

先端科学技術研究科 修士論文要旨

| | | | |
|---|--|-----|--------------|
| 所属研究室 (主指導教員) | ソフトウェア設計学 (飯田 元 (教授)) | | |
| 学籍番号 | 1911227 | 提出日 | 令和 3年 1月 25日 |
| 学生氏名 | 深澤 佑樹 | | |
| 論文題目 | A code summarization method using the structural information of abstract syntax trees 構文木の構造情報を用いたメソッド名推薦手法 | | |
| 要旨 | | | |
| <p>ソフトウェア開発において、識別子名の質はプログラム理解に多大な影響を及ぼす。しかし、適切な命名を行うことはソフトウェア開発に関する深い知識や経験が必要とするため、困難な作業となっている。そのため、識別子の命名を支援することを目的として、プログラムの内容から識別子名を推薦する様々な手法が提案されている。近年では、深層学習に基づく自然言語の機械翻訳手法、すなわち、ニューラル機械翻訳 (Neural Machine Translation, NMT) を応用した様々なメソッド名推薦の手法が提案されている。しかし、NMTで使用されるseq2seqを始めとする確率モデルの多くはシーケンススペースのモデルであり、構造が曖昧な自然言語のテキストとは対照的に、厳密な構造を有するソースコードの学習には適さない。本研究では、抽象構文木を走査することによりシーケンス化する手法であるSBT (Structure-Based Traversal) に着目し[1], SBTによりシーケンス化したソースコードからメソッド名を推薦する深層学習モデルを設計・実装した。提案手法は、11個のJavaプロジェクトを用いて評価した。推薦されたメソッド名のリスト内の単語が一部分でも実際に使われていたメソッド名に含まれていた場合に成功とする部分一致により評価した結果、各単一プロジェクト内のデータを用いてモデルを訓練し、該当プロジェクト内のメソッド名推薦を実施した場合、平均として約80%の精度で推薦が可能であると確認した。また、11個のプロジェクトのうち、1つのプロジェクトを検証用として選び、残りのプロジェクトのデータを横断的に用いてモデルを訓練した場合、検証用プロジェクトにおけるメソッド名推薦精度は約70%であることを確認した。</p> | | | |