

エージェントモデルによるマルチモーダル検索インタフェースの提案

瀬川 修 †, 田村 直良 ‡

†中部電力株式会社 電力技術研究所, ‡横浜国立大学 工学部 電子情報工学科

1 はじめに

情報検索のための知的インタフェースに求められる要件として、ユーザの検索要求を的確に把握し、検索タスクを代行できる“仲介者(サーチャー)”の存在がある。特にユーザが検索手段・方法について知識を持たない場合は“仲介者”の役割が重要である。本研究ではシステム化された情報検索タスクにおいて、このような“仲介者”をソフト的な知的エージェントとして実現することを目的としている。

ユーザ/システムの言語指向による対話手段としては、コマンドによるオペレーションが一般的であるが、計算機アーキテクチャに依存しない人間にとって普遍な自然言語による対話の試みもなされている。自然言語入力がコマンドに対して不利な点として、曖昧性、冗長性に起因して簡潔な論理表現が困難であること、機械表現とのミスマッチなどがあり、計算機との対話手段として用いるためには“仲介者”がこれらの意味的ギャップを解消しなければならない。すなわち、自然言語入力は“仲介者”が十分に知的であるという前提のもとに初めて成り立つ(実用に耐えうる)ものである。

一方、言語指向に加えてポインティングデバイス等による直接操作をうまく取り込んだマルチモーダルインタフェースに関する研究が盛んに行われている。これは言語入力とポインティングデバイスなどの非言語入力の併用によって、単一モードのウィークポイントを相互に補完するというコンセプトであり、自然言語インタフェースの有効な実装方法の一つであるといえる。

このような背景のもとに、本研究では検索タスクとして2次元視覚情報(街路図面)を用いた設備データ検索を設定し、言語指向と直接操作による検索条件指定が併用できることから、マルチモーダルインタフェースが有効に機能できる局面を設定した。そして、ユーザは知的な“仲介者”との対話の過程(マルチモーダルインタラクション)によって目的を達成するという検索スタイルのモデルを導入する。

本稿は、このような知的検索インタフェースにおけるユーザ/システムの動的なインタラクションを実現するために、マルチエージェントによるプランニングの枠組みを提案するものである。

2 マルチモーダル検索インタフェースの概要

2.1 検索タスクについて

データ検索のアプリケーション例として電力会社で用いられる地図情報システム、すなわち、電子化された街路図面をベースとしたビジュアルな入出力が可能な検索システムを対象としている。設備データ(電柱、変圧器、管路 etc.)は地図座標上のオブジェクトとして表示される。検索タスクの特徴として、視覚情報を用いることから言語入力とポインティングデバイスを用いた直接操作による検索条件指定が併用できることが挙げられる。図1に地図情報システムにマルチモーダル検索インタフェースを適用したイメージを示す。なお対象ユーザ(電力会社社員)としては以下の設定を行っている。

- 検索対象の設備に関する業務知識があることを前提とする。設備名や検索条件を(ある程度断片的でもよいから)自然言語で表現できる。
- データ構造、DBMS や検索言語に関する知識は前提としない。

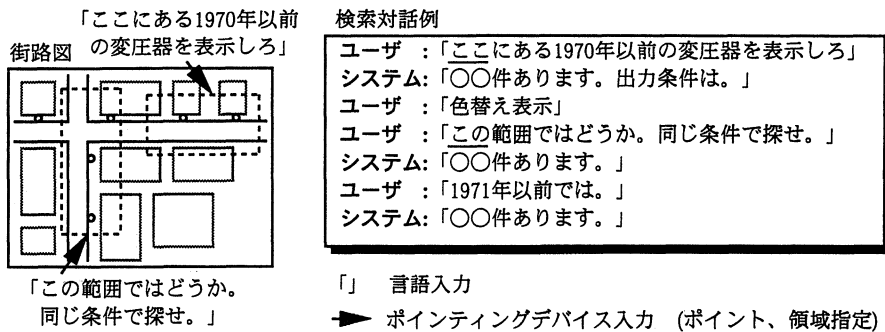


図 1: インタフェースの適用イメージ

2.2 検索スタイル

我々はユーザから見て“知的”な対話型インタフェースを介した情報検索のスタイルは、従来のシステム主導のメニュー型インタフェースによるものと比較して大きく異なると考えている。

自然言語入力は、その曖昧性に起因した論理表現の困難さから、システムに対し検索要求を1文、ないしは1度の検索ターンで条件として過不足なく与えるのは大変困難である。また言葉で表現するより直接操作のほうが効率がよい場合もある。すなわち、検索要求はユーザの1度の入力で決定されてしまうのではなく、比較的単純な(断片的な)条件の積み重ねをシステムとのマルチモーダル・インタラクションにより進めていき、ゴール(検索タスク)を達成するといった検索スタイルのモデルを用いる。

2.3 動的なプランニングの必要性

検索スタイルの変化によって、プランニングの枠組自体の見直しが必要である。検索対話におけるマルチモーダル・インタラクションの進展によって、システム側から見た外部環境は動的に変化し続けることが予想され、これに即応できる動的なプランニングの枠組みが求められる。

次節ではこの解決策としてマルチエージェントモデルによるプランニングの枠組を提案する。

3 エージェントモデルによるマルチモーダル解釈系

3.1 エージェントモデルの導入

エージェントには機能的に様々な分類がなされているが[1]、以下で提案を行うのはマルチエージェントによる処理系のモデルである。ユーザ側の視点からすれば、情報検索のための“仲介者”(インタフェースエージェント)のモデルであるが、仲介者エージェントは内部的には部分的な問題を分担する細分化されたエージェントから構成される。

3.2 モデルの概要

3.2.1 エージェントの機能的要素

エージェントを構成する機能的要素として、観測、環境モデル、プランニング、行為の4つがある[2]。エージェントは活動する計算機内部の世界(環境と呼ぶ)を観測し、自らの行動を決定するためのプランニングを行い、プランニングに基づいた行為を行う。

1. 観測 環境からの情報の獲得。

2. 環境モデル エージェントが把握する環境内の情報記述。
3. プランニング 自らの行為による環境の変化を予測し、適切な行為を決定する。
4. 行為 プランニングに基づいた環境内の情報記述の操作、エージェント間のインタラクション。

3.2.2 系の構成

系はエージェントと環境より構成される。環境モデルは黒板モデルに基づいており、エージェントは黒板による結果共有によって駆動される。環境は各エージェントが参照/書込ができる黒板と、処理系が共有して用いる知識源であるリソース(辞書、文法など)から構成される。図2(a)にモデルの全体構成を示す。

3.2.3 エージェントの構成単位と役割分担

(1) エージェントの設定 エージェントの知的機能の粒度として、高機能な熟考型と低機能な反応型の両極端な2つのタイプが考えられる[1]。熟考型は知識ベースと問題解決のための推論機構を有するタイプで、反応型は知識ベースや高度な推論機構を内部に持たないタイプである。

分担する処理によって両者のハイブリッド化も考えられるが、マルチモーダル対話の状況下での動的なプランニングの実現性から、本モデルでは後者の内部状態を持たない反応型エージェントによる系の構成について検討を行なった。

(2) エージェントの基本動作 図2(b)に反応型エージェントの仕組みを示す。個々のエージェントは環境にアクセスする機能(Sensor, Actuator)と、他のエージェントと交信する機能(Receiver, Transmitter)を持つ。エージェントはSensorとReceiverによる観測を行ない、内部に持つプランによりプランニングを行なう。そしてプランニングに基づきActuatorとTransmitterによる行為を行なう。

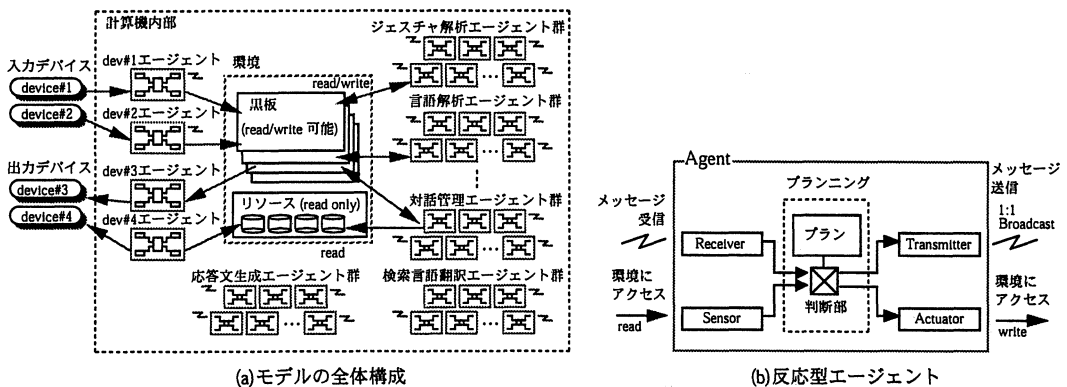


図 2: エージェントモデルによるマルチモーダル解釈系

3.3 プランニングの初期検討

マルチエージェントの一連の動作は、入力デバイスから計算機内部の環境に対して送られるイベントが発端となって生じられる。デバイスに固有の処理を分担するエージェントはイベント発生を検知すると、他のエージェント群に対し部分問題の処理を依頼するメッセージを送信する。これがトリガーとなってマルチエージェント間の相互作用による問題解決が行なわれる。

ここでは、エージェントのプランニングの初期検討として、ポインタ利用を伴う入力文解析系における指示詞の参照先同定の例を説明する。

一般に、話者が発話するコ、ソ、ア系指示詞のうち、コ系指示詞は話者の領域、ソ系指示詞は聞き手の領域、ア系指示詞¹は話者、聞き手両者から離れた領域を指すと説明されている [3]。検索結果はユーザが受理した段階でユーザの領域に同化されると仮定し、コ系指示詞でも参照可能と考える。

ユーザの発話内の指示詞の参照先は、地図データ中のポインタによって指定された領域または対象物、ユーザの発話内のある対象、システムの応答による検索結果またはシステムの応答内のある対象に限ることができる。ここで、発話時の話題候補の優先順序という概念を導入する。

発話時のポインタ指定 > それ以外の発話内の対象 > 直前の検索結果 > 以前のポインタ指定

指示詞の参照先同定の方略は、以下である。

1. コ系指示詞（コ系連体詞、コ系指示代名詞）

指示詞の参照先は、上記優先順序にしたがって同定される。

2. ソ系指示詞（ソ系連体詞、ソ系指示代名詞）

指示詞の参照先は、直前の検索結果以降で上記順序にしたがって同定される。

エージェントのプランニングによる実現は、例えば図3のようになる。

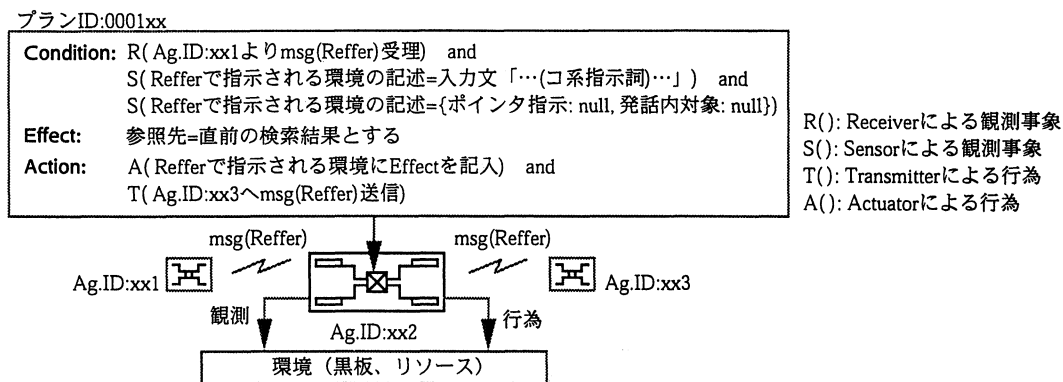


図3: 指示詞参照先同定のプランニングの例

上記プランのうち、**Condition:**はプラン適用条件である観測事象を、**Effect:**はプラン実行後の効果を、**Action:**は行為を記述する。

4 おわりに

本稿では情報検索タスクにおけるマルチモーダル入力の統合処理において、動的な環境変化に対する即応性を実現するための処理系の枠組みを提案した。各入力モダリティに固有の処理および、他の部分問題の処理をマルチエージェントの並列協調動作によりモデル化し、入力文中の参照先同定の処理についてプランニングの初期検討を行なった。

参考文献

1. 木下哲男: エージェントテクノロジーの応用とその課題, 信学技報 PFU95-149~156, Vol.95, No.364, pp.41-47
2. 山田誠二: エージェントのプランニング, 人工知能学会誌, Vol.10, No.5, pp.677-682
3. 石綿敏雄: 日本語表現の特色, 朝倉書店日本語新講座、運用 II, p.83, 朝倉書店 (1983)

¹ 今回の応用では、ユーザとシステムが対話をしながら対象を絞り込んでいく場面を想定しているため、ア系の指示詞の使用は検討対象外とする。