

# 構報探索雑談における自然なトピック遷移の実現

市川 宙<sup>†</sup> 徳永 健伸<sup>†</sup>

<sup>†</sup> 東京工業大学 大学院情報理工学研究科 計算工学専攻  
{ichikawa,take}@cl.cs.titech.ac.jp

## 1 概要

従来のタスク指向の情報探究型対話システムは、「来週、京都に旅行するための旅程を作りたい」といった、ユーザの具体的な要求を達成することはできるが、「何か面白い情報を知りたい」といったユーザの漠然とした要求に対応できない。本研究では、ユーザとの雑談を通してユーザの漠然とした要求に関連する情報を呈示する対話システムを提案する。本研究は特に、自然なトピック遷移の実現に焦点を当てた。まず、対話コーパスの分析結果(2節)を元に「トピック遷移の検出」(トピックが遷移したかどうかの判定)、「トピック遷移可能性の判定」(トピック遷移に適切なタイミングかどうかの判定)のアルゴリズムを作成し(3節)、上記コーパスを用いて評価した(4節)。更に、このアルゴリズムを用いてトピック遷移を行う対話システムを実装し、被験者実験を通じて評価した(5節)。

## 2 コーパスの分析

トピック遷移の検出・可能性判定の手がかりを見つけるために、人間同士の雑談を収録した Mister 0 コーパス<sup>1</sup>のうち 6 談話を人手で分析した。

まず各談話全体を読んでトピックに分割し、各トピック遷移について、「トピック遷移発話(トピック遷移直後の発話)」「トピック遷移前の数発話」に観測される特徴を分析した(表 1, 表 2)<sup>2</sup>。

## 3 アルゴリズム

### 3.1 特徴の判定

本手法では、表 1, 表 2 で挙げた特徴を機械的に判定するために、それぞれ以下の判定条件を用いる。

お互いに相槌を打つ 異なる話者による「相槌」または「感想」が連続した場合。ただし、「はい」「うん」などのキーワードを含み、内容語<sup>3</sup>を含まない発話を「相槌」とみなした。

<sup>1</sup>「びっくりしたこと」をテーマとする 2 者による 5 分間の対面対話の記録が 26 個収録されている。

<sup>2</sup>1 つのトピック遷移が複数の特徴に当てはまる場合は、両方にカウントした。

<sup>3</sup>ここでは「カタカナまたは漢字を含む、名詞または未知語」を「内容語」と判定した。

沈黙 発話終了後、1 秒以上次の発話が始まらない場合。  
感想 「びっくり」などのキーワードを含む場合。

明示的なトピック遷移の合図 「それより」「何かある」などのキーワードを含む場合。

初出の内容語を含む発話 それ以前に一度も出現していない内容語を含む場合。

イニシアティブの交代

- 直前 10 発話の間に片方の話者が「『私』『この前』」などのイニシアティブを示すキーワードを含む発話をした」または「相槌を全くしなかった」場合は、その話者にイニシアティブがあったと判定する。
- どちらでもない場合は、1 つ前の発話についてこの方法で判定した結果をそのまま用いる。上記いずれかの方法で「当該発話の話者が直前までイニシアティブを持っていなかった」と判断され、かつ当該発話がイニシアティブを示すキーワードを含む場合。

上記以外の特徴については、機械的な判定が困難なので、本研究では利用しないことにした。

### 3.2 トピック遷移検出・可能性判定

3.1 節の手法で判定した特徴を組み合わせて、トピック遷移検出・可能性判定を行う。

トピック遷移の検出では、表 2 の特徴のうち 3.1 節で判定アルゴリズムが実装できたもの (*sign, cont, init* の 3 つ) から 1~3 個の特徴を選び、選んだうちのどれかを満たす場合に、遷移を検出したとみなす<sup>4</sup>。どの特徴を選ぶべきかは、4 節の評価実験で検証する。トピック遷移の可能性判定についても同様の手法で、表 1 の *ack, sil, imp* を用いて判定する。

## 4 コーパスによる評価

### 4.1 評価方法

コーパスによる評価には、Mister 0 コーパスを用いる。このコーパスは 2 節の分析にも用いたが、ここで

<sup>4</sup>ただし例外として、トピック遷移検出後 3 発話の間は検出を行わないものとする。

表 1: トピック遷移前の数発話の特徴

特徴	略称	例	出現数
お互いに相槌を打つ	<i>ack</i>	「そっか」「そうそうそう」	14
沈黙(1秒以上)	<i>sil</i>		13
以前の発話を繰り返す		「超びっくりして」...「そう、びっくりした」	6
話の内容を一般化するような発話		特定のお化け屋敷について話したあとで 「お化け屋敷系びっくりするね」	4
感想	<i>imp</i>	「いいなあ」	4
上記の特徴のないトピック遷移			13
トピック遷移総数			45

表 2: トピック遷移発話の特徴

特徴	略称	例	出現数
明示的なトピック遷移の合図	<i>sign</i>	「ほかはないの？」 「あと、びっくりしたことだよね」	11
初出の内容語を含む発話	<i>cont</i>	「早稲田祭で大隈講堂に」	10
イニシアティブ(会話の主導権)の交代	<i>init</i>	相手の体験談のあとで 「あたし、CD屋さんでバイトしてるんだけど」	9
以前のトピックへの言及		びっくりした話のあとで別の話を挟んで 「びっくりするよね」	5
上記の特徴のないトピック遷移			11
トピック遷移総数			45

は分析に使用しなかった 20 談話を用いる。評価の手順は以下の通り。

- 対象となる 20 談話を人手でトピックに分割する。作業者は 3 人で、全員が全ての談話のトピック分割を行う。この結果得られたトピック分割の作業者間での一致率は、Fleiss の  $\kappa$  統計量 [1] で 0.561 だった。
- 対象となる 20 談話について、3.2 節の手法でトピック遷移の検出・可能性判定を行う。3.2 節で述べた特徴の全ての組み合わせ(検出と可能性判定でそれぞれ 7 通り)について評価を行う。
- 人手による結果を正解データとし、アルゴリズムによる検出・可能性判定の精度と再現率を求める。トピック遷移の検出では、検出したトピック遷移が、人手によるものと一致した場合に正解とする。トピック遷移可能性の判定は、判定後 2 発話以内に(人手で判定した)トピック遷移があれば正解とみなす。

ここで問題なのが、人手によるトピック分割には作業者によって揺れがあるという点である。そこで本研究では、作業者のうち少なくとも 1 人がトピック遷移と判定した箇所は「トピック遷移として許容できる箇所」、作業者全員がトピック遷移と判定した箇所は「確実にトピック遷移」とみなし、精度は「アルゴリズムがトピック遷移を検出・遷移可能と判定した箇所のうち、トピック遷移として許容できるものの割合」、再現率は「確実にトピック遷移であるような箇所のう

ち、アルゴリズムで検出・遷移可能と判定されたものの割合」として算出する。

## 4.2 結果

コーパスを用いた評価の結果を表 3、表 4 に示す<sup>5</sup>。精度  $p$ 、再現率  $r$  は、4.1 節の定義によるものである。F 値は、これらの値から  $F = 2pr/(p+r)$  という式で求めた。

表 3: トピック遷移検出の精度・再現率

判定条件	精度	再現率	F 値
<i>cont</i>	0.092	0.310	0.141
<i>init</i>	0.386	0.103	0.163
<i>init</i> $\vee$ <i>cont</i>	0.098	0.345	0.153
<i>sign</i>	0.429	0.103	0.167
<i>sign</i> $\vee$ <i>cont</i>	0.100	0.379	0.158
<i>sign</i> $\vee$ <i>init</i>	0.405	0.207	<b>0.274</b>
<i>sign</i> $\vee$ <i>init</i> $\vee$ <i>cont</i>	0.106	0.414	0.169

## 4.3 考察

### 4.3.1 トピック遷移の検出

表 3 の通り、検出の判定条件の中では *sign*  $\vee$  *init* (明示的なトピック遷移の合図またはイニシアティブの交代) の F 値が最大であり、他の判定条件の場合を

<sup>5</sup>判定条件には表 1、表 2 の記号を用いた。例えば *init*  $\vee$  *cont* とは、「イニシアティブの交代または初出の内容語があった場合に、トピック遷移があったと判定する」というものである。

表 4: トピック遷移可能性の判定の精度・再現率

判定条件	精度	再現率	F 値
<i>imp</i>	0.235	0.345	0.280
<i>sil</i>	0.230	0.310	0.264
<i>sil</i> ∨ <i>imp</i>	0.226	0.483	0.308
<i>ack</i>	0.254	0.586	<b>0.354</b>
<i>ack</i> ∨ <i>imp</i>	0.247	0.690	<b>0.364</b>
<i>ack</i> ∨ <i>sil</i>	0.238	0.759	<b>0.362</b>
<i>ack</i> ∨ <i>sil</i> ∨ <i>imp</i>	0.231	<b>0.793</b>	<b>0.358</b>

大きく上回っている。このことから、トピック遷移を検出するには *sign* ∨ *init* が有効だといえる。*cont* は再現率が高いが精度が非常に低く、有効な手がかりとは言いがたい。

しかし *sign* ∨ *init* の場合でも F 値は 0.274 であり、充分ではない。そこで、この条件でのトピック遷移の誤検出・見逃しの事例を調査したところ、キーワードによる「トピック遷移の合図」や「イニシアティブの交代」かどうかの判定が失敗しているところが多く見られ、これらの特徴の判定アルゴリズムが問題となっていることが分かった。

#### 4.3.2 トピック遷移可能性の判定

表 4 の通り、高い F 値を示したのは *ack*, *ack* ∨ *imp*, *ack* ∨ *sil*, *ack* ∨ *sil* ∨ *imp* の 4 つである。ここから、*ack* (お互いに相槌を打つ) がトピック遷移可能性の判定に有効な手がかりであることが示された。また、これと他の特徴を組み合わせることで、更に再現率を上げることができた。

精度は最高でも 0.254 と充分ではないが、今回の評価では、可能性判定の結果を「実際にトピック遷移が起きたか」で評価していたため、精度が低く見積もられた可能性がある<sup>6</sup>。可能性判定のより直接的な評価は、5 節の評価の中で行う。

## 5 対話システムによる評価

### 5.1 対話システムの構築

Web 上のニュース記事 [5, 7] を利用して人間と対話を行うシステムを実装した。ここでは個々の記事を 1 つのトピックとみなし、

- 3 節の手法で (ユーザ主導の) トピック遷移を検出した場合は、ユーザの発話に関連する記事を探索し、見つけた記事を新しいトピックとする。

<sup>6</sup> トピック遷移が可能なタイミングで、毎回実際にトピック遷移が起こるとは限らないため。

- 3 節の手法でトピック遷移が可能と判定した場合は、システム主導でそこまでの対話内容に関連する別のトピックに遷移する。

という方法でトピック (記事) を変えながら、それぞれの記事の内容を出力する。また、記事に関するユーザの質問に答えたり、相槌を打ったり、ユーザの発話を促すための質問をする機能も持つ。対話は、Web ブラウザを介してテキストチャット形式で行う。

記事の検索には [6]、質問への回答には [8, 9]、形態素解析・係り受け解析に [2, 4] を用いた。また、記事を適切な発話単位に分割するために [3] の手法を用いた。

### 5.2 評価手順と結果

5.1 節で構築した対話システムと人間との対話を行い、提案手法を対話システムに適用した際の、トピック遷移の自然さを評価する。

まず 9 人の被験者に対話システムと 20 分間対話してもらい、その対話を記録した。1 談話の平均発話数は 182 で、平均トピック遷移数は 7.0 だった。

次に、作成した対話記録からトピック遷移を抜き出し、5 セットに分割する。25 人のアンケート回答者に、この計 5 セットのうち 1 つを無作為に割り当て、各トピック遷移の前後<sup>7</sup>を提示し、各遷移について問 1 ~ 問 3 の設問に回答してもらう。

問 1 は「トピック遷移のタイミング (またはトピック遷移したという判断) の適切さ」、問 2 は「遷移先トピックの内容の自然さ」を評価する設問になっており、どちらも不自然 (正しくない) ~ 自然 (正しい) の 4 段階評価である。問 3 は、ユーザ主導のトピック遷移の検出漏れを列挙する設問である。

対話記録中のトピック遷移を、遷移の種類と検出した特徴 (表 1, 表 2) で分類し、それぞれについてアンケート評価を集計したところ、表 5 のようになった<sup>8</sup>。

### 5.3 考察

#### 5.3.1 トピック遷移の内訳

以下、トピック遷移の特徴には表 1, 表 2 の略称を用いる。

表 5 の出現数の分布を表 1, 表 2 と比較すると、特に *init* と *ack* の出現頻度が (Mister 0 コーパスに比べて) 少ないことが分かる。今回の対話はニュースを題材としたこともあり、ユーザが対話のイニシアティブを獲得する場面が少なく、結果的に *init* がほとん

<sup>7</sup> 遷移前のトピック全てと、遷移後のトピックの先頭 10 発話。

<sup>8</sup> 複数の特徴が同時に満たされた場合、例えば *sign* と *init* を同時に満たした場合は、*sign* と *init* の両方の行でカウントした。

表 5: 対話記録中のトピック遷移とそれに対するアンケート評価

遷移の種類	特徴	出現数	評価平均	
			問 1	問 2
ユーザ主導	<i>sign</i>	13	3.00	2.52
	<i>init</i>	2	3.25	3.25
	小計	14	3.07	2.62
システム主導	<i>ack</i>	6	2.77	2.55
	<i>sil</i>	45	2.59	2.30
	<i>imp</i>	3	3.07	2.72
	小計	49	2.62	2.31
全ての遷移		63	2.72	2.38

ど見られなかった。*ack* の頻度の差は、キーボードによる対話では対面対話に比べて相槌の利用が少ないことが原因だと思われる。

### 5.3.2 トピック遷移の評価

アンケート(表 5)からは、ユーザ主導遷移のほうがシステム主導遷移より問 2 の評価値がやや高いという結果が得られた。これは、ユーザ主導遷移ではユーザが具体的なトピックを指定している場合が多く、それを正しく検出できた場合に高い評価値を得ているためだと思われる。

ユーザ主導、システム主導ともに、特徴の出現数に非常に偏りがあったため、特徴の比較には不十分なデータとなった。特徴の比較を行うには、データ数を増やすなどの工夫が必要である。

また、アンケートによる評価が特に低かったものを分析したところ、以下のものが多く見られた。

- システム主導で遷移するときに、前のトピックやユーザ発話に関連するトピックが見つからず、全く別のトピックに遷移したケース
- システムが「～という、何を思い浮かべますか？」などの質問をしたあとで、それとは無関係なトピックに遷移したケース
- ユーザ主導遷移を検出できず、その後システム主導で別のトピックに遷移したケース

このような例をできるだけ減らす工夫が必要である。

### 5.3.3 検出できなかったトピック遷移

アンケートの問 3 の回答から得られた「本手法で検出できなかったユーザ主導のトピック遷移」を分析したところ、以下のパターンが多く見られた。

- コーパスによる評価でも多く見られた(4.3.1 節)、キーワードによる特徴判定のミス。
- 自分がそのトピックに詳しくないことを告げたり、現在のトピックに不満を示すことで、間接的にトピック遷移を要求するもの。人間同士の雑談にはあまり見られないパターンだが、システムとの対話ではしばしば見られた。

## 6 まとめと今後の課題

対話コーパスの分析に基づき、トピック遷移の検出・可能性判定を行う手法を提案し、評価した。検出には「トピック遷移の合図を示す手がかり語」「イニシアティブの交代」、可能性判定には「お互いに相槌」が有効な手がかりであることが示された。また、この手法を用いてトピック遷移を行う対話システムを構築し、評価した。

今後の課題として以下の項目が考えられる。

- キーワードの見直しやキーワード以外の判定手法の検討を行い、トピック遷移検出・可能性判定の精度を向上させる。
- 「人間同士の対話」と「人間とシステムとの対話」の違いを考慮し、対話システムに適したトピック遷移手法を模索する。

## 謝辞

本研究では、プロジェクト「科学研究費補助金・基礎研究(B)(1) 15320054 平成 15 年度～17 年度『アジアの文化・インターアクション・言語の相互関係に関する実証的・理論的研究』研究代表者 井出祥子日本女子大学教授」によって作成された「Mister 0 コーパス」を利用させていただきました。

## 参考文献

- [1] Fleiss, J. L. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, Vol. 76, No. 5 pp. 378-382, 1971.
- [2] Kudoh, T. and Matsumoto, Y. Japanese Dependency Analysis Based on Support Vector Machines. *EMNLP/VLC*, 2000.
- [3] 杉山聡, 堂坂浩二, 川端豪. 音声対話によるテキスト内容の伝達方法. *情報学論*, Vol. 41, No. 6, pp. 1883-1894, 2000.
- [4] 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛. 形態素解析システム『茶釜』 version 2.3. 3 使用説明書. 奈良先端科学技術大学院大学, 2003.
- [5] asahi.com <http://www.asahi.com/>
- [6] Google ニュース 日本語版 <http://news.google.co.jp/>
- [7] ITmedia <http://www.itmedia.co.jp/>
- [8] ウィキペディア日本語版 <http://ja.wikipedia.org/>
- [9] はてなダイアリーキーワード <http://d.hatena.ne.jp/keyword/>