

ユーザ発話の揺れに頑健な音声対話システムのための 条件付確率場を用いた発話意図の推定

大田 健統

諏訪東京理科大学システム工学部電子システム工学科

1. はじめに

コンピュータが自然な言語を理解するためには、**省略**や**倒置**、**あいまいさ**に対して頑健である必要がある。

話者の**発話意図を推定することにより**、**省略**や**倒置**、**あいまいさ**、さらには**音声認識誤り**に対して頑健に動作する音声対話システムの構築を目指す！！

2. なぜCRFか

I. 話者の発話意図を推定するには、**文脈**や**意図を理解**するうえで**重要な単語の出現位置**を考慮する必要がある。

II. 音声認識誤りにより出現形が変化しても、**品詞**や**構文情報**から意図を推定できる必要がある。

SVMやブースティングは多値分類への適用や、文脈を考慮するために拡張が必要になるので

識別モデルであるCRFは、生成モデルであるHMMとは異なり**柔軟な素性設計が可能**であり、かつMEMMの問題点である**ラベルバイアス問題を解決**できるので

3. 提案法概要

学習素性

素性は、Cabochaによって解析された結果を用いて付与した。**出現形**、**品詞**、**分節境界**および**主辞**、**機能語**、さらには**固有名**や「**どこ**」など**疑問詞**

| 出現形 | 品詞 | 構文情報 | 重要語タグ | ラベル |
|-----|---------------------|------|------------|-----|
| 諏訪湖 | (名詞-固有名詞-一般)// | 2 | <location> | 1 |
| へ | (助詞-格助詞-一般)// | 1 | 0 | 1 |
| 行く | (動詞-五段・カ行促音便/基本形) | 2 | 0 | 1 |
| に | (助詞-格助詞-一般)// | 1 | 0 | 1 |
| は | (助詞-係助詞)// | 0 | 0 | 1 |
| どう | (副詞-一般)// | 2 | <how> | 1 |
| したら | (動詞-自立/ラ変・スル/連用形) | 0 | <how> | 1 |
| いい | (助動詞/特殊・タ/仮定形) | 2 | <how> | 1 |
| です | (形動詞-自立/形動詞・イイ/基本形) | 0 | 0 | 1 |
| か | (助動詞/特殊・マス/基本形) | 0 | 0 | 1 |
| . | (助詞-副助詞/並立助詞/終助詞)// | 0 | 0 | 1 |
| . | (記号-句点)// | 0 | 0 | 1 |

1段目のCRFで付与 2段目のCRFで付与

Fig. 1 学習素性の例

発話意図の推定方法

発話意図ラベル推定のために2段階に分けてCRFを適用する。

1段目: 音声認識結果に対しCabochaで解析を行い、構文情報までを付与したデータに対し、CRFを適用し重要語タグを推定する。

2段目: 1段目で重要語タグを推定したデータに対し、CRFを適用し各形態素に対し発話意図ラベルを推定する。

| | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 縄文の湯に行くにはどうしたらいいですか。 車山のスキー場へはどうやって行ったらいいですか。 霧ヶ峰はどこにありますか。 ビラタスロープウェイはどこにありますか。 お土産はどこで買えますか。 新鮮な魚を食べるにはどこへ行けばいいですか。 スキー場はいつオープンしますか。 美術館がオープンする時間を教えてください。 ... | 同一のラベルを付与 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|

Fig. 2 想定質問の一例

4. 評価

音声認識率

音声認識エンジン: Julius4.0.1

言語モデル: CSJ+諏訪圏の観光情報(Webから収集)

音響モデル: CSJそのまま

テストデータ: 未登録語を含まない男性1名による27発話

発話単位で音声認識率は約81.5%

ラベルの推定結果

1段目: 形態素単位で正しく重要語タグを推定できたのは約94%、発話単位では約70.4%であった。

2段目: 正しくラベルを推定できたのは約92.6%であった。

1段目のタグ推定の誤りや音声認識誤りに対して、**頑健に発話意図ラベルを推定できていると考えられる。**

| 出現形 | 品詞 | 構文情報 | 重要語タグ | ラベル |
|-----|----------------------|------|------------|------------|
| 白樺湖 | (名詞-固有名詞-一般)// | 2 | <location> | 4/0.766293 |
| に | (助詞-格助詞-一般)// | 0 | 0 | 4/0.789750 |
| は | (助詞-係助詞)// | 1 | 0 | 4/0.788964 |
| 何か | (名詞-代名詞-一般)// | 2 | <what> | 4/0.769862 |
| かい | (助詞-副助詞/並立助詞/終助詞)// | 1 | 0 | 4/0.736743 |
| ます | (動詞-自立/五段・ワ行促音便/連用形) | 2 | 0 | 4/0.692395 |
| か | (助動詞/特殊・マス/基本形) | 0 | 0 | 4/0.639101 |
| . | (助詞-副助詞/並立助詞/終助詞)// | 0 | 0 | 4/0.573596 |
| . | (記号-句点)// | 0 | 0 | 4/0.532982 |

Fig. 3 音声認識誤りを含むがラベルが正しく推定できた例

5. 考察

ラベル推定に失敗した原因について

1段目: 重要語タグの付与の仕方の問題

例えば<location>と<mono>タグの品詞はともに名詞-一般のものが多かったため識別が困難になったと考えられる

2段目: 音声認識と重要語タグを同時に誤った場合に、最終的なラベルの推定に失敗する傾向がある。ただし、CRF++により計算される周辺確率が低くなるため検出可能?

未登録語がラベル推定に与える影響について

| 出現形 | 品詞 | 構文情報 | 重要語タグ | ラベル |
|-----|----------------------|------|---------|------------|
| 皆 | (名詞-代名詞-一般)// | 2 | <what> | 2/0.332239 |
| 言う | (動詞-自立/五段・ワ行促音便/基本形) | 2 | <what> | 2/0.413390 |
| 館 | (名詞-一般)// | 2 | <mono> | 2/0.720865 |
| は | (助詞-係助詞)// | 1 | 0 | 2/0.867201 |
| どこ | (名詞-代名詞-一般)// | 2 | <where> | 2/0.945051 |
| に | (助詞-格助詞-一般)// | 1 | <where> | 2/0.963058 |
| あり | (動詞-自立/五段・ワ行/連用形) | 2 | 0 | 2/0.968133 |
| ます | (助動詞/特殊・マス/基本形) | 0 | 0 | 2/0.964782 |
| か | (助詞-副助詞/並立助詞/終助詞)// | 0 | 0 | 2/0.966576 |
| . | (記号-句点)// | 0 | 0 | 2/0.917364 |

Fig. 4 未登録語を含む発話へのラベル推定例

6. まとめと今後の課題

条件付確率場を用いた発話意図の推定を行った。

少量のデータによる評価ではあるが、音声認識誤りや未登録語に対して頑健に動作することを確認した。

今後の課題は...

- 大規模なコーパスを用意し、検証を進める
- 素性やタグの付与の仕方を検討
- 対話システムとして完成させる etc.