

個人の概念空間を利用した興味の推定による情報提供

角 薫[†] 角 康之^{††} 間瀬 健二^{††} 中須賀真一^{†††}
堀 浩一^{†††}

Information Presentation by Inferring User's Interests Based on Individual Conceptual Spaces

Kaoru SUMI[†], Yasuyuki SUMI^{††}, Kenji MASE^{††}, Shin-ichi NAKASUKA^{†††},
and Koichi HORI^{†††}

あらまし 本論文では、コンピュータを情報提供者側とユーザとの情報の受け渡しの仲介役として位置づけ、ユーザの興味や知識に応じて個人化された情報提供を行うシステム、Takealookを紹介する。基本的なアイデアは、ユーザとコンピュータ間での質問-回答の対話を通してユーザの興味の推定を行い、それに基づいて情報提供を個人化していくものである。ユーザの興味の推定は、ユーザとのインタラクションが進むにつれて更新される概念空間を利用してなされる。この概念空間は、互いに関連度をもつキーワードの集合からなり、その関連度がユーザとのインタラクションを通してユーザの興味に適應していく。本論文では、手法の提案、基本的メカニズムの説明、評価実験の報告、及び考察を行う。

キーワード コミュニケーション支援、情報提供の個人化、ユーザ適應、概念空間

1. ま え が き

本論文では、ユーザの興味に応じて情報提供を個人化する手法を提案する。基本的アイデアは、ユーザとシステム間での質問-回答の対話を通してユーザの興味の推定を行い、それに基づいて情報提供を個人化していくものである。システムは、ユーザの情報探索が進むにつれてユーザの興味を推定し、ユーザが明示的に質問しなくてもユーザが興味をもつであろうものを先読みして詳細情報を提示したり、関連する他の情報コンテンツの推薦を行う。ユーザの興味の推定は、ユーザとのインタラクションが進むにつれて構築されるユーザ個別の概念空間を利用してなされる。概念空間は、互いに関連度をもつキーワードの集合からなり、その関連度がユーザとのインタラクションを通し

てユーザの興味に適應していくものである。

本論文では、個人の概念空間を利用することによる情報提供の個人化手法を提案し、それに基づいて試作されたシステム Takealookを紹介する。そして、Takealookを研究所見学における情報探索支援に適用して行われた試用実験について報告し、提案手法の有効性について考察する。

2. 本研究の目的と背景

本研究の目的は、ユーザとの対話を通してユーザの興味を推定し、その推定された興味に応じた情報提供を自発的に提供することでユーザの知的好奇心を促進するシステムを構築することである。

システム構築にあたって、我々は、個人の興味や知識はその人の質問から観察できる、という前提を採用する。認知科学の分野において、質問は人の興味や知識を外界から観察できる一手段であり、興味や知識との関連があるとの研究成果が報告されている [1], [2]。ユーザに負荷をかけないため、そしてユーザの無意識の興味を獲得するために、ユーザとシステム間の質問-回答の対話からユーザの振舞いを観察することとした。

更に、概念空間は個人により異なり、それはさまざまな部分的視点の組合せにより成り立っている、とい

[†] 郵政省通信総合研究所ブレイクスルー 21 西田プロジェクト、神戸市 Breakthrough 21 Nishida Project, Communications Research Laboratory, Ministry of Posts and Telecommunications, 588-2 Tenoka, Nishi-ku, Kobe-shi, 651-2401 Japan

^{††} (株) ATR 知能映像通信研究所、京都府

ATR Media Integration & Communications Research Laboratory, Seikisocho, Soraku-gun, Kyoto-fu, 619-0288 Japan

^{†††} 東京大学工学系研究科、東京都

Faculty of Engineering, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153 8904 Japan

う前提を採用する。概念空間とは、互いに意味的関連をもった複数概念の集合である。概念空間を形成する部分は各個人の視点に依存するものではあるが、そのいくつかは他人同士でも共有可能であり、それが人同士のコミュニケーションのよりどころとなる。このような考えは、人間の概念形成過程のモデルとして多くの認知科学研究者や哲学者によって採用されてきたものである [3], [4]。

以上の前提をもとにして、以下の仮説を立て、興味の推定と喚起のモデル化を行った。

- ユーザの質問からユーザの興味対象となる概念を特定し、それと関連する興味の対象（あるいはその潜在的なもの）をあらかじめ用意された複数人の概念空間の部分を利用して推定することができる。具体的には、個人がある概念について興味があるならばその概念と意味的に近いと他人が判断している概念についても興味がある可能性がある、と仮定することで、ユーザの未知の興味が推定する。

- システムが、推定された興味対象に対応する概念をユーザに提示することで、ユーザの興味を喚起することができる。

本研究ではこのモデルに基づいてシステムを構築し、その評価実験を通して本モデルの妥当性を議論する。

3. 関連研究

本研究が対象としている課題は、ユーザ個人の興味に応じて広大な情報空間内を効率的かつ創造的に案内することである。その際、対象となる情報空間は複数の情報提供者によって提供された情報から構成されるものを前提としている。

複数の情報提供者とユーザの間での知識流通を対象とした研究として、組織内での情報やノウハウ知識に関するユーザの質問に回答するシステムの構築を目指した研究が古くからなされてきた。例えば、ユーザの質問内容を判定して目標知識への到達を支援するために対象問題の関連ネットワークを整備する手法 [5] や、詳細な意味ネットワークを準備することで提供知識に意味構造を与える手法 [6] などが提案されており、対象とする情報空間の意味的構造を整備すればそれに見合った成果が期待できることがわかっている。

しかし我々が対象とするような、複数の情報提供者による情報空間を考えたとき、詳細な意味構造を準備するには大変なコストを伴うし、一度そのようなものを構築したらその変更は容易ではなく柔軟性に欠ける。

また、すべての概念間に一貫した意味属性を与えるにはどうしてもある特定の視点で作業を行う必要があるため、情報空間全体での平均的な性能向上には貢献するが、複数の情報提供者各自の視点を反映しユーザ各自の興味に対応することは難しい。そこで我々は、詳細な意味属性の導入はあえて避け、概念（キーワード）間の単純な関連リンク（関連の有無を表す）のみで情報空間の構造を扱うことにした。そのとき、各個人の視点や興味の違いを重んずるため、情報提供者各自に概念空間（概念間の関連リンクの集合）を提供してもらい、かつ、ユーザ各自の興味空間（情報提供者の概念空間をもとにして再構成される概念間の関連リンクの集合）を扱うこととした。

一方、最近のインターネットの普及に伴い、あらかじめ意味的な構造が与えられていない広大な情報空間を対象とした情報収集・組織化に関する研究が精力的に進められており [7]、本研究と関連がある。そこでは主に、単純なキーワード一致による検索を頑健な自然言語処理によって精度を上げるというアプローチと、ユーザの興味や探索行動のモデリングを行うことで情報推薦を行うアプローチがとられている。特に後者に属する、プロファイルが類似する他のユーザが利用した情報を推薦する協調フィルタリング（例えば [8]）手法がよく利用されている。また、提供情報のインデックスを利用する第 1 のアプローチと、ユーザのプロファイルを利用する第 2 のアプローチを統合することで情報推薦の性能を高めることができることが報告されている [9]。本研究もこの流れに属するものである。つまり、比較的単純な概念構造をもとに提供情報の探索を行い、かつ、ユーザの視点に近い視点をもつ情報提供者による概念空間に高い得点をつけてユーザの興味空間を精練化することで、情報提供の個人化の精度向上をねらっている。

4. 興味に基づいた情報提供の個人化システムの開発

図 1 は、本研究で提案するシステムの枠組みである。上部に情報提供者、下部にユーザ、中間部にそれらの仲介役をするシステムが位置している。

複数の情報提供者は、それぞれの発信情報として、情報ソースと概念空間を用意する。情報ソースは、基本情報となるテキスト情報・情報提供者各々が選択したキーワード・それに対応する質問-回答情報からなる。概念空間は、自分が選択したキーワードと概念と

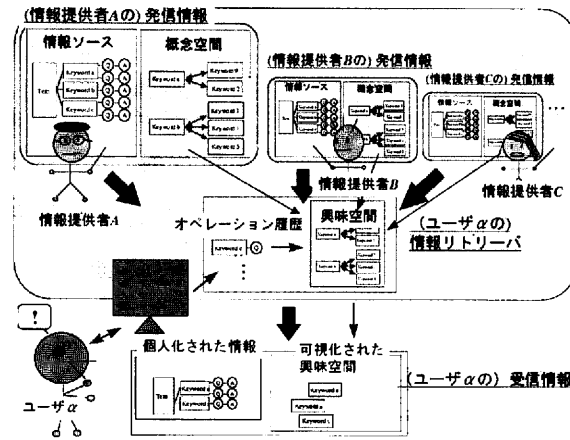


図1 システムの枠組み
Fig. 1 System framework.

して近いキーワードを定義したキーワードペアの集合で、情報提供者個別の情報である。

ユーザは、受信情報として個人化された情報と可視化された興味空間を得る。個人化された情報とは、ユーザが選択した情報ソース表示時にシステムがユーザの興味を推定することにより前もって表示される補助情報や、他の情報ソースの推薦のことである。興味空間とは、予測されるユーザの興味対象の空間のことであり、可視化された興味空間とは、それを可視化したものである。

システムは、情報リトリバを介することにより、蓄えられた情報提供者の発信情報をユーザごとの興味空間やオペレーションに合わせて選択し、提供する。

全体を通してみると、複数の情報提供者による情報を収集したデータベースがあり、ユーザは各自のために用意された情報閲覧インタフェースを利用して情報探索をするというようなことを行っている。ユーザが情報リトリバを介して情報提供者が用意した情報ソースを閲覧すると、システムは次にどのコンテンツを見ることがお薦めであるかといった情報推薦や個人の興味や知識に沿った情報提示を行い、更に、情報探索を視覚的にナビゲーションするために、ユーザの興味空間の可視化表示を提供する。

以下の節では、情報提供者側の情報の準備、興味の推定、情報の推薦、興味空間の可視化について順を追って説明する。

4.1 情報提供者側の情報の準備

情報提供者側から用意する情報としては、情報提供

者からユーザに発信する情報の源である情報ソースと情報提供者によってそれぞれ異なる概念空間がある。

情報ソースは、基本情報となるテキスト情報・情報提供者各々が選択したキーワード・それに対応する質問-回答情報から構成され、情報提供者各自で準備される。質問-回答情報の質問の種類は、“その意味は何ですか？”(What)という辞書的な意味を求めるもの、“それは具体的には何(例)ですか？”(What2)といった具体例を求めるもの、“どのようにそれは実現されているのですか？”(How)のような方法論などの詳細情報を求めるもの、“その理由は何ですか？”(Why)のような前述の方法論を適用した理由を求めるものの4種類を用意した。これらの質問は、What, What2, How, Whyの順に、よりその情報コンテンツに特化した内容になるということを想定して設定した。情報提供者によって選択されたキーワードに対する質問の決定は、各情報提供者に委ねられており、全種類の質問が用意されるとは限らない。

概念空間は、情報提供者が選択したキーワードと概念として近いキーワードを定義したキーワードペアの集合で、情報提供者個別の情報であり、これは情報提供者の視点の違い、つまり興味の違いやそれに伴った言葉の解釈の違いにより、異なってくるものである。具体的には、情報提供者が質問-回答情報を作る上で選択したキーワード各々に対し、他の全情報提供者が選択したものも含めた全キーワードの中から関連があると判断したものを選択し、リンクを張ったものである。例えば、“エージェント”と“擬人化エージェント”

といった包含的な関係、“コミュニケーション”と“インタラクティブ”といった連想的な関係、“テレビ”と“ラジオ”といった類似した関係がリンク付けの対象となる。これらの関係は可逆であり、どちらのキーワードからも興味が共起されるものと仮定し、ユーザの興味共起の推定に利用する。なお、キーワード間の関連付けの判断は各情報提供者の直感や連想に任せただけであり、特に厳密なガイドラインは定めなかった。

4.2 興味の推定

興味の推定は、ユーザ個別の興味空間を利用した情報リトリーバにより行われる。ユーザの興味空間は、ユーザがシステム上で質問の選択を行ったことをきっかけにして、複数の情報提供者による概念空間を利用することで自動構成される。その過程では、ユーザが質問したキーワードをユーザの興味対象とみなし、そのキーワードと興味が共起すると情報提供者が概念空間で表現しているもの、つまり対象キーワードとリンクが張られているキーワードをユーザの新しい興味対象の候補とする。興味空間中に存在するキーワード間のリンクには重み付け（関連度）が与えられている。関連度の値は 0.0（関連なし）から 1.0 までの値をとる。ユーザの興味推定の際には、これらの値に基づいて確率的に関連するキーワードを探索する。情報リトリーバは、推定された興味をもとに個人化した情報提供を行う。基本的に、興味の対象として推定されたキーワードが今見ている情報コンテンツ内にあったなら、興味の推定の際に用いられたもとのオペレーション履歴上の同種の質問に対する回答がユーザの興味対象の情報であると推定する。

このようにして、ユーザが興味をもつであろうと推定された情報は、システムからの補助情報として前もって自動的に表示される。興味推定の際には、キーワード間リンクを一つたどったもののみではなく、複数リンクの連鎖も対象となり、最高 2 段目までキーワードのリンクをたどるように設定した。情報探索における自動提示の際にはその理由づけの表示が重要だといわれているが [10]、補助情報の自動表示の際には、推定の理由として、どのように興味の推定がなされたかを説明するキーワード連鎖の軌跡を表示するようにしている。

そして、ユーザは自動的に提示される補助情報に対して、興味があるか否かの判定を行うように設計されている。その判定に応じて、推定に使われた連鎖中のキーワードリンクの関連度の調整が行われ、それによ

り興味推定に利用される興味空間のユーザ適応がなされる。具体的には、興味があつたか、なかつたかをユーザが選択しなければ提示された補助情報の窓を閉じることができず、すべての補助情報の窓を閉じなければ次へ進めないようにユーザインタフェースを設計した。

本システムは、広い概念空間の中からユーザの未知の興味対象を、比較的少ないインタラクションから推定することを目的としている。そのため、WWW 上の推薦システム [11] で成果をあげている協調フィルタリング [8], [9] に類似の方法を、情報リトリーバによる興味推定のユーザ適応に用いた。協調フィルタリングは、複数ユーザをグルーピングし、同じグループのユーザは同じ興味をもっているものと仮定することにより、情報の中身にふれることなく情報フィルタリングを行う手法である。しかしその短所として、複数ユーザの興味を表すデータ数が少ない場合にはうまく働かない。本システムの手法では、類似の興味をもった他人を探す際に用いる情報としてユーザ情報でなく情報提供者情報を利用し、その単位を個人ではなく個人の概念空間を構成する部分要素とすることで、その問題に対処した。

以下に興味推定のユーザ適応について具体的に例を示して説明する（図 2 参照）。

(1) 情報提供者の概念空間から興味空間へのキーワードの取込み：あるユーザ α が Keyword b に対して How の質問をした場合、複数の概念空間上で一つでも Keyword b とリンクのあるキーワード、例えば Keyword n, Keyword o, Keyword p, Keyword x, Keyword y が興味空間に取り込まれる。その際、現

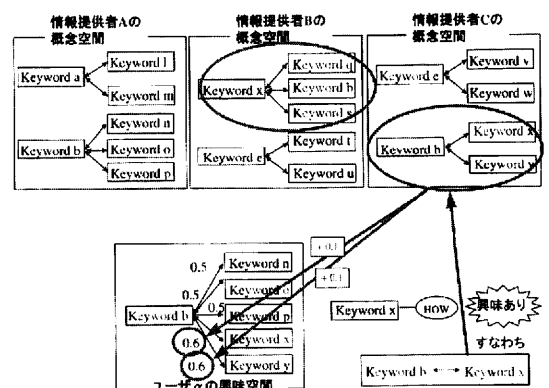


図 2 興味空間のユーザ適応方法
Fig. 2 User adaptation of interest spaces.

時点までに興味空間中に存在していなかったリンクについては、リンクの重み付けとして初期値(0.0と1.0の中間値である0.5)が与えられ、既に存在しているリンクについては、その時点の値をそのまま用いる。

(2) 提供情報の決定：興味の推定を行う情報リトリバは、Keyword bとリンクをもつこれら五つのキーワード候補から、今回情報の自動提示を行うべきキーワードを選択する。キーワードの選択は、リンクの関連度に応じて確率的に決定する^(注1)。本システムでは、キーワードリンクの探索を2段階、例えば、Keyword a ↔ Keyword b ↔ Keyword cといったリンクの連鎖を許している。このような場合には、それら2段階のリンクに与えられた関連度を掛け合わせた値がキーワード選択の際の確率として利用される^(注2)。確率選択の結果、ここでは例えばKeyword xが選択されたとする。ユーザαは、自分の質問に対する回答(Keyword bに関するHowの詳細情報)を得ると同時に、システムによる興味の先読みの結果として、Keyword xのHowに対する回答の情報を得ることになる。もしもKeyword xについてのHowに対する回答が情報ソースに存在しない場合は、次のキーワード候補が選択される。

(3) 興味空間の更新：自動的に提示された情報には、興味があるか否かについてユーザが応答することになっている。ユーザαが、自動提示されたKeyword xのHowに対する回答に興味があると応答したならば、興味の推定に用いられたキーワード連鎖であるKeyword b ↔ Keyword xに報酬が与えられる。報酬の値は0.1に設定した^(注3)。それと同時に、報酬が与えられたリンクが概念空間中存在している情報提供者を調べ、このリンクの右辺若しくは左辺が同じものに対してすべて関連度を高める。図2の場合、Keyword b ↔ Keyword xのリンクが存在する。情報提供者Bと情報提供者Cの概念空間が関連度の更新対象になる。まず、そのリンク自体(つまり、Keyword b ↔ Keyword x)の関連度を高める。同様に、その概念空間が左辺が同じだが右辺が異なるリンク(例えば、Keyword b ↔ Keyword y)が存在した場合、それについてもリンクの関連度を高める。このリンクは可逆のため、右辺が同じだが左辺の異なるKeyword x ↔ Keyword qやKeyword x ↔ Keyword sの場合についても同様の処理を行う。逆に、ユーザが興味のない情報だと評価したときには、リンクの関連度を低くする処理を行う。リンクを2段階利用した連鎖によって

関連キーワードが選択された場合も、同様に各々のリンク一つひとつについて関連度を調整する。

4.3 情報の推薦

本システムには、コンテンツ(情報ソース)ごとの推薦の度合を数値で表した推薦度を計算し、コンテンツ単位での推薦を行う機能がある。推薦度が計算されると、推薦度の高い順にコンテンツのタイトルを並べ、ユーザが選択しやすいようにしている。

推薦度は、基本的にはユーザがそれまでに質問したキーワードが該当の情報ソースにおいて共有されているかということに基づいており、そして、ユーザの興味空間を利用することにより、そのキーワードと興味が共起されると推定されるキーワードについても計算の対象としている。更に、興味として推定された対象のキーワードについて、ユーザの興味の対象となる質問の種類が用意されている情報ソースに高い点数が与えられるようになっている。これは、興味が推定されたキーワードが情報ソース内にあったとしても、興味の対象の質問-回答が情報ソースに用意されているとは限らないからである。

情報ソース*I*の推薦度*R(I)*は、以下のように表される。

$$R(I) = \sum_{i=1}^n \sum_{j=1}^l K_{ij} \cdot D_{ij}$$

ここで、*i*は現在のオペレーション*n*におけるキーワードのインデックス、*j*は情報ソース*I*におけるすべてのキーワード(その数は*l*)のインデックス、*K_{ij}*はユーザの興味空間中でのキーワード*i*から*j*への関連の重み付けで、0から1までの数値である。キーワード同士に関連がなければ0になり、同一キーワードであれば1になる。*D_{ij}*は、キーワードに対するある質問への回答が用意されているかどうかを表す因数である。つまり、ユーザが選択したキーワード*i*への質問に対する回答が、情報ソース*I*のキーワード*j*に

(注1)：対象となる(ここでは五つの)キーワードとのリンクの関連度を足し合わせたものを分母とし、各キーワードとのリンクの関連度を分子にした値で重み付けをした値を用いて確率選択を行う。

(注2)：つまり、Keyword a ↔ Keyword bとKeyword b ↔ Keyword cにそれぞれが0.5の関連度をもっているとすると、そのキーワード連鎖の値は0.5×0.5=0.25となる。この値は通常の直接リンクよりも小さい値となるので、相対的に2段階の連鎖によるキーワードは選択される確率が低くなる。

(注3)：複数報酬を与えられることによって関連度が1.0を超えることがあり得るが、その場合は1.0で飽和することにした。

についても用意されていれば1、用意されていない場合は0となる。

4.4 興味空間の可視化

ユーザの興味空間を、キーワード間の関連度が高いほど空間上の距離が近くなるように空間中の距離に対応させて、多次元尺度法を利用して2次元平面上に可視化する[12]。ユーザが質問した最新のキーワードを中心にし、ユーザにとって未知のキーワードを含め、空間配置を行いユーザに提示した。

この興味空間の可視化提示により、ユーザは自動的に提示された情報の理由を知ったり、それまでの利用履歴を概観することができるということを期待するものである。また、この可視化された興味空間には、キーワード表示のほかにそれらのキーワードに関連したコンテンツも同時に表示され、そこから対象のコンテンツに直接アクセスすることが可能である。可視化の立場としては、主に「ある現象があってそれを見る形に変換しようとする可視化」と「ある知的作業をするときにその作業を支援するための何らかの道具としての可視化」の2種類に分けられる[13]が、ここでは興味空間をユーザにわかりやすい形で提供するという意味では前者、可視化された空間からコンテンツの検索が可能であるという意味では后者であり、前者と後者を併せもった立場である。

5. 研究所見学における情報提供の個人化

本研究において開発されたシステム Takealook は、研究所見学をする際に見学者が研究所で行われている研究の内容を閲覧するため、それに先立って（若しくは見学後に）利用するという状況を想定して開発された。具体的には、ATR 知能映像通信研究所で行われている研究所見学の展示ガイドのプロジェクト[14]において試用を行った[15]。

図1で説明すると、ユーザは研究所の見学を行う見学者であり、情報提供者は研究者にあたる。情報提供者は、情報ソースとして5~6行の研究概要、その中から選択した5個から10個程度のキーワード、それに対する質問-回答、概念空間を用意する。

情報ソースの数は24、選択されたキーワードの総数は175、情報提供者らによってそれらキーワード間に張られたリンクの数は合計約3500である。

図3に開発したシステム Takealook のユーザインタフェースを示す。右側のウィンドウがメインウィンドウであり、大きく分けると三つの部分に分かれている。上部に情報ソースのタイトルが推薦度に応じてソートされて表示され(4.3で述べた情報推薦)、そのすぐ下に選択された情報ソースの概要が表示されるテキストエリア、その下にユーザの現時点の興味空間



図3 システム Takealook のユーザインタフェース
Fig. 3 User interface of the system Takealook.

が可視化されたものが表示されている(4.4で述べた興味空間の可視化)。テキストエリアの上部で、キーワードとそれに対する質問が選択できるようになっている。可視化された興味空間上にはユーザの興味である、過去に質問した興味のキーワードとそれから興味が共起される概念が表示され、キーワード間の距離はそれらの間の関連度を表している。一つの丸いアイコンは一つの情報ソースに対応しており、その情報ソース中のキーワードの重心に位置している。このアイコンをクリックすることにより、そのタイトルの研究の紹介の内容に移ることができる。

まず最初に、ユーザがタイトルから自分の興味のありそうなものの研究のタイトルにアクセスすると、テキストエリアに研究の概要が表示される。その研究概要を読み、ユーザはキーワード“マルチメディア”と、質問“その意味は何ですか?”(What)を選択し質問したとする。ユーザはあまりコンピュータになじみがなく、マルチメディアの辞書的な意味が知りたかったわけである。そうすると、システムはこの回答を図3の左の小窓、回答窓に表示する。そして、それと同時に他のキーワードの質問“その意味は何ですか?”(What)の回答も表示する。システムは、ユーザがキーワード“マルチメディア”の意味を尋ねたのだから、“音楽情報処理”や“インタラクティブ万華鏡”についても知らないだろう、そして、興味をもつであろうと推定し、前もってその情報を提示しているのである(4.2で述べた興味の推定)。また、それぞれ表示された理由として、推定の軌跡を回答窓の下部に示している。このような回答窓を得ることにより、ユーザの知識の補助となり、知識のないユーザの理解の助けとなることが期待される。

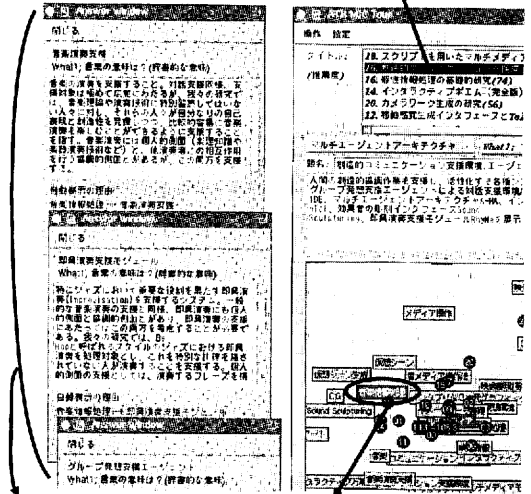
図4は、このユーザが次に別の情報ソースを閲覧したときの画面例である。この例では、上記の例において、ユーザが“音楽情報処理”について自動表示された情報に対して「興味がある」と返答したことを反映して、他の情報ソースを閲覧した際に、更に“音楽情報処理”に関連するキーワード“音楽演奏支援”と“即興演奏支援モジュール”についても情報の自動提示を行っている(4.2で述べた興味の推定)。

6. システムの試用評価

6.1 試用実験の概要

システムの効果を評価するために、複数人の被験者による試用実験を行った。被験者は38人の大学学部生

新しい情報ソースの閲覧選択



閲覧選択と同時に ユーザが過去に「興味がある」と判定したキーワード自動的に表示された回答窓

図4 関連キーワードについての回答窓の自動表示
Fig.4 Automatically presentation of answering windows about related keywords.

と大学院生である。そのうち25人は情報科学の専攻者であり、残りは材料科学、教育学などの専攻だった。

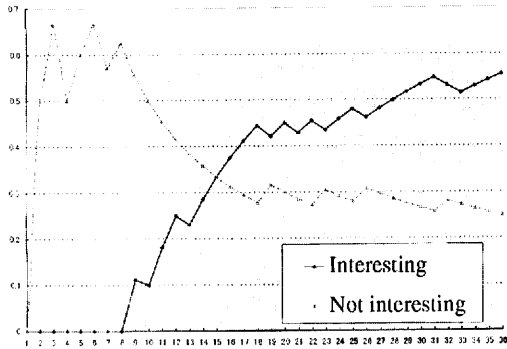
実験では、興味空間の適応結果を利用した情報提示を行うモード(Takealook 利用モード)と、興味空間の適応結果に関係なく、ユーザの選んだコンテンツ内に限ってランダムに情報提示を行うモード(ランダムモード)を用意し、各被験者には無作為にどちらかを利用してもらった。ユーザには、自分の興味に合わせて気に入った情報ソースから情報探索を始めてもらい、自分が満足するまでシステム利用を続けてもらった。システムの利用履歴を分析することにより、本システムにおける情報提示の精度の定量的評価を行い、被験者にシステム使用前とシステム使用後の両方にアンケートに答えてもらうことによりユーザの主観的評価を行った。そして、それらのデータや分析結果を合わせることで、システム応答とユーザの興味の变化の関係を評価した。

6.2 情報提示精度の定量的評価

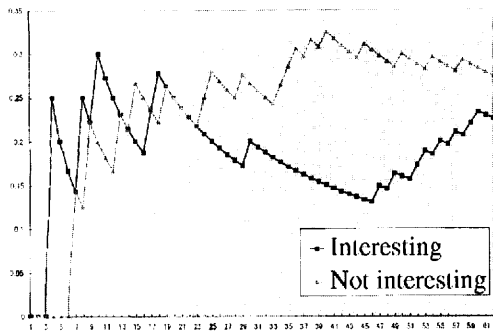
本システムの効果として最も期待していたことは、システム利用が進むにつれて、システムによってユーザの興味が徐々に獲得され、それが自動情報提示の精度を上げることである。このことを確認するために、システムが自動的に提示した回答窓に対して、ユーザ

が「興味がある」と判定したウィンドウの数と「興味がない」と判定したウィンドウの数の比率の変化を観察した。

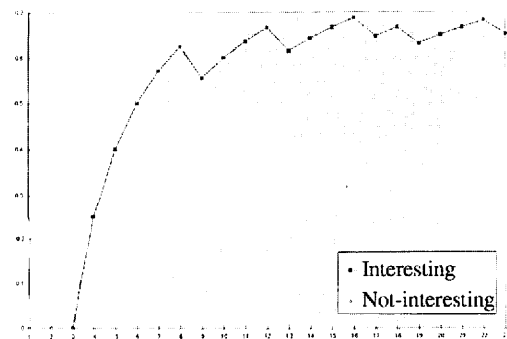
図 5(a)に、典型的なユーザ（これを Progress タイプユーザと呼ぶ）の場合の結果を示す。グラフの横軸は回答窓（ユーザの明示的な質問に対するものとシステムが自動的に表示したものを含む）の累積数、縦軸



(a) Progress タイプ
(a) Progress type.



(b) No-effect タイプ
(b) No-effect type.



(c) All-interesting タイプ
(c) All-interesting type.

図 5 ユーザによる自動情報提示に対する判定の推移
Fig. 5 Change in evaluation of automatically provided information.

はその中に占めるユーザの判定の比率を表している。この図を見ると明らかなように、最初はユーザにとって「興味がない」情報が多く提示されているが、ユーザの興味空間が徐々に適応されるにつれて、「興味がある」情報の割合が増えている。

Takealook 利用モードの実験を行った被験者 22 人中 21 人は、最終的に Interesting 曲線が Not-interesting 曲線よりも上回るという意味で、Progress タイプであるとみなすことができた。残りの 1 人はすべての提示情報に対して「興味がある」と判定をする、いわゆる「何でも興味がある」人だった。このことから、提案手法はユーザの興味が推定することに成功しており、的確な提供情報を選択するのに有効であるといえよう。

一方、ランダムモードの実験を行った 16 人の被験者については、7 人が No-effect タイプ、つまり Interesting 曲線の伸びが芳しくなく、最終的には Not-interesting 曲線の方が上回った結果となった(図 5(b) 参照)。しかし、Progress タイプの被験者も 7 人いた。これらの被験者に共通したのは、情報科学分野の学生である、かつ/または、該当分野に興味がある、ということであった。ランダムモードに Progress タイプがいたことは、該当分野にかかわりの高いユーザは、自ら該当分野に興味を見出し、徐々に該当分野での情報探索に順応していく能力があるためであると考えられる。

ランダムモードにおける残りの 2 人は「何でも興味がある」(All-interesting) タイプだった(図 5(c) 参照)。上記モードの 1 人を加えたこの All-interesting タイプの被験者に共通したのは、3 人も別分野の学生であったことである。これらの被験者は、該当分野とかかわりが低いため興味の対象が絞り込めず、しかしそれでも好奇心が旺盛なため、「何でも興味がある」という反応を示したものと考える。

6.3 ユーザの興味の変化とシステムへの応答との 相関に関する評価

システム使用前後でのユーザの興味の変化を調べるために、興味のあるカテゴリとキーワードを選択してもらってアンケート調査を行った。カテゴリは今回のドメインと合致しているという理由から(社)情報処理学会の研究会の分類を利用し、数に制限なく選択してもらった。キーワードは、今回扱ったコンテンツに含まれるキーワード 175 個である。Takealook がユーザの興味の変化に及ぼす影響を分析するために、各ユーザのオペレーション履歴と、システム使用前後にユーザに選択してもらった興味のあるキーワードの

結果を照らし合わせた。

以下、システム使用前後のいずれにおいても興味があるとして選択されたキーワードを（不変）、システム使用前には選択されてなかったが使用後に新たに興味対象として選択されたキーワードを（増）、システム使用前には選択されていたが使用後には選択されなかったキーワードを（減）と表現する。一方、ユーザのオペレーション履歴の中で、ユーザが自発的に質問を行ったキーワードを（自発的な質問）、自動的に情報提示され、それに対してユーザが「興味ある」と応答したものを（興味あり）、自動的に情報提示されたがユーザが「興味ない」と応答したものを（興味なし）と表現する。

これらのデータを照らし合わせて、該当する項目にあたるキーワードの数を Takealook 全被験者について平均した値を表 1 に示す。なお、表中で「なし」とあるのは、ユーザがアンケート調査において選択したにもかかわらず、システム利用中には閲覧されなかったキーワードを指す。同様に、ランダムモードの場合についても選択キーワードの数を全被験者について平均した値を表 2 に示す。

ここで、特に、「選択キーワード＝不変、ユーザの応答＝興味あり」と「選択キーワード＝増、ユーザの応答＝興味あり」の 2 値に注目する。なぜなら、前者は既存の興味に対するシステムからの情報提示の精度を表し、後者はユーザの気づかなかった興味をシステムが

喚起したものであることを表していると解釈できると考えるからである。以下、これらの値をそれぞれ「情報提示の精度の値」と「興味の促進の値」と呼ぶ。

表 1 にあるように、情報提示の精度の値及び興味の促進の値を Takealook 全被験者について平均した値はそれぞれ 3.95 と 3.76 である。これを、平均する以前の値を用いて、該当するキーワード数のユーザ分布を表したものが図 6 である。横軸が選択されたキーワード数の値であり、縦軸がそこに属するユーザの人数である。同様に、図 7 はランダムモード全被験者に関する分布を表したものである。

これらを比較すると、「情報提示の精度の値」と「興

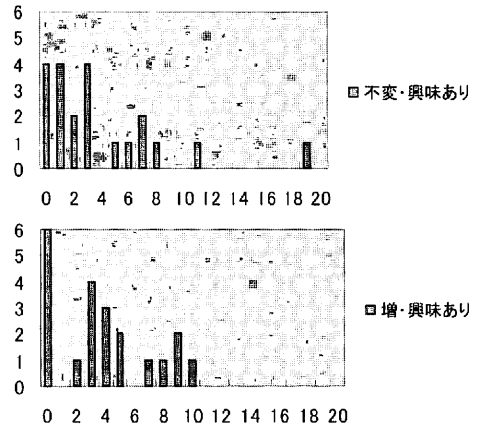


図 6 選択キーワード数のユーザ分布 (Takealook モード)

Fig. 6 User distribution against the number of selected keywords (Takealook mode).

表 1 Takealook 利用における興味の変化 (選択キーワード個数の全被験者平均)

Table 1 Change in users' interests during using Takealook (the average number of selected keywords).

選択キーワード	不変	増	減
ユーザの応答			
興味あり	3.95	3.76	1.67
興味なし	1.19	1.57	0.76
自発的な質問	1.14	1.71	0.43
なし	8.90	5.67	8.14

表 2 ランダムモードにおける興味の変化 (選択キーワード個数の全被験者平均)

Table 2 Change in interests of random-mode users (the average number of selected keywords).

選択キーワード	不変	増	減
ユーザの応答			
興味あり	2.19	2.75	1.00
興味なし	1.75	1.19	0.94
自発的な質問	6.13	6.56	2.00
なし	15.50	10.44	14.25

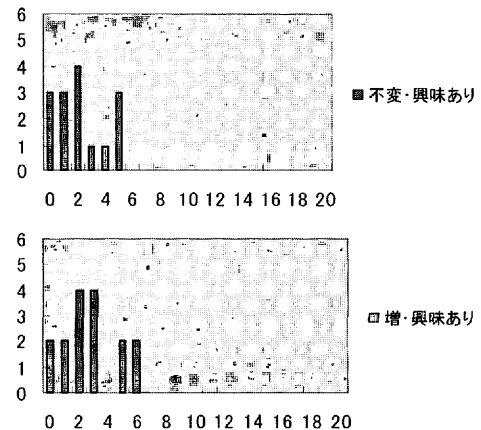


図 7 選択キーワード数のユーザ分布 (ランダムモード)

Fig. 7 User distribution against the number of selected keywords (random mode).

味の促進の値」のいずれについても、Takealookを使用した場合の方がランダムモードよりも大きい。そして、それらのユーザ分布を見ても、Takealook使用の場合には比較的値の大きいユーザが見られるがランダムモードの場合には見られない。以上のことから、Takealookによる自動的な情報提示の方がランダムモードのそれよりも、ユーザの既知の興味に沿ったものであり、かつ、ユーザが前もって気づいていなかった興味を喚起しているということが出来る。

ランダムモードの場合、全体としてはシステム前後のアンケートで選択されたキーワードの変化が多かった(表2中の(増)と(減)に対応するキーワード数が多いことに対応)にもかかわらず、「情報提示の精度の値」と「興味 の促進の値」が低かった。そして、応答=なし、つまりシステム使用中に閲覧されなかったキーワードの数がTakealookユーザに比べて高かった。このことは、ランダムモードでは実際にユーザの興味の変化がシステムによってもたらされたのかの信頼性が低いことを示している。

7. む す び

本論文では、ユーザのオペレーション履歴から個人の興味空間をユーザ適応し、それをを用いることにより情報の個人化を行う手法を提案した。興味空間のユーザ適応に際しては、複数の情報提供者によってあらかじめ準備された個人的視点を表す概念空間を利用した。

ユーザの利用履歴を定量的に分析することで、提案手法が、ユーザの興味を推定して情報提供を個人化する手法として適当であることを示すことができた。システム使用前後にユーザが興味を示したキーワードの変化と、システム使用中のユーザの応答を分析することで、ユーザの既知の興味と新たな興味への適応が可能であることも確認された。

今後の問題点としては以下のようなことが挙げられる。本研究の手法の本質的な限界の一つとして考えられるのは、ユーザがある1対のキーワード間に新しい関連を見出したとしても、それらの間に1人の情報提供者もリンクを張っていない場合には、自動情報提供のための概念間連鎖が起こり得ないことである。今回設定したような研究所見学等の場合、情報提供者の知識量が明らかにユーザに対して勝っているので、提供者の提供情報に頼るのは妥当であろう。しかしそれでも、ユーザが新たに発見したキーワード間の関連が提供者にとって新しい重要な視点をフィードバック

する可能性はある。そのような機構を提供することが今度の課題である。これは、従来の展示型情報提供の限界を越え、展示者と見学者の双方向のコミュニケーションを実現するものであると期待される。

また、情報コンテンツ収集の際の効率化の問題がある。キーワード間リンクの情報を準備する際、既存の辞典等から同義語・類義語などの候補を提示することで、情報提供者の負担軽減と見落とし防止をする方法が考えられる。また、情報ソースの準備の方法として、ネットワーク上に既に存在している情報(WWW上の情報等)を利用することが考えられる。その実用化のためには意味情報を含んだタグ付け技術を利用する必要があろう。

今後の展開として、システム利用を通してユーザの興味に適応した興味空間を、複数のコンテンツの情報探索の際に再利用することが考えられる。

謝辞 本研究の機会を与えて頂いたATR 知能映像通信研究所の中津良平社長、データ収集及び実験に御協力頂いた同研究所の皆様と東京大学工学系研究科堀・中須賀研究室の皆様、日ごろから御指導頂いているブレイクスルー21西川プロジェクトの西田豊明教授に感謝致します。

文 献

- [1] A. Flammer, "Towards a theory of question asking," *Psychological Research*, vol.43, pp.407-420, 1981.
- [2] N. Miyake and D.A. Norman, "To ask a question, one must know enough to know what is not known," *J. Verbal Learning and Verbal Behavior*, vol.18, pp.357-364, 1979.
- [3] R.A. Finke, T.B. Ward, and S.M. Smith, *Creative Cognition*, MIT Press, 1992.
- [4] M. Polanyi, *The Tacit Dimension*, Routledge & Kegan Paul Ltd., London, 1966. 佐藤敬三訳, 暗黙知の次元, 紀伊國屋書店, 1980.
- [5] M.S. Ackerman and T.W. Malone, "Answer Garden: A tool for growing organizational memory," *Proc. ACM Conf. Office Inf. Syst.*, pp.31-39, 1990.
- [6] 中山康子, 真鍋俊彦, 竹林洋一, "知識情報共有システム (advice/help on demand) の開発と実践— オフィス知識ベースとノウハウベースの構築," *インタラクション'97 論文集*, pp.103-110, 情報処理学会, 1997.
- [7] 武田英明, "ネットワークを利用した知的情報統合," *人工知能誌*, vol.11, no.5, pp.680-688, 1996.
- [8] U. Shardanand and P. Maes, "Social information filtering: Algorithms for automating "word of mouth"," *Proc. CHI'95*, pp.210-217, ACM, 1995.
- [9] M. Balabanović and Y. Shoham, "Fab: Content-based, collaborative recommendation," *Commun. ACM*, vol.40, no.3, pp.66-72, 1997.

- [10] B. Shneiderman, Designing the User Interface (Third edition), Addison-Wesley, 1997.
- [11] P. Resnick and H.R. Varian, eds., "Special section on recommender systems," Commun. ACM, vol.40, no.3, pp.56-89, 1997.
- [12] 角 康之, 堀 浩一, 大須賀節雄, "テキストオブジェクトを空間配置することによる思考支援システム," 人工知能誌, vol.9, no.1, pp.139-147, 1994.
- [13] 堀 浩一, "思考の可視化," 可視化情報学会誌, vol.19, no.72, pp.2-6, 1999.
- [14] 角 康之, 江谷為之, シドニーフォールズ, ニコラシモネ, 小林 薫, 間瀬健二, "C-MAP: context-aware な展示ガイドシステムの試作," 情処学論, vol.39, no.10, pp.2866-2878, 1998.
- [15] 角 薫, "個人の概念空間を利用したコミュニケーション支援," Ph.D. thesis, 東京大学大学院工学系研究科先端学際工学専攻, 1999.

(平成 11 年 3 月 1 日受付, 6 月 25 日再受付)



角 薫

東京理大・理・物理卒。筑波大大学院(経営システム科学)修士課程了。東大大学院(先端学際工学)博士課程単位取得退学。博士(工学)。その間、通信会社等でのシステムエンジニアリングの職務。(株)ATR 知能映像通信研究所研修研究員を経て、現在郵政省通信総合研究所関西支所ブレイクスルー 21 西田プロジェクト特別研究員。情報の個人化の研究に従事。人工知能学会、情報処理学会各会員。



角 康之 (正員)

1990 早大・理工・電子通信卒。1995 東大大学院(情報工学)了。同年より、(株)ATR 知能映像通信研究所研究員。博士(工学)。研究の興味は知識処理システムと CSCW。人工知能学会、情報処理学会、AAAI 各会員。



間瀬 健二 (正員)

1979 名大・工・電気卒。1981 同大大学院修士(情報)課程了。同年日本電信電話公社(現在 NTT)入社。以来、NTT 研究所にて画像情報システム、コンピュータグラフィックス及び画像処理、そのヒューマンインタフェースへの応用の研究に従事。1988~89 年米国 MIT メディア研究所客員研究員。1995 年 2 月より ATR。現在(株)ATR 知能映像通信研究所第二研究室長。IEEE, ACM, 情報処理学会各会員。博士(工学)。



中須賀真一

1983 東大・工・航空卒。1988 同大大学院工学系研究科博士課程了。同年、日本アイ・ビー・エム東京基礎研究所勤務。1990 東京大学航空学科講師。1993 助教授。1994 から先端学際工学専攻助教授。1998 より航空宇宙工学専攻助教授。現在に至る。工博。1996~1997 アメリカ、メリーランド大学コンピュータサイエンス学科客員研究員。1999 スタンフォード大学客員研究員。宇宙工学。特に宇宙機システムとその設計、航法誘導制御、宇宙機の知能化・自律化。宇宙ロボティクス、機械学習、群知能等の研究に従事。1991 日本航空宇宙学会奨励賞。1997 人工知能学会全国大会優秀論文賞受賞。1998 経営情報学会秋期大会優秀論文賞受賞。1999 同学術奨励賞(研究奨励賞)受賞。AIAA, IFAC, 計測自動制御学会、日本航空宇宙学会、日本ロボット協会に所属。



堀 浩一 (正員)

1979 東大・工・電子卒。1984 同大大学院博士課程了。同年、国文学研究資料館助手。1986 助教授。1988 東京大学助教授。1997 同教授。現在に至る。工博。この間、1989 年 9 月~1990 年 1 月仏国コンピュニユ大学客員助教授。現在、先端学際工学専攻所属。航空宇宙工学専攻と情報工学専攻を兼任。人工知能を中心とした情報処理システムの基礎から応用に至る広範囲の研究・教育に従事。最近の個人的な興味は創造活動支援システム。人工知能学会、情報処理学会、日本ソフトウェア科学会、日本認知科学会、IEEE, ACM 各会員。