$

*Proof* $\forall x\ (x \in CL\_N \rightarrow x \in CL)$

$\therefore\ CL\_N \subset CL$

**Definition 58** Let $S_{15}$ be a set of sentences in a sentence pattern.

**Lemma 20** $S_{15} := \langle CL\_N, V \rangle$ ; $S_{15} \subset S$

*Proof* $S_2 := \langle NP, V \rangle$ ; $S_2 \subset S$ from Definition 8

CL_N can function as NP from Definition 55

Therefore $S_{15} \subset S$

$\therefore\ S_{15} := \langle CL\_N, V \rangle$ ; $S_{15} \subset S$

**Definition 59** Let SUBJECT_CL be a set of relations from CL_N to V

$SUBJECT\_CL : CL\_N \rightarrow V$ ; CL_N occur before V in a sequence of the sentence.

**Definition 60** Let $S_{15'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 25** $S_{15'} := \langle CL\_N, SV \rangle$ ; $S_{15'} \subset S$

Relations of $S_{15'}$ are composed of $SUBJECT\_CL_1$

$SUBJECT\_CL_1 : CL\_N \rightarrow SV$ ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

*Proof* $S_{15} := \langle CL\_N, V \rangle$ ; $S_{15} \subset S$ from Lemma 20

$SV \subset V$ from Lemma 1

Therefore $S_{15'} \subset S_{15}$

$\therefore\ S_{15'} := \langle CL\_N, SV \rangle$ ; $S_{15'} \subset S$

$SUBJECT\_CL : CL\_N \rightarrow V$ from Definition 59

$\therefore\ SUBJECT\_CL_1 : CL\_N \rightarrow SV$ ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

**Lemma 21** $S_{16} := \langle CL\_N, V, NP_2 \rangle$ ; $S_{16} \subset S$

*Proof* $S_1 := \langle NP, V, NP_2\rangle$ ; $S_1 \subset S$ from Definition 8

CL_N can function as NP from Definition 55

Therefore $S_{16} \subset S$

$\therefore S_{16} := \langle CL\_N, V, NP_2\rangle$ ; $S_{16} \subset S$

**Definition 61** Let $S_{16'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 26** $S_{16'} \coloneqq \langle CL\_N, SV, NP_2\rangle$ ; $S_{16'} \subset S$

Relations of $S_{16'}$ are composed of $SUBJECT\_CL_1$ and $OBJECT_2$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

*Proof* $S_{16} := \langle CL\_N, V, NP_2\rangle$ ; $S_{16} \subset S$ from Lemma 21

$SV \subset V$ from Lemma 1

Therefore $S_{16'} \subset S_{16}$

$\therefore S_{16'} \coloneqq \langle CL\_N, SV, NP_2\rangle$ ; $S_{16'} \subset S$

SUBJECT_CL : CL_N $\rightarrow$ V from Definition 59

$\therefore SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

**Definition 62** Let $S_{17}$ be a set of sentences in a sentence pattern.

**Lemma 22** $S_{17} \coloneqq \langle CL\_N, V, ADJUNCT\rangle$ ; $S_{17} \subset S$

*Proof* $S_{15} \coloneqq \langle CL\_N, V\rangle$ ; $S_{15} \subset S$ from Lemma 20

V can be added information by ADJUNCT from Definition 5

Therefore $S_{17} \subset S$

$\therefore S_{17} \coloneqq \langle NP, V, ADJUNCT\rangle$ ; $S_{17} \subset S$

**Definition 63** Let $S_{17.1}{}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 27** $S_{17.1}{}' :=$ <CL_N, SV, PP> ; $S_{17.1}{}' \subset S$

Relations of $S_{17.1}{}'$ are composed of $SUBJECT\_CL_1$ and $MODIFYV_1$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

*Proof* $S_{17} :=$ <CL_N, V, ADJUNCT> ; $S_{17} \subset S$ from Lemma 22

SV $\subset$ V from Lemma 1

PP $\subset$ ADJUNCT from Definition 6

Therefore $S_{17.1}{}' \subset S_{17}$

$\therefore$ $S_{17.1}{}' :=$ <CL_N, SV, PP>; $S_{17.1}{}' \subset S$

SUBJECT_CL : CL_N $\rightarrow$ V from Definition 59

$\therefore$ $SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

**Definition 64** Let $S_{7.2}{}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 28** $S_{17.2}{}' :=$ <CL_N, SV, ADVP> ; $S_{17.2}{}' \subset S$

Relations of $S_{17.2}{}'$ are composed of $SUBJECT\_CL_1$ and $MODIFYV_2$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

*Proof* $S_{17} :=$ <NP, V, ADJUNCT> ; $S_{17} \subset S$ from Lemma 22

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $S_{17.2'} \subset S_{17}$

$\therefore$ $S_{17.2'} :=$ <CL_N, SV, ADVP> ; $S_{17.2'} \subset S$

SUBJECT_CL : CL_N $\rightarrow$ V from Definition 59

$\therefore$ $SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

**Definition 65** Let $S_{18}$ be a set of sentences in a sentence pattern.

**Lemma 23** $S_{18} :=$ <CL_N, V, $NP_2$, ADJUNCT> ; $S_{18} \subset S$

*Proof* $S_{16} :=$ <CL_N, V, $NP_2$> ; $S_{16} \subset S$ from Lemma 21

V can be added information by ADJUNCT from Definition 5

Therefore $S_{18} \subset S$

$\therefore$ $S_{18} :=$ <CL_N, V, $NP_2$, ADJUNCT> ; $S_{18} \subset S$

**Definition 64** Let $S_{18'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 29** $S_{18'} :=$ <CL_N, V, $NP_2$, ADVP> ; $S_{18'} \subset S$

Relations of $S_{18'}$ are composed of $SUBJECT\_CL_1$, $OBJECT_1$ and $MODIFYV_2$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

*Proof* $S_{18} :=$ <CL_N, V, $NP_2$, ADJUNCT>; $S_{18} \subset S$ from Lemma 23

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $S_{18}' \subset S_{18}$

$\therefore S_{18}' := \langle CL\_N, V, NP_2, ADVP \rangle$ ; $S_{18}' \subset S$

SUBJECT_CL : CL_N $\rightarrow$ V from Definition 59

$\therefore$ $SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the sentence and $MODIFYV_2 \subset MODIFYV$

**Definition 67** Let $S_{19}$ be a set of sentences in a sentence pattern.

**Lemma 24** $S_{19} := \langle CL\_N, V, NP_2, CL \rangle$ ; $S_{19} \subset S$

*Proof* $S_{16} := \langle CL\_N, V, NP_2 \rangle$ ; $S_{16} \subset S$ from Lemma 21

$S_{16}$ can be added information by CL_SUB

$NP_2$ can be added information by CL_REL

Therefore $S_{19} \subset S$

$\therefore S_{19} := \langle CL\_N, V, NP_2, CL \rangle$ ; $S_{19} \subset S$

**Definition 68** Let $S_{19}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 30** $S_{19}' := \langle CL\_N, SV, NP_2, CL\_SUB \rangle$ ; $S_{19}' \subset S$

Relations of $S_{19}'$ are composed of $SUBJECT\_CL_1$, $OBJECT_1$ and $COMPLETE_2$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

*Proof* $S_{19} := \langle CL\_N, V, NP_2, CL \rangle$ ; $S_{19} \subset S$ from Lemma 24

$SV \subset V$ from Lemma 1

$CL\_SUB \subset CL$ from Definition 24

Therefore $S_{19}' \subset S_{19}$

$\therefore S_{19}' := \langle CL\_N, SV, NP_2, CL\_SUB \rangle$ ; $S_{19}' \subset S$

$SUBJECT\_CL : CL\_N \rightarrow V$ from Definition 59

$\therefore SUBJECT\_CL_1 : CL\_N \rightarrow SV$ ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

**Definition 69** Let $S_{20}$ be a set of sentences in a sentence pattern.

**Lemma 25** $S_{20} := \langle CL\_N, CL\_REL, V, NP_2 \rangle$ ; $S_{20} \subset S$

*Proof* $S_{16} := \langle CL\_N, V, NP \rangle$ ; $S_{16} \subset S$ from Lemma 21

NP in CL_N can be added information by CL_REL

Therefore $S_{20} \subset S$

$\therefore S_{20} := \langle CL\_N, CL\_REL, V, NP_2 \rangle$ ; $S_{20} \subset S$

**Definition 70** Let MODIFYN_CL be a set of relations from NP in CL_N to CL_REL

$MODIFYN\_CL : NP_{in\ CL\_N} \rightarrow CL\_REL$ ; CL_REL occur after CL_N in a sequence of the sentence.

**Definition 71** Let $S_{20}'$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 31** $S_{20'} := \langle CL\_N, CL\_REL, SV, NP_2 \rangle$ ; $S_{20'} \subset S$

Relations of $S_{20'}$ are composed of $SUBJECT\_CL_1$ $MODIFYN\_CL_1$, and $OBJECT_1$

$SUBJECT\_CL_1 : CL\_N \rightarrow SV$ ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$MODIFYN\_CL_1 : NP_{\text{in } CL\_N} \rightarrow CL\_REL$ ; CL_REL occur after CL_N in a sequence of the sentence and $MODIFYN\_CL_1 \subset MODIFYN\_CL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

*Proof* $S_{20} := \langle CL\_N, CL\_REL, V, NP_2 \rangle$ ; $S_{20} \subset S$ from Lemma 25

$SV \subset V$ from Lemma 1

Therefore $S_{20'} \subset S_{20}$

$\therefore S_{20'} := \langle CL\_N, CL\_REL, SV, NP_2 \rangle$ ; $S_{20'} \subset S$

$SUBJECT\_CL : CL\_N \rightarrow V$ from Definition 59

$\therefore SUBJECT\_CL_1 : CL\_N \rightarrow SV$ ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$MODIFYN\_CL : NP_{\text{in } CL\_N} \rightarrow CL\_REL$ from Definition 70

CL_REL can occur to add information to a noun phrase in a noun clause.

$\therefore MODIFYN\_CL_1 : NP_{\text{in } CL\_N} \rightarrow CL\_REL$ ; CL_REL occur after CL_N in a sequence of the sentence and $MODIFYN\_CL_1 \subset MODIFYN\_CL$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

**Definition 72** Let $S_{21}$ be a set of sentences in a sentence pattern.

**Lemma 26** $S_{21} := \langle CL\_N, CL\_REL, V, ADJUNCT \rangle$ ; $S_{21} \subset S$

*Proof* $S_{17} := \langle CL\_N, V, ADJUNCT \rangle$ ; $S_{17} \subset S$ from Lemma 22

CL_N can be added information by CL_REL

Therefore $S_{21} \subset S$

$\therefore S_{21} := \langle CL\_N, CL\_REL, V, ADJUNCT \rangle$; $S_{21} \subset S$

**Definition 73** Let $S_{21'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

**Theorem 32** $S_{21'} :=$ <CL_N, CL_REL , SV, PP> ; $S_{21'} \subset S$

Relations of $S_{10'}$ are composed of $SUBJECT\_CL_1$, $MODIFYN\_CL_1$ and $MODIFYV_1$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

$MODIFYN\_CL_1$ : $NP_{in\ CL\_N} \rightarrow$ CL_REL ; CL_REL occur after CL_N in a sequence of the sentence and $MODIFYN\_CL_1 \subset MODIFYN\_CL$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

*<u>Proof</u>* $S_{21} :=$ <CL_N, CL_REL, V, ADJUNCT>; $S_{21} \subset S$ from Lemma 26

SV $\subset$ V from Lemma 1

PP $\subset$ ADJUNCT from Definition 6

Therefore $S_{21'} \subset S_2 1$

$\therefore$ $S_{21'} :=$ <CL_N, CL_REL , SV, PP>; $S_{21'} \subset S$

SUBJECT_CL : CL_N $\rightarrow$ V from Definition 59

$\therefore$ $SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset SUBJECT\_CL$

MODIFYN_CL : $NP_{in\ CL\_N} \rightarrow$ CL_REL from Definition 70

CL_REL can occur to add information to a noun phrase in a noun clause.

$\therefore$ $MODIFYN\_CL_1$ : $NP_{in\ CL\_N} \rightarrow$ CL_REL ; CL_REL occur after CL_N in a sequence of the sentence and $MODIFYN\_CL_1 \subset MODIFYN\_CL$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the sentence and $MODIFYV_1 \subset MODIFYV$

**Definition 74** Let $S_{22}$ be a set of sentences in a sentence pattern.

**Lemma 27** $S_{22} :=$ <CL_N, CL_REL, V, $NP_2$, CL>; $S_{22} \subset S$

*Proof* $S_{19} \coloneqq$ <CL_N, V, NP, CL> ; $S_{19} \subset S$ from Lemma 24

CL_N can be added information by CL_REL

Therefore $S_{22} \subset S$

$\therefore$ $S_{22} \coloneqq$ <CL_N, CL_REL, V, $NP_2$, CL> ; $S_{22} \subset S$

Definition 75 Let $S_{22'}$ be a set of sentences with two consecutive serial verbs in a sentence pattern.

Theorem 33 $S_{22'} \coloneqq$ <CL_N, CL_REL, SV, $NP_2$, CL_SUB>; $S_{22'} \subset S$

Relations of $S_{22'}$ are composed of $SUBJECT\_CL_1$, $MODIFYN\_CL_1$, $OBJECT_1$ and $COMPLETE_2$

$SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset$ SUBJECT_CL

$MODIFYN\_CL_1$ : $NP_{\text{in CL\_N}}$ $\rightarrow$ CL_REL ; CL_REL occur after CL_N in a sequence of the sentence and $MODIFYN\_CL_1 \subset$ MODIFYN_CL

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset$ OBJECT

$COMPLETE_2$ : $SV_{NP}$ $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset$ COMPLETE

*Proof* $S_{22} \coloneqq$ <CL_N, CL_REL, V, $NP_2$, CL>; $S_{22} \subset S$ from Lemma 27

SV $\subset$ V from Lemma 1

CL_SUB $\subset$ CL from Lemma 4

Therefore $S_{22'} \subset S_{22}$

$\therefore$ $S_{22'} \coloneqq$ <CL_N, CL_REL, SV, $NP_2$, CL_SUB>; $S_{22'} \subset S$

SUBJECT_CL : CL_N $\rightarrow$ V from Definition 59

$\therefore$ $SUBJECT\_CL_1$ : CL_N $\rightarrow$ SV ; CL_N occur before SV in a sequence of the sentence and $SUBJECT\_CL_1 \subset$ SUBJECT_CL

MODIFYN_CL : $NP_{\text{in CL\_N}}$ $\rightarrow$ CL_REL from Definition 70

CL_REL can occur to add information to a noun phrase in a noun clause.

$\therefore$ $MODIFYN\_CL_1$ : $NP_{\text{in CL\_N}}$ $\rightarrow$ CL_REL ; CL_REL occur after CL_N in a sequence of the sentence and $MODIFYN\_CL_1 \subset$ MODIFYN_CL

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

COMPLETE : V $\rightarrow$ CL_SUB from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore$ $COMPLETE_2$ : $SV_{NP}$ $\rightarrow$ CL_SUB ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_2 \subset COMPLETE$

**Definition 76** Let AUX is a set of auxiliaries.

$AUXP = AUX^+$

AUXP occur before V in a sentence

**Theorem 34** AUXP can occur before SV in all sentence patterns with two consecutive serial verbs.

*Proof* $SV \subset V$ from Lemma 1

AUXP occur before V in a sentence from Definition 76

$\therefore$ AUXP can occur before SV in all sentence patterns with two consecutive serial verbs.

**Definition 77** Let $NP_{REL}$ be a set of noun phrases which occur before relative clauses in sentences.

**Definition 78** Let $NCL\_REL_1$, $NCL\_REL_2$ be a set of relative clauses and noun phrases which occur before relative clauses in different clause patterns.

$NCL\_REL_1 \coloneqq < NP_{REL}, PR\_REL, V, NP_2>$

$NCL\_REL_2 \coloneqq < NP_{REL}, PR\_REL, V>$

**Definition 79** Let SUBJECT_REL be a set of relations from $NP_{REL}$ to V

SUBJECT_REL : $NP_{REL}$ $\rightarrow$ V ; $NP_{REL}$ occur before the relative clause in a sentence.

**Definition 80** Let $NCL\_REL_1'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 35** $NCL\_REL_1' \coloneqq < NP_{REL}, PR\_REL, SV, NP_2>$

Relations of $NCL\_REL_1'$ are composed of $SUBJECT\_REL_1$ and $OBJECT_1$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

*Proof* $NCL\_REL_1 := < NP_{REL}, PR\_REL, V, NP_2>$ from Definition 78

$SV \subset V$ from Lemma 1

Therefore $NCL\_REL_1{}' \subset NCL\_REL_1$

$\therefore NCL\_REL_1{}' := < NP_{REL}, PR\_REL, SV, NP_2>$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset$ SUBJECT_REL

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

**Definition 81** Let $NCL\_REL_2{}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 36** $NCL\_REL_2{}' := < NP_{REL}, PR\_REL, SV>$

Relations of $NCL\_REL_2{}'$ are composed of $SUBJECT\_REL_1$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

*Proof* $NCL\_REL_2 := < NP_{REL}, PR\_REL, V>$ from Definition 78

$SV \subset V$ from Lemma 1

Therefore $NCL\_REL_2{}' \subset NCL\_REL_2$

$\therefore NCL\_REL_2{}' := < NP_{REL}, PR\_REL, SV>$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

**Definition 82** Let $NCL\_REL_3$ be a set of relative clauses and noun phrases which occur before relative clause in a clause pattern.

**Lemma 28** $NCL\_REL_3 \coloneqq \langle NP_{REL}, PR\_REL, V, ADJUNCT \rangle$

*Proof* $NCL\_REL_2 := \langle NP_{REL}, PR\_REL, V \rangle$ from Definition 78

V can be added information by ADJUNCT from Definition 5

$\therefore$ $NCL\_REL_3 := \langle NP_{REL}, PR\_REL, V, ADJUNCT \rangle$

**Definition 84** Let $NCL\_REL_{3.1}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 37** $NCL\_REL_{3.1}' \coloneqq \langle NP_{REL}, PR\_REL, SV, PP \rangle$

Relations of $NCL\_REL_{3.1}'$ are composed of $SUBJECT\_REL_1$ and $MODIFYV_1$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset MODIFYV$

*Proof* $NCL\_REL_3 := \langle NP_{REL}, PR\_REL, V, ADJUNCT \rangle$ from Lemma 28

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

Therefore $NCL\_REL_{3.1}' \subset NCL\_REL_3$

$\therefore$ $NCL\_REL_{3.1}' \coloneqq \langle NP_{REL}, PR\_REL, SV, PP \rangle$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore$ $SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

PP can occur to add information to verb in a relative clause.

$\therefore$ $MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset MODIFYV$

**Definition 85** Let $NCL\_REL_{3.2'}$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 38** $NCL\_REL_{3.2'} := < NP_{REL}, PR\_REL, SV, ADVP>$

Relations of $NCL\_REL_{3.2'}$ are composed of $SUBJECT\_REL_1$ and $MODIFYV_2$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

*Proof* $NCL\_REL_3 := <NP_{REL}, PR\_REL, V, ADJUNCT>$ from Lemma 28

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $NCL\_REL_{3.2'} \subset NCL\_REL_3$

$\therefore NCL\_REL_{3.2'} := < NP_{REL}, PR\_REL, SV, ADVP>$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a relative clause.

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

**Definition 86** Let $NCL\_REL_{3.3'}$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 39** $NCL\_REL_{3.3'} := < NP_{REL}, PR\_REL, SV, ADVP, PP>$

Relations of $NCL\_REL_{3.3'}$ are composed of $SUBJECT\_REL_1$, $MODIFYV_2$ and $MODIFYV_1$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; NP occur before SV in a sequence of the relative clause and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$MODIFYV_2$ : SV → ADVP ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV → PP ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset$ MODIFYV

*Proof* $NCL\_REL_3$ := <$NP_{REL}$, PR_REL, V, ADJUNCT> from Lemma 28

SV $\subset$ V from Lemma 1

ADVP $\subset$ ADJUNCT from Definition 15

PP $\subset$ ADJUNCT from Definition 6

Therefore $NCL\_REL_{3.3}' \subset NCL\_REL_3$

$\therefore$ $NCL\_REL_{3.3}'$ := < $NP_{REL}$, PR_REL, SV, ADVP, PP>

SUBJECT_REL : $NP_{REL}$ → V from Definition 79

$\therefore$ $SUBJECT\_REL_1$ : $NP_{REL}$ → SV ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset$ SUBJECT_REL

MODIFYV : V → ADJUNCT from Definition 13

ADVP can occur to add information to verb in a relative clause.

PP can occur to add information to verb in a relative clause.

$\therefore$ $MODIFYV_2$ : SV → ADVP ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV → PP ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset$ MODIFYV

**Definition 87** Let $NCL\_REL_{3.4}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 40** $NCL\_REL_{3.4}'$ := < $NP_{REL}$, PR_REL, SV, PP, ADVP>

Relations of $NCL\_REL_{3.4}'$ are composed of $SUBJECT\_REL_1$, $MODIFYV_1$ and $MODIFYV_2$

$SUBJECT\_REL_1$ : $NP_{REL}$ → SV ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset$ SUBJECT_REL

$MODIFYV_1$ : SV → PP ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset$ MODIFYV

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

*Proof* $NCL\_REL_3 := <NP_{REL}, PR\_REL, V, ADJUNCT>$ from Lemma 28

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $NCL\_REL_{3.4}' \subset NCL\_REL_3$

$\therefore NCL\_REL_{3.4}' := < NP_{REL}, PR\_REL, SV, PP, ADVP>$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

PP can occur to add information to verb in a relative clause.

ADVP can occur to add information to verb in a relative clause.

$\therefore MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset MODIFYV$

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

**Definition 88** Let $NCL\_REL_4$ be a set of relative clauses and noun phrases which occur before relative clause in a clause pattern.

**Lemma 29** $NCL\_REL_4 := <NP_{REL}, PR\_REL, V, NP_2, ADJUNCT>$

*Proof* $NCL\_REL_1 := < NP_{REL}, PR\_REL, V, NP_2>$ from Definition 78

V can be added information by ADJUNCT from Definition 5

$\therefore NCL\_REL_4 := <NP_{REL}, PR\_REL, V, NP_2, ADJUNCT>$

**Definition 89** Let $NCL\_REL_{4.1}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 41** $NCL\_REL_{4.1}' := < NP_{REL}, PR\_REL, SV, NP_2, ADVP>$

Relations of $NCL\_REL_{4.1}'$ are composed of $SUBJECT\_REL_1$, $OBJECT_1$ and $MODIFYV_2$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

*Proof* $NCL\_REL_4 \coloneqq \langle NP_{REL}, PR\_REL, V, NP_2, ADJUNCT \rangle$ from Lemma 29

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $NCL\_REL_{4.1}' \subset NCL\_REL_4$

$\therefore NCL\_REL_{4.1}' \coloneqq \langle NP_{REL}, PR\_REL, SV, NP_2, ADVP \rangle$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a relative clause.

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

**Definition 90** Let $NCL\_REL_{4.2}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 42** $NCL\_REL_{4.2}' \coloneqq \langle NP_{REL}, PR\_REL, SV, NP_2, ADVP, PP \rangle$

Relations of $NCL\_REL_{4.2}'$ are composed of $SUBJECT\_REL_1$, $OBJECT_1$, $MODIFYV_2$ and $MODIFYV_1$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

$MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset MODIFYV$

*Proof* $NCL\_REL_4 := \langle NP_{REL}, PR\_REL, V, NP_2, ADJUNCT \rangle$ from Lemma 29

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

$PP \subset ADJUNCT$ from Definition 6

Therefore $NCL\_REL_{4.2'} \subset NCL\_REL_4$

$\therefore NCL\_REL_{4.2'} := \langle NP_{REL}, PR\_REL, SV, NP_2, ADVP, PP \rangle$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a relative clause.

PP can occur to add information to verb in a relative clause.

$\therefore MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the relative clause and $MODIFYV_2 \subset MODIFYV$

$MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the relative clause and $MODIFYV_1 \subset MODIFYV$

**Definition 91** Let $NCL\_REL_5$ be a set of relative clauses and noun phrases which occur before relative clause in a clause pattern.

**Lemma 30** $NCL\_REL_5 := \langle NP_{REL}, PR\_REL, V, CL\_SUB \rangle$

*Proof* $NCL\_REL_2 := \langle NP_{REL}, PR\_REL, V \rangle$ from Definition 78

CL_REL can be added information by CL_SUB

$\therefore$ $NCL\_REL_5 \coloneqq <NP_{REL}, PR\_REL, V, CL\_SUB>$

**Definition 92** Let $NCL\_REL_5{'}$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 43** $NCL\_REL_5{'} \coloneqq < NP_{REL}, PR\_REL, SV, CL\_SUB>$

Relations of $NCL\_REL_5{'}$ are composed of $SUBJECT\_REL_1$ and $COMPLETE_1$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$COMPLETE_1 : SV \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the sentence and $COMPLETE_1 \subset COMPLETE$

*Proof* $NCL\_REL_5 \coloneqq <NP_{REL}, PR\_REL, V, CL\_SUB>$ from Lemma 30

$SV \subset V$ from Lemma 1

Therefore $NCL\_REL_5{'} \subset NCL\_REL_5$

$\therefore$ $NCL\_REL_5{'} \coloneqq < NP_{REL}, PR\_REL, SV, CL\_SUB>$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore$ $SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a sentence.

$\therefore$ $COMPLETE_1 : SV \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the relative clause and $COMPLETE_1 \subset COMPLETE$

**Definition 93** Let $NCL\_REL_6$ be a set of relative clauses and noun phrases which occur before relative clause in a clause pattern.

**Lemma 31** $NCL\_REL_6 \coloneqq <NP_{REL}, PR\_REL, V, NP_2, CL>$

*Proof* $NCL\_REL_1 \coloneqq < NP_{REL}, PR\_REL, V, NP_2>$ from Definition 78

CL_REL can be added information by CL_SUB

$\therefore$ $NCL\_REL_6 \coloneqq <NP_{REL}, PR\_REL, V, NP_2, CL>$

**Definition 94** Let $NCL\_REL_{6.1}{}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 44** $NCL\_REL_{6.1}{}' := \langle NP_{REL}, PR\_REL, SV, NP_2, CL\_SUB \rangle$

Relations of $NCL\_REL_{6.1}{}'$ are composed of $SUBJECT\_REL_1$, $OBJECT_1$ and $COMPLETE_2$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the relative clause and $COMPLETE_2 \subset COMPLETE$

*Proof* $NCL\_REL_6 := \langle NP_{REL}, PR\_REL, V, NP_2, CL \rangle$ from Lemma 31

$SV \subset V$ from Lemma 1

$CL\_SUB \subset CL$ from Definition 24

Therefore $NCL\_REL_{6.1}{}' \subset NCL\_REL_6$

$\therefore NCL\_REL_{6.1}{}' := \langle NP_{REL}, PR\_REL, SV, NP_2, CL\_SUB \rangle$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$COMPLETE : V \rightarrow CL\_SUB$ from Definition 29

CL_SUB can occur to complete information in a relative clause.

$\therefore COMPLETE_2 : SV_{NP} \rightarrow CL\_SUB$ ; CL_SUB occur after SV in a sequence of the relative clause and $COMPLETE_2 \subset COMPLETE$

**Definition 95** Let $NCL\_REL_{6.2}{}'$ be a set of relative clauses and noun phrases which occur before the relative clause with two consecutive serial verbs in a clause pattern.

**Theorem 45** $NCL\_REL_{6.2'} := \langle NP_{REL}, PR\_REL, SV, NP_2, CL\_REL \rangle$

Relations of $NCL\_REL_{6.2'}$ are composed of $SUBJECT\_REL_1$, $OBJECT_1$ and $MODIFYN_2$

$SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$MODIFYN_2 : NP_2 \rightarrow CL\_REL$ ; CL_REL occur after $NP_2$ in a sequence of the relative clause and $MODIFYN_2 \subset MODIFYN$

*Proof* $NCL\_REL_6 := \langle NP_{REL}, PR\_REL, V, NP_2, CL \rangle$ from Lemma 31

$SV \subset V$ from Lemma 1

$CL\_REL \subset CL$ from Lemma 6

Therefore $NCL\_REL_{6.2'} \subset NCL\_REL_6$

$\therefore NCL\_REL_{6.2'} := \langle NP_{REL}, PR\_REL, SV, NP_2, CL\_REL \rangle$

$SUBJECT\_REL : NP_{REL} \rightarrow V$ from Definition 79

$\therefore SUBJECT\_REL_1 : NP_{REL} \rightarrow SV$ ; $NP_{REL}$ occur before the relative clause with two consecutive serial verbs and $SUBJECT\_REL_1 \subset SUBJECT\_REL$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the relative clause and $OBJECT_1 \subset OBJECT$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore MODIFYN_2 : NP_2 \rightarrow CL\_REL$ ; CL_REL occur after $NP_2$ in a sequence of the relative clause and $MODIFYN_2 \subset MODIFYN$

**Definition 96** Let $CL\_SUB_1$, $CL\_SUB_2$ be a set of subordinate clauses in different clause patterns.

$CL\_SUB_1 := \langle C\_SUB, V, NP_2 \rangle$ ; $CL\_SUB_1 \subset CL\_SUB$

$CL\_SUB_2 := \langle C\_SUB, V \rangle$ ; $CL\_SUB_2 \subset CL\_SUB$

**Definition 97** Let $CL\_SUB_1{'}$ be a set of subordinate clauses with two consecutive serial verbs in a clause pattern.

**Theorem 46** $CL\_SUB_1{'} \coloneqq$ <C_SUB, SV, $NP_2$> ; $CL\_SUB_1{'} \subset CL\_SUB$

Relations of $CL\_SUB_1{'}$ are composed of $OBJECT_1$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

*Proof* $CL\_SUB_1 \coloneqq$ <C_SUB, V, $NP_2$> from Definition 96

$SV \subset V$ from Lemma 1

Therefore $CL\_SUB_1{'} \subset CL\_SUB_1$

$\therefore$ $CL\_SUB_1{'} \coloneqq$ <C_SUB, SV, $NP_2$>; $CL\_SUB_1{'} \subset CL\_SUB$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore$ $OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

**Definition 98** Let $CL\_SUB_3$ be a set of subordinate clauses in a clause pattern.

**Lemma 32** $CL\_SUB_3 \coloneqq$ <C_SUB, V, ADJUNCT> ; $CL\_SUB_3 \subset CL\_SUB$

*Proof* $CL\_SUB_2 :=$ <C_SUB, V> from Definition 96

V can be added information by ADJUNCT from Definition 5

Therefore $CL\_SUB_3 \subset CL\_SUB$

$\therefore$ $CL\_SUB_3 :=$ <C_SUB, V, ADJUNCT> ; $CL\_SUB_3 \subset CL\_SUB$

**Definition 99** Let $CL\_SUB_{3.1}{'}$ be a set of subordinate clauses with two consecutive serial verbs in a clause pattern.

**Theorem 47** $CL\_SUB_{3.1}{'} \coloneqq$ <C_SUB, SV, PP> ; $CL\_SUB_{3.1}{'} \subset CL\_SUB$

Relations of $CL\_SUB_{3.1}{'}$ are composed of $MODIFYV_1$

$MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the subordinate clause and $MODIFYV_1 \subset MODIFYV$

*Proof* $CL\_SUB_3 :=$ <C_SUB, V, ADJUNCT> from Lemma 32

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

Therefore $CL\_SUB_{3.1}{'} \subset CL\_SUB_3$

$\therefore$ $CL\_SUB_{3.1'}$ $:=$ <C_SUB, SV, PP> ; $CL\_SUB_{3.1'} \subset$ CL_SUB

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a subordinate clause.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the subordinate clause and $MODIFYV_1 \subset$ MODIFYV

**Definition 100** Let $CL\_SUB_{3.2'}$ be a set of subordinate clauses with two consecutive serial verbs in a clause pattern.

**Theorem 48** $CL\_SUB_{3.2'}$ $:=$ <C_SUB, SV, ADVP, PP> ; $CL\_SUB_{3.2'} \subset$ CL_SUB

Relations of $NCL\_REL_{3.2'}$ are composed of $MODIFYV_2$ and $MODIFYV_1$

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the subordinate clause and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the subordinate clause and $MODIFYV_1 \subset$ MODIFYV

*Proof* $CL\_SUB_3$ := <C_SUB, V, ADJUNCT> from Lemma 33

SV $\subset$ V from Lemma 1

ADVP $\subset$ ADJUNCT from Definition 15

PP $\subset$ ADJUNCT from Definition 6

Therefore $CL\_SUB_{3.2'} \subset CL\_SUB_3$

$\therefore$ $CL\_SUB_{3.2'}$ $:=$ <C_SUB, SV, ADVP, PP> ; $CL\_SUB_{3.2'} \subset$ CL_SUB

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a subordinate clause.

PP can occur to add information to verb in a subordinate clause.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the subordinate clause and $MODIFYV_2 \subset$ MODIFYV

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the subordinate clause and $MODIFYV_1 \subset$ MODIFYV

**Definition 101** Let $CL\_SUB_4$ be a set of subordinate clauses in a clause pattern.

**Lemma 33** $CL\_SUB_4$ $:=$ <C_SUB, V, $NP_2$, ADJUNCT> ; $CL\_SUB_4 \subset$ CL_SUB

*Proof* $CL\_SUB_1$ $:=$ <C_SUB, V, $NP_2$> from Definition 96

V can be added information by ADJUNCT from Definition 5

Therefore $CL\_SUB_4 \subset CL\_SUB$

$\therefore$ $CL\_SUB_4 := \langle C\_SUB, V, NP_2, ADJUNCT\rangle; CL\_SUB_4 \subset CL\_SUB$

**Definition 102** Let $CL\_SUB_4{}'$ be a set of subordinate clauses with two consecutive serial verbs in a clause pattern.

**Theorem 49** $CL\_SUB_4{}' := \langle C\_SUB, SV, NP_2, ADVP\rangle$ ; $CL\_SUB_4{}' \subset CL\_SUB$

Relations of $CL\_SUB_4{}'$ are composed of $OBJECT_1$ and $MODIFYV_2$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

$MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the subordinate clause and $MODIFYV_2 \subset MODIFYV$

*Proof* $CL\_SUB_4 := \langle C\_SUB, V, NP_2, ADJUNCT\rangle$ from Lemma 33

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $CL\_SUB_4{}' \subset CL\_SUB_4$

$\therefore$ $CL\_SUB_4{}' := \langle C\_SUB, SV, NP_2, ADVP\rangle$ ; $CL\_SUB_4{}' \subset CL\_SUB$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore$ $OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

ADVP can occur to add information to verb in a subordinate clause.

$\therefore$ $MODIFYV_2 : SV \rightarrow ADVP$ ; ADVP occur after SV in a sequence of the subordinate clause and $MODIFYV_2 \subset MODIFYV$

**Definition 103** Let $CL\_SUB_5$ be a set of subordinate clauses in a clause pattern.

**Lemma 34** $CL\_SUB_5 := \langle C\_SUB, V, NP_2, CL\rangle$ ; $CL\_SUB_5 \subset CL\_SUB$

*Proof* $CL\_SUB_1 := \langle C\_SUB, V, NP_2\rangle$ from Definition 96

$NP_2$ can be added information by CL_REL

Therefore $CL\_SUB_5 \subset CL\_SUB$

$\therefore$ $CL\_SUB_5 := \langle C\_SUB, V, NP_2, CL\rangle$ ; $CL\_SUB_5 \subset CL\_SUB$

**Definition 104** Let $CL\_SUB_{5'}$ be a set of subordinate clauses with two consecutive serial verbs in a clause pattern.

**Theorem 50** $CL\_SUB_{5'} \coloneqq \langle C\_SUB, SV, NP_2, CL\_REL \rangle$ ; $CL\_SUB_{5'} \subset CL\_SUB$

Relations of $CL\_SUB_{5'}$ are composed of $OBJECT_1$ and $MODIFYN_2$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

$MODIFYN_2 : NP_2 \rightarrow CL\_REL$ ; CL_REL occur after $NP_2$ in a sequence of the subordinate clause and $MODIFYN_2 \subset MODIFYN$

*Proof* $CL\_SUB_5 \coloneqq \langle C\_SUB, V, NP_2, CL \rangle$ from Lemma 34

$SV \subset V$ from Lemma 1

$CL\_REL \subset CL$ from Lemma 6

Therefore $CL\_SUB_{5'} \subset CL\_SUB_5$

$\therefore$ $CL\_SUB_{5'} \coloneqq \langle C\_SUB, SV, NP_2, CL\_REL \rangle$ ; $CL\_SUB_{5'} \subset CL\_SUB$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore$ $OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

CL_REL can occur to add information to a noun phrase.

$\therefore$ $MODIFYN_2 : NP_2 \rightarrow CL\_REL$ ; CL_REL occur after $NP_2$ in a sequence of the subordinate clause and $MODIFYN_2 \subset MODIFYN$

**Definition 105** Let $CL\_SUB_6$ be a set of subordinate clauses in a clause pattern.

**Lemma 35** $CL\_SUB_6 \coloneqq \langle C\_SUB, V, ADJUNCT, CL \rangle$ ; $CL\_SUB_6 \subset CL\_SUB$

*Proof* $CL\_SUB_3 \coloneqq \langle C\_SUB, V, ADJUNCT \rangle$ from Lemma 32

ADJUNCT can be PP added information by CL_REL

Therefore $CL\_SUB_6 \subset CL\_SUB$

$\therefore$ $CL\_SUB_6 \coloneqq \langle C\_SUB, V, ADJUNCT, CL \rangle$ ; $CL\_SUB_6 \subset CL\_SUB$

**Definition 106** Let $CL\_SUB_{6'}$ be a set of subordinate clauses with two consecutive serial verbs in a clause pattern.

**Theorem 51** $CL\_SUB_{6'} \coloneqq \langle C\_SUB, SV, PP, CL\_REL \rangle$ ; $CL\_SUB_{6'} \subset CL\_SUB$

Relations of $S_6{}'$ are composed of $MODIFYV_1$ and $MODIFYN_3$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the subordinate clause and $MODIFYV_1 \subset$ MODIFYV

$MODIFYN_3$ : $NP_{in\ PP} \rightarrow$ CL_REL ; CL_REL occur after PP in a sequence of the subordinate clause and $MODIFYN_3 \subset$ MODIFYN

*Proof* $CL\_SUB_6 \coloneqq$ <C_SUB, V, ADJUNCT, CL> from Lemma 35

SV $\subset$ V from Lemma 1

PP $\subset$ ADJUNCT from Definition 6

CL_REL $\subset$ CL from Lemma 6

Therefore $CL\_SUB_6{}' \subset CL\_SUB_6$

$\therefore$ $CL\_SUB_6{}' \coloneqq$ <C_SUB, SV, PP, CL_REL> ; $CL\_SUB_6{}' \subset$ CL_SUB

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a subordinate clause.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the subordinate clause and $MODIFYV_1 \subset$ MODIFYV

MODIFYN : NP $\rightarrow$ CL_REL from Definition 32

$PP = P \bullet N^+$ and $NP = N^+ \bullet (P \bullet N^+)^*$ from Definition 6

PP can be $P \bullet NP$ when $NP = N^+$

CL_REL can occur to add information to a noun phrase in a prepositional phrase.

$\therefore$ $MODIFYN_3$ : $NP_{in\ PP} \rightarrow$ CL_REL ; CL_REL occur after PP in a sequence of the subordinate clause and $MODIFYN_3 \subset$ MODIFYN

**Definition 107** Let $CL\_N_1$, $CL\_N_2$ be a set of noun clauses in different clause patterns.

$CL\_N_1 \coloneqq$ <NP, V, $NP_2$> ; $CL\_N_1 \subset$ CL_N

$CL\_N_2 \coloneqq$ <NP, V> ; $CL\_N_2 \subset$ CL_N

**Definition 108** Let $CL\_N_1{}'$ be a set of noun clauses with two consecutive serial verbs in a clause pattern.

**Theorem 52** $CL\_N_1{}' \coloneqq$ <NP, SV, $NP_2$> ; $CL\_N_1{}' \subset$ CL_N

Relations of $CL\_N_1{}'$ are composed of $SUBJECT_1$ and $OBJECT_1$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

$OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in a sequence of the noun clause and $OBJECT_1 \subset OBJECT$

*Proof* $CL\_N_1 := <NP, V, NP_2>$ from Definition 107

$SV \subset V$ from Lemma 1

Therefore $CL\_N_1{}' \subset CL\_N_1$

$\therefore CL\_N_1{}' := <NP, SV, NP_2>$ ; $CL\_N_1{}' \subset CL\_N$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

$OBJECT : V \rightarrow NP$ from Definition 9

$\therefore OBJECT_1 : SV \rightarrow NP_2$ ; $NP_2$ occur after SV in occur after SV in a sequence of the subordinate clause and $OBJECT_1 \subset OBJECT$

**Definition 109** Let $CL\_N_3$ be a set of noun clauses in a clause pattern.

**Lemma 36** $CL\_N_3 := <NP, V, ADJUNCT>$ ; $CL\_N_3 \subset CL\_N$

*Proof* $CL\_N_2 := <NP, V>$ from Definition 107

V can be added information by ADJUNCT from Definition 5

Therefore $CL\_N_3 \subset CL\_N$

$\therefore CL\_N_3 := <NP, V, ADJUNCT>$ ; $CL\_N_3{}' \subset CL\_N$

**Definition 110** Let $CL\_N_3{}'$ be a set of noun clauses with two consecutive serial verbs in a clause pattern.

**Theorem 53** $CL\_N_3{}' := <NP, SV, PP>$ ; $CL\_N_3{}' \subset CL\_N$

Relations of $CL\_N_3{}'$ are composed of $SUBJECT_1$ and $MODIFYV_1$

$SUBJECT_1 : NP \rightarrow SV$ ; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the noun clause and $MODIFYV_1 \subset MODIFYV$

*Proof* $CL\_N_3 := <NP, V, ADJUNCT>$ from Lemma 36

$SV \subset V$ from Lemma 1

$PP \subset ADJUNCT$ from Definition 6

Therefore $CL\_N_{3'} \subset CL\_N_3$

$\therefore$ $CL\_N_{3'} := <NP, SV, PP>$ ; $CL\_N_{3'} \subset CL\_N$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

PP can occur to add information to verb in a subordinate clause.

$\therefore$ $MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the noun clause and $MODIFYV_1 \subset MODIFYV$

**Definition 111** Let $CL\_N_4$ be a set of noun clauses in a clause pattern.

**Lemma 37** $CL\_N_4 := <NP, V, NP_2, ADJUNCT>$ ; $CL\_N_4 \subset CL\_N$

*Proof* $CL\_N_1 := <NP, V, NP_2>$ from Definition 107

V can be added information by ADJUNCT from Definition 5

Therefore $CL\_N_4 \subset CL\_N$

$\therefore$ $CL\_N_4 := <NP, V, NP_2, ADJUNCT>$ ; $CL\_N_4 \subset CL\_N$

**Definition 112** Let $CL\_N_{4'}$ be a set of noun clauses with two consecutive serial verbs in a clause pattern.

**Theorem 54** $CL\_N_{4'} := <NP, SV, NP_2, ADVP>$ ; $CL\_N_{4'} \subset CL\_N$

Relations of $CL\_N_{4'}$ are composed of $SUBJECT_1$, $OBJECT_1$ and $MODIFYV_2$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

$OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the sentence and $OBJECT_1 \subset OBJECT$

$MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the noun clause and $MODIFYV_2 \subset MODIFYV$

*Proof* $CL\_N_4 := <NP, V, NP_2, ADJUNCT>$ from Lemma 37

$SV \subset V$ from Lemma 1

$ADVP \subset ADJUNCT$ from Definition 15

Therefore $CL\_N_{4'} \subset CL\_N_4$

$\therefore$ $CL\_N_{4'} := \langle NP, SV, NP_2, ADVP \rangle$ ; $CL\_N_{4'} \subset CL\_N$

SUBJECT : NP $\rightarrow$ V from Definition 9

$\therefore$ $SUBJECT_1$ : NP $\rightarrow$ SV; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

OBJECT : V $\rightarrow$ NP from Definition 9

$\therefore$ $OBJECT_1$ : SV $\rightarrow$ $NP_2$ ; $NP_2$ occur after SV in a sequence of the noun clause and $OBJECT_1 \subset OBJECT$

MODIFYV : V $\rightarrow$ ADJUNCT from Definition 13

ADVP can occur to add information to verb in a sentence.

$\therefore$ $MODIFYV_2$ : SV $\rightarrow$ ADVP ; ADVP occur after SV in a sequence of the noun clause and $MODIFYV_2 \subset MODIFYV$

**Definition 113** Let $CL\_N_5$ be a set of noun clauses in a clause pattern.

**Lemma 38** $CL\_N_5 := \langle NP, V, ADJUNCT, CL\_REL \rangle$ ; $CL\_N_5 \subset CL\_N$

*Proof* $CL\_N_3 := \langle NP, V, ADJUNCT \rangle$ from Lemma 36

ADJUNCT can be PP added information by CL_REL

Therefore $CL\_N_5 \subset CL\_N$

$\therefore$ $CL\_N_5 := \langle NP, V, ADJUNCT, CL\_REL \rangle$ ; $CL\_N_5 \subset CL\_N$

**Definition 114** Let $CL\_N_{5'}$ be a set of noun clauses with two consecutive serial verbs in a clause pattern.

**Theorem 55** $CL\_N_{5'} := \langle NP, SV, PP, CL\_REL \rangle$ ; $CL\_N_{5'} \subset CL\_N$

Relations of $CL\_N_{5'}$ are composed of $SUBJECT_1$, $MODIFYV_1$ and $MODIFYN_3$

$SUBJECT_1$ : NP $\rightarrow$ SV ; NP occur before SV in a sequence of the noun clause and $SUBJECT_1 \subset SUBJECT$

$MODIFYV_1$ : SV $\rightarrow$ PP ; PP occur after SV in a sequence of the noun clause and $MODIFYV_1 \subset MODIFYV$

$MODIFYN_3 : NP_{in\ PP} \rightarrow CL\_REL$ ; CL_REL occur after PP in a sequence of the noun clause and $MODIFYN_3 \subset MODIFYN$

*Proof* $CL\_N_5 := \langle NP, V, NP_2, CL\_REL \rangle$ from Lemma 38

$SV \subset V$ from Lemma 1

Therefore $CL\_N_5' \subset CL\_N_5$

$\therefore CL\_N_5' := \langle NP, SV, NP_2, CL\_REL \rangle$ ; $CL\_N_5' \subset CL\_N$

$SUBJECT : NP \rightarrow V$ from Definition 9

$\therefore SUBJECT_1 : NP \rightarrow SV$; NP occur before SV in a sequence of the sentence and $SUBJECT_1 \subset SUBJECT$

$MODIFYV : V \rightarrow ADJUNCT$ from Definition 13

$\therefore MODIFYV_1 : SV \rightarrow PP$ ; PP occur after SV in a sequence of the noun clause and $MODIFYV_1 \subset MODIFYV$

$MODIFYN : NP \rightarrow CL\_REL$ from Definition 32

$PP = P \bullet N^+$ and $NP = N^+ \bullet (P \bullet N^+)^*$ from Definition 6

PP can be $P \bullet NP$ when $NP = N^+$

CL_REL can occur to add information to a noun phrase in a prepositional phrase.

CL_REL can occur to add information to a noun phrase.

$\therefore MODIFYN_3 : NP_{in\ PP} \rightarrow CL\_REL$ ; CL_REL occur after PP in a sequence of the noun clause and $MODIFYN_3 \subset MODIFYN$

**Theorem 56** AUXP can occur before SV in all clause patterns with two consecutive serial verbs.

*Proof* $SV \subset V$ from Lemma 1

AUXP occur before V in a clause from Definition 76

$\therefore$ AUXP can occur before SV in all clause patterns with two consecutive serial verbs.

# APPENDIX D

# MEANING OF SEMANTIC BINDING PATTERNS

## D.1 Meaning of Semantic Binding Patterns.

Subject: NP
Verb: SV
Object NP

Agent: subject
Action: SV
Patient: object

Agent: subject(PER)
Action: SV
Patient: object

Agent: subject
Action: SV
Patient: object
Instrument :subject

Subject NP
Verb: SV
Modifyv1: PP

Agent subject
Action: SV
Location: modifyv1

Agent subject
Action: SV
Coordination: modifyv1

Agent subject
Action: SV
Purpose: modifyv1

Agent subject
Action: SV
Means: modifyv1

Agent subject
Action: SV
Source or Destination: modifyv1

Subject: NP
Verb: SV
Modifyv2 : ADVP

Agent subject
Action: SV
Manner: modifyv2

Agent subject
Action: SV
Cooperating or Separating: modifyv2

Agent subject
Action: SV
Time or Frequency: modifyv2

Agent subject
Action: SV
Quantity: modifyv2

Subject: NP
Verb: SV
Object NP
Modifyv2 : ADVP

Agent subject
Action: SV
Patient: object
Manner: modifyv2

Agent subject
Action: SV
Patient: object
Cooperating or Separating: modifyv2

Agent subject
Action: SV
Patient: object
Time or Frequency: modifyv2

Agent subject
Action: SV
Patient: object
Quantity: modifyv2

Subject: NP
Verb: SV
Modifyv1 : PP
Modifyv2 : ADVP

Agent subject
Action: SV
Location: modifyv1
Time or Frequency: modifyv2

Agent subject
Action: SV
Coordination: modifyv1
Manner: modifyv2

Agent subject
Action: SV
Purpose: modifyv1

Agent subject
Action: SV
Means: modifyv1
Time or Frequency: modifyv2

Agent subject
Action: SV
Source or Destination: modifyv1
Time or Frequency: modifyv2

Figure D.1: Meaning of semantic binding patterns.

Subject: NP
Verb: SV
Modifyv2: ADVP
Modifyv1 : PP

Agent subject
Action SV
Manner modifyv2
Location modifyv1

Agent subject
Action SV
Manner modifyv2
Source or Destination modifyv1

Agent subject
Action SV
Cooperating or Separating: modifyv2
Location modifyv1

Agent subject
Action SV
Cooperating or Separating: modifyv2
Source or Destination modifyv1

Agent subject
Action SV
Time or Frequency: modifyv2
Purpose: modifyv1

Agent subject
Action SV
Time or Frequency: modifyv2
Coordination modifyv1

Agent subject
Action SV
Quantity. modifyv2
Location modifyv1

Agent subject
Action SV
Quantity. modifyv2
Source or Destination modifyv1

Subject NP
Verb : SV
Object: NP
Modifyv2: ADVP
Modifyv1: PP

Agent: subject
Action SV
Patient object
Manner modifyv2
Location modifyv1

Agent: subject
Action SV
Patient object
Manner modifyv2
Source or Destination modifyv1

Agent: subject
Action SV
Patient object
Cooperating or Separating modifyv2
Location modifyv1

Agent: subject
Action SV
Patient object
Cooperating or Separating modifyv2
Source or Destination modifyv1

Agent: subject
Action SV
Patient object
Time or Frequency : modifyv2
Purpose: modifyv1

Agent: subject
Action SV
Patient object
Time or Frequency : modifyv2
Coordination modifyv1

Agent: subject
Action SV
Patient object
Quantity. modifyv2
Location modifyv1

Agent: subject
Action SV
Patient object
Quantity. modifyv2
Source or Destination modifyv1

Figure D.1: Meaning of semantic binding patterns (cont.).

Subject: NP
Verb: SV
Object NP
Modifyv2: ADVP

Agent subject
Action: SV
Manner: modifyv2
Instrument :subject

Agent subject
Action: SV
Cooperating or Separating modifyv2
Instrument :subject

Agent subject
Action: SV
Patient object
Time or Frequency: modifyv2
Instrument :subject

Agent subject
Action: SV
Patient object
Quantity modifyv2
Instrument :subject

Subject NP
Verb: SV
Object NP
Modifyv2: ADVP

Agent subject(person)
Action: SV
Manner: modifyv2

Agent subject(person)
Action: SV
Patient: object
Time or Frequency: modifyv2

Agent subject(person)
Action: SV
Patient: object
Quantity modifyv2

Subject: NP
Verb: SV
Object NP
Modifyv2: ADVP
Modifyv1 : PP

Agent subject
Action: SV
Patient object
Instrument :subject
Manner: modifyv2
Location: modifyv1

Agent subject
Action: SV
Patient object
Instrument :subject
Cooperating or Separating modifyv2
Location: modifyv1

Agent subject
Action: SV
Patient object
Instrument :subject
Cooperating or Separating modifyv2
Source or Destination: modifyv1

Agent subject
Action: SV
Patient object
Instrument :subject
Quantity modifyv2
Location: modifyv1

Subject NP
Verb: SV
Object NP
Modifyv2: ADVP
Modifyv1 : PP

Agent subject(person)
Action: SV
Patient: object
Manner: modifyv2
Location: modifyv1

Agent subject(person)
Action: SV
Patient: object
Manner: modifyv2
Source or Destination: modifyv1

Agent subject(person)
Action: SV
Patient: object
Time or Frequency: modifyv2
Purpose: modifyv1

Agent subject(person)
Action: SV
Patient: object
Quantity modifyv2
Location: modifyv1

Figure D.1: Meaning of semantic binding patterns (cont.).

Subject: NP
Verb: SV
Modifyv1 : PP

Agent subject(person)
Action SV
Location modifyv1

Agent subject(person)
Action SV
Coordination modifyv1

Agent subject(person)
Action SV
Means modifyv1

Agent subject(person)
Action SV
Source or Destination modifyv1

Subject: NP
Verb: SV
Modifyv2 : ADVP

Agent subject(person)
Action SV
Manner modifyv2

Agent subject(person)
Action SV
Cooperating or Separating modifyv2

Subject: NP
Verb: SV
Modifyv2 : PP
Modifyv1 : ADVP

Agent subject(person)
Action SV
Location modifyv1
Time or Frequency modifyv2

Agent subject(person)
Action SV
Means modifyv1
Time or Frequency modifyv2

Subject: NP
Verb: SV

Agent -
Action SV
Patient subject

Subject NP
Verb: SV
Modifyv1 : PP

Agent -
Action SV
Patient subject
Source or Destination modifyv1

Subject: NP
Verb: SV
Modifyv2: ADVP

Agent: -
Action SV
Patient subject
Manner modifyv2

Agent: -
Action SV
Patient subject
Cooperating or Separating: modifyv2

Agent: -
Action SV
Patient subject
Time or Frequency modifyv2

Agent: -
Action SV
Patient subject
Quantity modifyv2

Subject: NP
Verb: SV
Modifyv2: ADVP
Modifyv1 : PP

Agent -
Action SV
Patient subject
Quantity modifyv2
Location modifyv1

Subject: NP
Verb: SV
Complete1 : Subordinating Clause

Agent: subject(person)
Action SV

Figure D.1: Meaning of semantic binding patterns (cont.).

# APPENDIX E
# TWO CONSECUTIVE THAI SERIAL VERBS

## E.1 Two Consecutive Thai Serial Verbs

Some two consecutive Thai SVCs,which its first element is the top ten most found serial verbs in the ORCHID corpus, are shown in Table E.1.

Table E.1: Two consecutive Thai serial verbs in the ORCHID corpus.

| | 1st verb | 2nd verb | | 1st verb | 2nd verb | | 1st verb | 2nd verb |
|---|---|---|---|---|---|---|---|---|
| 1 | ใช้ | ออกแบบ | 39 | ใช้ | ลาก | 77 | ทำหน้าที่ | เชื่อมต่อ |
| 2 | ใช้ | แสดง | 40 | ใช้ | ติดต่อ | 78 | ทำหน้าที่ | อ่าน |
| 3 | ใช้ | ทด | 41 | ใช้ | ประมวลผล | 79 | ทำหน้าที่ | ประมวลผล |
| 4 | ใช้ | อธิบาย | 42 | ใช้ | ถ่ายทอด | 80 | ทำหน้าที่ | ผลิต |
| 5 | ใช้ | สอน | 43 | ใช้ | สนับสนุน | 81 | ทำหน้าที่ | ถอดรหัส |
| 6 | ใช้ | พิจารณา | 44 | ใช้ | ตรวจรับ | 82 | ทำหน้าที่ | พิจารณา |
| 7 | ใช้ | ควบคุม | 45 | ใช้ | ตรวจจับ | 83 | ทำหน้าที่ | ปล่อย |
| 8 | ใช้ | บันทึก | 46 | ใช้ | แลกเปลี่ยน | 84 | ทำหน้าที่ | จัด |
| 9 | ใช้ | บริการ | 47 | ใช้ | ดำ | 85 | ทำหน้าที่ | แสดงผล |
| 10 | ใช้ | วัด | 48 | ใช้ | เปลี่ยน | 86 | ทำหน้าที่ | แสดง |
| 11 | ใช้ | ดู | 49 | ใช้ | บรรจุ | 87 | ทำหน้าที่ | แทน |
| 12 | ใช้ | แก้ไข | 50 | ใช้ | ประเมินผล | 88 | ทำหน้าที่ | วิเคราะห์ |
| 13 | ใช้ | ค้น | 51 | ใช้ | เรียก | 89 | ทำหน้าที่ | สังเคราะห์ |
| 14 | ใช้ | พิมพ์ | 52 | ใช้ | เลือก | 90 | ทำหน้าที่ | รวม |
| 15 | ใช้ | ส่ง | 53 | ใช้ | ต่อ | 91 | ทำหน้าที่ | แก้ไข |
| 16 | ใช้ | รับส่ง | 54 | ใช้ | ทดแทน | 92 | ทำหน้าที่ | ลด |
| 17 | ใช้ | ทำ | 55 | ใช้ | กำหนด | 93 | ทำหน้าที่ | พิมพ์ |
| 18 | ใช้ | แสดงผล | 56 | ทำหน้าที่ | ประสานงาน | 94 | ทำหน้าที่ | ติดตั้ง |
| 19 | ใช้ | สลับ | 57 | ทำหน้าที่ | ฝึกอบรม | 95 | ทำหน้าที่ | เชื่อมโยง |
| 20 | ใช้ | กำหนด | 58 | ทำหน้าที่ | เก็บ | 96 | ทำหน้าที่ | สั่งงาน |
| 21 | ใช้ | ป้อน | 59 | ทำหน้าที่ | เชื่อม | 97 | ทำหน้าที่ | วิเคราะห์ |
| 22 | ใช้ | สื่อสาร | 60 | ทำหน้าที่ | หา | 98 | ต้องการ | เรียน |
| 23 | ใช้ | เก็บ | 61 | ทำหน้าที่ | ติดต่อ | 99 | ต้องการ | เพิ่ม |
| 24 | ใช้ | ตรวจสอบ | 62 | ทำหน้าที่ | จัดการ | 100 | ต้องการ | ใช้ |
| 25 | ใช้ | จัดการ | 63 | ทำหน้าที่ | สร้าง | 101 | ต้องการ | ยกเลิก |
| 26 | ใช้ | ตรวจเช็ค | 64 | ทำหน้าที่ | เปลี่ยน | 102 | ต้องการ | ค้นหา |
| 27 | ใช้ | อ่าน | 65 | ทำหน้าที่ | แปลง | 103 | ต้องการ | ดู |
| 28 | ใช้ | ทดสอบ | 66 | ทำหน้าที่ | แยก | 104 | ต้องการ | ปรับแต่ง |
| 29 | ใช้ | วิเคราะห์ | 67 | ทำหน้าที่ | สลับ | 105 | ต้องการ | วัด |
| 30 | ใช้ | สร้าง | 68 | ทำหน้าที่ | ปรับ | 106 | ต้องการ | ควบคุม |
| 31 | ใช้ | แยก | 69 | ทำหน้าที่ | แปล | 107 | ต้องการ | พัฒนา |
| 32 | ใช้ | รักษา | 70 | ทำหน้าที่ | ควบคุม | 108 | ต้องการ | แก้ไข |
| 33 | ใช้ | บอก | 71 | ทำหน้าที่ | รับ | 109 | ต้องการ | หา |
| 34 | ใช้ | พัฒนา | 72 | ทำหน้าที่ | ตรวจสอบ | 110 | ต้องการ | พิมพ์ |
| 35 | ใช้ | ปลูก | 73 | ทำหน้าที่ | ตีความ | 111 | ต้องการ | สร้าง |
| 36 | ใช้ | ตั้ง | 74 | ทำหน้าที่ | ขยาย | 112 | ต้องการ | ถ่าย |
| 37 | ใช้ | เชื่อม | 75 | ทำหน้าที่ | หาร | 113 | ต้องการ | ส่ง |
| 38 | ใช้ | ช่วย | 76 | ทำหน้าที่ | เลือก | 114 | ต้องการ | เลิก |

Table E.1: Two consecutive Thai serial verbs in the ORCHID corpus (cont.).

| | 1st verb | 2nd verb |
|---|---|---|
| 115 | ต้องการ | เคลื่อน |
| 116 | ต้องการ | ใช้งาน |
| 117 | ต้องการ | ติดต่อ |
| 118 | ต้องการ | เปิด |
| 119 | ต้องการ | ติดตั้ง |
| 120 | ต้องการ | เก็บ |
| 121 | ต้องการ | ปรึกษา |
| 122 | ต้องการ | สนทนา |
| 123 | ต้องการ | เข้าสู่ |
| 124 | ต้องการ | แลกเปลี่ยน |
| 125 | ต้องการ | ศึกษา |
| 126 | ต้องการ | กระตุ้น |
| 127 | ต้องการ | เพิ่ม |
| 128 | ช่วย | ออกแบบ |
| 129 | ช่วย | ลด |
| 130 | ช่วย | สนับสนุน |
| 131 | ช่วย | พิจารณา |
| 132 | ช่วย | สอน |
| 133 | ช่วย | ประหยัด |
| 134 | ช่วย | สรุป |
| 135 | ช่วย | เสริม |
| 136 | ช่วย | ทำให้ |
| 137 | ช่วย | แบ่งเบา |
| 138 | ช่วย | สร้าง |
| 139 | ช่วย | วิเคราะห์ |
| 140 | ช่วย | กำหนด |
| 141 | ช่วย | แก้ปัญหา |
| 142 | ช่วย | ประมวล |
| 143 | ช่วย | พัฒนา |
| 144 | ช่วย | กด |
| 145 | ช่วย | ติดตั้ง |
| 146 | ช่วย | แปลง |
| 147 | ช่วย | ผลักดัน |
| 148 | ช่วย | ค้นหา |
| 149 | ช่วย | ใช้ |
| 150 | ทดลอง | สร้าง |
| 151 | ทดลอง | ให้ |
| 152 | ทดลอง | ทำ |
| 153 | ทดลอง | บันทึก |
| 154 | ทดลอง | ออกแบบ |

| | 1st verb | 2nd verb |
|---|---|---|
| 155 | ทดลอง | ใช้งาน |
| 156 | ทดลอง | ติดตั้ง |
| 157 | ทดลอง | อ่าน |
| 158 | ทดลอง | สั่งซื้อ |
| 159 | ทดลอง | แลกเปลี่ยน |
| 160 | ทดลอง | ป้อน |
| 161 | ทดลอง | เรียก |
| 162 | ทดลอง | ใช้ |
| 163 | ไป | ควบคุม |
| 164 | ไป | อ่าน |
| 165 | ไป | เก็บ |
| 166 | ไป | ตรวจจับ |
| 167 | ไป | ตอบสนอง |
| 168 | ไป | รบกวน |
| 169 | ไป | แทรก |
| 170 | ไป | จำหน่าย |
| 171 | ไป | รับ |
| 172 | ไป | ตรวจ |
| 173 | ไป | ตรวจสอบ |
| 174 | ไป | กำหนด |
| 175 | ไป | สนับสนุน |
| 176 | ดำเนินการ | สร้าง |
| 177 | ดำเนินการ | พัฒนา |
| 178 | ดำเนินการ | ดูแล |
| 179 | ดำเนินการ | เชื่อมโยง |
| 180 | ดำเนินการ | จัด |
| 181 | ดำเนินการ | จัดหา |
| 182 | ดำเนินการ | ตอบสนอง |
| 183 | ดำเนินการ | ทด |
| 184 | ดำเนินการ | จัดตั้ง |
| 185 | ดำเนินการ | จัดซื้อ |
| 186 | ดำเนินการ | ติดตั้ง |
| 187 | ดำเนินการ | ประยุกต์ |
| 188 | พยายาม | กำหนด |
| 189 | พยายาม | ศึกษา |
| 190 | พยายาม | ใช้ |
| 191 | พยายาม | ค้นหา |
| 192 | พยายาม | หา |
| 193 | พยายาม | ทำ |
| 194 | พยายาม | พัฒนา |

| | 1st verb | 2nd verb |
|---|---|---|
| 195 | พยายาม | ทำให้ |
| 196 | พยายาม | ติดต่อ |
| 197 | พยายาม | สังเกต |
| 198 | พยายาม | ทำให้เกิด |
| 199 | พยายาม | เลี้ยง |
| 200 | พยายาม | ออกแบบ |
| 201 | พยายาม | จัด |
| 202 | พยายาม | เปรียบเทียบ |
| 203 | มา | ช่วยเหลือ |
| 204 | มา | ใช้ |
| 205 | มา | ทำ |
| 206 | มา | ช่วย |
| 207 | มา | ต่อ |
| 208 | มา | สนับสนุน |
| 209 | มา | สาธิต |
| 210 | มา | ขยาย |
| 211 | มา | รับ |
| 212 | มา | เก็บ |
| 213 | มา | ใช้งาน |
| 214 | มา | พัฒนา |
| 215 | มา | ประชุม |
| 216 | มา | อ่าน |
| 217 | มา | ยืม |
| 218 | มา | ใช้ |
| 219 | เลือก | เอา |
| 220 | เลือก | สร้าง |
| 221 | เลือก | เก็บ |
| 222 | เลือก | พิมพ์ |
| 223 | เลือก | เล่น |
| 224 | เลือก | ดู |
| 225 | เลือก | ค้น |
| 226 | เลือก | อ่าน |
| 227 | เลือก | สั่ง |
| 228 | เลือก | ปฏิบัติการ |
| 229 | เลือก | ตอบ |
| 230 | เลือก | พัฒนา |

## E.2 Semantic Interpretations

The semantic interpretation of some two consecutive SVCs which its first element is the top ten most found serial verbs in the ORCHID corpus are shown in Table E.1, including types of these two consecutive SVCs from related research.

**Table E.2: The first verb in SVCs and semantic interpretations.**

| No. | First Verbs | Semantic Interpretation | (Thepkanjana,1986) | (Sudmuk, 2005) |
|---|---|---|---|---|
| 1 | ใช้/chai/ | Purpose | Purpose simultaneous serial verbs | Use-SVCs, Purposive |
| 2 | ช่วย/chui/ | Complement | Complements of modality verbs | Open class SVCs, Simultaneous |
| 3 | ดำเนินการ/dam-noen-kan/ | Complement | Complements of modality verbs | Open class SVCs, Simultaneous |
| 4 | ต้องการ/tong-kan/ | Complement | Complement of modality verbs | Open class SVCs, Simultaneous |
| 5 | ทดลอง/thot-long/ | Complement | Complement of modality verbs | Open class SVCs, Simultaneous |
| 6 | ทำหน้าที่/tham-na-thi/ | Complement | Complements of modality verbs | Open class SVCs, Simultaneous |
| 7 | ไป/pai/ | Purpose | Purpose simultaneous serial verbs | Motion SVCs, Purposive |
| 8 | พยายาม/pha-ya-yam/ | Complement | Complement of modality verb | Open class SVCs, Simultaneous |
| 9 | มา/ma/ | Purpose | Purpose simultaneous serial verbs s | Motion SVCs, Purposive |
| 10 | เลือก/leuak/ | Purpose | Purpose simultaneous serial verbs | Open class SVCs, Purposive |

E.3 Tag definitions in Thai consecutive SVCs lexicon

Table E.3: Tag definitions in Thai consecutive SVCs lexicon

| Tag | Definition |
|---|---|
| <item></item> | No. of words |
| <wd></wd> | Word |
| <syn></syn> | Syntactic pattern |
| <cat></cat> | Word's category |
| <scat></scat> | Particle word's category |
| <vp></vp> | Syntactic verb pattern |
| <sem></sem> | Semantic pattern |
| <intpt></intpt> | Semantic Interpretation |
| <rel></rel> | Relation pattern |
| <agt></agt> | Agent of action |
| <pat></pat> | Patient of action |
| <ins></ins> | Instrument of action |
| <manner></manner> | Adverb of manner |
| <coop></coop> | Adverb of cooperating or separating |
| <time></time> | Adverb of time or frequency |
| <deg></deg> | Adverb of degree |
| <loc></loc> | Preposition for location |
| <coor></coor> | Preposition for coordination |
| <pur></pur> | Preposition for purpose |
| <mean></mean> | Preposition for mean |
| <sorc></sorc> | Preposition for source or destination |

# BIOGRAPHY

| | |
|---|---|
| Name | Jutapuck  Pugsee |
| Sex | Female |
| Date of Birth | November 27,1978 |
| Place of Birth | Prachenburi |
| Education: | |
| 2008 | Ph.D. in Computer Engineering, Chulalongkorn University |
| 2001 | M.Sc. in Computer Science, Chulalongkorn University |
| 1996 | B.Sc. in Computer Science, Chulalongkorn University |