

# マルチタスク転移学習による 小規模教師データを用いた意図理解

Intention Understanding with Small Training Data Sets

by Utilizing Multi-Task Transfer Learning

城光英彰 内出隼人 小路悠介 大塚貴弘

Hideaki Joko, Hayato Uchiide, Yusuke Koji and Takahiro Ohtsuka

三菱電機株式会社 情報技術総合研究所  
Information Technology R&D Center, Mitsubishi Electric Corporation

**Abstract:** In this research, we propose intention understanding method utilizing multi-task transfer learning. Our method improves intention understanding accuracy using data of different kind of domain as source domain. As source domain's training data, we use Japanese-English translation data (translation task) and Japanese Wikipedia data (sentence prediction task). As target domain's training data, we use transcribed utterance data of voice control of equipment. In this data, each utterance has one intention label. As an experimental result, we found that proposed method provides a performance improvement over previous transfer learning method in the case of small training data (the number of data for each intention label are 1, 3, 5, 10 and 30).

## 1. はじめに

機器の音声操作において、ユーザの発話文の意図を推定する意図理解は重要である。意図理解において Deep Neural Network (DNN) による方式の有効性が知られているものの [1]、適用分野（目標ドメイン）ごとに大規模教師データを作成する必要があり、コストがかかる問題が知られている。この問題に対し、大規模データが得られやすいドメイン（元ドメイン）のデータを活用し、目標ドメインに必要な教師データ数を削減する「転移学習」の有効性が報告されている [1]。筆者らも日英翻訳を元ドメインとした転移学習による意図理解方式を提案し、目標ドメインでのデータが小規模の場合の有効性を確認した [2]。この方式において、目標ドメインでの意図理解の正解率は、元ドメインのデータやタスクに強く依存しており、正解率を高めるためには、目標ドメインと「似ている」性質のドメインを元ドメインとすることが望ましいことが知られている [3]。しかし、このような性質のドメインの数は限られており、加えて、大規模教師データが手に入らない場合も多く問題である。

特定のドメインのデータやタスクに強く依存しない学習方式として、複数ドメインのデータやタスク

を同時に一つの学習機で学習するマルチタスク学習がある [4]。これを元ドメインでの学習に利用することで、目標ドメインでの精度向上が狙える可能性がある。

目標ドメインでの精度向上のために、マルチタスク学習を元ドメインでの学習に利用した研究には、Subramanian et al. [5] のものがある。Subramanian et al. は、元ドメインにマルチタスク学習を用いた転移学習（マルチタスク転移学習）が、同義文判定タスクなどに有効であることを示した。しかし、Subramanian et al. が目標ドメインの学習に用いた教師データ数は同義文判定（二値分類）タスクに対し 1,000 から 25,000 データと、容易に収集できる量ではない。加えて、意図理解タスクでの検証もしていない。そこで、本研究では、マルチタスク転移学習による意図理解方式を提案し、教師データが小規模な場合の意図理解タスクでの有効性の検証をする。

提案方式に使用した学習機は Encoder-Decoder Model [6] である。元ドメインとして日英翻訳データ（日英翻訳タスク）および日本語 Wikipedia（文予測タスク）を、目標ドメインとして各発話文に一つの意図ラベルが付与されている機器操作データ（意図理解タスク）を用い、教師データ数と意図理解正解率の関係をもとめた。

## 2. 提案方式

提案方式は、日英翻訳タスクおよび文予測タスク（元ドメイン）により学習した DNN のパラメータを意図理解（目標ドメイン）に活用することによって少数データでも意図理解の精度を高める方式である。なお、文予測タスクとは、文書（本実験では日本語 Wikipedia）を文に分割したとき、文書の  $t$  番目の文  $s_t$ （対象文）から、次の文  $s_{t+1}$  と前の文  $s_{t-1}$  の予測をするタスク [7] である。

元ドメインおよび目標ドメインの学習に用いた DNN は、どちらも 1 層の Embedding Layer（300 次元）と 1 層の Hidden Layer（150 次元）からなる Bi-Directional Long Short Term Memory (Bi-LSTM) [8] を備えた Attention 構造 [9] を持つ Encoder-Decoder Model である（図 1）。転移学習の方式としては、元ドメインで学習したパラメータを、目標ドメインでの学習機の初期値として利用する INIT (Parameter Initialization) 方式 [1] を用いた。Embedding Layer は 300 次元、Hidden Layer は 150 次元である。

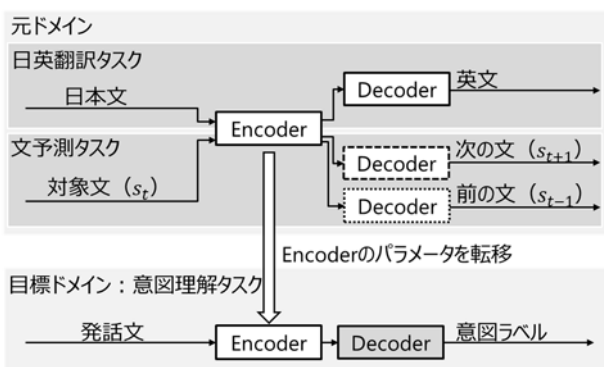


図 1: マルチタスク転移学習のイメージ図。元ドメインとして日英翻訳データ（日英翻訳タスク）および日本語 Wikipedia（文予測タスク）を、目標ドメインとして各発話文に一つの意図ラベルが付与されている機器操作データ（意図理解タスク）を用いた。

表 1: 使用した教師データ

教師データ	使用ドメイン	教師データ数	入力文の語彙数
Tanaka Corpus	元ドメイン	25,176 対	8,875
日本語 Wikipedia	元ドメイン	21,791 対 (次の文) 21,703 対 (前の文)	20,249
機器操作データ	目標ドメイン	5,600 対	20,247

## 3. 実験

### 3.1. 使用データ

元ドメインの教師データには、日英翻訳データとして Tanaka Corpus<sup>1</sup>を、日本語 Wikipedia として、Wikipedia 日本語版のダンプ<sup>2</sup>を用いた。目標ドメインとして各発話文に一つの意図ラベルが付与されている機器操作データを用いた（表 1）。意図ラベルの種類は 56 種類であり、各意図ラベルには 100 個の発話文が対応している。前処理として、Tanaka Corpus の、意味的にあいまいな対、文字長が 100 以上の対の修正や削除をした。また、計算時間削減の目的で、Tanaka Corpus と日本語 Wikipedia とともに、入力文の単語の 8 割が、機器操作データに出現する単語で構成されているもののみを、教師データとして使用した。

表 2: 目標ドメインで使用した教師データ数: 合計の教師データ数は「意図の数」と「各意図ラベルに対する教師データ数」の乗算値である

意図の数	各意図ラベルに対する教師データ数	合計の教師データ数
56	1	56
	3	168
	5	280
	10	560
	30	1,680
	45	2,520
	90	5,040

### 3.2. 実験内容

実験条件を表 3 に示す。パラメータの転移手法は Embedding Layer のパラメータを転移するもの (EMB)、Embedding Layer と Hidden Layer を転移するもの (ALL) の二種類である。元ドメインのタスクとしては、翻訳、文予測、翻訳と文予測を同時に行うマルチタスクの三種類がある。そのため、実験条件は合計で六種類となる。

実験内容を次に示す。

**実験 1:** まず、教師データ数を 10 としたときの各実験条件の意図理解正解率を算出し、提案方式が小規模教師データにおいて有効であることを示す。

**実験 2:** 次に、従来方式と提案方式の各々の中でもっとも正解率が高い実験条件について、意図ラベルごとの教師データ数を 1, 3, 5, 10, 30, 45, 90 と変化させたときの意図理解正解率を算出し、教師データが小規模なほど、提案方式が有効なことを示す。

実験 1、実験 2 とともに、正解率は 10 分割交差検定

<sup>1</sup> [http://www.edrdg.org/wiki/index.php/Tanaka\\_Corpus](http://www.edrdg.org/wiki/index.php/Tanaka_Corpus)

<sup>2</sup> <https://dumps.wikimedia.org/jawiki/>, 2017 年 11 月取得

により算出する。また、算出結果の信頼性を高めるため、各交差検定について、評価用のデータ 10 個以外の全てのデータ 90 個を少なくとも一回は教師データとして使用するよう複数回の学習をする。具体的には、意図ラベルごとの教師データ数が 1 のときは 90 回、3 のときは 30 回、5 のときは 18 回、10 のときは 9 回、30 のときは 3 回、45 のときは 2 回、90 のときは 1 回の学習をする。つまり、例えば教師データ数が 10 の場合は、合計 90 回 (9×10) の学習および意図理解正解率の算出をすることになる。

表 3: 元ドメインにおける学習の実験条件。EMB は Encoder の Embedding Layer のパラメータの転移を、ALL は Encoder の全てのパラメータの転移を表す。

実験条件の名称	元ドメインの教師データ		転移するパラメータ	
	日英翻訳データ	日本語 Wikipedia	Embedding Layer	Hidden Layer
従来方式: 翻訳(EMB)	使用	-	転移	-
従来方式: 翻訳(ALL)				転移
提案方式: マルチタスク(EMB)	使用	使用	転移	-
提案方式: マルチタスク(ALL)				転移
文予測(EMB)	-	使用	転移	-
文予測(ALL)				転移
転移学習なし	-	-	-	-

### 3.3. 実験結果

**実験 1** : まず、教師データ数を 10 としたときの各実験条件の意図理解正解率をもとめた。結果を図 2 に示す。全体として、Embedding Layer のみの転移(EMB)の方が、Encoder の全てのパラメータを転移した場合(ALL)よりも正解率が高いことがわかる。また、全実験条件の中でもっとも正解率が高い実験条件は「マルチタスク(EMB)」であることがわかる。この、「マルチタスク(EMB)」と、各他実験条件の意図理解正解率について、等分散の仮定の下で右片側二標本 t 検定を適用したところ、「マルチタスク(EMB)」の正解率が、各他実験条件と比べ有意に高いことが確認できた(有意水準  $\alpha = 0.05$ )。

**実験 2** : 次に、従来方式と提案方式の各々の中でもっとも正解率が高い実験条件である、「翻訳(EMB)」と「マルチタスク(EMB)」について、意図ラベルごとの教師データ数を 1, 3, 5, 10, 30, 45, 90 と変化させたときの意図理解正解率をもとめた。結果を図 3 に示す。各意図ラベルに対する教師データ数が 1, 3, 5, 10, 30 と少ないとき、提案方式の意図理解正解率は従来方式と比較し高くなることがわかる。この結果について実験 1 と同様の条件で検定を行っ

たところ、教師データ数が 1, 3, 5, 10, 30 と少ないとき、「マルチタスク(EMB)」の正解率が、「翻訳(EMB)」と比べ有意に高いことが確認できた。

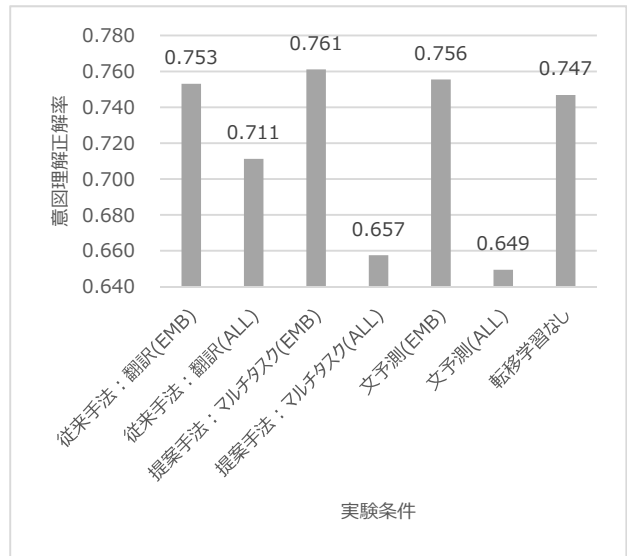


図 2: 教師データ数を 10 としたときの、各実験条件の意図理解正解率。「提案方式: マルチタスク(EMB)」の正解率が、他手法と比べ有意に高い。

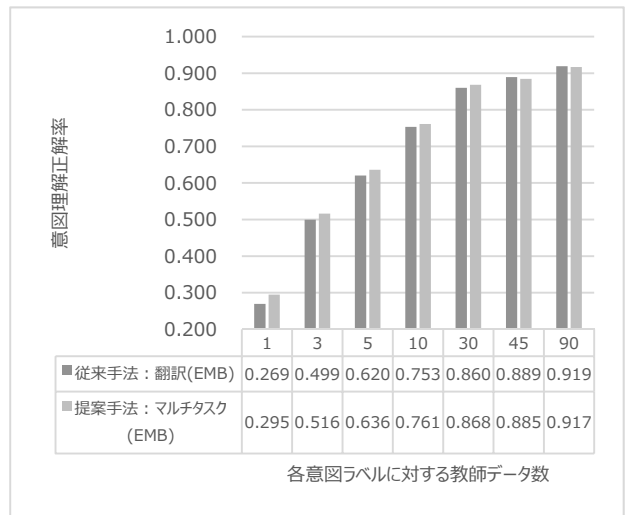


図 3: 従来方式と提案方式における目標ドメインの教師データ数を変化させたときの意図理解正解率。各意図ラベルに対する教師データ数が 1~30 と少ないとき、提案方式の意図理解正解率は従来方式と比較し有意に高い。

また、提案方式は教師データ数が少ないほど有効なこともわかった。特に、各意図に対する教師データ数が1の場合では、正答率は従来手法「翻訳(EMB)」の0.269から提案手法「マルチタスク(EMB)」の0.295へと、2.6ポイント向上した。この向上の理由は、元ドメインにマルチタスク学習を適用したことにより、元ドメインにおいて、特定のドメインのデータやタスクに強く依存しない学習ができたからと考えられる。

#### 4. まとめ

本研究では、小規模教師データを用いたマルチタスク転移学習方式を提案し、意図理解正解率の評価を行った。元ドメインとして日英翻訳データ(日英翻訳タスク)および日本語Wikipedia(文予測タスク)を、目標ドメインとして各発話文に一つの意図ラベルが付与されている機器操作データ(意図理解タスク)を用い、教師データ数と意図理解正解率の関係をもとめた。その結果、各意図ラベルに対する教師データ数が1, 3, 5, 10, 30と少ないとき、提案方式の「マルチタスク(EMB)」が、従来方式である「翻訳(EMB)」の意図理解正解率を有意に上回り、目標ドメインでの教師データ数が少ない場合において提案方式が有効であることがわかった。今後は、元ドメインのデータおよびタスクを追加することで、意図理解正解率の向上を目指す。また、目標ドメインのタスクを追加し、元ドメインで学習したパラメータの汎用性の評価も行う予定である。

#### 参考文献

- [1] Mou, L., Meng, Z., Yan, R., Li, G., Xu, Y., Zhang, L., and Jin, Z.: How transferable are neural networks in nlp applications?, In EMNLP, (2016)
- [2] 城光英彰, 内出隼人, 小路悠介, 大塚貴弘: 転移学習による小規模教師データを用いた意図理解, 電子情報通信学会 全国大会, (2018)
- [3] 神嶋敏弘: 転移学習, 人工知能学会誌, Vol. 25, No. 4, pp. 572-580, (2010)
- [4] Luong, M.-T., Le, Q. V., Sutskever, I., Vinyals, O., and Kaiser, L.: Multi-task sequence to sequence learning, In International Conference on Learning Representations, (2015)
- [5] Subramanian, S., Trischler, A., Bengio, Y., and Pal, C. J.: Learning General Purpose Distributed Sentence Representations via Large Scale Multi-task Learning, In ICLR, (2018)
- [6] Sutskever, I., Vinyals, O., and Le, Q. V.: Sequence to sequence learning with neural networks, In Advances in neural information processing systems, pp. 3104-3112, (2014)
- [7] Kiros, R., Zhu, Y., Salakhutdinov, R. R., Zemel, R., Urtasun, R., Torralba, A., and Fidler, S.: Skip-thought vectors, In Advances in neural information processing systems, pp. 3294-3302, (2015)
- [8] Ma, X. and Hovy, E.: End-to-end sequence labeling via bi-directional LSTM-CNNsCRF, In ACL, (2016)
- [9] Bahdanau, D., Cho, K. and Bengio, Y.: Neural machine translation by jointly learning to align and translate, In International Conference on Learning Representations, (2015)