

物語からの動作・発話ベースの構築と類推による応答生成

後藤 伸男[†] 原田 実[‡]

青山学院大学大学院 理工学研究科 知能情報コース[†] 青山学院大学 理工学部 情報テクノロジー学科[‡]

1. はじめに

近年、自然言語による応答生成システムの研究が盛んに行われている。例えば、質問を6種に分け回答を抽出する観光案内システム[1]や推論ルールと領域知識を全て自然語文で表現して推論する対話システム[2]などがある。一方、対話生成用知識構築の研究では、物語中に起きた出来事のみを対象に時間順に並べるシステム[3]や人物情報を抽出するシステム[4]などが提案されている。

本研究では、人間と計算機の自然な対話を類推を用いて実現することを目的に、類推による応答生成システム EVE2010 の研究開発を行った。類推には意味グラフ照合[5]を用いる。また、対話知識を小説から自動構築する。図1に示すように、発話と主体及び相手、更に発話の状況を意味グラフとして蓄積したものを対話知識とする。小説はハヤカワ文庫<SF1748>「月は無慈悲な夜の女王」(ロバート・A・ハインライン著 矢野徹訳)を用いた。

発話ベース	主体と相手		状況
	主体	相手	purpose . . .
「アダム・セレーネは憶えておきます」	アダム・セレーネ	マヌエル・ガルシア・オケリー	
「一番目のカテゴリーを二番目のと区別するのは、どういふところですか？どうか定義してください」	アダム・セレーネ	マヌエル・ガルシア・オケリー	それにアダム・セレーネはそうあっさり別れを告げるには機械と心を通わせるようになるには

図1 対話知識の概要

2. 応答生成システム EVE2010 の概要

EVE2010 は、小説から対話知識を自動構築するシステムと、対話知識を用いた類推による応答生成システムからなる。

2-1. 対話知識の自動構築の流れ

対話知識の自動構築システムは図2に示すように、小説を意味解析システム[6]で意味解析し、結果を意味グラフに展開する。この結果を用いて、まず各発話の主体と相手を決定して、次に各発話が発せられた状況を決定する。

1) 人名決定の流れ

Response generation based on analogy and construction of action-dialog base from story

[†]Graduate School of Science and Engineering, Intelligence and Information Course, Aoyama Gakuin University

[‡]Faculty of Science and Engineering, Department of Integrated Information Technology, Aoyama Gakuin University

まず、略称等が行われる人名を小説内で統一する。人名を、表1に示す8つの素性で学習した決定木で識別する。人名を示す名詞を、表2に示す5つの素性で学習した決定木で抽出する。最後に、識別した人名によって、人名の略称、及び人名を示す名詞を統一する。

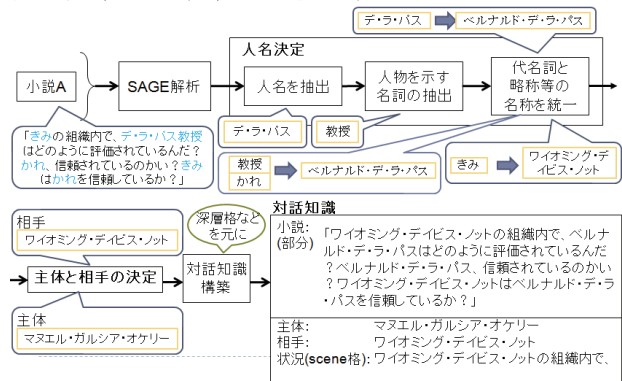


図2 対話知識の自動構築概要

表1 人名抽出のための素性

固有名詞	姓	名	普通名詞	未定義
出現割合	EDR概念辞書性状・性向3f9871に属する		EDR概念辞書人間30f6b0に属する	

表2 人名を示す名詞抽出のための素性

接尾語	人名の前に存在するか	EDR概念辞書組織30f746に属する
EDR概念辞書ある個体を基準とした関係で捉えた人間3aae71に属する		
EDR概念辞書職業,肩書,役割で限定した人間 444df2に属する		

2) 主体と相手の決定の流れ

発話の主体と相手は、文節間の役割的關係を表す深層格及び照応關係を表す照応辺[7]と発話順によって決定する。発話文では、表3に示す規則で決定する。一方、発話文以外では、表4に示す規則で決定する。さらに、主体と相手のどちらかが未決定の発話があれば、その前後の発話の主体と相手から決定する。

3) 状況の決定の流れ

図3に示すように発話が発せられる状況を決定する。

2-2. 類推による応答生成の流れ

類推による応答生成システムは図4に示すように、入力された発話の内容と、話し相手と状況が最も類似している発話ベース(n)を、対話知識から検索し、発話ベース(n+1)から類推を用いて応答を生成する。

入力発話: 「マヌエルは凡庸な人間だから、正確に残っているかは誰もわからないよ。」

状況：「期間」

発話ベース (n)：「やめろよ、ミルドレッド。開拓者の社会に結婚許可証などあるものか」

発話ベース (n+1)：「でもマヌエル・ガルシア・オケリーがたは記録を残しておられるんでしょう？」

この場合、図 5 に示すようにグラフ照合により入力発話と発話ベース (n) を検索し、次に両者の相違点を求める。さらに、発話ベース (n) と発話ベース (n+1) の対応点を求める。最後に、発話ベース (n+1) から類推によって応答「残っているのは誰も人間だから、わからないよ。」を生成する。

表 3 主体と相手の決定規則 1

主体	相手	
人名 + agent	人名 + main	人名 + goal
人名 + o-agent	人名 + agent(照応辺)	EDR概念辞書“呼びかけている相手” 3cea37に属する
人名 + o-agent(照応辺)	人名 + co-agent	

表 4 主体と相手の決定規則 2

直前の発話の相手かつ直後の発話の主体
人名 + agent + EDR概念辞書“相手に発話で伝える”3aa952に属する
人名 + object + EDR概念辞書“人が音を出すこと”3f971fに属する
人名 + agent + 動詞 (子音動詞ワ行)
人名 + o-agent + 動詞 (子音動詞ワ行)

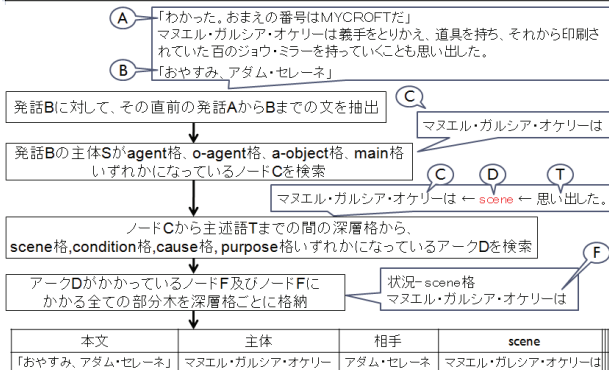


図 3 状況の決定の概要

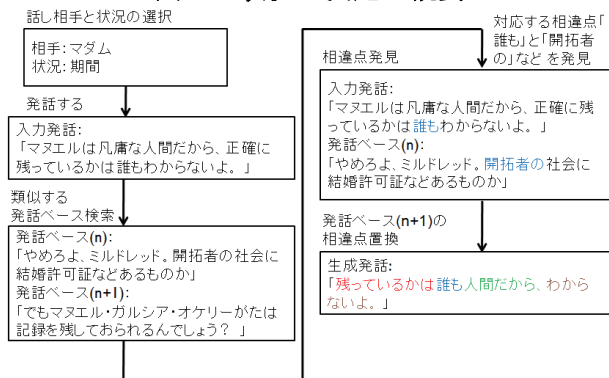


図 4 類推による応答生成概要

3. 応答事例

図 6 に示すように、一連の応答を生成でき、「小説の対話や状況の変遷に強く依存するが、発話の内容にある程度あった応答を返してい

る」という主観的評価をえた。

ノード類似度: 17.90			ノード類似度: 15.02		
入力発話	類似度	発話ベース(n)	発話ベース(n)	類似度	発話ベース(n+1)
「マヌエルは	0.35	ある			
凡庸な	0.00	ミルドレッド。			
人間だから、	0.50	結婚許可証など	結婚許可証など	0.72	記録を
正確に	0.00	社会に			
残っているかは	0.01	「やめろよ、	「やめろよ、	0.00	「でも
誰も	0.39	開拓者の	開拓者の	0.33	マヌエル・ガルシア・オケリーがたは
わからないよ。」	0.00	ものか」	ものか」	0.00	残しておられるんでしょう？」

図 5 意味グラフ照合で得られた対話の相違点

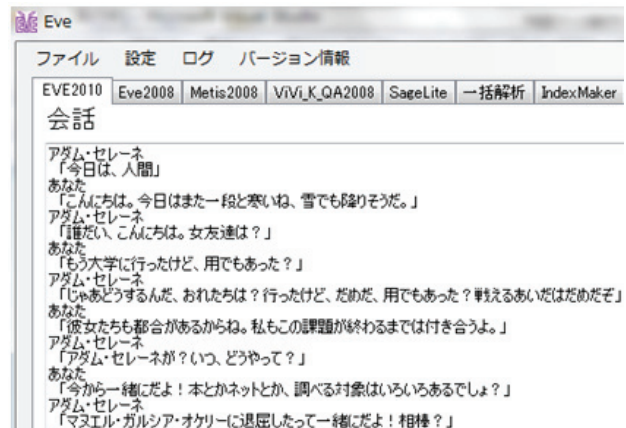


図 6 システムと人間の対話事例

4. 結論

自然言語ベースの対話システムを実現することを目標に、類推による応答生成システム EVE2010 の研究開発を行った。小説から対話知識を自動構築でき、類似した対話から、類推によって新しい応答を生成することができた。

5. 参考文献

- [1] 翠輝久, 河原達也: “質問応答技術を利用したインタラクティブな観光案内システム”, 言語処理学会第 13 回年次大会講演論文集, pp. 1160-1163 (2007).
- [2] 後藤伸男, 中川侑美, 原田実: “自然語ルールと知識をベースとした対話システム Eve の開発”, 情報処理学会研究報告, Vol. 2009-NL-191, No. 13, pp. 1-8 (2009).
- [3] 大石顕祐, 小方孝: “物語テキストからのストーリー抽出について”, 人工知能学会第 22 回全国大会論文集, 1C2-2 (2008).
- [4] 馬場こづえ, 藤井敦: “小説テキストを対象とした人物情報の抽出と体系化”, 言語処理学会第 13 回年次大会発表論文集, D3-3, pp. 574-577 (2007).
- [5] Minoru Harada, Yuhei Kato, Kazuaki Takehara, Masatsuna Kawamata, Kazunori Sugimura, and Junichi Kawaguchi: “QA System Metis Based on Semantic Graph Matching”, Proc. of the 6th International Conference on NII Test Collection for IR Systems (NTCIR6), Tokyo, Japan, pp. 448-459 (2007).
- [6] 原田実, 水野高宏: “EDR を用いた日本語意味解析システム SAGE”, 人工知能学会論文誌, Vol. 16, No. 1, pp. 85-93 (2001).
- [7] 南旭瑞, 原田実: “語意の類似性を用いた照応解析システムの開発 Anasys”, 情報処理学会第 64 回全国大会論文集, 3M-06 第 2 分冊, pp. 53-54 (2002).