

「高度メディア社会の生活情報技術」
平成12年度採択研究代表者

橋田 浩一

(産業技術総合研究所サイバーアシスト研究センター 副センター長)

「人間中心の知的情報アクセス技術」

1. 研究実施の概要

誰でもいつでもどこでも欲しい情報を手際良く利用できる環境を実現するため、意味と状況の理解を人間と人工物との間で共有する方法を技術的に具体化することによって人間中心の高度な知的情報アクセスを実現する方法を研究する。すなわち、コンテンツを意味構造化する作業のコストを抑制し、また情報サービスの品質を高めるために、構造化作業およびコンテンツ利用者である人間の特性に合わせたコンテンツの処理とインタフェースについて研究する。

平成12年度には、意味構造を用いた高性能の情報検索、経済コンテンツの可視化と提示、ユーザモデルの構築とその理論的基礎、および位置情報の取得等について検討を行なった。その結果、コンテンツの意味構造の利用法や意味記述の標準化に関する基礎的な知見を得るとともに、ユーザの置かれた文脈に関する情報の取得についても実証的および理論的な見通しを立てた。特に、インタラクティブな情報検索を高い精度で行なうための方法を考案する等の成果を挙げている。

今後は、各研究項目に関する研究を深めるとともに、コンテンツの意味構造の作成および利用の両面にわたるユーザモデルの取得と利用について研究を進める予定である。

2. 研究実施内容

誰でもいつでもどこでも欲しい情報を手際良く利用できる環境を実現するには、さまざまな情報コンテンツに対しその形式ではなく意味内容と利用の文脈に即してアクセスする技術を確立せねばならない。それには、意味と状況の理解を人間と人工物との間で共有する必要がある。本研究課題では、その方法を技術的に具体化することによって人間中心の高度な知的情報アクセスを実現する方法を研究する。すなわち、コンテンツを意味構造化する作業のコストを抑制し、また情報サービスの品質を高めるために、構造化作業およびコンテンツ利用者である人間の特性に合わせたコンテンツの処理とインタフェースについて研究する。

平成12年度には以下のような研究を行なった。
意味構造を用いた情報検索

言語の係り受けや代名詞の照応などの意味構造を情報検索を用いて、高精度かつ高効率の情報検索を行なう方法について検討した。情報検索が難しいのは、検索質問と検索対象の間の表現の差異を理めて意味的な類似性を検出する推論が難しいからであり、一般にはそのような推論を完全自動的に行なうことは現在の技術では不可能である。しかし、意味構造を手掛かりとして人間とコンピュータとがインタラクションを行なうことにより、この推論を半自動的に行なうことは可能と思われる。

意味構造はラベル付きグラフと見なすことができるので、このグラフを用いてインタラクティブに検索質問を拡張して検索範囲を絞り込む方法を実装し、被験者を使って実験を行なったところ、既存の情報検索の方法に比べて遜色のない正解率と再現率が得られたが、既存の方法よりも明らかに優れているという結果は得られなかった。被験者からの意見の聴取と実験結果の分析を行なったところ、この実装法では検索範囲の絞り込みが性急過ぎて、次々に絞り込んで行くとしばしば検索に失敗してしまうことがわかった。

この問題を解決するには、検索質問の改訂は検索範囲の絞り込みではなく正解候補に関するスコアの計算の調整を引き起こすとするのが良いとの結論に達した。またその他に、意味構造を用いたインタラクションとして、各語をシソーラス拡張した結果の適格性の判定をユーザに行なわせ、しかもその結果を以降の検索過程において保持することが効果的だろうとの見通しを得た。平成13年度は、これらを生かした実装を行ない、意味構造に基づく検索の有効性を実証する予定である。また、この方法に基づく検索サービスを一般利用者向けのデモンストレーションとして公開するとともに、さまざまな応用を探りたい。

経済コンテンツの提示

経済システムのマルチエージェントシミュレーションの結果をユーザに提示する技術に関して予備的な調査研究を行った。その結果、この分野でのシミュレーションは各研究者・システム開発者が個別にシステム開発を行っており、他の研究者の実験の内容を検証したり、システムを改良して新たな実験を行ったりするのが困難な状況にあることが判明した。またインターネットでの情報伝達を前提としたデータの表現形式についての配慮が必ずしも十分ではなく、大規模なシステム作成やネットワークを経由したデータの交換が容易にできるような環境が整っていないことが明らかとなった。

この予備的な研究から明らかになったことは、1) 経済主体を表現するエージェントの標準的な表現を用意するよりも、エージェント間の通信に着目して、通信プロトコルの標準化を進めるという標準的な手法が経済システムシミュレーションでも有効だと思われる 2) 通信量は比較的少ないが、質の点においては数値・

自然言語の混在系である 3) ミクロな経済エージェントのレベルとマクロなシステム全体の階層は比較的明瞭に分離されている、などである。

これらの結果から、この分野のシステムの標準化の方針として、1) XMLのような既に標準として確立している表現形式でミクロなエージェント間の通信プロトコルを定義する 2) それと並行して、マクロな経済システムに関しては国民経済 (SNA) のようなデータを XML 表現するプロトコルについて考察する、の 2 つのアプローチが有効との感触を得た。

ユーザモデルの構築

ユーザモデルを取得するには、情報サービスを受けるユーザの状況の把握や提供された情報に対するユーザの行動・要求を推測する必要がある。高度なサービスを実現するために、まず、WWW上の検索システムの現状と、実際のユーザに何が必要とされているかを明らかにするために、インターネット検索の利用動向に関するアンケート調査を実施し、約1000人から回答を得た。当該アンケートの自由記述回答を自動解析し、選択式回答との相関を抽出するためのテキストマイニング手法について検討し、実現の見通しを得た。

現在のインターネットにおける情報検索では、単にキーワードによるマッチングとキーワードの出現頻度しか使っていないので、個人のニーズに合った情報が即座に得られるとは限らない。そこで、複数の情報リソースからの情報を取捨選択して提供できる情報メタデータを開発しつつある。平成12年度は、情報メタデータを構築するためのシステム全体の枠組みを検討し、ユーザの状況を推定するために必要となる、挙動を計測するシステムについて検討した。

ユーザモデルの理論的基礎

統計的な推論と記号的推論を組み合わせた深いユーザモデリングを可能にする技術的基盤として、ベイジアンネットワークに注目し、調査を行なった。

ベイジアンネットワークは、有限次元の結合分布をある独立性の仮定の下に有向グラフとして表現したものであり、観測結果 e が与えられたときに事象 f が生起する確率 $P(f|e)$ を計算することができる。統計学のグラフィカルモデリングと親類関係にあるが、相関関係よりも因果関係の矢印によるグラフ的表現を重視し、また計算量などコンピュータサイエンスから見た計算の側面に焦点を当てている所が異なる。

計算の側面を見ると、一般の $P(f|e)$ の計算はNP困難であり、singly connect networks というサブクラスが多項式時間で計算出来る事がわかっている。現状では、もとのネットワークをjoin treeという新しい木構造に変換して、メッセージパッシングで計算する方法が主流のようである。また、ベイジアンネットワークの表力は、命題論理と同等であり、親子関係など表わすことができないので、

たとえば関係データベースにおける不確定性を確率によって表現するのは困難である。

ベイジアンネットワークを述語論理ベースのものに拡張するため、論理型言語に統計的要素を加味した記号的統計モデリング言語PRISMを開発中である。PRISMの意味論的基盤はおよそ整ったが、実働化の方は、現在徐々に進行中である。一方、自然言語コーパスなど実データを使ったモデリング実験の準備も行なっている。

また、事前には明示的に記述しにくいユーザの深層的な嗜好性を表層に表われるデータから推定する方法について検討した。これはグラフ構造を持つ確率モデルを頻度データから学習するものであるが、とくに隠れ変数を導入し、これを推定するためにEMアルゴリズムの適用可能性を検討した。こうしたグラフ構造を持つ確率モデルをデータから効率的に構築するために、ベイジアンネットワークの構築システムの開発と、ユーザモデリングへの適用について検討を進めた。

位置情報の取得と利用

適切なユーザモデルを構築するには、ユーザが置かれた状況に関する情報が必要であり、その中でも物理的な位置は特に重要と考えられる。屋外での位置情報の取得にはGPSなどが実用化されているが、屋内用の位置取得方式は、ジャイロ方式、超音波方式など様々な手法が試みられているが、まだ実用化されていない。

われわれは、その手法のひとつとして、室内レーザレーダシステムを開発してきた。これはレーザ光で対象の位置の同定と通信の両方の役割を実現するシステムである。情報のダウンロードはレーザ光の変調で行ない、アップロードは対象側鏡の反射率を変化させることによって行なう。すでに位置同定実験には成功しており、通信が今後の課題である。平成12年度には、情報アップロード用の光反射通信制御ターゲットの評価システムを構築し、反射率変調による受動空間光通信のための液晶光シャッタの動作特性を明らかにした。液晶光シャッタには多孔質ポリマに液晶を浸透させた光散乱制御方式の素子を採用し、その3dB変調周波数帯域はキャリア周波数500Hz、1kHz、10kHzにおいていずれも約30Hzであった。

また、本年度に購入した大型ディスプレイを利用して、街角掲示板の可能性を検討した。これは上記の室内レーダを含む近距離通信に基づき、携帯デバイスを持ったユーザが近付くとそれを関知し、そのユーザに対し適切な情報提示を行なうシステムである。

3. 主な研究成果の発表（論文発表）

- 伊藤、山本、山本、吉田、阿部、大宮司、原、真野、松田、西山、下田、中村、伊藤：室内低消費電力無線通信ターゲット用レーザレーダシステム、電総研彙報、

Vol.64, No.3, pp.79-85, 2000.

- H.Itoh, S.Yamamoto, Y.Yamamoto, M.Iwata, Nishiyama, T.Matsuda, M.Watanabe, 1.06um Optically addressable Liquid Crystal Spatial Light Modulator for an "HV target" of the Indoor Laser Radar System. Proc. of Optics in Computing SPIE, Vol.4089, pp.736-742, 2000.
- Sato,T.: Parameterized Logic Programs where Computing Meets Learning. Proc. of FLOPS2001, LNCS 2024 Springer, pp.40-60, 2001