

機械学習による HOL の自動証明に関する研究

大分高専専攻科 電気電子情報工学専攻 佐藤 貴一 (指導教員 西村俊二)

1. はじめに

プログラムの正当性を確かめる方法の一つに定理証明があるが、プログラムを形式的に証明することは多くの労力を必要とする。証明の自動化が可能となれば、プログラムの正当性を確かめる手間を大幅に減らすことができる。証明は多くの中間的な論理ステップから構成されており、その中間的な論理ステップ予測することで定理証明の自動化が可能になると考えられている[1]。

定理証明を自動化する手段として、本研究では近年特に注目を集めている機械学習手法を適用する。具体的には、推論対象である多くの中間的な論理ステップから構成されている証明に対して、機械学習を用いることで学習した中間的な論理ステップが有用であるか否かの判別を行う。論理式は構文規則に従う性質をグラフで表すことが可能であるため、入力がグラフ構造の機械学習モデルである Graph Convolution Neural Network(GCNN)を用いて行う。

2. 先行研究

定理証明器 HOL の証明を機械学習を用いて自動化する先行研究として、Kaliszyk ら [1]は高階述語論理を用いた証明のデータセットの導入と、そのデータセットを用いて既存の機械学習モデルに学習させ、証明の中間的な論理ステップが推論対象の証明に対して有用であるかを判別している。加えて、機械学習モデルのベンチマークを行っている。使用されている機械学習モデルはロジスティック回帰と Convolution Neural Network(CNN), Recurrent Neural Network(RNN)である。データセットの特性が中間的な論理ステップの有用性の判別に最も役立つか調べるために実験を行っている。しかし、Kaliszyk らは、これらの機械学習モデルでは証明の複雑な構造を捉えることは難しいという問題があると述べている。

3. 提案手法

論理式は構文規則に従うので、グラフで表すことが可能であり、かつ、変数の名前を変更しても意味は変わらないという性質があることが Kusumoto らのプレプリント[2]で述べられている。これらの性質を利用するために、機械学習モデルの一つである GCNN を用いて証明の中間的な論理ステップの学習を行う。

4. データセット

本研究で用いたデータセットは定理証明器 HOL で記述された証明と、ケプラー予想の形式的証明から構築されている。学習データが 9,999 個、テストデータが 1,411 個が含まれている。これらには中間

的な論理ステップが関連付けられており、合計 2,209,076 の推測文の組がある。

5. Graph Convolution Neural Network

論理式は構文規則に従うので、グラフで表すことが可能であり、かつ、GCNN はこのグラフ構造を直接処理することで隠れた特徴を抽出することが可能である。加えて、論理式には変数の名前を変更しても意味は変わらないという性質があるため、変数名を変更しても不変である隠れた特徴を抽出できる。

GCNN は同じ意味の論理式に対して常に同じ値を出力することが可能である。

6. 評価手法

本研究では、Kaliszyk ら[1]が導入した定理証明のための機械学習データセットを用いて証明の中間的な論理ステップの有用性を判別を行う。GCNN を用いた場合と Kaliszyk らが用いた機械学習モデルの場合のベンチマークを行う。

7. 結果

Model name	CNN	RNN	Logistic regression	GCNN
accuracy	0.82	0.83	0.71	0.73

表1: 実験結果

表1に Kaliszyk らが用いたモデルと GCNN の結果を示す。GCNN に関して、CNN と RNN よりも精度が低かった。これは論理式をグラフとして表した入力データに対して 5. で示した隠れた特徴を抽出する処理が十分に行えていないことが原因と考えられる。

8. おわりに

本研究では学習した中間的な論理ステップが推論対象である証明に対して有用であるか否かの判別を機械学習モデルの一つである GCNN を用いて行った。実験の結果から、論理式をグラフとして表した入力データを有意に活用できないことが示唆された。今後の課題として、新たなモデルアーキテクチャを提案し、有用性を検証する。

参考文献

- [1] Cezary Kaliszyk, Francois Chollet, Christian Szegedy, "HOLSTEP: A Machine Learning Dataset For Higher-Order Logic Theorem Proving," Published as a conference paper at ICLR 2017.
- [2] Mitsuru Kusumoto, Keisuke Yahata, Masahiro Sakai, "Automated Theorem Proving in Intuitionistic Propositional Logic by Deep Reinforcement Learning," <https://arxiv.org/pdf/1811.00796.pdf>, 2018.