

概念空間を利用した情報提供の個人化

小林 薫 ††, 角 康之 †, 間瀬 健二 †, 堀 浩一 †

†ATR 知能映像通信研究所

†東京大学 工学系研究科

本稿では、ユーザの興味に基づき、個人化された情報の提供をする手法を提案する。具体的には、オペレーション履歴、個人化された概念空間、それらを制御する知識から構成される情報リトリバにより、個人ごとの情報提供と情報推薦が行われるものである。ここでは、研究所見学という枠組みの中で、情報提供者としての研究者と情報の受け手として見学者がコンピュータを介した質問回答のインタラクションの履歴を利用することにより見学者の興味や知識のレベルを推定し、それにそった情報提供を行うことを考える。

Personalized Information Presentation using Individual Spaces

Kaoru Kobayashi ††, Yasuyuki Sumi †, Kenji Mase †, and Koichi Hori †

†ATR Media Integration & Communications Research Labs.

†The University of Tokyo

In this paper, we propose a method for presenting information based on individual user interests and knowledge. In the proposed method, the system acquires the user's implicit and/or explicit interests and knowledge from a question-and-answer interaction and presents the user with focus-of-interest information based on them. For personalizing the presentation and recommendation of information, the system provides each user with a information retriever using the user's operation records, a personal conceptual space, and control knowledge.

1 はじめに

情報を発信し伝達しようとする際には、さまざまなコミュニケーションの形態がある。直接言い伝えるような同期的なコミュニケーションの場合には相手の反応を見て、それに応じて伝達する情報を変えることもできるが、文章などの非同期的な場合では、情報発信側からの一方的な情報になってしまいがちである。現代のように情報が膨大に溢れている世の中では、情報側が個人の視点に合わせて変化してくれることはありがたい。

膨大な情報からの検索を支援するため、ユーザの過去の検索履歴などからユーザの興味を抽出し、それに基づいてインターネットのホームページを推薦する研究が盛んである。

しかし、これまでの研究はあらかじめ与えられたカテゴリへの分類分けやキーワードが含まれるか含まれないかなどの方が多い [Balabanovic 97], [Caglayan 97], [Shardanand 95]。カテゴリへの分類分けは、前もって固定のカテゴリの決定や分類分けが必要であり、キーワードの共起による方法だけでは、興味を表現するのに不十分であることが多い。

そして、情報提供の個人化を考える場合、カテゴリ

やキーワードレベルのユーザの興味の表現だけでなく、そのキーワードのどのようなことに対して個人が興味を示しているかや個人がどのくらいのレベルの知識を持っているかという個人による違いも考慮に入れ、それに応じて情報が変化する必要がある。

本研究では、情報技術を情報提供者側と情報受け手側のコミュニケーションの仲介役としてシステム側から人間への歩み寄りを考え、個人の興味と理解に合わせて情報を変化させる情報提供を行うシステムの構築を目指す。

基本的アイデアは、ユーザとシステムの質問-回答の対話を通して、ユーザの興味の獲得をする。そして、過去のユーザの興味をもとに、複数の情報提供者の異なる視点を利用し、ユーザとシステムとのインタラクションを通して後述の collaborative filtering に類似の方法で学習する推論エンジンによる興味の共起の推論を行い情報提供するというものである。この問題は、広い概念空間の中から未知であるユーザの興味対象を、情報提供者の複数の視点を利用しながら、比較的少ないインタラクションで推論するというに言い換えることができる。

本章以降では、2章において研究の背景と動機、3章で質問-回答対話における情報提供の個人化手法の概要を述べる。4章でシステムの実装と動作シナリオを示し、5章でまとめをする。

連絡先: 小林 薫

〒 619-0288 京都府相楽郡精華町 ATR 知能映像通信研究所

phone: 0774-95-1401, fax: 0774-95-1408

e-mail: kaoru@mic.atr.co.jp

2 研究の背景と動機

一般に情報は、テキスト情報、画像情報、音声情報など他にも様々な情報があり、それらの集まりから構成されている。そしてさらに、テキスト情報を例としてあげると、言葉の意味、動作の説明、その動作の理由など様々な意味情報の組み合わせにより構成されている。個人が求める情報は、持っている興味や知識により異なっており、これらの情報をその興味や知識によって異なる要素で組み合わせることにより情報提供の個人化は可能であると考えている。

興味や知識といった個人情報を獲得する上で、ユーザに負荷のない方法としてユーザの質問-回答の対話を観察する方法を考えた。個人の興味や知識がその個人の発した質問から判断できるということは過去の研究において報告されている。好奇心は、概念知識やそれにもとづく期待と新しく得られた情報との間にずれがあるときや、認知構造のなかに空隙や矛盾があることが認知されたときに引き起こされるということを [稲垣 82] は言っており、[Flammer 81] は、質問は質問者が持っていない情報を求めるものであると言っている。また、質問の質と知識の度合いの関連の実験として、[Flammer 81] や [Miyake 79] の研究が知られている。これらのことから、質問はユーザの興味や知識を外界から観察できる手段であると言えよう。そして、好奇心と知識と質問については関連があることが推測され、質問の質から知識のレベルは観察することができそうである。

[波多野 73] と [稲垣 82] によると、人が情報収集をする際の好奇心には、2種類あると言われている。一つは、退屈ないし情報への飢えより生じ、はっきりした方向性をもたず、幅広く情報を求める傾向は拡散的好奇心 (diverse curiosity) で、もう一つは、特定の環境の、特定の特性や関係の代表する情報の取得のみを追求する選択的な情報収集の傾向は特殊的好奇心 (specific curiosity) とよばれる。拡散的好奇心の際には、広く情報を検索したいと思ひ、特殊的好奇心ではある特定のものに関連した情報を得たいと思うであろう。

そこで、好奇心をさまざまにユーザの主体的な情報の検索、好奇心を刺激するユーザの興味に合わせた情報のシステム側からの提示といったユーザと情報とのかかわりの過程の中で、ユーザの好奇心である質問の観察により、ユーザ情報を獲得し情報の組み合わせで情報を個人化を行うこととした。

3 情報提供の個人化の枠組み

本稿で提案するシステムでは、ユーザの質問-回答の対話を繰り返すことで徐々にユーザの興味や知識を推論し、それに応じてユーザによって異なる情報を付加して提示することにより情報提供の個人化をはかる。

図 1 は、本稿で提案するシステムの枠組みである。システムは、複数の情報提供者から提供された発信情報をユーザ固有の情報リトリバを介して受信情報としてユーザに提供する。

矢印 (大) は提供される情報の流れ、矢印 (小) はユーザ情報の流れを表している。

以下の節では、情報提供者側の情報の準備、情報リトリバの機能、推薦値の計算、ユーザへの情報の提供について順を追って説明する。

3.1 情報提供者側の情報の準備

最初に、テキスト情報、キーワード、それに対する質問-回答 (Q&A) から構成される情報ソースを情報提供者が用意する。

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や、“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や、“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や、“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もとめるもののように、数種類用意できるようになっている。

この質問は、What, What2, How, Why の順に知識のレベルの高くなり、および興味の深さのレベルが深くなるということを考慮して設定した。選択された全てのキーワードに対して全種類が用意されるわけではなく、あるキーワードについては2種類、あるキーワードについては1種類など情報提供側に委ねられている。

情報提供者側で用意する概念空間は、各情報提供者の視点にあたるもので、情報提供者が Q&A を作る上で使ったキーワードに対してそれぞれ、他の全情報提供者の選択した全キーワードで興味の共起がおこりそうなものについて結び付きをつけたものである。

例えば、“エージェント”ならば“擬人化エージェント”などという、包含的な関係や、“コミュニケーション”ならば“インタラクティブ”などという連想的な関係、“テレビ”ならば“ラジオ”といった類似した関係も対象となり、この関係は可逆とする。

3.2 情報リトリバの機能

情報リトリバは、個々のユーザ毎に用意され、個人化された情報を提供する。

ユーザが質問をしたということはその対象について興味を持っていると見なし、選択したキーワードと質問の情報はオペレーション履歴としてシステムに蓄積される。そのため、ユーザのオペレーション負荷なく質問-回答の対話を通してユーザの興味や知識を獲得する。

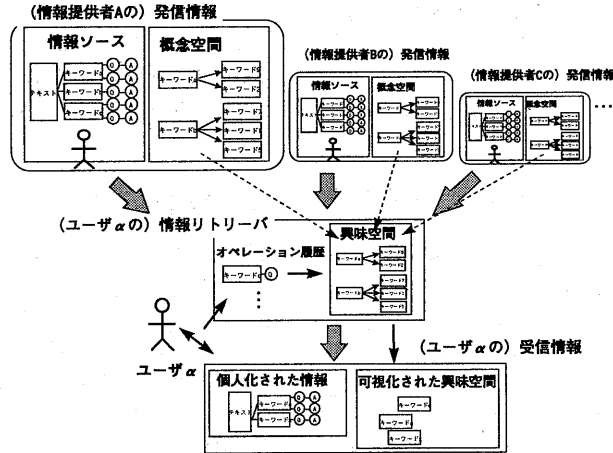


図 1: システムの枠組み

情報リトリバの推論エンジンでは、各情報提供者の概念空間を情報提供者固有の視点として利用し、CF (Certainty Factor) 値付きのプロダクションルールにより、興味の共起の推論を行っている。ユーザーのオペレーション履歴からある一定の閾値より新しいものでプロダクションルールの発火を行う。推論された結果のキーワードおよび質問(発火元のオペレーション履歴の質問をそのまま引き継ぐ)が情報ソースに存在していれば、ユーザーの個人化情報の候補の一つとなる。

質問は、What, What2, How, Why の順に知識のレベルの高くなり、および興味の深さのレベルが深くなるということ considering 設定されているので、What などの知識レベルの低い質問があった場合、推論したキーワードに高い知識レベルのものもシステム側で提示することも考えられるが、拡散的好奇心であった場合を考慮して、システム側からは推論したキーワードに対して適用する質問は、同レベルものということにした。このように同レベルのものを引き継ぐ方法は、特殊的好奇心を促進するという考えられると思っている。

また、このように推論された結果提示された情報をユーザーに興味があったか否かを評価してもらうことで、ルールの重み付けの調整が行われ洗練化される。

広い概念空間の中から未知であるユーザーの興味対象を、情報提供者の複数の視点を利用しながら、比較的少ないインタラクションで推論するという問題において、ユーザーの評価を最大限に利用するために、人間の特性に基づいたフィルタリングの手法である collaborative filtering に類似の方法で情報リトリバの推論エンジンの強化学

習を行う。collaborative filtering は、ユーザー情報を利用することにより情報をフィルタリングする手法で、ユーザーをグルーピングし、同じグループのユーザーは同じ興味を持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フィルタリングを行っている。本手法では、ユーザーの視点は部分的な視点が複数組み合わせられているものとみなし、その部分的な視点をフィルタリングの単位とする。

例えば、「概念 α ならば概念 β に興味がある」というプロダクションルールを概念 $\alpha \rightarrow$ 概念 β と表現すると、推論されたルールの発火の軌跡である概念 $\alpha \rightarrow$ 概念 $\beta \rightarrow$ 概念 $\gamma \rightarrow$ 概念 δ は、オペレーション履歴の概念 α から概念 δ が導き出されたということである。例えば、ユーザーが概念 δ について興味のある情報だと評価した場合には、このルールの発火の軌跡のそれぞれのルールに相当するものが概念空間に存在しているものを調べ、その概念空間内で該当ルールの右辺に相当する視点の部分の単位でプロダクションルールの重みを全て加える。例えば上記の例では、概念 $\alpha \rightarrow$ 概念 β で、概念 $\alpha \rightarrow$ 概念 β に相当するものがある概念空間を探し、その概念空間が右辺が同じだが左辺が異なる概念 $\alpha \rightarrow$ 概念 i 概念 $\alpha \rightarrow$ 概念 κ などのルールに相当するものがあったなら、それらについてプロダクションルールの重みを加えるという意味である。

ルールは可逆のため、左辺についても同様の処理を行う。これは、ユーザーの全視点のうち、ある部分的な視点について強めて取り入れたことに相当する。逆に、ユーザーが興味のない情報だと評価した時には、同様のルール

の重みを減らす処理を行う。このような強化学習を繰り返すことにより、情報リトリバの推論エンジンを次第に洗練化させる。

3.3 推薦値の計算

推薦値の値は研究ごとに計算し、推薦値の高い順にタイトルの部分に表示する。

推薦値は、基本的に過去のオペレーション履歴からユーザが質問したキーワードが該当研究において共有しているか、そして、そのキーワードそのものが共有しているかだけでなく、個人の興味空間の興味共起されるキーワードについても計算の対象とする。もし、複数の研究概要にそのキーワードが含まれていたとすると、それに対してユーザの興味のあるであろう質問に対する回答を含んでいたものに高い点数が与えられるようになっている。これは同じキーワードを含んだ別の研究があっても、情報提供者がそれら共に同種の質問と回答を用意しているとは限らないからである。

3.4 ユーザへの情報の提供

このようにしてユーザに提供される情報は、共通のテキスト情報と前もって提示されたQ&Aの答えに相当する個人化情報とユーザの興味空間である。

ユーザの個人化情報は別窓に補助情報として表示されるため、発火ルールの重みの積の数値が大きい順に、認知容量の理論 [Miller 56] に基づいて最大6自動的に表示させることとした。この別窓には、推論の理由として推論の軌跡を表示している。

ユーザの個人化情報は表示されると同時に、ユーザの興味空間が表示され、自動的に表示された情報の推論の理由として示されるようになっている。推論エンジンのプロダクションルールの重みがキーワードのペアの距離に対応しており、多次元尺度法を利用した興味空間を可視化したものである。最初に発火したキーワード、すなわち過去にユーザの質問があったものとの興味共起の関連が距離で表現されている。

4 システムの実装と動作シナリオ

本システムは、研究所見学をする際、ユーザがその前後にオフサイトでアクセスできる研究所の研究を紹介するシステムとして、研究所見学のガイダンス、*Context-aware Mobile Assistant Project (C-MAP)* [角 98] において適用した [小林 98], [Kobayashi 98]。

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ものである。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7個のキーワード、それに対する質問-回答、概念空間を用意する。

情報提供の種類は約25展示物、全部で約175のキーワード、学習前の推論エンジンルール数は最大約200種類である。

4.1 情報提供の個人化

以上のように準備されたものにより、実際どのように情報提供されるかを順を追って説明する。

図2において、右側のウィンドウがメインウィンドウで大きく分けると3つの部分に分れている。上部にタイトルが推薦度と一緒に表示され、そのすぐ下に研究の概要が表示されるテキストエリア、その下にユーザの興味空間が表示されている。タイトルは推薦度(後述)の高い順に推薦度とともに表示される。タイトルを選択すると、そのタイトルの研究の紹介の内容にうつることができる。

研究の概要が表示されたテキストエリアでは、上部分がキーワードと質問が選択できるようになっていて、ユーザが質問ができるようになっている。

興味空間はユーザの興味である、過去に質問した興味キーワードとそれから共起される概念が表示され、キーワード間の距離は興味共起の確からしさ、で表されている。ここで、四角いアイコンはキーワードで、丸いアイコンは研究を表しており、その研究に関連するキーワードの重心に位置している。このアイコンはメインウィンドウの上部のタイトルを選択すると同様、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うつることができる。

まず最初に、ユーザがタイトルから自分の興味のあるも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などになじみがなく、マルチメディアの辞書的な意味が知らなかったわけである。そうすると、システムではこの回答を図2の左の小窓、アンサーウィンドウに表示する。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”についての意味を知らないのなら、“マルチメディア技術”や“CG”についても知らないだろうと推論し、前もってその情報を提示しているのである。ここでは他にキーワード“音楽情報処理”や“ダンスインストゥルメント”などもあり、それぞれ表示された理由として、推論の軌跡をアンサーウィンドウの下部に示している。このようなアンサーウィンドウを得ることにより、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成ることが可能である。

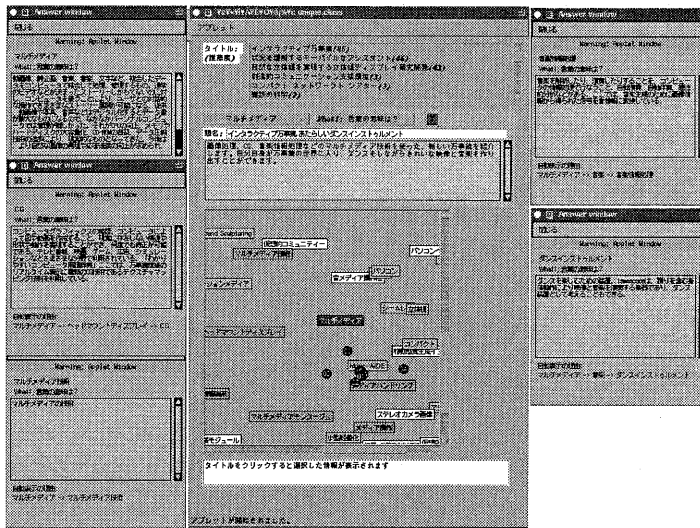


図 2: 提示された研究概要内のキーワードに関する質問 - 回答の対話 (マルチメディア)

また、次に計算機に興味があるユーザの例をあげる。ユーザは計算機の機種を問うためキーワード“携帯情報端末”“それは具体的には何(例)ですか”(What2)の質問をして、それから他の研究情報にアクセスした場合に、システムは自動的にキーワード“パソコン”における質問“それは具体的には何(例)ですか”(What2)の回答を自動的に表示した。これは、このような特定の興味を持ったユーザに対しても、情報提供の個人化ができ、ユーザの特殊的好奇心を助長するようにシステム側からの suggestion が行われていると言える。

また、システムは情報リトリバの推論をもとに情報提供を行うが、自動的に表示されたアンサーウィンドウをユーザが閉じる際には興味がある内容か否かを入れなければ閉じれないようになっており、その結果により情報リトリバは学習される。次にその学習の例を述べる。

これは、ユーザが“擬人化エージェント”というキーワードにひかれてその研究にアクセスしたのだが、その研究は認知科学の分野の研究であり、ユーザの求めていた映像的な研究ではなかった場合の例である。

まず、ユーザが研究情報表示後、メインウィンドウで、キーワード“擬人化エージェント”において“具体的には何(例)ですか?”(What2)という質問を行ったとすると、キーワードと質問がオペレーション履歴として追加され、それをもとに情報リトリバが個人化した情報を自動表示する(図 3)。興味空間には、キーワード“擬人化エージェント”のまわりに興味と共に起される概念が表示され、現在開いている研究情報にアンサーウィンドウが表示さ

れる。

ユーザは、認知科学的なものに興味を持っていないので、“対人社会心理的反応”、“対人的つきあい”などのアンサーウィンドウには「興味がない」と回答する。その結果、図 4 のような認知科学的なキーワードの距離が離れて表示された興味空間となり、認知科学的なキーワードは表示されなくなるようにシステムは学習される。そして、“擬人化エージェント”から共起される興味の概念である“エージェント”などが表示されている。

このように、学習された結果を見てみると、特に前もって分野別に分類して学習させているわけではないにもかかわらず、興味のない分野については情報が表示されなくなり、興味がある分野については表示される。情報提供者の単位でのフィルタリングではなく、キーワードの連鎖を行って情報提供者の視点の部分を使ったきめ細かいフィルタリングをした結果、比較的早く学習を行うことができた。

5 おわりに

本稿では、提供者側からのテキスト情報、質問 - 回答の情報、および提供者の視点である概念空間を利用し、ユーザのオペレーション履歴から個人の興味空間を学習し、それ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。本手法により、ユーザは全ての情報を検索することなく、そしてオペレーション付加なく比較的早いシステムの学習で興味がある情報にたどり着くことが可能であると考える。

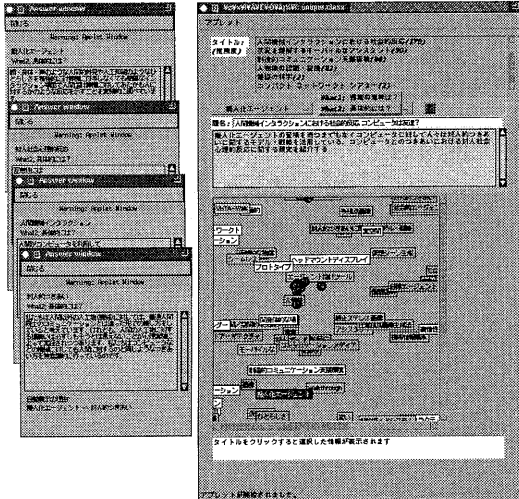


図3: キーワード“擬人化エージェント”に関する“What2?”質問の適用とそれに対する回答

本手法では、従来推薦システムなどにおいて行われていたような単純にキーワードの有無等によりテキストを分類したようなものよりも、より意味的に深く踏み入れたキーワードの背後に隠れたユーザの興味を考慮し、そして領域を超えた推薦が可能であると考えられる。

また、今後の拡張として、学習したユーザの興味空間を違うコンテンツに対してのナビゲーションへ利用する事や、ユーザ自身のコンテンツ生成などの知的作業支援、さらには概念空間同士の比較を利用した情報提供者とユーザまたはユーザ同士の人間同士の出会いの支援も考えられる。

参考文献

[Balabanovic 97] Balabanovic, M. and Shoham, Y.: Fab: Content-based collaborative recommendation, *Communications of the ACM*, 40, 3 (1997).

[Caglayan 97] Caglayan, A.K. and Harrison, C.G.: *Agent Source Book*, Willey Computer Publishing (1997).

[Flammer 81] Flammer, A.: Towards a theory of question asking, *Psychological Research*, 43, 407-420 (1981).

[稲垣 82] 稲垣 佳代子: 認知への動機づけ, 認知心理学講座 4 学習と発達, 東京大学出版会 (1982).

[波多野 73] 波多野 諄余夫, 稲垣 佳世子: 知的的好奇心, 中公新書 (1973).

[小林 98] 小林 薫, 角 康之, 間瀬 健二: 個人の興味と知識に基づく情報提供の手法, 知的教育システム研究会 知識ベース

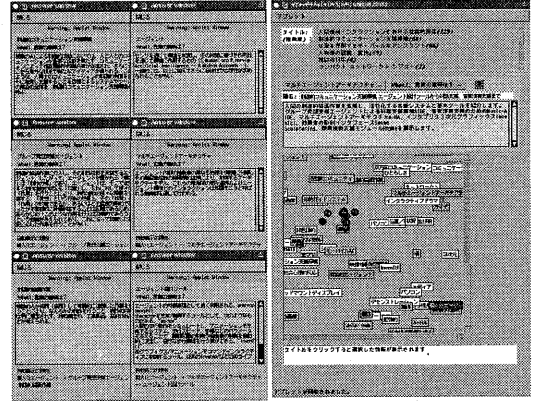


図4: 他の情報への遷移後の説明の自動的な表示

システム研究会共催 特集テーマ「知識の相互伝達」人工知能学会資料, pp.53-pp.60 (1998).

[Kobayashi 98] Kaoru Kobayashi, Yasuyuki Sumi, and Kenji Mase: Information Presentation based on Individual User Interests, KES'98 Second International Conference on Knowledge-based Intelligent Electronic Systems, IEEE, pp.375-pp.383 (1998).

[Miyake 79] Miyake, N. and Norman, D.A.: To ask a question, one must know enough to know what is not known, *Journal of Verbal Learning and Verbal Behavior*, 18, 357-364 (1979).

[Miller 56] Miller, G.A.: The magical number seven plus or minus two: Some limits on our capacity for processing information, *Psychological Review*, 63, 81-97 (1956).

[Shardanand 95] Shardanand, U. and Maes, P.: Social information filtering: Algorithms for automating 'word of mouth', *Proceedings of the CHI-95 Conference*, Denver, CO, ACM Press (1995).

[角 98] 角 康之, 小林 薫, 江谷 為之, 間瀬 健二: 展示情報空間における出会い/情報共有支援情処研報, ヒューマンインターフェース, this volume (1998).