

ルールに基づく Python メソッドの要約コメント自動生成

1190364 久川 竜正 【ソフトウェア検証・解析学研究室】

1 はじめに

プログラムにコメントを記述することは、ソースコードを理解しやすくする効果がある。しかし、世の中のコードにはコメントが少ないものもあり、特にメソッド一つ一つに対してコメントが付いていることは少ない。プログラムにコメントがない場合ソースコードのみから内容を理解する必要があり、プログラミング初学者や扱い慣れていない言語の場合だと理解が難しい。そのため、ソースコードから自動的にコメントを生成することは有用だと考えられる。

Sridhara らの研究では、Java メソッドの動作を要約するコメントを自動的に生成する手法が提案されている [1]。本研究では、Sridhara らの手法を元に Python メソッドのソースコードから、メソッドの動作を要約するコメントを、ルールに基づいて自動生成することを目的とする。

2 設計と実装

本研究では、Python で記述されたメソッドのテキストを入力として、メソッド呼び出し中の各語へのタグ付け、コメント生成に必要な文の特定、テンプレートを用いたテキスト生成の三つの手順を行うことでコメントを生成する。また、提案システム自体も Python で実装した。

2.1 メソッド呼び出し中の各語へのタグ付け

メソッド呼び出しの各語に対して、action, theme, secon の 3 つのタグ付けを行う。例えばメソッド呼び出し `list.append(num)` があれば、`append` を action, `num` を theme, `list` を secon としてタグ付けを行う。実装手順としては、ast モジュールを用いて得られた構文木を解析し、メソッド呼び出しの位置を特定する。特定されたメソッド呼び出しの “.” や “()” に注目して分割しタグ付けを行う。

2.2 コメント生成に必要な文の特定

プログラム中の全てのコードに対してコメントを生成すると冗長なコメントとなり、メソッドの動作を要約しているとは言えない。そのため、メソッド内のどの文がコメント生成に必要なのか特定する必要がある。コメント生成に必要な文として、制御フローにおける最終行、制御フロー式 if, while, for のある文などをコメント生成に用いる。これら文の特定には、抽象構文木および、上記 2.1 で行ったタグ付けの情報を使用する。

2.3 テンプレートをを用いたテキスト生成

コメント生成のために特定された文に対して、プログラム文からコメントとしてのテキストを生成するため

に、様々なプログラム文に対応したテンプレートに基づいて、メソッド呼び出しの各語に付けられたタグを当てはめていくことでコメントテキストを生成する。

3 実験結果

12 個のメソッドに対してコメント生成を行ったところ、メソッドの動作を説明するのに十分なコメントが生成できたものが 4 個あった一方、生成の上手くいかないものも多く見られる結果となった。コメント生成が成功した例としては、ソースコード 1 にサンプルコードを、図 1 に生成されたコメントのテキストを示す。このサンプルコードでは、メソッドの動作を説明するのに必要だと思われるコメントのみが生成されている。コメント生成の上手くいかないプログラムからは、コメントの文脈として正しくないテキストが生成されることが多く見られた。そのため、多くのコメント生成が成功しなかった理由として、テキスト生成のテンプレートがプログラム文の行う様々な動作やその書き方に対応しきれていないことが考えられる。

ソースコード 1	サンプルコード
<pre> 1 def sumEven_num(list): 2 try: 3 even_list = [] 4 max_num = 0 5 for ln in list: 6 if ln % 2 == 0: 7 even_list.append(ln) 8 max_num = max(list) 9 sum_even = sum(even_list) 10 return sum_even 11 except TypeError: 12 print(list, 'is not an list')</pre>	<pre> 1 Sum even num takes list as input 2 "[" get assign to "even list" 3 Iterate the contents of list as ln 4 If "ln % 2 == 0 ", 5 Append ln to even list 6 Sum even list and get assign to "sum even" 7 Return sum even</pre>

図 1 コメントテキスト

4 まとめ

本研究では、Python で記述されたメソッドに対して、コメントを生成するルールを適用することでコメントテキストの自動生成を行うシステムを構築した。今後の課題としては、テキスト生成をより多くのプログラム文に対応したものへと拡張することなどが考えられる。

参考文献

- [1] G. Sridhara, E. Hill, D. Muppaneni, L. Pollock, and K. Vijay-Shanker. “Towards Automatically Generating Summary Comments for Java Methods.” ASE '10, pp. 43–52, 2010.