

英語・中国語の2言語による知識グラフの推論を用いた対話状態推定のための素性抽出

Inferred Knowledge Graph Feature for Dialog State Tracking on English-Chinese Cross-Language Corpus

村瀬 行俊^{1*} 吉野 幸一郎^{1,2} 中村 哲¹
Yukitoshi Murase¹ Kouichirou Yoshino^{1,2} Satoshi Nakamura¹

¹ 奈良先端科学技術大学院大学

¹ Nara Institute of Science and Technology

² 科学技術振興機構

² Japan Science and Technology Agency

Abstract: Dialog State Tracking (DST) is one of the most important subtasks for task-oriented dialog systems. The module of DST predicts machine readable state of dialog from a current user utterance for the dialog manager. Last few years, Dialog State Tracking Challenge (DSTC) has been held as an annual shared task for dialog state tracking. In DSTC4, the focus of the challenge moves to tracking of state frames in English human-human conversation, which is more complicated than human-machine conversation. DSTC5 provides Chinese dialogs and translations to build a cross-language dialog state tracker. We proposed a method that creates inferred knowledge graph feature to improve the accuracy by using cross-lingual general knowledge graph such as Wikidata.

1 はじめに

対話状態推定は、対話の任意の時点におけるユーザ発話や対話履歴から、ユーザの意図や発話内容をユーザの状態として推定するタスクである [1]。対話システムは、これによって推定した状態を基に対話管理を行い、次に行うべきシステムの応答を選択している。そのため、対話状態推定の精度向上はシステムの応答の質を高めるうえで重要なタスクであると考えられている [2]。

近年、対話状態推定のシェアードタスクとして Dialog State Tracking Challenge (DSTC) がコンペティション形式で行われている。DSTC4では人間同士の旅行案内対話コーパスを用いて、旅行者とガイドの英語による対話内容を状態として推定するタスクの精度を競っている [3]。これに対し DSTC5では、DSTC4の対話コーパスに中国語の発話文と機械翻訳文を加えることで、言語横断対話状態推定のタスクを定義している [4]。これらのチャレンジで提案された対話状態推定の手法としては、人手によるルールベース [5]の他に分散表現を用いた畳み込みニューラルネットワークを用いる手法が

高い精度を実現している [6]。

対話状態推定を行うために外部の知識ベースをグラフとして構築し、そのグラフ上の推論結果を推定した状態とする手法が提案されている [7]。知識ベースにはオープンソースで利用可能な Wikidata [8] 等があり、データベースのような構造で構成されている。こうした知識ベースは一般のユーザにより編集可能なサービスとなっている場合も多く、Wikidata の場合は Wikipedia を基に自動更新も行われている。Wikidata では多数の言語において横断的に知識ベースが構成されており、DSTC5 のような複数の言語を対象とするタスクに相性が良いと考えられる。我々はこうした知識ベースを利用して部分グラフを作成し、その上で推論を行う事で特徴量の抽出を行い対話状態推定に利用する手法を提案しており [9]、本研究ではこれを多言語に拡張して利用した。

先に述べたように DSTC5 ではコーパスに英語と中国語の発話文が含まれており、Shi ら [6] の実験では両言語の発話文から分散表現による特徴量を抽出することで対話状態推定の精度向上を確認している。そこで本研究では、DSTC5 のタスクで Wikidata を用いて両言語の特徴量抽出を行い、多層パーセプトロンで複数ラベルの状態推定を行った。これを1つの言語での特徴量の推論結果のみを学習に用いるトラッカーと比較

*連絡先： 奈良先端科学技術大学院大学
〒 630-0192 奈良県生駒市高山町 8916 番地 5
E-mail: y-murase@is.naist.jp

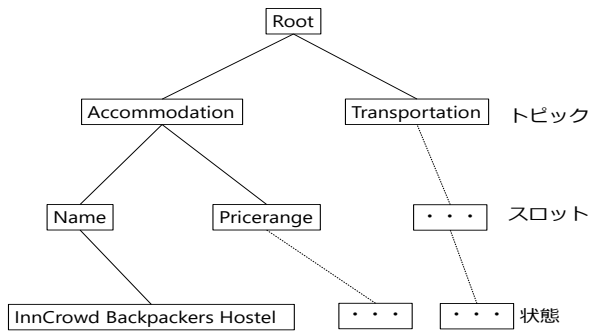


図 1: オントロジーの例

し推定精度の差を確認した。

2 実験

本研究では DSTC5 のデータセットのドメインにより Wikidata から部分知識グラフを作成し、知識グラフ上でラベル伝搬法の推論により特徴量の抽出を行う。これで抽出した特徴量の次元は大きいため、次元削減した物を多層パーセプトロンの入力とし、複数ラベル分類問題として対話状態推定を行う。

2.1 データセット

本研究で用いる DSTC5 のデータセットの特徴として英語と中国語どちらかの発話文と、もう一方の機械翻訳文が与えられている [4]。トレーニングセットでは英語による発話文が与えられており、これをもとに中国語へ機械翻訳された文が与えられている。これに対して開発・テストセットでは発話文が中国語、機械翻訳文が英語で与えられているため、これを考慮した対話状態推定を行う必要がある。データセットのサイズはトレーニングセットが 35 対話であり、これは DSTC4 の全てのデータセットである。また、開発セットは 2 対話でテストセットは 12 対話となっている。加えて、利用可能な知識としてオントロジーが与えられている。この簡略化した例を図 1 に示す。オントロジーではルートノードの子に 5 種類のトピックがあり、トピックの子にスロットがある。各スロットには複数の値が子として存在しており、この値が発話文から予測されるべき状態である。アノテーションされた対話状態は、これらオントロジーのリーフノードの値をもとに対話区間ごとに与えられている。

2.2 特徴量の次元圧縮

我々の提案する特徴量の抽出では Wikidata から部分グラフを作成し、グラフ上でラベル伝搬法を用いることで、観測ノードから対話中には生じなかった未観測ノードの特徴量を推定している [9]。ただし、このように推定した特徴量は作成した部分グラフのノードの数だけ次元があるため、モデルのサイズが大きくなり、計算時間も長くなる。そこで、[9] の手法で得た特徴量に主成分分析を用いることで次元の圧縮を試みた。その結果、英語による特徴量のサイズは 3100、中国語では 1900 程度で累積寄与率が 1 になり、精度を損なうことなくモデルの縮小と計算時間の短縮ができることが解った。

2.3 英語・中国語・両言語と多層パーセプトロンによる推定

本実験では、[9] のモデルをもとに英語と中国語それぞれで得られた特徴量を多層パーセプトロンの入力として、推定精度に差が出るかを検証した。また、両言語を同時に入力することによる精度の向上を検証した。多層パーセプトロンのハイパーパラメタは学習率が 0.0000005、ウェイトディケイが 0.000001、ドロップアウト率が 0.4、エポック数は英語、中国語、両言語、に対して各々 46, 90, 34 である。

3 結果

英語・中国語・両言語を入力として作成したモデルの F 値による評価を表 1 に示す。この結果から、英語と中国語では英語の特徴量を用いた場合のほうが精度が優れていた。これは、英語の知識ベースのエンティティが多く、特徴量のサイズが中国語の物より大きいことで、より多くの特徴を捉えることができたのではないかと考えられる。これに対して、両言語を入力としたモデルでは単言語を入力とした場合よりも 5~8% ほど F 値が上がっており、それぞれの言語で得られる特徴量が違う事から、精度の向上がみられると考えられる。最後に、予測結果の例を図 2 に示す。この例では英語を入力とした場合に "Chinatown" が推定され、中国語では "Suvenir shop" が推定されており、それぞれ正解状態に含まれている。また、両言語を使用することで、それぞれの言語から予測できなかった状態の組み合わせが推定された。

	英語	中国語	両言語
Precision	0.442	0.519	0.467
Recall	0.215	0.175	0.272
F-measure	0.290	0.262	0.340

表 1: F 値による各モデルの結果

入力発話文:

対話区間番号	215
中国語発話文	在牛车水%啊 %呃好, 又可以买纪念品
英語翻訳文	in Chinatown. er, good, you can buy souvenirs again.

推定結果:

正解状態	英語	中国語	両言語
Souvenir shop	Chinatown	Souvenir shop	Souvenir shop
Chinatown			Chinatown

図 2: 入力発話文と予測結果

4 まとめ

本研究の結果により, 多言語の知識グラフ上で推論を用いる特徴量抽出を行う事で, 単言語のみを用いる場合よりも, 両言語を用いた場合のほうがF値が向上することを確認し, 推定される状態が変化することが分かった. この要因として, 言語ごとに異なった特徴が推論されていることが考えられる. この結果をもとに, 推定された両言語間の特徴量の違いを分析し, どのような特徴量がこの差に影響を与えているのかを分析する必要がある.

謝辞

本研究開発の一部は総務省SCOPE(受付番号152307004)の委託を受けたものです.

参考文献

- [1] Williams, J., Raux, A., Ramachandran, D. and Black, A.: The dialog state tracking challenge, *Proceedings of the SIGDIAL 2013 Conference*, pp. 404–413 (2013)
- [2] LEE, S.: Extrinsic Evaluation of Dialog State Tracking and Predictive Metrics for Dialog Policy Optimization, *Proceedings of the SIGDIAL 2014 Conference*, pp. 310–317 (2014)
- [3] Kim, S., D’Haro, L.S., Banchs, R.E., Williams, J.D., Henderson, M.: The Fourth Dialog State Tracking Challenge, *Dialogues with Social Robots*, pp. 435–449 (2017)
- [4] Kim, S., D’Haro, L. F., Banchs, R. E., Williams, J., Henderson, M.: The Fifth DialogState Tracking Challenge, *Proceedings of the 2016 IEEE Workshop on SLT* (2016)
- [5] DERNONCOURT, F., LEE, J. Y., BUI, T. H., BUI, H. H.: Robust Dialog State Tracking for Large Ontologies, *Dialogues with Social Robots*, pp. 475–485 (2017)
- [6] Shi, H., Ushio, T., Endo, M., Yamagami, K., Horii, N.: Convolutional Neural Networks for Multi-Topic Dialog State Tracking, *Dialogues with Social Robots*, pp. 451–463 (2017)
- [7] Ma, Y., Crook, P. A., Sarikaya, R. and Fosler-Lussier, E.: Knowledge Graph Inference for Spoken Dialog Systems, *2015 IEEE ICASSP*, pp. 5346–5350 (2015)
- [8] Vrandečić, D., Krotzsch, M.: Wikidata: a Free Collaborative Knowledgebase, *Communications of the ACM*, pp. 78–85 (2014)
- [9] Murase, Y., Yoshino, K, Mizukami, M, Nakamura, S.: Feature Inference Based on Label Propagation on Wikidata Graph for DST, *Proceedings of IWSDS* (2017)