

## 要求絞り込みのための知識収集の検討

平 博順      牧野 俊朗      飯田 敏幸

NTT (株) コミュニケーション科学研究所

**概要:** われわれが提唱している情報提供システム「知識プロバイダ」の機能のひとつとして、利用者と対話を行ない、漠然とした要求をより具体的なものにするのを助ける機能がある。この機能の実現のために必要となる知識のうち、われわれが選択項目と呼んでいるものについて、その収集方法を検討した。選択項目には人が無意識のうちにも考慮するものも含まれているので、意識下の知識を引き出すトリガとして否定理由の回答を強要する状況を設定した。この状況設定に基づく「否定的回答のアンケート」を提案し、実際にアンケート調査を行なった。そして、得られた選択項目について分析を行ない、「知識プロバイダ」の持つべき知識や機能について検討した。

## Acquisition of Knowledge for Specifying Users' Requirements

Hirotoishi Taira      Toshiro Makino      Toshiyuki Iida  
NTT Communication Science Laboratories

**Abstract:** We propose an information system called "Knowledge Provider". The system has the ability to interactively specifying users' ambiguous requirements through dialogue. We tried collecting users' choice criterias, one type of knowledge useful for realizing the ability. But since this includes unconscious knowledge, it requires a situation where respondents are prompted to reply on questionnaires, as to the reason why they rule out some choices, as a trigger for drawing out their knowledge. We designed "Questionnaires for Refusing" based on assumptions about the situation, and we carried out a questionnaire survey. We analyzed the choice criterias extracted from the replies and examined what kind of knowledge and functions the system needs.

### 1 はじめに

ネットワーク上に散在する情報の中から、利用者の要求に応じて情報を抽出し、整理して、構成した知識を利用者に提供するシス

テムである「知識プロバイダ」をわれわれは提唱している [1][2][3]。「知識プロバイダ」は、従来の知識提供システムとは異なり、利用者に関する情報や利用者のおかれた環境に関する情報などを利用して発話を理解し、

利用者と円滑なコミュニケーションを実現する。さらに、人間同士のコミュニケーションの場においても、その時の話題にあった知識を提供することにより、コミュニケーションの支援・活性化を行なうことを目指している。

このシステムにおいて、利用者の持つ意図を利用者についての個人情報やネットワークなどの外部情報源から得られる情報に応じて適切に把握する要素を「意図理解部」と名付けている。「意図理解部」の機能のひとつとして、利用者の漠然とした要求をより具体的なものにする機能がある [2]。この機能は、利用者が「週末、どこかへ遊びに行きたい」というような要求を持っている場合に、システムが利用者へ「スキー」などのより具体的な遊びの候補を提示することによって、利用者が自分の要求を具体化することを助ける機能である。

ところで、人間同士の会話の場合、相手の漠然とした要求に対して、相手のおかれた状況を適切に判断して、その状況にあった適当な候補を挙げるということが自然と行なわれている。例えば、「週末、どこかへ遊びに行きたい」と誰かが言った時に現在の季節が冬である場合には、「海水浴」という候補は挙げず、「スキー」などの候補を挙げる。われわれは、「知識プロバイダ」にこのような自然な応答をさせたいと考えている。本稿では「知識プロバイダ」に自然な応答を行なわせるために必要な知識のうち、選択項目について検討した。

## 2 知識プロバイダの概要

「知識プロバイダ」は、利用者の要求を対話を通じて適切に把握して、有用な情報を利用者へ分かりやすく提供するシステムである。

図1にシステム構成を示す。「知識プロバイダ」は、会話制御部、意図理解部、知識オーガナイザ、情報収集部、表現変換部からな

る。会話制御部は、利用者との対話を行なって利用者の要求を受け取る。意図理解部は、要求の具体性を判断する。要求が漠然としていると判断した場合には、利用者の個人情報やネットワークなどの外部情報源から得られた情報に応じて、要求の具体化を促すような対話を行なう。情報収集部は、具体化された要求について外部情報源から情報収集を行ない、知識オーガナイザは、情報の整理を行なう。表現変換部は、整理された情報を利用者にとって分かりやすい形に変換して表示する。

例えば、本稿のはじめに述べた「週末、どこかへ遊びに行きたい」といった要求を利用者が持っている場合、「知識プロバイダ」はこの要求は漠然としていると判断し、対話をしながら、利用者の要求が「スキーに行く」程度まで具体化するのを助ける。その上で、おすすめのスキー場やスキーツアーのプランなどの情報を利用者へ提供する。

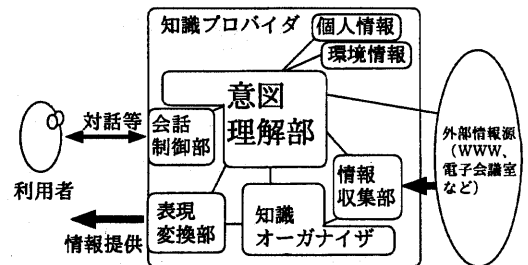


図1: 知識プロバイダの構成

## 3 自然な応答のために必要な知識

自然な応答の実現のために、下に挙げる [知識 A] ~ [知識 C] の知識をシステムに持たせようと考えている。

### [知識 A]

「遊び」には「スキー」、「海水浴」、「美術館」などがあるという知識

この知識は、物を階層的に分類したツリー状のもので上位のものほど抽象度が高いものになっている。「スキー」、「海水浴」、「美術館」などは「遊び」のひとつとして分類され、「遊び」をルートとする部分木の葉に位置している。この知識を利用することにより、要求を具体化させる際にシステムが提示する妥当な候補を得ることができる。

**[知識 B]**

「海水浴」は「夏に行なうもの」といった知識

この知識は表1に示すように、「スキー」、「海水浴」などの具体的な候補となりうるものに関する常識的知識のうち、費用や屋内・屋外などの活動場所の別や、必要人数などのような、各候補を行なう上での制約となるような知識である。

表 1: [知識 B]

	必要人数	費用	季節	屋内外
スキー	1	20000	冬	屋外
海水浴	1	1000	夏	屋外
美術館	1	800	年中	屋内
遊園地	1	3000	年中	屋外
マーじゃん	4	3000	年中	屋内

**[知識 C]**

「現在の季節が冬なので、夏に行なうものは不適當」といった知識

[知識 B] と利用者の個人情報や外部情報源などから得られる情報との間の対応関係を取り、要求の候補が現在の状況において適切であるか否かを判断するための知識である。例えば、ネットワーク上から「天候」が「雨」という情報が得られた場合、[知識 B] の「屋内外」の値が「屋外」のもの是不適當であることを表す知識などを想定して

いる。

以上の知識を利用して、システムは次に示すように利用者の要求の具体化を助けるための自然な応答を行なう。例えば、利用者から「週末、どこかに遊びに行きたい」と言われた場合を考える。まず、システムが [知識 A] を調べ「遊び」を頂点とする部分木を見つけ、具体的な「遊び」の候補として「スキー」、「海水浴」、「美術館」などがあることを見つける。次に、これらの候補について、その候補を挙げることが現在の状況において妥当かどうかを判断する。例えば「スキー」に関して [知識 B] を調べることにより、スキーにはどのくらい費用がかかるかや、スキーは屋外であるものであることなどが分かる。これらと利用者の個人情報から得られる利用者の予算やネットワーク上などから得られる天候などの情報とをくらべ、[知識 C] を使うことにより、現状において「スキー」を候補として挙げることが適切か否かを判断する。もしも「週末の天候」が“雨”の場合、「屋内外」の項目の値が“屋外”になっている「スキー」は不適當だと判断し候補から取り除く。このように各候補に関して不適の判断を行ない最終的に適切と思われる候補「美術館」について「美術館に行ったらどうですか?」といった発話をする。

3種類の知識の中で [知識 A] についてはソーラス [4] を利用することになっている。また、[知識 C] は [知識 B] を作成しないと作成できない。

さらに [知識 B] は表1のように「必要人数」、「費用」、「季節」、「屋内外」といった属性と「スキー」の「季節」が“冬”だといった属性値に分かれる。われわれは前者を「選択項目」、後者を「選択項目の値」と呼んでいる。選択項目の値は選択項目がないと作成できない。そこで、まず選択項目を収集することにした。

## 4 否定的回答のアンケートの検討

### 4.1 検討の背景

われわれが収集しようとする選択項目は、人が無意識のうちに考慮するものも含まれる。例えば、夏には遊びの候補としてスキーを考えないし、今日の午後遊びに行こうという時には、海外旅行に行こうなどとは思わない。この時、われわれは無意識に今の季節を考慮しているし、遊びにかけることのできる時間がどのくらいあって、それぞれの候補がその時間内で可能か否かの判断をしている。こういった判断を行なうための知識を人はあたり前に持っているため、書物などに改めて記述されることは稀であり、コーパスや電子化辞書からの獲得は困難である。そこで、われわれはこの選択項目を、アンケートを用いて直接、人から獲得することを考えた。

ところが、以前からエキスパートシステムの知識獲得の際にも問題にされてきたように[5]、意識下の知識を引き出すためには何らかのトリガとなるものが必要である。そこでわれわれは、トリガとして、選択項目を意識するような状況を設定した。

### 4.2 否定理由の回答の強要

われわれが設定した状況は、ある候補を否定するという状況でその際の原因を回答してもらうことである。例えば、「いいアルバイトを探す」という題材について、以下のような二人の対話の状況を設定する。

A: 「いいアルバイトないかな」

B: 「ハンバーガーショップで働いてみれば？」

A: 「いやだよ○○○だから」

ここで、アンケート回答者には、Aの立場になってもらう。この際に、相手の提案を断るにあたって、理由を言う必要が生じる。そこで、Aはなぜ、それを選択しないかという理由を意識せざるを得なくなるわけである。

このアンケートは、人に勧められたり候補として提示されたものを、利用者が選ばないとき、その理由にはどんなものがあるかを知るために行います。ですから、あなたは、友達勧めや誘いを断る立場だとして、断る理由を考えてください。また、断る理由が複数考えられるときは、複数書いてください。

次のやり取りの( )の中に断る理由を記入してください。理由が複数あると思われるときは複数書いてください。( )の数がたりなければ、余白に書いてください。

あなた：どこかに遊びに行きたいね。

友 達：スキーに行こうか？

あなた：スキーはいやだ。( ) から。  
( ) から。  
( ) から。

あなた：どこかに遊びに行きたいね。

友 達：温泉に行こうか？

あなた：温泉はいやだ。( ) から。  
( ) から。  
( ) から。

図 2: 否定的回答のアンケート

### 4.3 アンケートの概要

以上のような考えに基づき、実際に否定理由の回答を収集するためのアンケートを行なった。被験者 50 人に対し、図 2 のような形式で行なった。

題材としては、要求を具体化した時にたくさん候補が想定されるものの方が、さまざまな種類の回答が得られると考え、「どこかへ遊びに行きたい」を題材としてアンケートを行なった。遊びの候補としては、「スキー」「温泉」「ドライブ」「テニス」「マージャン」「映画」「魚釣り」「美術館」「カラオケ」「ゲームセンター」

の 10 種類とした。回答は複数回答を許し、否定理由を回答してもらった。

ここで、他の候補と比べてその候補を選択しない理由を考えてしまうと、ここに挙げた候補の中で共通する選択基準しか得られない可能性が高い。そこで、各候補自身について

否定理由を回答してもらった。

## 5 アンケート回答の分析

### 5.1 回答からの選択項目の設定

この回答からわれわれが得たいのは、遊び一般について、人が選択を行なう際に共通する選択項目である。遊びの10種類の候補は、その選択項目を調べるための例であり、ここで挙げなかった候補についても共通するようなものを考慮している。

選択項目は、回答の前提としている選択項目を、われわれが考えて設定した。例えば「寒い」という回答は「寒さ」、「お金がかかる」という回答は「費用」という選択項目を前提としていると考えた。

### 5.2 否定理由を回答してもらったことによる効果

例えば、「道具を持っていない」や「身体に悪い」といった回答が得られた。道具を持っていることや、身体に悪い影響がないということは、ほとんど意識されていないが、否定理由を回答してもらった方法をとったことにより、このような選択項目も得ることができた。

### 5.3 選択項目の分類

アンケート回答から得られた選択項目をシステムを構築する立場から以下の2つで分類した。

#### (1) 多数の候補にまたがった共通性があるか

システムを構築する際に、システムに持たせる選択項目はできるだけ多い方が望ましいが、記憶容量の制限がある。そこでシステムに、特にどのような選択項目を優先して持たせるべきかという点が問題になる。われわれは、対話による漠然とした要求の絞り込みが効果的に行なえる選択項

表 2: 共通性の高い項目と低い項目

分類	選択項目 (括弧内は遊びの候補名)
共通性の高い項目	身体的なもの 「疲労しやすさ」(ドライブ、スキー、テニスなど) 「日焼けしやすさ」(テニスなど)
	量的なもの 「費用」(スキー、テニス、映画など) 「必要人数」(マージャン)
	公共手段 「交通手段が必要か」(スキー、ドライブ、温泉)
	自然環境 「天候に影響されるか」(スキー)
共通性の低い項目	「歌える曲の種類」(カラオケ) 「コート予約の有無」(テニス) 「釣ざおの有無」(魚釣り)

目を優先させることにする。例えば、「費用」という選択項目は、多数の候補に共通する選択項目であり、ある候補について、利用者から費用に関する情報を得れば、この情報は他の候補を落とす判断材料にもなる。これに対して、「スキー道具を持っているか否か」という選択項目は、スキー以外の候補を落とす判断材料にはほとんどならない、個別的な選択項目である。

回答から分析された選択項目の中で、複数の候補について共通性の高い項目と共通性の低い項目とに分類したものを表2に示す。

#### (2) 時間変化の周期が短いか

「知識プロバイダ」は個人情報や外部情報源から得た情報などを利用して情報提供を行なうが、時間変化の周期が長いものに関しては、外部から頻繁に情報を収集する必要はない。そこで、選択項目と関連する情報について、時間変化の周期の長さについて分析を行なった。

表3に、選択項目の中で、時間変化の周期の短い情報に関連したものを挙げる。

表 3: 時間変化の周期の短い選択項目

選択項目 (括弧内は遊びの候補名)	関連する情報
「けがをしても可能か」 (テニスなど)	けがの有無
「体調が悪くても可能か」 (スキー、魚釣りなど)	体調
「費用」 (スキー、テニス、映画など)	所持金
「交通手段が必要か」 (スキー、ドライブ、温泉)	交通情報
「天候に影響されるか」 (スキー)	天候

## 6 システムに必要な知識と機能

### 6.1 システムに必要な知識

#### 6.1.1 [知識 C] と外部情報

表 2 に示したような共通性の高い選択項目を利用するために、システムには次のような [知識 C] と外部からの情報が必要であることが分かった。

- 利用者の身体に関するもの
 

「疲労を感じるものであるか否か」といった利用者の身体に関する選択項目を使用するための、「利用者の好まない生理現象や感覚を伴うものは不適當」という [知識 C] と「利用者の生理的現象や感覚についての許容度」に関する情報。
- 量的なもの
 

「費用」、「必要人数」という選択項目を使用するための、「所持金が  $x$  円未満であれば、 $x$  円以上のものは不適當」、「人数が  $y$  人未満であれば、必要人数が  $y$  人以上のものは不適當」といった [知識 C] と「所持金」、「人数」などの量的な情報。
- 公共手段に関するもの
 

候補を実現する上で必要な公共手段に関する選択項目を使用するための、「道

路が混雑している場合には、バス・タクシーが必要なものは不適當」といった [知識 C] と「道路混雑状況」「鉄道運行状況」などの公共手段に関する情報。

- 自然環境に関するもの

候補を実現する上で必要な自然環境に関する選択項目を使用するための、「雨が降る場合には、屋外でやるものは不適當」といった [知識 C] と「天候」などの自然環境に関する情報。

#### 6.1.2 時間変化の周期についての知識

表 3 に示したような時間変化の周期の短い選択項目に関連した外部情報の時間変化の周期についての知識であり、この知識を持つことにより、要求を具体化する際の候補を選択するために参照する、個人情報や外部情報源からの情報について、どのくらいの頻度で情報を更新すべきかが分かる。例えば「利用者の所持金については、一日程度の情報更新でいい」といった知識である。

## 6.2 システムに必要な機能

### 6.2.1 どの情報をいつ更新するかを判断する機能

上で述べたような外部情報の時間変化の周期に応じて、システムが外部情報の収集を行なう時期を判断する機能が必要である。例えば、利用者の所持金についての情報収集を、一日ごとに情報収集部に行なわせる機能である。この機能により、利用者のおかれた状況について、より正しく把握することが可能になる。

### 6.2.2 外部情報源からその時の状況に関するさまざまな情報を収集する機能

上で述べたように、時間変化の周期の短い情報に関しては、リアルタイムで最新の情報を収集しないと、その時の状況を適切に把握することができない。したがって、外部

情報源からさまざまな情報を収集する必要がある。時間変化の周期の短い情報には、天候、道路の混雑状況、利用者の所持金、利用者の体調などがある。このうち、天候に関しては、現在ネットワーク上で情報提供がなされているし、道路の混雑情報なども近い将来に、ネットワーク上から得ることができると思われる。また、利用者の所持金についても将来的に電子マネーが普及すれば、ネットワーク上から情報を収集することが可能になるとと思われる。そこで、どのような情報がネットワーク上のどこにあるかを把握し、必要に応じてその情報を収集する機能があれば、これらの情報をリアルタイムで収集することができる。

### 6.2.3 利用者の発話から情報を収集する機能

時間変化の周期の短い情報の中で、利用者の体調などは、ネットワークから収集することは困難であるし、利用者が毎日その時の体調を入力するのは、利用者に負荷がかかり過ぎるので望ましくない。そのため、利用者の発話をモニタして、その中から必要な情報を自動的に抽出する機能が必要である。例えば、「いま調子が悪い」といった発話から利用者の個人情報として「体調」が「悪い」ことを把握する。

## 7 おわりに

「知識プロバイダ」の持つ利用者の漠然とした要求を具体化する機能を実現するために必要な選択項目に関して検討した。選択項目には人が通常意識しないものも含まれるのでそれを引き出すための手法として「否定的回答のアンケート」を提案した。さらに、回答から得られた選択項目に関して候補をまたがった共通性の高さを選択項目に関係する情報の時間変化の周期の長さという2つの観点で分類を行ない、「知識プロバイダ」に必要なと思われる知識や機能について検討した。

本稿では「遊ぶ」を題材とした場合について述べたが、他にも漠然とした要求として「学びに行きたい」「働きに行きたい」などがある。これらについても同様のアンケートを行なったが、これらの分析結果については別の機会に報告したい。システムにこれらの知識や機能を持たせることで、利用者の個人情報や環境情報などを利用して発話を理解し、利用者と円滑なコミュニケーションを実現することができると考えている。

### 参考文献

- [1] 八巻、飯田、松澤、湯川、牧野：「知識プロバイダの提案 — 情報検索から知識創生へ —」, 情報処理学会第 55 回 (平成 9 年後期) 全国大会, 3AF-2, pp. 2-380 - 2-381 (1997).
- [2] 牧野、飯田：「知識プロバイダにおける意図理解法」, 情報処理学会第 55 回 (平成 9 年後期) 全国大会, 3AF-3, pp. 2-382 - 2-383 (1997).
- [3] 湯川、松澤：「知識プロバイダにおけるオントロジ自動獲得」, 情報処理学会第 55 回 (平成 9 年後期) 全国大会, 3AF-4, pp. 2-384 - 2-385 (1997).
- [4] 池原 他：「日本語語彙体系」, 岩波書店, (1997).
- [5] Feigenbaum, E.: "The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering", IJCAI 5, pp. 1014 - 1029 (1977).