Towards automated proof of HOL by machine learning using graph structure

Takakazu Sato, Shunji Nishimura

Oita, Japan
National Institute of Technology, Oita College, Japan

1. Abstract

Formal proving of programs requires a lot of labor. If automation of formal proving can be realized, it is possible to greatly reduce the trouble of checking the legitimacy of the program. Since the proof consists of many intermediate logic steps, it is thought that automation can be realized by predicting them. In this research, machine learning is applied as a means to realize automatic certification. Since logical expressions can graphically show the properties conforming to syntactic rules, we choose Graph Convolutional Network (GCN), whose input is a machine learning model of graph structure. The accuracy of the model implemented in this experiment was 0.74. This result suggested that the input data representing the logical expression as a graph cannot be used significantly.

2. Introduction

Theorem proving is a method of verifying the validity of a program, but it requires much effort. If automation of verification can be realized, it is possible to greatly reduce the trouble of checking the legitimacy of the program. Proof is composed of many intermediate logic steps, and it is thought that it becomes possible to automate theorem proving by predicting intermediate logical steps [1].

In order to automate theorem proving, this study applies the machine learning method which attracts particular attention in recent years. Specifically, we use machine learning to determine whether a certain intermediate logical step is useful for the statement which is demanded to prove. Since logical expressions can graphically show the properties conforming to syntactic rules, input is made using Graph Convolutional Network (GCN) which is a machine learning model of graph structure.

3. Previous research

Kaliszyk and colleagues [1] have learned the data set of the higher order predicate logic to the existing machine learning model, and the intermediate logical steps of the proof are inferred. It is judged whether it is useful for the proof of. The machine learning model used is logistic regression, Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Kaliszyk and colleagues conduct experiments to find out whether the characteristics of the dataset are most useful for discrimination of the usefulness of intermediate logical steps, but it is difficult to capture the complex structure of proof with these machine learning models.

4. Proposed method

It is stated in Kusumoto et al.'s preprint [2] that logical expressions conform to syntactic rules, so that they can be represented graphically and that the meaning does not change even if the names of variables are changed. In order to utilize these properties, learning of intermediate logical steps of certification is performed using GCN which is one of machine learning models.

4.1. Graph Convolutional Network

GCN is a model that defines a convolution operation in graph structure and behaves similarly to CNN. Since logical expressions conform to syntax rules, they can be represented by graphs, and GCN can extract hidden features by directly processing this graph structure.

5. Result

In this study, we evaluated the usefulness of intermediate logical steps of proof using the data set of higher order predicate logic introduced by Kaliszyk et al. [1], and compared the cases using GCN and the machine learning model used by Kaliszyk et al.

Table 1 shows the model used by Kaliszyk et al. And the results of our GCN. Regarding GCN, accuracy was lower than CNN and RNN. This is considered to be due to the fact that processing for extracting hidden features described in 3.2 is not sufficiently performed on input data that represented logical expressions as graphs.

Model name	CNN	RNN	Logistic regression	**GCN**
Accuracy	0.82	0.83	0.71	**0.74**

Table1：Experiment result

6. Conclusion

In this study, we determined whether or not the learned intermediate logic step is useful for the statement that we want to prove by using GCN which is one of machine learning models.

From the experimental results, it was suggested that the input data representing the logical expression as a graph cannot be used significantly. For future work, we propose a new model architecture and verify its usefulness.

Acknowledgements

In promoting this research, I am deeply grateful to everyone at Nishimura Laboratory who gave me many knowledge and suggestions.

References

1. Cezary Kaliszyk, Francois Chollet, Christian Szegedy, “"HOLSTEP": A Machine Learning Dataset For Higher-Order Logic Theorem Proving,” ICLR2017.
2. Mitsuru Kusumoto, Keisuke Yahata, Masahiro Sakai, “Automated Theorem Proving in Intuitionistic Propositional Logic by Deep Reinforcement Learning,” 2018.