

## 発表概要

# 非成文化規則としてのプログラムの意味論に向けて： コンパイラにさようなら

堀江 幸生<sup>1,a)</sup>

2020年10月29日発表

昨今、機械学習によって自然言語処理が革新的に進化したことに相まって、自然言語処理と同様な手法によってプログラムの振舞いが予測できるようになった。この振舞いのことを本発表では意味と呼ぶ。この振舞いは非成文化規則によって定義できる。本発表では、金田の研究を引き継ぎ、形式的なプログラムの意味論とは異なった観点で、プログラムの意味を可能な限り直観的に自然言語処理、機械学習の観点から見直すことを試みた。その過程で、我々はある仮説を提示する。機械学習は、ソースコードとバイナリの意味を理解したように振る舞う。さらにクエリからソースコードの検索も可能になった。それらを組み合わせることにより、我々は不完全なソースコードまたはクエリを入力することで、それに対応するバイナリを生成できるのではないかとこの仮説にいきついた。これが可能であれば、コンパイラやデバッガの存在する意味は変化する。現在、我々は小規模であるがそれを立証しつつある。ただし、我々の最終的な目標であるプログラムの意味については、この領域の圏論による大きく広がる可能性を示唆しながらも、体系的に扱える見通しは今のところ得られていない。計算機が、小さな意味の積み重ね（四則演算、ループ）を、その上位概念としての、たとえば、悪意のあるソフトウェアであれば“悪意”、をどうやって獲得するに至ったのか、我々の研究途上では不確かだからである。結果は、このようなものとなったが、我々は本発表での研究がプログラムの意味について1つ踏み込めるきっかけを提示できた。

## Presentation Abstract

### The Linguistic Approach to the Program Semantics: Good bye Compiler

SACHIO HORIE<sup>1,a)</sup>

Presented: October 29, 2020

Recently, the deep learning has revolutionized natural language processing. It has made it possible for us to predict the behavior of the program in the method of natural languages. We call this behavior the semantics in this presentation. In this presentation, we found papers that the similarity of source code (malicious source code and benign source code) can be predicted with doc2vec and consider the semantics from the viewpoint of natural language processing, without expressing the behavior of the program in the form that can be treated mathematically from a viewpoint different from the semantics of the program. By doing so, we will continue Kaneda's research and consider the meaning of the program. Finally, we are preparing to demonstrate an innovative hypothesis. It is as follows. The machine learning can understand semantic of the source code, it also will be able to understand the semantic of the binary, furthermore it is able to search the source code from the query. If we input the imperfect source code or the query, we can realize the corresponding binary (this binary is regenerated), however we do not need the compiler and the debugger. We want to realize this. What is the semantics of the program then?, we must indicate this question.

---

This is the abstract of an unrefereed presentation, and it should not preclude subsequent publication.

<sup>1</sup> 名古屋大学情報学研究科社会情報学  
Graduate School of Informatics, Nagoya University, Nagoya,  
Aichi 464-8601, Japan

<sup>a)</sup> horie.sachio@f.mbox.nagoya-u.ac.jp