

発話対モデルに基づいた応答生成における文脈の利用について

高野 敦子 平井 誠 北橋忠宏
 兵庫大学 経済情報学部 松下電器産業(株) 中央研究所 大阪大学 産業科学研究所

1. はじめに

自然言語インターフェイスを構築する上で、ユーザフレンドリーな応答の生成は重要な課題の1つである。従来の発話生成を扱った研究は、その大半がユーザが明確な目的を持つ事を前提としたタスク遂行型対話を対象としていた^[1]。実際、ユーザの目的を用いずに従来の意味でユーザフレンドリーな応答を導くことは困難である。しかし、近年急増しているインターネットなどを使った情報検索作業におけるユーザインターフェイスの実現のためには、明確な目的を持たないユーザを考慮せざるをえない^[2]。そこで、本研究は、明確な目的を持たないユーザに対してもユーザフレンドリーな応答生成の実現を目指す。明確な目的を持たないユーザに対しては、タスク遂行よりユーザの発想支援に重点を置いた積極的な関連情報の提供が望ましいというのが、本研究の基本的考え方である。本稿では、明確な目的を持たないユーザの情報要求に対する応答に焦点を絞って述べる。

我々は既に発話認識を目的とした発話対モデルを構築している^[3]。それは収集した対話例の分析による多様な応答パターンの分類に基づいて構築したものであり、その中には、明確な目的を持たないユーザの発想を支援する可能性を持った応答パターンも含まれている。我々はまず、そのモデルを応答生成の観点から捕らえ直し、応答生成モデルを構築した。ただしこれはあくまで認知的モデルであり、実際にインターフェイスを構築する際に必要な計算モデルとするためには対象となるシステム固有の情報を用いて翻訳する必要がある。しかしこの認知的モデルからは、単に個々のインターフェイスにおける応答生成フローが導けるだけでなく、一般にユーザフレンドリーな応答を生成するために必要となる情報やその獲得方法及び応答生成への利用に対する知見を得ることができる。

生成モデルが認識モデルと最も異なる点は、分岐点で

の意味であり、生成モデルの場合それは必然的に決まるものと任意性を持つものに分かれる。さらに、発話内容についても任意性が生じる。本稿は、まず、応答生成モデルにおける分岐点での選択を左右する要因を分析する。その結果、基本要素として、従来から用いられる‘ユーザの目的’に加えて、‘ユーザの興味’及び‘ユーザの対話形態に対する嗜好’を導く。次にその情報を獲得し、実際に応答内容を決定する方法について議論する。その結果、認識処理においては、文脈情報を用いない、対をなす2発話のみの分析が有効であることを主張したが、応答生成の場合には文脈情報が不可欠であることがわかる。ここで、後半の議論のためには対象となるシステム固有の情報が必要となるため、ペーババックデータベースに対する情報検索システムを取り上げて具体的に議論を行う。

2. 認知的応答生成モデル

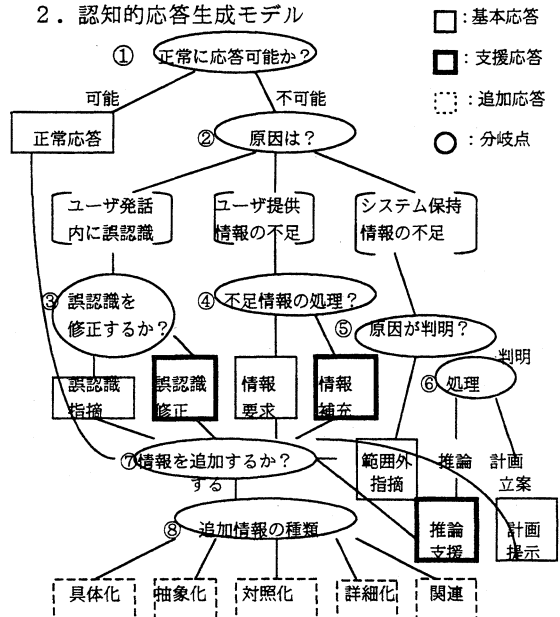


図1. 応答生成モデル

既に提案している発話対モデルを生成の立場から再構

築した応答生成モデルを図1に示す。基本的な構造は[ユーザの情報要求に対する直接的応答(8タイプ)+追加情報の応答(5タイプ)]である。後半の追加情報の応答を行うか否かは分岐点⑦においてインターフェイスシステムがなんらかの基準で判断する。それに対して分岐点①, ②, ⑤はシステム本体の処理の結果から必然的に進行方向が決まる。また、生成される応答パターンもその内容が必然的に求まるものと任意性を持つものがある。

2. 1 分岐点の分析

任意性を持つ分岐点での選択を左右する要素としてまず挙げられるのは、ユーザの目的である。例えば、分岐点③において、システムがユーザの持つ明確な目的に対するプラン構造知識を持っていれば、その知識と照合することによって、ユーザの誤認識を特定し修正して応答することが可能である。一方、例1はペーパバックデータベースの検索システムにおける明確な目的を持たないユーザの情報要求発話例である。

例1 U1: 重厚な文体の小説が読みたいです。

U2: イギリスミステリーはありますか？

U3: アガサ・クリスティーの書いたものなどありますか？

ここで、‘重厚な文体の小説’と‘アガサ・クリスティーの書いたもの’は矛盾し、ユーザが誤って認識していることがわかる。分岐点③では単に誤認識の指摘に止めるか、2つの条件のどちらかを修正してそれに対する応答を生成するか、を選択することになる。次に、ユーザの発話がU1のみで、データベースシステムが検索条件が不足していると判断した場合は分岐点④の例であり、インターフェイスシステムは不足している条件を問い合わせる情報要求応答をするか、システム自身が追加条件を推測して再検索し情報補充応答を返答するかを選択することになる。上述の2つの選択は基本的には応答の正確さを重視して、必要最低限の応答に止めるのか、情報の取捨選択はユーザに任せることを前提にユーザの発想を支援し得る情報を積極的に提供するかを選択といえる。前者のタイプの応答を基本応答、後者のタイプを支援応答と呼ぶことにし、図1にその分類を示した。追加情報の応答も積極的な情報提供という意味で支援応答に

準じ、支援応答と追加情報の応答をあわせて発想支援応答と呼ぶ。

任意性を持つ分岐点を分析した結果、我々は各選択の拠り所となる基本要素としてユーザの目的以外に、(1) 提供し得る情報のユーザへの貢献度、(2) ユーザの対話形態に対する嗜好、の二要素を導いた。(1)を具体化すると、生成する発想支援応答がユーザの興味に適合している可能性が高いかどうかと言える。この時興味としては、概念自身に対する興味のみでなく、概念間の関係というメタ興味も扱う必要がある。例えば、追加情報応答の内容を選択するためには概念間の上下関係や原因・結果の関係などのメタ興味の所在が必要となってくる。(2)はユーザが発想支援の可能性を持つ応答を望んでいるかどうかということである。冗長でも積極的に関連情報を提供するインターフェイスを望むユーザもいれば、簡潔な応答を望むユーザもいるはずである。

3. ユーザの認識状態の記述

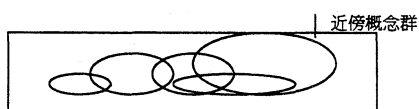
前章で導いた基本要素を扱うために、ユーザの興味や嗜好を管理してユーザの認識状態を記述する仕組みを提案する。

ユーザの認識状態を記述するために、まずシステムが理解可能な、自然言語で記述された概念(キーワードと呼ぶ)全体を構造化する。これはユーザの内部で概念がどのように関連付けられているかをモデル化する試みである。関連性という漠然とした要素に基づくという本研究の特徴から、次の2項目を満足する、より単純な構造が適していると考えられる。

- 1) 自然言語で表された個々のキーワードに対する理解には個人差があり、かつ状況によっても変化するため、雑音の影響を避けるために、処理の情報単位としては、個々のキーワードではなく、関連性の高いキーワードをグループ化したものを採用する。このグループ化には大局的整合性を求めず、また1つのキーワードが複数のグループに含まれることを許す。さらにこのグループ化はユーザの観点到に依存する。ここでは観点を各時点でのユーザのメタ興味と捕らえ、その所在に応じて動的に変化させる。

2) 個々のキーワード間には、概念の階層的分類上の上下関係、部分・全体関係、連想関係など様々な関係が生じ得るが、グループ間の関係としては、関係の種類までは扱えないため単に関係の強さのみを扱う。以上の要件を満たすグループを近傍概念群と名付け、それによって領域（具体的にはキーワード全体）を図2のように覆う。

図2. 近傍概念群による構造化



対話の履歴からユーザのメタ興味の所在と対話形態に対する嗜好を獲得し、それに基づいて近傍概念群を用いて構造化した領域上でユーザの近傍概念群に対する興味の度合いを与えたものをユーザの認識状態と考える。

4. ユーザの認識状態の獲得方法

以降の具体的な議論は対象となるシステムに依存するため、ペーパバックデータベースの検索というタスクを設定して議論を進める。議論の焦点を絞るために、ユーザの情報要求文はキーワードを用いて表現された検索条件 C_i の積に翻訳できているとする。

4.1 メタ興味の獲得

メタ興味としてはデータベース検索という問題の性格を考慮して、以下の5種類の概念間の関係に対する興味を扱う。

- [1] 概念の階層的分類上の上下関係
- [2] 同一カテゴリー関係
- [3] 対照関係
- [4] 対象とその形容表現関係
- [5] 連想関係

ここで、カテゴリーとしては、作品、作者、作品の形容表現、人間（作者、読者）の形容表現、ジャンル、トピックスの7個を扱う。

上記の関係に対するユーザの興味の所在を管理するために、まず全キーワード対の間に上記の5種類の関係がそれぞれ成立しているかどうかを与えておく。キーワード K_i と K_j の関係を $R_{ij}(n)$, $n=1, \dots, 5$ で表し、上

記の関係が成立する場合には1、しない場合には0を与えておく。例えば、 K_i を「イギリスミステリー」、 K_j を「重厚な文体」とすると、 $R_{ij}(1)=R_{ij}(2)=R_{ij}(3)=0, R_{ij}(4)=R_{ij}(5)=1$ となる。この処理は類語辞典を用いて、基本的にはキーワードを構成する単語間のマッチングによって行うが、まだ実験段階であるので、ここでは詳細は述べない。

関係 n 毎に、同一または連続する発話内に現れた2つのキーワード K_i と K_j に対する $R_{ij}(n)$ の値を累積し、その値が低い順にメタ興味度 $MI(n)$, $n=1 \dots 5$ に1から5の評価値を与える。

4.2 対話形態に対する嗜好の獲得

ユーザがインターフェイスに対して発想支援応答を望んでいるか否かはユーザが発想支援応答に対して反応する割合を用いて測る。これを発想支援支持率 (asr) と呼び、発想支援応答が新たに導入したキーワードが属する近傍概念群と有意に関係する近傍概念群に属するキーワードがその応答に後続するユーザの発話内で言及された回数を発想支援応答の回数で割った値とする。ここで‘有意に関連する’とは近傍概念群間の関連度に対する統計量を用いて関連が強いと判断するという意味であり、具体的定義は4.3で述べる。asrと発想支援応答回数 N に応じて設定した2個のしきい値 $asr_{tl}(N)$, $asr_{th}(N)$ ($asr_{tl}(N) < asr_{th}(N)$) との関係によって以下のように嗜好度を決定する。

$asr < asr_{tl}(N)$ の時: 弱い
 $asr_{tl}(N) \leq asr < asr_{th}(N)$ の時: 中程度
 $asr_{th}(N) \leq asr$ の時: 強い

ここで、嗜好度を3段階に分ける根拠は、追加情報の応答に比べて支援応答はより高い嗜好度を必要とすると考えたからである。

4.3 近傍概念群の生成方法

4.1で定義したキーワード K_i と K_j の間の関係 $R_{ij}(n)$ を用いて、各時点でのユーザのメタ興味の所在を考慮したキーワード間の関連度 Rd_{ij} を計算し、それを用いて近傍概念群を形成する。まずキーワード間の関連度を以下の式で計算する。

$$Rd_{ij} = \sum_{n=1, \dots, 5} MI(n) R_{ij}(n)$$

この値がすべてのキーワードの組み合わせに対する関連度の中で上位 $R \cdot dt\%$ に入っている時、その組み合わせは有意に関連すると呼ぶ。全キーワードに対して、同一近傍概念群に属する2つのキーワードも有意に関連するように、また逆に有意に関連する任意の2つのキーワードは同一近傍概念群に属するように構造化する。

さらに、近傍概念群 N_i と N_j 間の非関連度 ND_{ij} を以下の式で求める。ここで、各キーワードがユーザの発話内で言及された回数を累積した値をそのキーワードの活性値と呼ぶ。

$$ND_{ij} = \left(\sum_{(Kw_k \in N_i \cap N_j)} Kw_k \text{の活性値} + \sum_{(Kw_l \in N_i \cap N_j)} Kw_l \text{の活性値} \right) / \sum_{(Kw_m \in N_i \cap N_j)} Kw_m \text{の活性値}$$

キーワードの場合と同様にすべての近傍概念群の組み合わせに対する非関連度の中で下位 $R \cdot dt\%$ に入っている時、その組み合わせは有意に関連すると呼ぶ。

4. 4 興味の管理方法

近傍概念群に対する興味度をそれに属するキーワードの活性値の和とする。興味度についても全ての近傍概念群に対する興味度の中で上位 $N \cdot it\%$ に入っている時、有意に興味度は高いとみなす。

5. 応答生成におけるユーザの認識状態の利用

前章で導入したユーザの認識状態を用いて応答生成モデルにおける分岐点での選択や発想支援応答内容の生成を行う方法について述べる。

図1において任意性を持つ分岐点のうち、③、④、⑥は支援応答をするかどうかを選択することになる。支援応答を選択するための条件は、まず発想支援応答の嗜好度が強いこととする。それに加えて、③では修正するキーワードが、それが属する近傍概念群と有意に関連する近傍概念群に含まれるキーワードによって修正できる場合とする。⑦において追加情報の応答を行うのは、発想支援応答の嗜好度が中程度以上の時で、直接応答で焦点となっているキーワードが属する近傍概念群と有意に関連する近傍概念群の中に、メタ興味をもっとも高い関係にあるキーワードが存在するときとし、そのキーワードを用いた追加情報を生成する。

5. 1 応答生成例

簡単な応答生成例として、2. 1で示した例1に対するシステムの応答生成について述べる。

ユーザ発話内に誤認識が認められたため、分岐点③の選択を行うことになる。ここで、ユーザの発想支援応答の嗜好度が強いとする。さらに、図3における近傍概念群 N_1 と N_3 が有意に関連している場合、S5のような誤認識修正応答が生成できる。

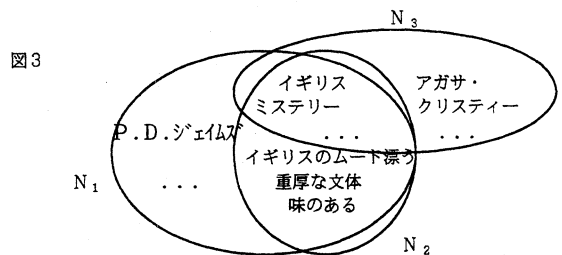


図3

S4: アガサ・クリスティーの小説は重厚な文体ではありません。
S5: P. D. ジェイムズの“女には向かない職業”はどうですか？

さらに発想支援応答の嗜好度が強いとしているので追加情報を応答する条件も満たしており、この時点での最も高いメタ興味が[4]対象とその形容表現関係であるとすると、S6のような追加情報の応答が生成できる。

S6: イギリスのムード漂う味のある作品です。

6. まとめ

明確な目的を持たないユーザに対する発想支援を狙った応答が有効となるための条件とその内容の選択方法について議論した。その際に必要となる文脈情報は概念の関連性に注目したユーザモデルと考えられ、本稿ではその獲得方法と処理方法に対するアイデアを提案した。今後は実験によってその理論の洗練化および検証をしていく予定である。

参考文献

- [1] Allen, J.F.: "Recognizing intention from natural language utterances", in Computational Models of Discourse Cambridge MA: MIT Press(1983)
- [2] 武田: "ネットワーク利用した知的情報統合", 人工知能学会誌 Vol.11 No5, pp.680-688
- [3] 高野, 平井, 北橋: "発話間の意味的結束性のモデル化", 電子情報通信学会論文誌 D[II], Vol. J79-ZD-II No.12 pp.2146-2153