

個人知識を考慮してカスタマイズ可能な音声対話システム

Development of a dialogue system considering personal knowledge

吉田 繭^{*1}
Mayu Yoshida

桐山 伸也^{*2}
Shinya Kiriyama

北澤 茂良^{*2}
Shigeyoshi Kitazawa

^{*1} 静岡大学大学院情報学研究科
Graduate school of Information, Shizuoka University

^{*2} 静岡大学情報学部
Faculty of Information, Shizuoka University

A dialogue system dealing with the daily conversation in a car was constructed. The flow of the conversation is customized by analyzing not only the situation in a car but also the content of the utterance. We propose a method of dialogue management which enables users to obtain smart actions.

1. はじめに

近年、音声対話システムに対して「認識精度・利便性・操作性」だけでなく、「ユーザにとってより自然であること」が求められている。ユーザからは音声対話に対して「ハンズフリーで入力ができる・うまく認識しない」というだけでなく、合成音声の不自然さや画一的な応答から「人間的でない・気が利かない」という声も挙がる。音声での対話は、人間の対話において最も自然な対話方法であり、音声対話の自然性が機械での音声対話の不自然さを強調しているといえる。

また、私たちが人間同士で対話をするときには自然と相手に合わせた返答をするものである。相手の好みや現在の様子を、しぐさや声色、発声速度、これまでの様子などから読み取り、互いに相手に気を遣いながら対話を進める。しかし現在の音声対話システムでは、一方的に人間側がシステムに合わせて合わせなければならず、ユーザ側がシステムが聞き取れるように発声し、システムが理解できそうな言葉を選んで発声している。感情を理解できればより自然な対話が進められるのではという観点から音声から感情を読み取る試み[原 2003][刀根 2000][川波 1997]はあるが、感情だけでは状況を理解するには足りない。本研究では音声対話システムのこの不自然さに注目し、音声だけでなく周囲の状況も考慮しながら、気の利いた対話を進めることのできるシステムの開発を目指す。

2. システムの特徴

2.1 個人知識の利用

本研究ではシステムの応答の際に個人知識を考慮する。上で述べたように、人間は相手によって反応を変えながら対話を進めている。気の利いた対話を実現するためには、システムの反応も相手によって変化するものでなければならない。そこで、本研究では相手の好みや習慣についての情報を個人知識として蓄え、動作の決定に利用する。個人知識を考慮し生成した応答は、人間の会話の場合と同じように相手や状況に合わせた反応に近づくと考えられる。

2.2 さりげない動作

音声対話システムの利用における一般的なコミュニケーションが「ユーザ(自分)-システム」であるのに対して、人間のコミュニケーションは「自分-周りの人間すべて」であるといえる。そこで本システムでは、より自然なコミュニケーションのためにシステムに向けて発話が行われることを前提にせず、「発話者-(周り

の人間すべてとシステム)」の形での対話を対象にする。システムも周囲の人間の一人として話者の話を聞き、行動するのである。それはつまり、システムは言われたことばかりやるのではないということの意味する。ユーザ同士の会話に現れる感情や意図を汲み取り、周りの状況に合わせて動作する。ユーザにとって心地良い環境を作り出すためのさりげない気遣いができるシステムを目指す。

3. システムの設計

3.1 対象とする対話

本研究では機器・場所の場を車内対話に絞ることとする。場を限定することで空間・状況の把握がし易くなり、動作もより的確に決定することができる。また、車という我々に身近な題材を扱い、普段馴染みの薄い音声対話へ親しみを持ってもらおう。本システムはあたかも車内にもう1人人間が存在しているかのように振舞う。自動車の車内という限られた空間でどのような状況でどのようなサポートができるかを検証する。今回の検討では、さらにユーザが発話した時のみに場を限定してシステムを設計した。ユーザの発話を前提にするため、「暑い」や「眠い」などのユーザの欲求と結びつきやすく、行動を起こしやすいような発話の例に対してユーザ適応を行った。

3.2 場面の例

次に示すのは、本研究が想定する場面の例である。

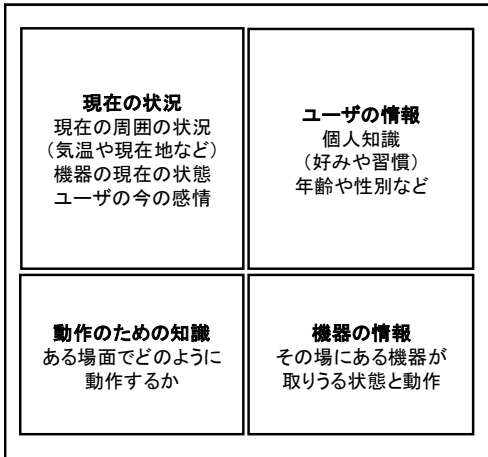
状況:「ちょっと暑いね」「そうだねー」 要求:車内を涼しくしたい システムの知識:エアコンをつける 窓を開ける 涼しげな音楽をかける ユーザの好み:エアコンを効かせるのが好き 必要な情報:窓の開閉状態 エアコンの運転状況

状況:「あー喉が渇いたー」 要求:何か飲みたい システムの知識:飲み物があれば飲む なければ買いに行く ユーザの好み:コーラが好き 必要な情報:カップホルダーの状況 現在地・最寄のコンビニ
--

3.3 情報の収集

(1) 用いる情報

本研究で扱うシステムは音声対話システムであるが、必ずしも音声によるのみで対話が進められるわけではない。人間の対話と同じように、音声と共に相手の動作や周囲の状況も入力となる。また、出力においても同様である。よってシステムの動作の決定には次の4種類の知識を使用する。



<図1: 情報の種類>

(2) 情報の獲得・収集

まず、前項で挙げた情報の種類のうち「現在の状況」をどう獲得するかについて述べる。現在の状況は刻々移り変わる。そのため、その都度情報を獲得する必要があるが、現在の車内の情報を知るには車内の機器との連携が必要である。カーナビゲーションシステムからは現在地を、エアコンからは運転状態をというように必要な情報を車内の機器から収集する。また、状況をさらに詳しく知るためにセンサ情報も考慮する。たとえば、車内に飲み物があるかどうかといった情報はカップホルダーにとりつけたセンサから得られる。

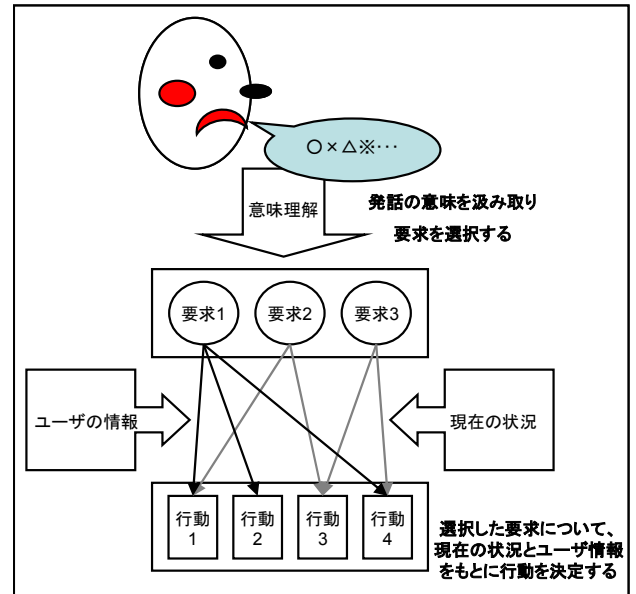
次に、ユーザの情報と動作のための知識の獲得について述べる。現在の状況と違ってこれらは変化が少なめである。また、ユーザの情報においては、年齢や性別などは1度の入力ですぐに利用できる固定された情報である。しかし、本研究で「個人知識」として位置づけた情報は好みや習慣など1度きりの入力では推定しにくいものであり、長期間の比較的ゆっくりした変化も現れる。また、動作のための知識についてもある場面に対する動作は1つだけではなく候補は無数に考えられる。このことから、人間の持つ知識を収集して利用することを考えた。収集する項目としては

- どのような気遣いがあるか
- そのときの判断材料
- 話者の感情や考え

などであり、場面を区切って広くユーザの入力を得る。ユーザ別に整理したものを個人知識と捉え、個人の好みや傾向が現れていると考える。また、全体の入力から得られた知識を一般的な知識(常識)として扱う。

3.4 個人知識を考慮した動作の決定

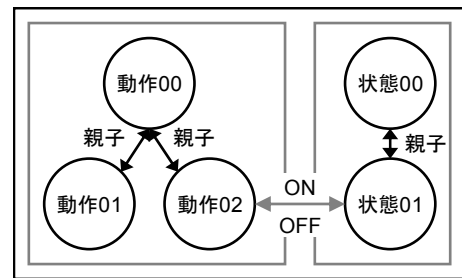
次に示した図は、システムが動作を決定するときの動きである。



<図2: 動作決定の流れ>

要求から動作を決定するまでについて、本研究では個人知識を考慮して特定のユーザ用にカスタマイズできるようにするため、どの段階で個人知識を利用するかが重要になる。このシステムでは、行動を選択する際に行動に優先順位を与えている。要求が与えられた際に、優先順位の高いものから順に参照し、現在の状態と照らし合わせ実行する。ユーザ用に個人性を適応したものは、この優先順位がユーザごとに違うことになる。ユーザの好みによって選択する行動を変え、優先順位を入れ替えることによってカスタマイズする。

図3は本システムで用いる情報の構造を示したものである。左側の2つが動作についての情報、右側が現在の状態についての情報である。



<図3: データ構造(動作:状態)>

4. キーボードによる情報収集

個人知識と動作のための知識を獲得するために、キーボード入力によるユーザからの知識の収集を行った。下に収集時に表示した質問と回答の例を示す。

こんなときあなたはどうしますか?
 ☆眠くなってきたが、どうしても寝てはいけなとき
 →(栄養ドリンクを飲む)
 →(ガムを噛む)
 →(どこかに停車して休憩)
 →(何か飲む)

