SIGMA - Normalization

Isaac Griffith and Rosetta Roberts
Empirical Software Engineering Laboratory
Informatics and Computer Science
Idaho State University
Pocatello, Idaho, 83208
Email: {grifisaa@isu.edu, roberose@isu.edu}

***Abstract*—Introduction:**
Objective:
Methods:
Results: What are the main findings? Practical implications?
Limitations: What are the weaknesses of this research?
Conclusions: What is the conclusion?
***Index Terms*—Island Grammars, Automated Grammar Formation, Software Language Engineering**

I. Introduction

REWORD

Modern software development practice has led to the creation of software systems using multiple languages. As an example, the modern web application might use 5 or more languages (e.g. SQL, Java, TypeScript, HTML, CSS). Such multilingual codebases present a difficult challenge to the development and maintenance of code analysis tools [1]. To address this challenge current source code analysis tools utilize multiple parsers (one per supported language supported).

Outline:

1. Mention current SIGMA tool.
2. Normalization section is untested.

RG Evaluate and test each step of the normalization proccess used by SIGMA to understand the effectiveness of each step to improve the normalization process of SIGMA.

Organization

The remainder of this paper is organized as follows. Sec. II discusses the theoretical foundations of this work while also discussing other related studies. Sec. IV details the design of experiments which evaluate the normalization steps of SIGMA. Sec. V details the threats to the validity of this study. Finally, this paper is concluded in Sec. VI.

II. Background and Related Work

A. Theoretical Foundations

Context-free grammars are defined as $G = (V, \Sigma, P, S)$, where V is the set of non-terminal symbols, Σ is the set of terminal symbols, P is the set of productions, and S is the start production [@haoxiangLanguagesMachinesIntroduction1988].

Island Grammars are a type of grammar designed for extracting out features of interest. Island Grammars are composed of two types of rules: land/island rules which identify the features of interested, and water rules which are catch all rules for everything that is not of interest [@moonenGeneratingRobustParsers2001]. Island grammars are advantageous to regular grammars for several reasons, including being easier/faster to write, lower complexity, and better tolerance of errors in the input [@moonenGeneratingRobustParsers2001]. Because of this, island grammars have been used for the development of multilingual parsers [@synytskyyRobustMultilingualParsing2003].

In this paper, we evaluate the normalization process of SIGMA. SIGMA uses a unique ## Related Studies {#sec:related}

B. Research Contributions

III. Approach

IV. Experimental Design

A. Goals, Hypotheses, and Variables

Goals: - Evaluate each step of normalization

Hypotheses: - Null hypothesis: each step has no effect on the halstead effort for MCC.

Variables: - Size of grammar as chosen in SIGMA - Blocking, - Halstead/MCC - The steps that the normalization goes through. (2^#numSteps)

B. Design

Blocked factorial design

Reasons - Effect from size found in SIGMA - Interactions from size also seen in SIGMA

C. Experimental Units

D. Data Collection Procedures

E. Analysis Procedures

F. Evaluation of Validity

Conclusion Validity:
Internal Validity:
Construct Validity:
External Validity:

V. Threats to Validity

VI. Conclusions and Future Work

The timeline to complete this study is as follows:

We intend to publish these results at one of the following conferences:

ACKNOWLEDGEMENTS

REFERENCES

[1] Z. Mushtaq, G. Rasool, and B. Shehzad, "Multilingual Source Code Analysis: A Systematic Literature Review," *IEEE Access*, vol. 5, pp. 11 307–11 336, 2017, bibtex: mushtaqMultilingualSourceCode2017.