

「話の展開」を推測する方式について

A method for smoothing man-machine communication

高澤 千鶴^{*1}
Chiduru Takazawa

松澤 和光^{*1}
Kazumitsu Matsuzawa

神奈川大学大学院 工学研究科^{*1}
Graduate School of Engineering, Kanagawa University Graduate School

1. はじめに

自然言語処理技術の発達により、人とコンピュータとの会話を用いた様々なシステムやサービスが実現されつつある。しかし、これらは主として一方向的な情報伝達を目的としたものであり、単純で機械的な「会話」が実現されているに過ぎない。これを人同士の相互的なコミュニケーションのレベルに高めることが出来れば、より高度で豊かなサービスの実現につながると考える。このための技術課題として、本報告では「話の展開の推測」に着目した。人同士の会話では普通、相手の話の自然な展開を推測しながら話を進める。しかしコンピュータはそれが出来ないため、人とコンピュータの会話が不自然で冗長・単調なものになっているのではないかと。そこで「感情」や「意思」を表すデータを用い、話の展開を推測する方式を提案する。このような方式が実現できれば、人とコンピュータの会話を人間味あふれる楽しいものにできると考える。[1]

2. 「話の展開」推測方式の提案

2.1 コミュニケーションのモデル

コミュニケーションとは単なる情報伝達ではなく、情報の交換を通じてお互いの理解を深め、相互の関係をより快適なものに調整していく行為である。この立場から考えれば、会話における「話の展開の推測」とは、過去の会話から必然的に導かれる帰結を「予想」することでは決してない。そんな予想が可能ならばそもそも会話を進める意味が無いし、人間も展開を「当てる」つもりで推測しているわけではないのではないかと。ここでは人が会話において話の展開を推測する目的を、以下を相手に表明するためと捉えた。

- 1) 相手の発話を受け取る「主体」が確かに存在すること
- 2) その「主体」が相手の発話を能動的に受け取っている、つまり発話に反応する意思があること
- 3) 相手の発話の捉え方を示すことにより、「主体」自体を理解してくれるよう相手に促すこと

これらの反応があるからこそ、発話者はコミュニケーションすべき相手の存在を確認することができ、会話を進める意欲が生じるのではないかと。だとすれば、話の展開の推測は必ずしも発話者の意図に沿う必要は無く、推測を行っていること、それ自体を表明できれば良い。そしてその推測結果が、相手の理解できる範囲内であるという意味で「自然な」ものでありさえすれば、まさに会話は展開していくと思われる。

2.2 推測方式の基本的な構想

前節の考察の下で、話の展開を推測する以下の方式を提案する。

相手の発話を文型データと照合
照合した文型データAを感情データBに変換
感情データBによって意思データCを選択
文型データAに関連する文型データの中から意思データCに適合するDを選択
文型データD Aとして に戻る

～ を繰り返すことで話の展開を推測し続けることができると考えた。

なお、文型データとは動詞を指標として典型的な文構造をパターン化したもので、例えば「SがRをVする」のような形式のデータを想定している。「話の展開」は上記の方式で生成される文型データの連鎖で表現する。

また、感情データとは文型データの動詞Vに対する主観的な心の価値付けで、例えば「面白い」「難しい」などである。これが2.1節2)で述べた「相手の発話への反応」を表す。

一方、意思データとは次の行動・動作へ繋がる何かをしようとする考えで、「食べたい」「助けたい」のようなデータを考えている。これが2.1節3)で述べた「相手の発話の捉え方」を表す。

2.3 実現性の見込み

- 1) 文型データ
NTTのALTパターン辞書[2]やEDRの日本語動詞共起パターン副辞書[3]等、自然言語処理用に様々なものが開発・提供されている。
- 2) 感情データ
筆者らは先に物語文の自動生成を目的とした研究[4]において、文型データに対応する感情のデータベースを構築した。これは約9000個程度の規模であるが、これを手がかりにさらに大規模なデータを構築できる可能性がある。
- 3) 意思データ
まだ未検討であるが、上記の感情データと同様にすれば構築できる可能性はあると考える。
- 4) 文型データ間の関連性
文型データに付随する「文例」を利用し、同じまたは類似の名詞を持つ文型データを検索する方法、あるいは新聞コーパス等から共起頻度の高い文型データを抽出する方法などが考えられる。

3. 推測方式実現に向けた試み

3.1 データの準備

今後の自動構築に向け、データの基礎的な性質を明確にする目的で、今回は文型データと、文型データから文を生成するための単語データを人手により実験的に作成した。なお、感情データと意思データは文型データに含めた。

文型データ

Code Temp Type V Sent Feel Will Temp-No

単語データ

Code Temp W1 W2 W3 W4

文型データの Code は文型の動詞で分類した番号、Temp は文型で分類した番号である。Type は文型パターン、V はその動詞、Sent は Type の文例である。Type と単語データの W1 ~ W4 を使うことにより自然な文を生成することが出来る。今回はこのデータに、Feel と Will、Temp-No のデータを追加した。Feel は V に対しての感情データ、Will は意思データである。Temp-No は Type、Feel、Will から繋がる文の Temp である。

単語データの Code と Temp は文型データと同一のもので、文型データから単語データを呼び出すときに用いる。W1、W2 ... は Temp に対応した Type の<語 1>、<語 2>... に当てはめる名詞である。

なお、文型データの Code、Temp、Type、V、Sent は EDR の日本語動詞共起パターン副辞書を利用し、Feel、Will、Temp-No および単語データの W は手入力で用意した。

3.2 実験システム

提案方式の実現に向けた見通しを得るため、2.2 節の構想を元に人間の入力文から推測した文を出力する、ごく簡単なシステムを作成した。

ユーザの入力文 N と同じものが文型データにあるか検索する。在ればそのデータをそのまま用いる。無ければ、N の V と同じものを含む文例 Sent を呼び出す。複数あれば N の V と近い意味で使われている Sent を 1 つ選択し、N' とする。

「N と Feel、Feel から Will。」

次に、Code と Temp-No から単語データの W1 ~ W4 を取得する。

Temp-No から単語データの Temp を呼び出して、取得した W を当てはめて推測文 S を生成する。

「Will ので、S。」

Temp-No が 2 つ以上のときはこれらの処理を繰り返す。

「S と、S'。」

N から S' まで繋げれば

「N と Feel、Feel から Will、Will ので、S、S と、S'。」

このような推測結果が得られるはずである。

3.3 実験結果

一例として入力文 N 「戸ががたがたする」に対する結果を示す。

[入力文 N]

戸ががたがたする

[推測結果]

ドアががたがたと困る。

困るから修理したい。

修理したいので、私が修理を専門家に依頼する。

私が修理を専門家に依頼すると、専門家が修理依頼を請負う。

まず、文型データに N が無かったため、N の文末動詞 V と同じ文例 Sent 「ドアががたがたする」を文型データから呼び出す。

183 6 <語 1>が<語 2>がたがたする がたがたする ドアががたがたする 困る 修理したい 280 11436

データ Feel、Will を用いることで「ドアががたがたと困る。困るから修理したい。」まで推測できる。次に、Code183 と Temp-No280、11436 から単語データを呼び出す。

183 280 私 修理 専門家

183 11436 専門家 修理依頼

Temp-No の順に W を Type に当てはめる。Type はそれぞれ

<語 1>が<語 2>を<語 3>に依頼する

<語 1>が<語 2>を請負う

最後に、これらを繋げることで「修理したいので、私が修理を専門家に依頼する。私が修理を専門家に依頼すると、専門家が修理依頼を請負う。」までスムーズな推測の結果が出た。

4. 今後の課題

今回の実験で「話の展開」を推測する方式の実現に向けて、ある程度の見通しが得られた。今後は、2.3 節で述べたように各データの自動構築について検討を進め、より大規模なデータで方式の検証を行っていく予定である。

参考文献

- [1] 松澤和光、天沼博:「コミュニケーション知能」に関する工学研究の一提案,FIT2004,F-017,pp249-250,2004.
- [2] 池原悟、宮崎正弘、横尾昭男:日英機械翻訳のための意味解析用知識とその分解能、情処論 Vol.34,No.8,pp1692-1704,1993.
- [3] EDR 電子化辞書
http://www2.nict.go.jp/kk/e416/EDR/J_index.html
- [4] 高澤千鶴、松澤和光:「起承転結」構成の物語によるコミュニケーション、人工知能学会第 17 回ことば工学研究会 (SIG-LSE-A401), pp83-89,2004.