

意味、画像をとり扱える データベースシステム

打浪清一*

データベースシステム

計算機が大型化、高速化され、それを利用するソフトウェア技術の進歩と共に、事務計算、科学技術計算機械にすぎなかった計算機が、より高度な使用に耐えるようになり、データベースシステムなど知的作業への応用が盛んになってきた。

データベースシステムとは、情報基地ともいふべきもので、種々のデータ、情報を有機的に関連づけて統一的に蓄積、管理し、多角的な観点からの検索や、データの維持更新を行うと共に、人間の知的相談相手となるものである。蓄積されたデータ群をデータベース、これを管理するソフトウェアをデータベース管理システムこれらを統合化したものをデータベースシステムと呼ぶ。

初期の段階では、データベースの利用も、事務計算や、在庫管理等が主であったが、少し前の文部省の特定研究「学術情報の組織化」や最近開始された通産省の「第5プロジェクト」等により、その適用範囲が学術用データ、画像、音声、知識へと拡張されつつある。著者は意味的情報、画像情報を扱うデータベースシステムを研究、開発してきた。以下では、この2種のデータベースシステムについてのべる。

位相情報空間型データベースシステム

既存のデータベース管理システムにおいては、意味的な遠近関係、包含関係を忠実に記述することができない。何故ならデータベースに蓄えるデータ構造のモデルとして、木構造、網構造、表構造を採用しているものが殆どで、これらの構造では、連続的な遠近関係や包含関

係を忠実に表現できないからである。

そこで情報を記述するモデルとして位相情報空間モデルという、意味地図を基本とするモデルを提案し、その数学的性質の解明と、そのモデルに基づくデータベース管理システムの設計と試作を行ってきた。この意味地図は位相情報空間生成文法により定義される多次元位相空間である。これは先ず部分意味地図間の関係がマクロに定義され、次いで各部分意味地図を構成する軸の性質（数値で表現される軸か、順序だけが意味をもつ軸か、循環性をもつか否か等）が定義され、最後にその各軸に目盛あるいは属性値が入れられる。

このようにして知識の入るわく組が作られると、個々のデータはこの空間中の点あるいは部分空間を占めることになる。この空間を多次元の立方体（ファセット）に分割し、実際のデータはその該当するファセット内に蓄えられる。検索においては、さがしたい事物の各属性値を指定することにより、空間内の部分空間が定まり、その部分空間内のファセット内のデータをさがすことにより行なわれる。これは頁に分割された地図を読むことに相当する。またある一つの概念から次々と連想により関連概念を求める操作は、この空間内の、その概念が位置するファセットから始めて近くにあるデータを次々と求めてゆくことにより行なわれる。また意味的には近くないが何等かの関連がある語は、地図における航空路に類似した関連パスをたどることによりたどってゆく。結局ある概念の意味を知るということは、この空間内のその点の位置と他の概念の位置との関係、空間内の相対的位置等を知ったり、その概念を構成する各属性の値を知ることによりなされる。

知識には、外延的知識と、内包的知識があり、前者は構成要素を列挙する方法、後者は構

*打浪清一 (Seiichi UCHINAMI), 大阪大学, 工学部, 通信工学科, 手塚研究室, 助手, 工学博士

成要素群が共通にもつ属性を列挙する方法で概念を説明したものである。これらの解説は、空間内のある概念領域に含まれる概念要素名の列挙、あるいは、概念領域の各軸の値の指示により行なわれる。

また知識は、宣言的知識（定義やいいかえによるもの）と手続的知識（アルゴリズムやプログラムによるもの）があり、その何れも概念間の類似性、包含性を反映する形で空間内に配置される。宣言的知識の場合は、該当領域の各軸の値を読む方法でデータが読み出されるが、手続的知識の場合はその概念にアクセスした後そのアルゴリズムが駆動される。

本データベースシステムは、この知識のユニバースが表現されている辞書としての全情報空間と、ここから得た知識と手持ちのデータから思考、創造活動を行う場としての概念スキーマフリー・データベースシステムからなる。後者はメモを次々と蓄え、これを整理分析し、仮説を作り、それを検定、新理論やシステムの設計を行うことに使われる。意味地図に関しては、MT 化された三省堂明解国語辞典などをもとに作成すべく作業中である。

画像データベースシステム

LANDSAT（資源探査衛星）による写真や医学における各種 X 線写真、美術資料など各種の画像、図形が数多く収集され、これらの有効な活用が期待されているが、これらの蓄積されたデータベースにおいて、画像や図形の内容から検索できるシステムはまだ実用化されていない。これら画像、図面につけられたメモをもとに検索している現状である。画像内容には(1)被写体、(2)図形（被写体を意識せず同質部分を一つの単位としての形状的認識）、被写体(3)これらの位置関係、(4)複数枚の画像を解析することにより得られる事実、がある。これらの画像内容からの検索を可能にするために、画像生成文法を抄録に利用する新しい方法を提案した。本方法においては、画像中の被写体および図形を認識し、これを構図解析する。そして基本単位図形とそれらの位置関係を表す生成規則を抄録とする。

検索時には、会話型でこれと同様な操作を行い、検索したい画像を構図解析し、基本図形やその位置関係をシステムが求め、抄録を参照することにより、意味的に関連のある画像を検索する。

ケーススタディとして植物図鑑を選び、実験システムを構成しその有効性を確認した。北隆館の原色植物検索図鑑を対象に、植物採集してきた植物の同定を行う場合を想定してシステムを作成した。植物画像、図形を生成する文法を定義したが、本図鑑掲載の575種の植物は、これら生成規則により172種に分類され、約400種の生成規則で記述できた。本システムを用いて植物の検索実験を行ったところ大体7～8回の特徴入力で植物の同定が可能であった。この検索画面例を写真1に示す。

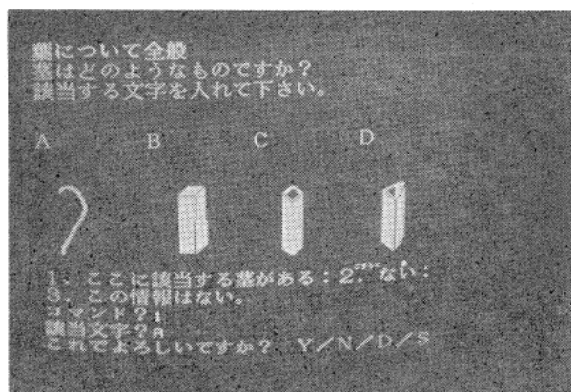


写真 1

画像には無限の変化がありこれをうまく抄録するのはむづかしいが、本方法ではこれを典型的なパターンとその接続情報で抄録したことになり、圧縮率の高い抄録となっている。

むすび

画像データベースシステムでは、個々の個体の部分の特徴による抄録や検索法であったが、これら個体間の類似性を反映させて位相空間内に配置すれば、位相情報空間型データベースシステムで取り扱える。換言すれば、個々の個体内の特徴による抄録、検索と個体間の関係の抄録、検索とは互に可換である。それ故意味地図を種々の超平面に射影してブラウン管上に表示しながら検索してゆけるよう拡張を考えている。

最後に日々御指導戴く手塚教授、本稿を草する機会を与えられた滑川教授に感謝致します。