

Automatic Requirement Simplification for Informal Requirement Specifications

Woo Sung Jang
Department of Software and
Communications Engineering
Hongik University
Korea
jang@selab.hongik.ac.kr

R. Young Chul Kim
Department of Software and
Communications Engineering
Hongik University
Korea
bob@hongik.ac.kr

ABSTRACT

In requirements engineering, the correct definition of requirements is an important issue. Correct redefinition of incorrect requirements are very necessary, but difficult. In addition, complex Korean language requirements are difficult to read and understand. And automatically generation of test cases with wide coverage from Korean requirements is difficult. However, related studies on Korean language requirements are very scarce. To solve the problem, we propose a simplification of Korean requirements. The simplified requirements are divided into conditional clauses, result clauses, and conjunction clauses (AND, OR). The three clauses can be easily generated as a test case.

CCS CONCEPTS

• **Software and its engineering** → **Software creation and management** → **Designing software** → **Requirements**; Software organization and properties → Software functional properties → Correctness → Functionality

KEYWORDS

Requirement analysis, Semantic analysis, Korean requirement, Requirement Specifications

1 INTRODUCTION

Most software organizations have a lot of testing techniques. However, in requirements-based testing, they have difficulties in performing high-quality tests [1]. As a result, if test cases are automatically generated and executed from requirements, test efficiency can be increased. Semantic analysis of Korean-based requirement and generation of test case is very difficult [2]. Related studies are also very rare. We study the process of automatically generating test cases from Korean-based requirements. If a requirement is converted into a Cause-Effect Graph, a Cause-Effect Graph is converted into a Decision Table, and a Decision Table is converted into a test case, the test case guarantees 100% functional requirement coverage [3]. However, automatically identification of Cause and Effect from the Korean-based requirements is difficult [4]. We propose the Korean requirements simplification method for automatic identification of Cause and Effect from Korean

requirements. Korean requirements are divided into cause clauses, result clauses, and conjunction clauses. The cause clause becomes a Cause. The result clause becomes an Effect. A conjunction clause is a combination of Cause. As a result, a Cause-Effect Graph can be generated through simplified clauses.

This paper is organized as follows. Chapter 2 describes related studies. Chapter 3 describes automatic simplification method of Korean language requirements. Chapter 4 describes environment of Korean requirements automatic analyzer. Chapter 5 describes the analysis results and future research.

2 RELATED STUDY

If the knowledge about the relationship of sentence component is identified in the sentence, the semantic analysis of the sentence is possible. However, the structure of Korean sentences is complex. Identification of knowledge about the relationship in the natural sentence is difficult. Sentence normalization through corpus normalization rule is necessary. The difficulty of semantic analysis is lowered for sentences to which sentence normalization is applied [5, 6].

The corpus normalization rule consists of “restoration of case movement”, “unification of duplicated forms of a compound word”, “recovery of an abbreviated subject of conjunct sentences”. In corpus normalization rule, Restoration of case movement identifies morphemes in sentences, identifies corpus in identified morphemes, and rearranges misplaced cases in identified corpus. The rearranged sentences are easily recognized automatically. The part of process is shown in Table 1.

Table 1. Process of restoration of case movement

No.	Process
1	Restoration of an active sentence from a passive sentence. Passive means that the subject is changed by external factors. Active means an action or action performed by one's own strength.
2	Restoration of a Judong sentence from a Sadong sentence. Sadong means that the subject acts due to the actions of external factors. Judong means to do some action or action by yourself.

3 AUTOMATIC SIMPLIFICATION METHOD OF KOREAN REQUIREMENTS

The structure of Korean requirements automatic analyzer is shown in Figure 1. Sentence Segmentation divides a sentence, recovers the missing subject from the divided sentence, and identifies the type of clause (Cause, Result, Conjunction). Syntax Analysis converts passive sentences into active sentences using sentence normalization. Finally, Simplify complex auxiliary verbs.

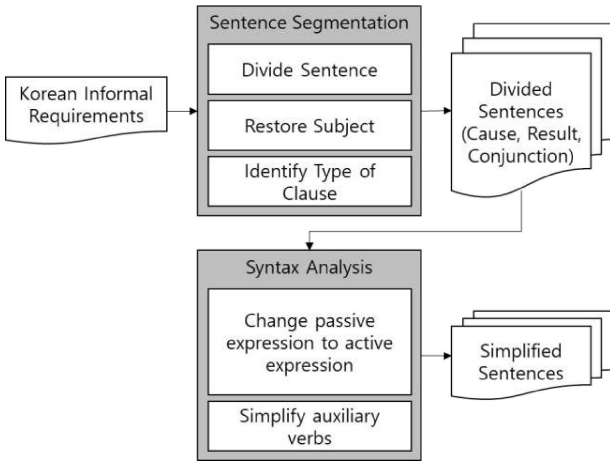


Figure 1. Korean requirements Simplification Process

The Cause clause, Result clause, and Conjunction clause in the sentence are distinguished through the identification of the connective ending. Connective ending is only Korean morphemes. A connective ending is added at the end of a clause. Korean uses connective endings to identify types of clause. Table 2 shows the types of connective endings for identification of conjunction clauses.

Table 2. Combination Types between Conjunctions [5,6,7]

Conjunction Type	Connective Ending
Causality Relationship	'-느라고(Nerago)' Etc.
Sequential relationship	'-고(Go)(서(Seo))' Etc.
Parallel relationship	'-고(Go)', '-(으(Eu))며(Myeo)', '-(으(Eu))면서(Myeonseo)' Etc.
Contrast Relationship	'-(으(Eu))나(Na)', '-지만(Jiman)' Etc.
Selection Relationship	'-거나(Geona)', '-든지(Deunji)' Etc.
Time relationship (concurrency)	'-고(Go)', '-(으(Eu))며(Myeo)', '-(으(Eu))면서(MeonSeo)', Etc.
Time relationship (sequentiality)	'-고(Go)', '-(아(Ah)/어(Eo))서(Seo)', '-자(Ja)(마자(Maja))' Etc.

Cause and Reason Relationship	'-아(Ah)/-어(Eo)', '(아(Ah)/어(Eo))서(Seo)', Etc.
Concession Relationship	'-(아(Ah)/어(Eo))도(do)', '더라도(Deolado)', Etc.
Purpose and Intention Relations	'-(으(Eu))러(lu)', '-고자(Goja)', '-(으(Eu))려고(Lyugo)' Etc.
Conversion Relationship	'-다가(Daga)' Etc.

Table 3 shows the types of connective endings for identifying condition clauses and result clauses. The clause before the connective ending is a condition clause. The clause after the connective ending is the result clause.

Table 3. Conditional clause identification method based on connective ending type[8]

Connective ending type	Connective ending form
Simple type	-면(Myun), -거든(Geodeun), -어야(Uya)
Complex type	-다면(Damyeon), -은들(Eundeul), -다가(Daga), -든지(Deunji), -려면(Lyeomyeon), -거르면(Geodeumyeon), -거들랑(Geodeullang), -대서야(Daeseoya), -고야(Goya), -을라치면(Eullachimyeon), ...

If a clause of a sentence is identified, one clause is converted into a sentence. Then, sentence normalization is applied to each converted sentence. Table 4 shows part of the sentence normalization methods. If the sentence has the passive sentence form of Rule1, the case marker in the sentence is converted into the case marker in active sentence form of Rule1. Case marker is only Korean morphemes. It is added after the noun and defines the role of the noun. In English, the role of a noun is determined according to the position of the noun (subject, object, etc.). However, in Korean, the role of a noun is determined according to the type of case marker after the noun.

Table 4. Methods of changing passive expression to active expression [5,6]

Rule	Positive expression	Active expression
1	에게/한테/에/에 의해 (ege/hante/e/e euhae)	가(ga)
	가(ga)	를(leul)
2	에게/에 의해(ege/e euhae)	가(ga)
	가(ga) 에서(eseo)	로(lo) 를(leul)

Table 5. Combination Types between Conjunctions

Conjunction Type	Relationship between A, B, and C	Logical Expression
Causality Relationship	If B=true for A=true, then C=true.	AND
Sequential relationship	If B=true after A=true, then C=true.	AND
Parallel relationship	If A=true and B=true at the same time, then C=true.	AND
Contrast Relationship	If A=false and B=true, then C=true.	OR
Selection Relationship	If A=true or B=true, then C=true.	OR
Time relationship (concurrency)	If A=true and B=true at the same time, then C=true.	AND
Time relationship (sequentiality)	If B=true after A=true, then C=true.	AND
Cause and Reason Relationship	If B=true for A=true, then C=true.	AND
Concession Relationship	If B=false after A=true, then C=true.	OR
Purpose and Intention Relations	If A=true after B=true, then C=true.	AND
Conversion Relationship	If B=true after A=true, then C=true.	AND

Table 6. Part of Auxiliary Verb Simplification

Before Simplification		After Simplification	
Korean	English	Korean	English
하여야 한다(Hayeoya handa)	Should do	한다(Handa)	Do
해야 한다(Haeya handa)	Must do	한다(Handa)	Do
할 수 있다(Hal soo itda)	Can do	한다(Handa)	Do

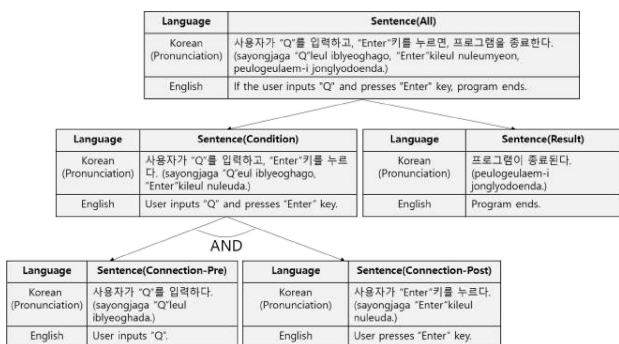


Figure 2. Simplified Korean requirement.

Conjunction clauses are classified as AND or OR according to the conjunction type. The classification by conjunction type is shown in Table 5.

A part of the auxiliary verb simplification process is shown in Table 6. Simplify various forms of auxiliary verbs.

Korean requirements automatic simplification process is shown in Figure 2.

4 KOREAN REQUIREMENTS AUTOMATIC ANALYZER

The structure of the implemented Korean requirements automatic analyzer is shown in Figure 3. Analyze the morpheme of a sentence using the MeCab-Ko library. Sentence Segmentation and Syntax Analysis algorithms are implemented in JAVA. Libraries and implemented programs run in a Linux environment.

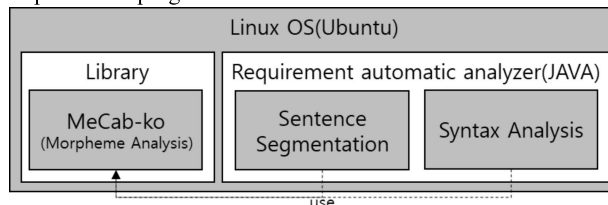


Figure 3. Korean requirements automatic analyzer

5 CONCLUSIONS

We show the results of sentence analysis using Korean requirements automatic analyzer. The 78 sentence was correctly simplified with 96.1%. 3.9% of sentences have no cause clause and no result clause. In the future, we will automatically convert the simplified clauses into a Cause-Effect Graph. And we will consider the analysis of sentences without cause and effect clauses.

ACKNOWLEDGMENTS

The research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R111A305040711) and the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education(MOE, Korea) (No.2021R111A305040711).

REFERENCES

- [1] Byung-Sun Cho, Seok-Won Lee, A Comparative Study on Requirements Analysis Techniques using Natural Language Processing and Machine Learning, Journal of The Korea Society of Computer and Information, ACM 25, 7 (2020), 27-37.
- [2] Hyunsang Park, Kyunghee Choi, Kihyun Jung, A strategy of detecting feasible test cases for requirement based testing, The Korean Institute of Information Scientists and Engineers, ACM 36, 1B (2009), 90-95.
- [3] Gary E.Mogyorodi, B. Math. 2005. Requirements-Based Testing-Cause-Effect Graphing, Software Testing Services.
- [4] Myers, G. L. 1979. The Art of Software Testing, Wiley-Interscience.
- [5] Jeong Mi Cho, Young Hwan Cho, Gil Chang Kim, A Corpus Formalization for Extracting the Syntactic Relations, The 8th Annual Conference on Human & Cognitive Language Technology, (1996), 207-215.
- [6] Jeong Mi Cho, Gil Chang Kim, A Corpus Formalization for Extracting the Syntactic Relations, The Korean Society for Cognitive Science, ACM 7, 2 (1996), 39-56.
- [7] Kim Kisung, A Comparative study on conjoined sentence between modern Mongolian and Korean, The Korean Association for Mongolian Studies, ACM 27, 0 (2009), 151-186.
- [8] Jeong Moon Ha. 2007. A contrastive study on Korean conditional connective ending and Chinese conditional conjunction expression, Kyunghee University.