

## 一般的な概念辞書を用いたなぞなぞ質問文生成システム

金久保 正明†

静岡理工科大学総合情報学部人間情報デザイン学科†

## 1. はじめに

近年、なぞなぞ等「言葉遊び」をコンピュータにより自動生成する研究が盛んになり、謎解きとして駄洒落を用いるなぞなぞ作成システムも幾つか成功事例がある[1]。文献[1]では、或る名詞Aに対する動詞Xの因果関係Yに着目し、「AをXするとYとなりますが、XしてもXしてもYとならないものは何でしょう」といったテンプレートに当てはめる事により、質問文の生成に成功している。ここでは、動詞Xを例えば「溶かす」と「梳かす」等の同音異義語を用いる事で、駄洒落を基盤としたなぞなぞを成立させている。テンプレートを用いている分、精度の高い質問文の生成に成功しているが、テンプレートに合致した特定のパターンのなぞなぞ生成に限定され、なぞなぞ生成用に特化した自前の単語データを用いている。一方、従来から一般的な自然言語処理用の様々な概念体系や格フレームが用意されている。本論文では、このような言語資源の一般的な構造からなぞなぞの生成を試みる。例えば、下位概念Xがなぞなぞの答えであるとき、Xを伏せてその上位概念や、関係を持つ単語を繋げる事により、なぞなぞ質問文を生成する。

## 2. 提案システムの概要

## 2.1 生成するなぞなぞの基本的な構造

提案システムでは、「問い：完治可能な病気は？」「答え：化膿」のような、駄洒落となる二語のうち一方（可能）が質問文に出て、他方（化膿）が上位概念（病気）に隠れる構造となるなぞなぞを生成する。駄洒落の関係にある二語のうち、質問文に示す語を「明示語」、上位概念に隠れる語を「秘匿語」と呼ぶ。

上位概念を使って隠した下位概念の中に、明示語と駄洒落になる単語が多数あると、その中から「答え」として秘匿語だけを提示するのは無理がある。そこで、下位概念の中に該当する

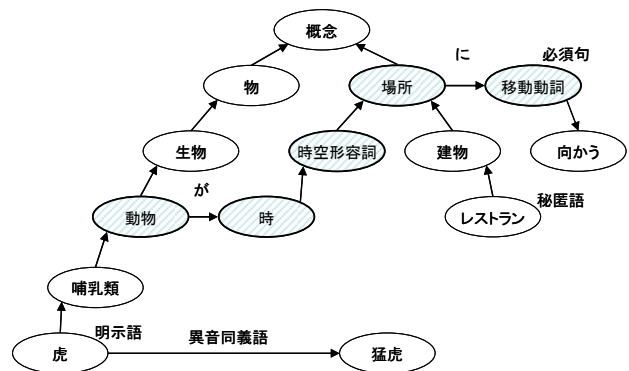


図1 提案手法によるなぞなぞ生成の概略

語が秘匿語しかない場合のみ、質問文を生成する。質問文を構成するには、駄洒落を成す二語以外を必要とする場合もある。この単語と上位概念で伏せた下位概念の中に別の駄洒落が出来ること、こちらも解答として考えられるのでそれを避ける工夫も行った。

図1に、提案システムによるなぞなぞ質問文生成の概略を示す。「猛虎が向かう場所は？答え：レストラン」「理由：れす（とら）ん、だから」という例となる。明示語の「虎」を、異音同義語である「猛虎」で隠している。太い円で囲った「動物が+時+時空形容詞+場所に+移動動詞」が共通文型であり、答えとなる秘匿語（レストラン）の上位概念は「場所」となり、これが質問文に使われる。

異音同義語の他にも、人間の作成するなぞなぞでは、時々英語読みを使って難しくする。「問い：国連が制定した地球の日はいつ？」「答え：明日（アース）」は、日本語読み的一致だけで考えると答えが出ない。従って、明示語の英語読みと秘匿語の日本語読み、明示語の日本語読みと秘匿語の英語読みが駄洒落になる場合も、駄洒落であるとした。

## 2.2 提案システムによるなぞなぞ生成

提案システムでは、単語ごとに（1）単語表記（2）日本語読み（3）英語読み（4）上位概念（5）下位概念（6）異音同義語（7）異音同義語の日本語読み、をデータベースとして保持する。これとは別に、例えば「具体物は場所に在る」といった接続可能な上位概念の組み合わせを基本文型として定義し、これを複数用

意する。ユーザが任意の単語を入力すると、駄洒落になる別の語を検索し、双方の上位概念を辿って両方を含む共通文型からなぞなぞを生成する。入力語を明示語とし、駄洒落を成す語（秘匿語）の上位概念と、他の文を構成する語により質問文を生成する。質問文の生成には規則的な語の並び替えや最後に「？」を付す等の操作も行う。

### 3. 試験評価

基本文型はサンプルとして、以下の4文型のみを扱った。+は、単語の接続を示す。(1) 具体形容詞+可動物が+時空形容詞+場所に+存在動詞 (2) 動物が+時+時空形容詞+場所に+移動動詞 (3) 人が+時+具体形容詞+物を+対物動詞 (4) 人が+時+具体形容詞+可動物を+可動動詞。「物」は具体物、「可動物」は動物が動かせる具体物、「具体形容詞」は具体物の形容詞、「時空形容詞」は時刻、空間の形容詞、「存在動詞」は「在る」、「移動動詞」は「行く」等の場所に接続できる動詞、「対物動詞」は「物を」に接続出来る動詞、可動動詞は「可動物を」に接続出来る動詞、「時」は時刻又は或る時間帯を指す語、「場所」は「野原」のような空間や、「警察署」のような建物を指す単語群である。単語は、約7500語の基本語を収録した日本語単語集[2]から、共通文型で用いる各概念カテゴリーに含まれる単語を全て登録した。

#### 3.1 生成された駄洒落数及びなぞなぞ数

登録した647語の全ての単語に付き、総当たりで生成された駄洒落は350。共通文型等の条件から生成されたなぞなぞは156個だった。このうち16個のなぞなぞと、インターネット上に公開されている人気のあるなぞなぞのサイト[3]から、駄洒落を謎解きとするものを16個選び両者をランダムに混ぜて試験評価を行った。

#### 3.2 試験評価の概要

被験者は16~23歳の一般的な大学生及び高校生14人で、普段はあまり、なぞなぞやらない被験者が多かった。一問につき一分の制限時間を設け、解答して貰った。事前に被験者には謎解きが駄洒落である事は知らせ、英語読み、異音同義語等が使われる場合がある事などは知らせしていない。解答後、各問題につき、納得感、意外性、面白さを問う5段階評価のアンケートを実施した。

#### 3.3 試験評価の結果

全被験者の正解率の平均は、コンピュータ制作問題に対し38.8%、人間の制作問題に対し59.3%で、コンピュータが作成した方が難しか

った。自由記述でも「問題が難し過ぎる」という感想が目立った。少なくともなぞなぞとして成立するレベルの難しさはコンピュータでも制作出来たと考えられる。コンピュータ制作問題に対し、コンピュータが制作したと答えた確率は51.3%、人間制作問題に対し人間が制作したと答えた確率も53.1%で、チューリングテストの結果も双方ともほぼ半数であった。

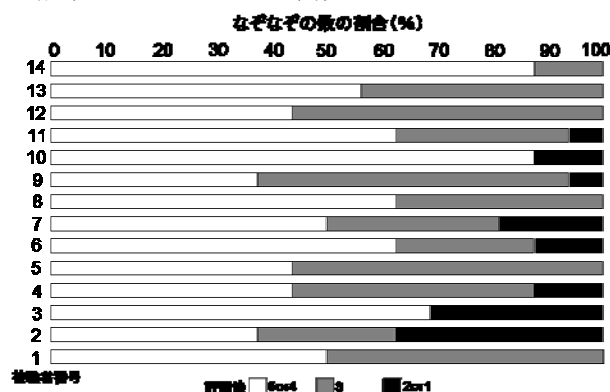


図2 面白さに対する評価

図2に、コンピュータ制作問題に対する被験者別・評価値別の面白さに対する評価結果(なぞなぞの数の割合)を示す。肯定的評価と否定的評価の割合を比較するため、5段階評価値を「5か4」(肯定的評価)、「3」(どちらでもない)、「2か1」(否定的評価)の3段階に分け、割合を示している。「肯定的評価のなぞなぞの数の割合」と「否定的評価のなぞなぞの数の割合」には差が無いという帰無仮説を立て、有意水準5%でt検定を行った。t値は2.19で、帰無仮説は棄却された。

以下は、肯定的評価の多かった提案システム制作のなぞなぞである。

よく知られた植物は? 答え「梅」理由: ゆ(うめ)い、だから

帽子を見付ける人は? 答え「消防士」理由: しょう(ぼうし)、だから

#### 参考文献

[1] 濱田真樹, 鬼沢武久: 同音異義語の意味の多様性を構造にもつなぞなぞの生成, 日本知能情報ファジイ学会誌, Vol. 20, No. 5, pp. 696-708, (2008).  
 [2] 高橋恭子, 高橋 寛: 日本語いつでもどこでも単語集, 大盛堂書房, (2008).  
 [3] <http://zonazona.at.infoseek.co.jp/index2.html>