

対話システムにおける「答」の統計的決定

1H-2

藤本 和則

松澤 和光

日本電信電話(株) コミュニケーション科学研究所

1 はじめに

人間と自然に対話するシステムのもつ重要な機能の一つとして“質問から答を導く推論”がある。答を導く推論においては、各命題の真偽の導出よりも、それを答として提示することが質問者にとって有効かどうかの判定が重要となる(例えば[1])。我々は、このように情報の有効性を評価できる枠組として「統計的決定理論[3]」を採用し、これに基づく推論システムの構築を進めている。

本稿では、まず、対話システムでの“質問に対する答”の性質について述べ、統計的に“答”を決定するシステムの構成を示す。統計的な決定を行うシステムの構築には、決定の効果程度を表す「主観効用値」の獲得が必要となるが、本稿では、これを実現する方法として、(1) 人間が記述した知識から獲得する方法と、(2) 質問された“状況”を利用して生成する方法の二つの方法を示す。

2 対話での“答”

「James が電線にとまっている雀を撃ちました。どうなるでしょう?」と質問されたとき、人間は「雀が逃げる。」や「雀が死ぬ。」など、雀について答えることが多い。しかしながら、質問に対して「論理的に正しい解」は、雀についてのものだけではなく、銃について「銃から弾が出る。」や、James について「James が雀を見る。」など、数多く存在する。人間は日常の対話活動において、複数の「論理的に正しい解」の中から適切な一つを選んで“答”として回答している。我々は、この“答の選択”は質問者と回答者の間の自然な対話を実現する一つの重要な要因であると考え、この選択機能の機械的

“Statistical Decision Making for Attentions in Dialogue Systems”, Kazunori FUJIMOTO, NTT Communication Science Laboratories, 1-2356 Take, Yokosuka-shi, Kanagawa 238-03, Japan.

な実現の研究を進めている。以下では、こうした“答の選択機能”をもつ対話システムの構築へ向けての、我々のアプローチについて述べる。

3 “答”の統計的決定

まず、適切な“答”を選び出すには、どういった答が適切で、どういった答が適切でないかを規定する必要がある。我々は、ある“答”が適切かどうかは、その“答”が質問者にとって望まれるかどうかで決まると考えた。機械と人間との自然な対話を実現するにあたっては、質問者に望まれる答を提示することが必要である。したがって、複数の正しい解のうち、質問者に望まれるものを回答するメカニズムを実現する。

こうしたメカニズムを実現する推論法としては、さまざまな形式のものが考えられるが、我々は、質問者にとっての「答の効用」の程度を数量化して扱うことのできる「統計的決定理論」の枠組を採用して、選択メカニズムの構築を進めている。

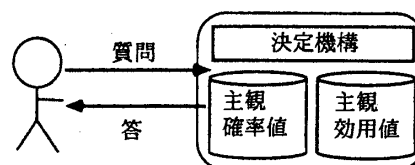


図1: 対話システム

統計的決定理論に基づく対話システムの構成を図1に示す。図1において、人間から質問が投入されると、決定機構は各答候補の期待効用値を計算し、その値が最大になる答候補を答として決定する。

こうした「統計的に答を決定するシステム」の実現の難しさは、本質的には「効用値をどのように獲得するか」にある。効用値の獲得法として、我々は二つのアプローチをとって研究を進めている。以下では、この二つのアプローチについて説明する。

3.1 人間の記述に基づく獲得

最も単純な獲得方法は、人間が直接的に主観効用値を記述する方法である(図2(a))。しかしながら、人間の不得意な数値化作業に基づく獲得法では、大量の主観効用値の獲得は困難である。そこで、我々は、人間にとって記述の容易な形式で集められた「主観情報」から、主観効用値を推定して獲得するという枠組の研究を進めている(図2(b))。我々は、これまでに、主観効用値を獲得するさい、概念ベース [4] から自動的に算出される「命題間の類似度」を用いて、人間が記述する労力を削減する方法 [2] を提案した。

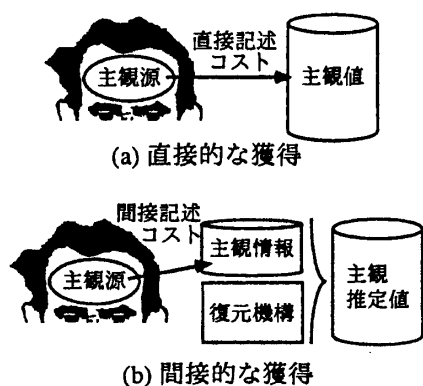


図2: 主観値の獲得

3.2 質問された“状況”に基づく獲得

我々は、上記のように、人間が記述した知識に基づく方法とは別に、「質問がなされた状況」から質問者の望む“答”を推定する方法の研究を進めている [1]。ここでは、答の決定問題の一つとして、分類問題において答の抽象度を決定する問題を取り上げ、質問者の注目する属性から質問者の望む“答”の抽象度を推定する方法について述べる。

図3に示すように、「この動物なに?」と質問されたとき、正しい答としては、「鳥です。」、「鶏です。」、「尾長鶏です。」など複数の抽象度の答が存在する。人間と自然に対話するシステムでは、これら正しい解のうち、質問者の望む答を選んで回答してやる必要がある。

我々は、こうした抽象度を決定する方法として、「質問者の注目する属性」を利用する方法の研究を進めている。この方法では、質問者が回答を望む抽象度は、「質問者の注目する属性で最も良く分類されている抽象度」であるという仮説に基づいて決定を行う。具体的には、まず、各抽象度がその属性のもとに分類されてい

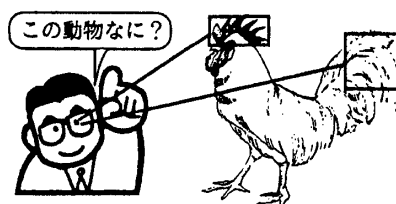


図3: 答の抽象度決定問題

るか否かの程度を表す「分類度」を各抽象度について計算し、その値が最も大きくなる抽象度に着目する(図4)。そして、着目した抽象度の中で確からしいカテゴリを選択することにより、適切な答を決定するという方法である。

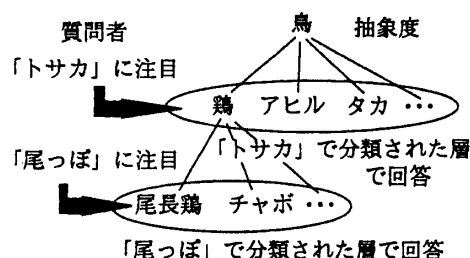


図4: 抽象度決定の原理

4 おわりに

本稿では、まず、対話システムでの“質問に対する答”の性質について説明し、統計的に“答”を決定するシステムの構成を示した。統計的な決定を行うシステムの構築には、決定の効果程度を表す「主観効用値」の獲得が必要となるが、本稿では、これを実現する方法として、(1)人間が記述した知識から獲得する方法と、(2)質問された状況を利用して生成する方法の二つの方法を示した。

参考文献

- [1] 藤本, 湯川, “質問の情報量に着目した回答の抽象度決定法”, 情報処理学会第50回全国大会, Vol. 3, pp. 51-52, 1995.
- [2] 藤本, 村松, “統計的決定理論に基づく大規模推論システムの構築”, 人工知能学会第10回全国大会, pp. 317-320, 1996.
- [3] 矢島 他, “自然科学の統計学”, 東京大学出版会, 1992.
- [4] Kasahara, K., et.al., “Viewpoint-based measurement of semantic similarity between words”, *Pro. AI & Statistics '95*, pp.292-302, 1995.