

常識判断システムを用いた知的質問生成方式

Intelligent Question Generation for Natural Conversation Using Commonsense Judgment System

西澤優† Yu Nishizawa 渡部 広一† Hirokazu Watabe 河岡 司† Tsukasa Kawaoka

1. はじめに

コンピュータの普及は急速に進み、現在では人間の生活において非常に便利な存在であり必要不可欠なものである。そして、今後は人間のパートナーとしての役割がこれまで以上に期待されると考えられる。そのためにはコンピュータをより人に優しい使いやすいものにする必要がある。そこで人間とコンピュータとの双方向会話によるコミュニケーションをインターフェースとして取り入れることが注目されている。

会話は文章がどのような意図(質問文、命令文、etc...)、意味かを理解し、適切な応答(返答、質問)を返すことで展開されていく。本研究は会話の応答処理、特に質問に重点を置いている。

会話は様々な質問文により大きく展開していくことが多い。常識判断システムを用い、語を連想することで様々な展開が考えられる質問文の生成が可能である。しかし、その文章が最適かどうかの判断は難しい。そこで、本稿では、様々な質問文を生成し、その中から文章を導出する質問文生成方式を提案する。

2. 常識判断システム

常識判断は意味理解、質問文生成における語句の変換、連想で必要不可欠なシステムである。以下にその代表的なものについて述べる。

2.1 職種判断

職種判断とは、要求の入力(例えば、「本を買いたい」など)に対しそれを満たすことのできる職種(例えば、「本屋へ行けばよい」など)と、キーとなる動詞を出力するシステムである。

2.2 場所判断

場所判断とは、ある単語が場所に関する単語かどうかを判断し、そこに何が存在し(主体語)、何をやる場所(目的語)なのかを連想するシステムである。

2.3 感覚判断^[1]

感覚判断とは、名詞から感覚(視覚・嗅覚・聴覚・味覚・触覚)を連想するシステムである。

3. 意味理解

質問生成方式では文章を構成する語句の意味を把握することから始まる。本研究では、その方法として常識判断を用い置換処理や連想処理を行うことによって、文の内容を意味フレームに格納する意味理解システムを利用している。用意されている情報フレームは、5W1H(Who, What,

When, Where, Why, How)と「誰に」の Whom, および用言から成る。

4. 質問生成方式

先に述べた常識判断システム、意味理解を利用することにより一つの文章に対し様々な質問文を生成することが可能である。次節では、本方式の中で用いる質問文生成法について述べる。

4.1 質問文生成法

質問文生成には4つの手法を用いている。具体的には5W1Hの意味フレームを用い、その空きフレームに対し5W1Hの質問を行う生成方式(以下、5W1H生成法)、感覚判断、シソーラス、概念ベース^[2]を用い、特定条件の目的語から連想された名詞に対し質問文を生成する方式(以下、知的オウム返し法^[3])、そして、場所判断、職種判断、概念ベースを用い名詞から動詞を連想し、5W1Hと組み合わせる、もしくは、主語を目的格の語に置き換えた文章と連想された動詞の進行形を組み合わせた質問文の生成方式(以下、5W1H-連想動詞法、Whom主語-連想動詞法)がある。

4.2 質問文選択法

4つの質問文生成法により得られる文章は、そのすべてにおいて正しいものが得られるわけではなく、状況に応じての使い分けが必要であると考えられる。そこで、常識判断システムの特性を考え、生成法を使い分けることにより質問文を選択するアルゴリズムを提案する。

この選択法は役割の違いから大きく2つに分けられる。また、次節では質問文選択法の前半部について説明する。

4.2.1 質問文選択法・前半部

質問文選択法・前半部を図1に示す。この前半部では主に次の2つの処理を行う。

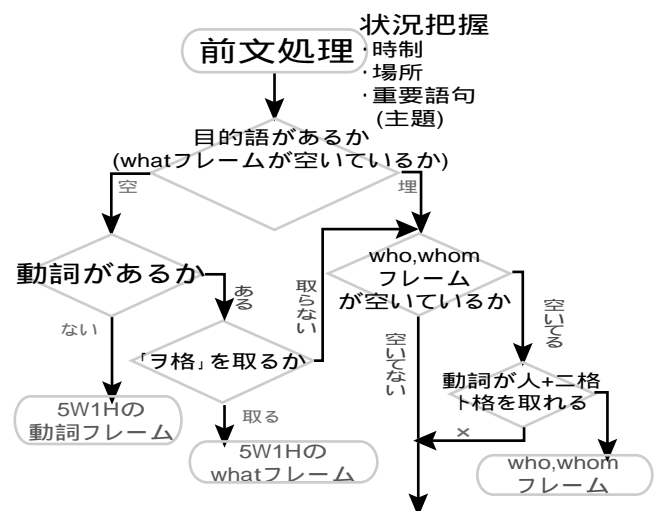


図1 質問文選択法・前半部

†同志社大学大学院工学研究科
Graduate School of Engineering, Doshisha University

前文処理

実際の会話では数多くの語が省略され展開している。つまり、一つの文章に対し処理では、省略された語に対する質問文生成がなされ、会話の流れにおいて不適切な文生成がなされる。そこで、前文処理において、“主語”、“目的語”、“時制”、“場所”を対称となる文章の意味理解に補うことで不適切な文生成を防ぐだけでなく、4つの質問生成手法の選択の可能性を広げる。

特定の5W1H法による文生成

人は会話において、ある一文中に他動詞を伴う目的語が省略ではなく存在していない場合はその文を理解することは難しく、疑問に感じる。また、“人+に+他動詞”の形をとる場合に“人”が存在していない文章も同様に感じる。

そこで、前文処理後の入力文の意味理解において What, Whom 意味フレームに語が格納されない場合、これらに対する質問文を生成する。

また、この方法では動詞の型により質問文選択を行っているが、これには日本語語彙体系の構文体系を元に作成されたシソーラスの文型パターンを利用している。このシソーラス文型パターンは、各質問文生成法や質問文選択法の後半部においても使用し、意味的に謝った文章の生成を防止している。

4.2.2 質問文選択法・後半部

前半部において質問文が選択されなかった場合、後半部の処理に移行する。これを図2に示す。

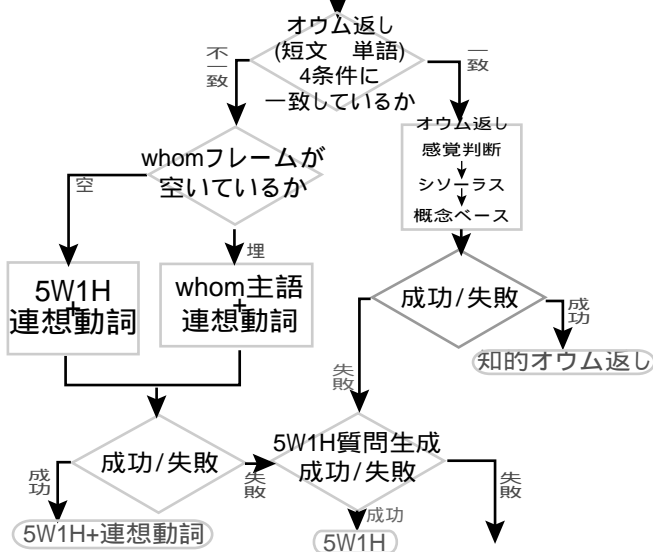


図2 質問文選択法・後半部

後半部では、始めに知的オウム返し法の選択が可能かを判断する。これは知的オウム返し法に示す条件に一致した特定の複合語である場合においてのみ連想語を獲得し、文生成を行うことができる手法であり、より具体的な名詞を連想することができること、また、条件一致の確率が低いことから判断を優先させている。

1	形容詞-自立 + 名詞(例:赤い食べ物)
2	名詞-形容動詞語幹 + 名詞(有名な料理)
3	名詞間に助詞-連体化(緑の野菜)
4	名詞間に助詞-並立助詞(林檎と蜜柑)

図3 オウム返し法4条件

次に連想動詞を用いた2つの文生成法について判定を行う。この手法の判定は常識判断システムにより動詞を連想できるかに大きく依存する。連想に成功した場合、その動詞の取りうる助詞と意味フレームの Whom に対する語の有無から5W1H-連想動詞法, Whom 主語-連想動詞法の選択を判断する。

上記の方法で質問文生成が行えなかった場合、5W1H法を選択する。

5. 評価

提案した選択法に対し評価を行った。テストデータとして、TOEICの参考書から対話文形式の質問文80組を用いた3段階(○:適切, □:適切ではない, ×:不適切)で評価を行った。

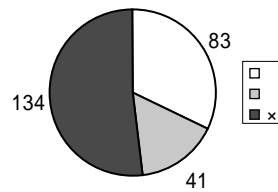


図4 評価結果

また、5W1Hなどの複数文生成の可能性がある場合では、全ての文に対し評価を行った。図4のとおり、80組のテストデータから258の質問文が生成され、適切な文83と、適切ではない文41を合わせて、精度46%が得られた。

この結果は意味理解システムの影響が大きいと考えられる。同じテストデータにおいて意味理解の段階での精度が55%であったことからわかる。意味フレームに正確に格納された状況下での質問文生成方式の評価を行う必要がある。

6. おわりに

本論文では、常識判断システムを用いた質問文生成法を適切に組み合わせ、最適な質問文を導出する質問生成方式を提案した。今回は評価によりその有効性を十分には証明できなかったが、4つの質問生成法から得られる文をランダムに選択する場合に比べ確実に精度は高いと考えられる。今後、前文処理、各質問生成法の拡張、精度向上を行っていくことでより知的で円滑な会話を展開できるようになることが期待できる。

本研究は文部科学省からの補助を受けた同志社大学の学術フロンティア研究プロジェクト「知能情報科学とその応用」における研究の一環として行った。

参考文献

[1] 渡部広一, 河岡 司: 常識的判断のための概念間の関連度評価モデル, 自然言語処理, Vol.8, No.2, pp.39-54 (2001)
 [2] 渡部広一, 河岡司: 常識的感覚判断システムにおける名詞からの感覚想起手法, 人工知能学会誌, vol.19, No.2, pp.73-82 (2004)
 [3] 岩下直人, 渡部広一, 河岡司: ”会話モデルにおける言い換え方式--知的おうむ返し”, 言語処理学会第8回年次大会発表論文集, C2-1, pp.315-318 (2002)