

## 双方向マルコフ連鎖を用いた文章自動生成

渡辺 由貴 中田 豊久

新潟国際情報大学

### 1. はじめに

近年、ユーザの発言を理解していなくても適切な発言を返すことが出来るという人工無脳が身近になりつつある。この人工無脳とは、コンピュータと人間が一見会話をしているように見せかけることによって、コンピュータの「人間らしい」発言を楽しむというチャットプログラムの総称である[1]。しかし、この人工無脳には「発言の固定化」、「文脈や話題のズレ」、「感情の不安定さ」等の問題が挙げられ、「人間らしい」会話を行うにはまだ不安定な部分が多い。本研究は、上記の問題の一つである「発言の固定化」と「人間らしさの追求」に対して、双方向マルコフ連鎖を提案する。

### 2. 関連研究

人工無脳におけるマルコフ連鎖とは、返答文の再構成に用いられている。仕組みは、返答文候補に対して形態素解析を行い、品詞ごとに分割、文頭の語を接頭辞とし、接頭辞と一致する語を返答文候補から検索、一致した語に続く語を任意の数だけ再構成文章に追加し、追加文の文末にある語を新たに接頭辞として上書きし、一致する語を返答文候補から検索する。以上を返答文候補の文末まで繰り返すことによって、一つの返答文章から複数の返答文章を生成するというものである[2]。人工無脳における「発言の固定化」を解消する提案の一つであるが、一致語を検索する際に同じ語を何度も発見してしまい、再構成文章が同じ言葉を繰り返し、結果的に長く意味の分からない文が生成されることも珍しくない。また、同じ助詞が連続する文が再構成されることもあり、語の繋がりの不自然さや違和感から、人間らしい会話文ではない、という問題も挙がった。

### 3. 双方向マルコフ連鎖

従来のマルコフ連鎖や人工無脳における問題を解消するために、本研究では双方向マルコフ連鎖を提案する。これは探索に双方向性を持たせることで、連鎖のスタート地点を文頭から自由に指定できるようになったものである。一致語の探索を前方から後方のみだけでなく、後方から前方への探索を追加することで、文の中心を指定、話題のズレを抑えることにした。

また、同じ語を繰り返し発見することで文が無意味に長くなることを防ぐために、乱数と探索範囲の自動調整を追加した。マルコフ連鎖は一致語を複数発見した時、一つを選択するために乱数を使用することが多い。同じ語が再度接頭辞となった時、返答文候補の文頭から文末までを新しく探索し直し、一回目の探索と同じ位置の語、またはそれより前の位置の語を乱数で選択してしまうことでループが発生し、長文化が発生する。探索範囲の自動調整はこのループを回避するためのもので、一度選択した語以前の範囲は次の探索範囲から除外するよう設定されている。

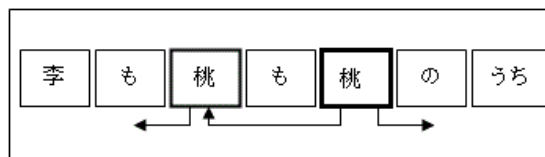


図1：双方向マルコフ連鎖の流れ

### 4. チャットプログラム

人工無脳における双方向マルコフ連鎖の有効性を検証するため、双方向マルコフ連鎖型チャットプログラムを作成した。本プログラムは、SNS サイト「Twitter」にて行われた会話の約 600 パターンを辞書として保持しており、返答文はユーザ側の入力文に依存する。これは、始めにユーザの入力文に対して形態素解析を行い、品詞ごとに語を分割、次に分割した品詞のうち、指定品詞（名詞、動詞、副詞、感動詞、形容詞、形容動詞）の語を抽出し、辞書内の抽出語を最も多く含む発言上

位三文を取得，その内の一文を乱数にて選択し，返答文候補とする．更にその返答文候補に対して形態素解析を行い，品詞ごとに文を分割，分割した語のうち，指定品詞を抽出し，その内の一語を文の中心として指定し，残りの上位二文と合わせて双方向マルコフ連鎖を行う，という仕組みである．

## 5. 評価実験

チャットプログラムを実際に使ってもらうことで，会話における人工無脳（双方向マルコフ連鎖型）の有効性を検証した．主な検証点は，人工無脳であるチャットプログラムにどれだけ人間を感じたか，どれだけ人間らしい会話を行うことが出来たかの二つである．

また，他人工無脳プログラムとの比較として，ログ型チャットプログラムを用意した．このログ型というものは，あらかじめ大量の会話ログを保持しておき，会話の際，ユーザの発言に最も近い会話を探し出し，その次に返答として発言された文をコンピュータ側の返答として出力するものである．

### 5-1. 実験手順と結果

男女 21 名に，ログ型チャットプログラムと，双方向マルコフ連鎖型チャットプログラムをそれぞれ 2 分間ずつ使ってもらった．

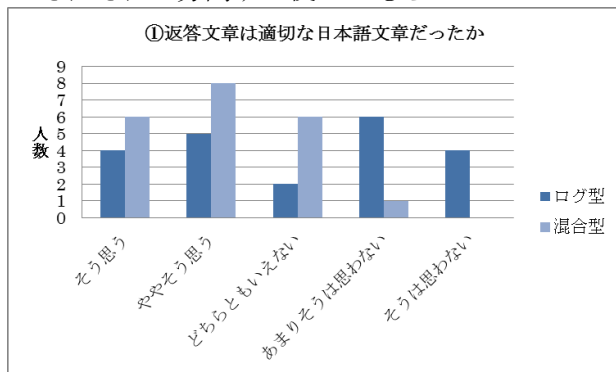


図 2：返答文章は適切な日本語文章だったか

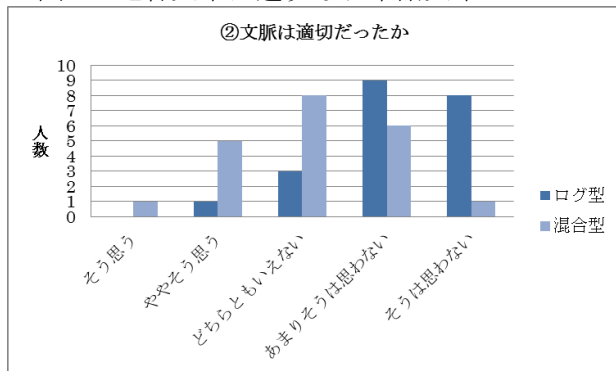


図 3：文脈は適切だったか

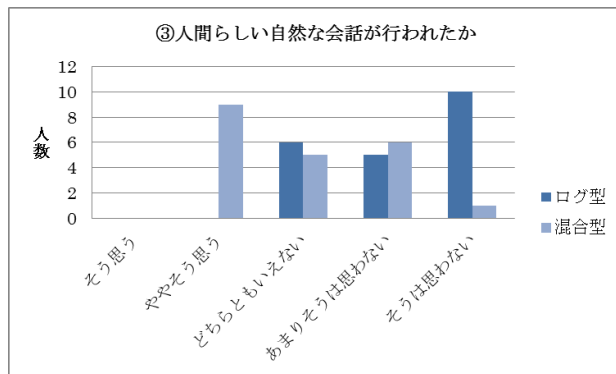


図 4：人間らしい自然な会話が行われたか

評価項目は，①自然な日本語文だったか，②文脈は適切だったか，③人間らしい自然な会話が行われたかの三項目である．

アンケート結果より，一つ目のプログラム(ログ型)に比べて，二つ目のプログラム(ログ型とマルコフ文生成型の混合)の方が，比較的人間らしい自然な会話を行ったという結果となった．会話が續かない，文脈が会話に沿っていない，という意見もいくつかあったが，その場合，入力文が会話ではなく極端に短い単語のみという傾向にあり，結果，辞書内の会話ログ検索がうまくいかない，また，会話ログ内容の偏りがある，ということが問題として挙げた．逆に，会話が繋がり，楽しかったという意見もあり，ユーザの入力文章に対応する会話ログ辞書の内容次第であると感じた．

## 6. おわりに

実験の結果，完成したプログラムは対応する会話パターンが少ない，偏っている，という問題もあったが，マルコフ文生成型とログ型の混合チャットプログラムの基盤として完成し，従来の人工無脳にある「発言の固定化」の問題解決を示唆した．今後は，口調や語尾を自動で統一し，キャラクター付けを行うことで，人間らしさの追及に努めたい．

## 参考文献

- [1] 人工無脳は考える  
(<http://www.ycf.nanet.co.jp/~skato/muno/>)
- [2] マルコフ連鎖による文章の自動生成  
(<http://blog.kentarok.org/entry/20040415/1081998210>)