

2R-3

履歴データベースのための  
視覚的操作言語の設計と開発

田中克己 大本英徹

(神戸大学・工学部)

1.まえがき

履歴データベースの操作言語には、関係データモデルに基づくTQuel<sup>[1]</sup>やHTQuel<sup>[2]</sup>などがあるが、これらは、(a)レコード形式や時間属性や属性毎の時刻印を陽に意識して質問を作成する必要があり、(b)意味データモデルで重要な汎化(is-a)関連がサポートされていない等の問題がある。また、自然言語による時間に関する質問も質問が複雑化することにつれて必ずしも最適な表現方法とは考えられない。一方、関係データベースの視覚的操作言語として、例によるfill in the blank方式のQBEがあるが、ここでは履歴そのものをチャート形式で表現し、これをfill in the blank方式で入力することで質問を作成する操作言語HistoryChartを開発した。本稿では、このHistoryChartの概要と特徴について概要を報告する。この言語は(a)言葉やコマンド形式では表現困難な履歴検索質問の視覚的生成、(b)検索結果の視覚化による直観的把握が可能、(c)汎化階層を用いた一般化情報の検索・表示等の特徴を有している。

2.履歴検索質問の視覚化

例えば、我々がAとBの両方のデータが同時に存在している区間の履歴を考える場合、図1のようなチャート表現をししばしば用いる。

これは図1のように時間区間を3つに分け(I)Aだけが出現している区間、(II)A,B両方出現している区間、(III)Bだけが出現する区間という情報のみを表している考えるのが自然であり、それぞれの区間の長さやデータの出現・消失が最初なのか最後なのかといったことには特に意識しない限り注意を払わないと考える。すなわちI,II,IIIと3つの部分に分割することが最も本質的であり、その他の情報は付随的であると考え。今回、開発した視覚的履歴データベース操作言語

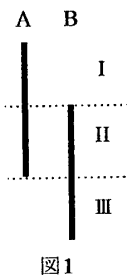


図1

HistoryChartでは、この人間の自然な発想のインターフェイス上における実現を目指している。

3.HistoryChart

HistoryChartは基本的に表形式をイメージしており、ディスプレイ画面上の空表に履歴を表すパターンを埋めることにより質問が生成される。例えばAとBが同時に出現している区間のCの履歴を求める質問は、図2のように表現される。これにより、適合する時間区間における項目Cのデータが表示される。

A	B	目標(C)
	×	
	×	
		?
		?
		?
×		
×		

図2

表の1行目の各欄には、条件項目が入力されるが、これは、関係データベースの属性名のみが対応するのではないことに注意を要する。本言語

では、すべての属性名とその値は汎化(is-a)階層木の形で構造化されており、その木構造の任意の節点を条件項目として入力できる。さらに、これらの値をandやorを用いて結合し(AandB)or(CandDandE)orFなどという複雑な項目名も用いることも可能である。また、パターンは時間長が不定な時間区間を表す5種類の記号の組合せから構成される。各記号は次のように条件項目の値の存在・非存在を表す。

- ① | :データが連続存在する区間、
- ② × :データが連続非存在した区間、
- ③ T :時間系列の最初にデータが存在した区間、
- ④ ⊥ :時間系列の最後にデータが存在した区間、
- ⑤ ? :存在・非存在に特に着目しない区間(これを用いてデータの存在・非存在に注目しない区間(don't care)を検索条件として用いることができる。)

目標(...)を項目名とする列は出力結果となる履歴パターンを記述する。これらの記号は、その列の項目名に対応するデータの時間区間における状

態を表現するが、その時間的長さは規定されていない。ただし、同一行に並ぶそれぞれの記号は同一時間区間の状態を示す。ここで時間長さ不定であるから、連続する2行を比較してまったく同一ならば1行にまとめても良いことになる。本システムは圧縮された検索条件をPrologの質問に変換しデータベースを検索する。

条件項目の指定はメニュー形式で汎化階層木をたどり節点選択するだけで行える。パターン入力にはマウスを用いて容易な操作を可能にしている。Aの最初の存在時点以後であり、その存在・非存在にかかわらずとにかく、Aの最後の存在時点までのBの履歴を求める質問例を図3に示す。

A	目標(B)
T	?
?	?
?	?
?	?
⊥	?

図3

4.汎化階層を用いた情報の抽象的扱い

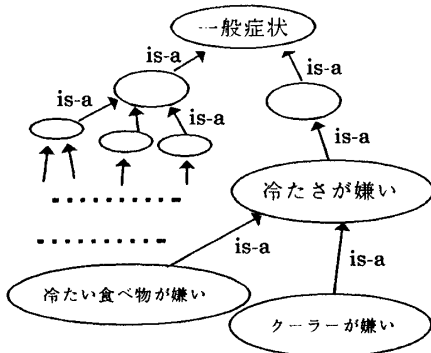


図4

我々の意識する「概念」は常に詳細ではない。例えば、図4のような汎化階層を例にとると本システムで、汎化階層木中の節点Xを質問の条件項目中にいれる場合の意味は、Xより下位のすべての節点集合を意味するのではなく、その上位概念がXであるような少なくとも1つのデータ(項目)を意味している。例えば、図4の例では「冷たさが嫌い」という項目を入力することは「クーラーが嫌い」や「冷たい食べ物が嫌い」といった概念を全て同時に有するデータを検索することではなく、それらのうちのどれかを有するデータを検索することを意味する。

また、この言語は、データをチャート化して出力する機能を持ち情報の視覚化及び、その傾向の直観的把握を可能としている。具体的にはメニューでチャート化出力機能選択後、通常と同様の項目指定を行うと、それに適合するデータの存在・非存在

が先の|,×記号を用いて時刻印(日付)と並行して出力される。

5.システムの構成

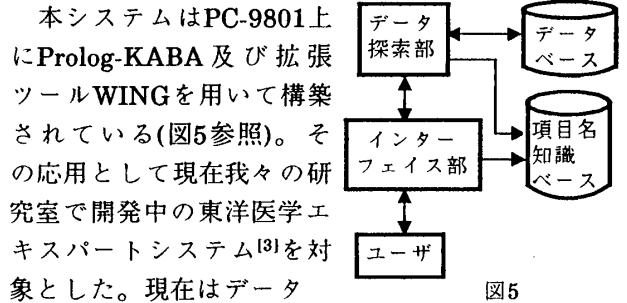


図5

本システムはPC-9801上にProlog-KABA及び拡張ツールWINGを用いて構築されている(図5参照)。その応用として現在我々の研究室で開発中の東洋医学エキスパートシステム<sup>[3]</sup>を対象とした。現在はデータベースはProlog事実節として記述され時系列に並べられて履歴データベースを構成している。これをプログラムとしてシステム内に読み込んでパターンマッチング及びバックトラックを用いたprimitiveな方法を用いて検索を行っている。また項目汎化階層知識ベースはインターフェイスにおけるメニュー生成及びデータ探索に際し必要に応じて参照される。

6.あとがき

以上、我々の開発した履歴データベースの視覚的操作言語について報告した。今後に残された課題として、(1)データベースの適当な構造化、探索手法の改良などによる検索速度の向上(2)数値情報、時間スケージングの取扱い(3)日本語による検索質問確認をふくめた対話的システム操作機能(4)HistoryChartの表現能力の理論的検討等が考えられる。

尚、本研究は一部、文部省科学研究費重点領域研究(2)#63633518による。

最後に、御討論いただいた本学北村研究室の諸氏に感謝致します。

参考文献

[1] R.Snodgrass,ACM PODS,March 1984  
 [2] S.K.Gadia et al.,ACM PODS,March 1985  
 [3]Tanaka et al.,COMPSAC,Oct. 1987