

外部情報を用いたプログラム手続き生成の Word2Vec による fine-tuning

Program procedure generation by fine-tuning with Word2Vec using external information

大西 朔永^{*1}, 椎名 広光^{*2}

Sakuei ONISHI^{*1}, Hiromitsu SHIINA^{*2}

^{*1} 岡山理科大学大学院総合情報研究科

^{*1} Graduate School of Informatics, Okayama University of Science

^{*2} 岡山理科大学

^{*2} Okayama University of Science

Email: i20im01os@ous.jp

あらまし : プログラムの理解の補助を目的に, ソースコードから手続きの自動生成を行っている. 従来ソースコードと手続きの対訳ペアを Encoder-Decoder モデルに学習させていたが, 課題の問題文とプログラムを用いることで手続きを生成したほうが良いと考えられる. 本研究では, 問題文をコーパスとして, コメント側の分散表現を Word2Vec で学習することで, 問題文の単語等を活かした手続き生成が可能となっている.

キーワード : アルゴリズム学習, 手順生成, ニューラルネットワーク, fine-tuning

1. はじめに

近年, AI やビックデータなど IT 技術の発展に伴い, プログラミングに関心が集まっている. これらの影響により, プログラミングの学習者が増加傾向にある. プログラミングの学習では, プログラミング言語の文法の学習だけでなく, プログラムに必要な処理とその手順を理解することが重要である.

手続き学習の研究として, 新開ら⁽¹⁾の手作業によるアルゴリズムの学習についての研究がある. 手続き学習には, ソースコードをプログラム手続きに変換する必要があるが, 人手で各プログラムの手続きを作成するのは, コストが高く, 困難である. そのため, プログラムのソースコードからプログラム手続きを自動生成する手法の開発を行っている. 先行研究としては, ニューラル機械翻訳をソースコードとコメントの各行に適用する手法がある⁽²⁾.

本研究では, 外部情報のドメインに適したコメント生成を目的に, ニューラル機械翻訳を対訳データに適用する手法(sep2sep モデル)に加えて, 外部情報をコーパスとした Word2Vec を用いたコメント側の分散表現の fine-tuning を試みている. 外部情報として大学の情報科学科の講義で使用されたプログラムの問題文を用いているため, 大学1年生向けのコメントが生成されると考えている. また, 生成コメントに対し, 人手と BLEU による評価をしている.

2. 手続き学習

プログラムには, プログラムの目的を達成するために, 必要な処理とその処理の順序を正しく記述する必要がある. 必要な処理を正しい順序に並び替える問題を解くことにより, プログラムの手順を理解することが手続き学習の目的である. 手続き学習の例を図1に示す.

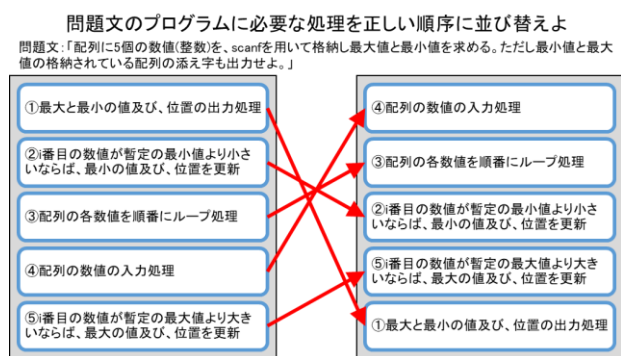


図1 手続き学習の例

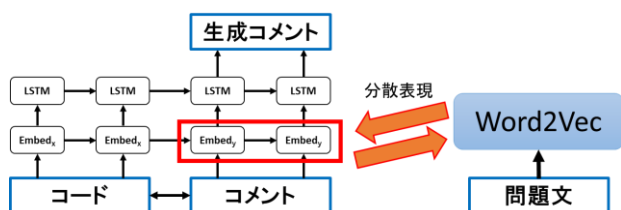


図2 fine-tuning の概要

3. 外部情報を用いた fine-tuning

従来のプログラムのコメント生成では, プログラムのソースコードとコメントを対訳データとして, 行単位で seq2seq モデルに入力することで, ソースコードとコメントの関係性を学習させる. 学習済みモデルにソースコードを入力することで, ソースコードに応じたコメントが生成される.

本研究では, 従来手法に対して, 外部情報を用いた fine-tuning を加えた手法を提案している. fine-tuning の概要を図2に示す. 最初に, 学習済みモデルのコメント側の分散表現を取り出し, Word2Vec の分散表現の初期値とする. 次に, 外部情報をコーパスとして Word2Vec の学習を行う. 最後に, 外部情

表 1 コメントの生成例(評価は人手による 6 段階評価)

行	ソースコード	sep2sep		fine-tuning			
		生成コメント	評価(1-6)	BLEU(0-1)	生成コメント	評価(1-6)	BLEU(0-1)
2	#define NUMBER 5	文字列を表示する	3.167	0.088	配列の大きさを定義づける	5.333	0.307
7	int vx [NUMBER];	整数型の配列の宣言	5.333	0.089	整数型の配列の宣言	5.333	0.089
8	for (i = 0; i < NUMBER; i++){	定義づけた回数分ループ	5.000	0.537	定義づけた回数分ループ	5.000	0.537
9	printf(" vx [%d]=", i);	配列のすべての添え字内に数値を入力	4.167	0.367	配列の添え字内に数値を入力	4.833	0.467
10	scanf("%d",& vx [i]);	整数を入力	5.500	0.841	整数を入力	5.333	0.841
11	}	そうでない場合	2.833	0.000	変数	2.833	0.000
14	max = vx [0];	maxの初期値をvx[0]とする	5.667	1.000	maxの初期値をvx[0]とする	5.833	1.000
15	max_loc = 0;	maxのあり場所を0	3.833	1.000	maxのあり場所を0	4.000	1.000
16	for (i = 1; i < NUMBER; i++){	定義づけた回数分ループ	5.500	0.537	定義づけた回数分ループ	5.500	0.537
21	if (max < vx [i]){	maxの方が小さい場合	5.167	1.000	maxの方が小さい場合	5.333	1.000
22	max = vx [i];	max(はいつも一番大きく	3.167	0.000	max(はいつも一番大きく	3.167	0.000
23	max_loc = i;	maxのあり場所を0	3.167	0.180	最大値の場所を記憶	5.000	1.000

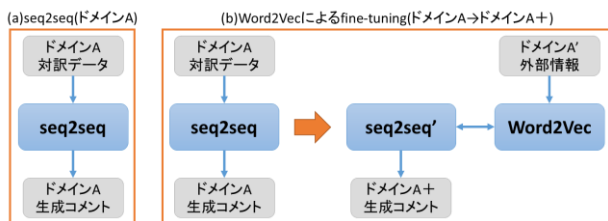


図 3 fine-tuning による生成コメントの変化

報で調整された分散表現をモデルのコメント側の分散表現に戻す。その後は、fine-tuning した学習済みモデルにソースコードを入力することで、外部情報のドメインに適したコメントが生成される。

fine-tuning による生成コメントの変化を図 3 に示す。提案手法の利点は、外部情報のドメインに適したコメント生成である。(a)対訳データによる従来手法(sep2sep モデル)の学習に、(b)Word2Vec によって外部情報における言葉の使われ方を分散表現に学習させる fine-tuning を加える提案手法(fine-tuning モデル)は、モデルが持つドメイン A に関する情報を豊富にするため、生成コメントの質の向上に繋がると考えている。また、外部情報のみに含まれる言葉が、Word2Vec で学習され、コメント生成時に出力可能となる利点もある。さらに、学習データの作成コストの低減も利点として考えられる。従来手法(sep2sep モデル)では、ドメイン A の対訳データを大量に作成する必要がある。しかし、提案手法では、作成コストの低い外部情報(問題文等)も利用しているため、対訳データの必要量が少なく、データを用意するコストが低く抑えられると考えている。

4. 生成コメントの評価

評価するモデルは、従来手法(sep2sep モデル)、外部情報を用いて fine-tuning を行うモデルの 2 種類である。対訳データとして、大学の情報科学科の講義で使用された C 言語のプログラム 56 個を用い、外部情報として、講義で使用されたプログラムの問題文を用いた。評価方法は、クローズドテストで生成されたコメントを対象とした 6 人の大学 4 年生に

表 2 評価の平均

モデル	人手による評価	BLEU
sep2sep	4.632	0.667
fine-tuning	4.897	0.727

よる 6 段階評価(1~6)と BLEU(0~1)である。

人手による評価の平均と BLEU を表 2 に示す。従来手法(sep2sep モデル)と比べ、問題文を用いて fine-tuning したモデルでは、人手による評価が約 0.26、BLEU が約 0.06 とある程度向上している。

コメントの生成例を表 1 に示す。sep2sep モデルと fine-tuning モデルの大きな差は、2 行目と 23 行目である。2 行目は 7 行目の配列の宣言で用いる定数を定義するコードである。sep2sep モデルでは無関係のコメントが生成されているが、fine-tuning モデルではこのプログラムに適したコメントを生成できている。23 行目は 15 行目から 1 文字だけ変化したコードである。sep2sep モデルは 15 行目と同じコメントを生成しているが、fine-tuning モデルは 23 行目のコードに適したコメントを生成できている。

5. まとめ

外部情報による sep2sep モデルの fine-tuning を行い、外部情報のドメインに適したコメントの生成を試みた。生成されたコメントの評価では、人手による評価と BLEU で fine-tuning の効果を確認できた。

現段階では、プログラムの行単位のコメント生成であるが、ブロック単位のコメント生成や疑似コードの生成を本研究の最終目的としている。

参考文献

- (1) 新開他: “手作業の学習教材を活用したアルゴリズム教育の試み”, 日本教育工学会第 31 回全国大会, pp.407-408 (2015)
- (2) 高橋他: “問題文に合わせたプログラミングコメントの自動生成”, 教育システム情報学会, 特集論文研究会, Jsis Research Report Vol. 33, no.7 (2019-3), pp.51-55, (2019)