

## 知識工学の医療への応用

東京大学 医学部 助教授  
病院情報処理部

開原 成 允

東京都老人総合研究所

小山 照 夫

## 1 はじめに

医師が病気を診断する過程は、これまで医学以外の領域の人々からも興味をもって研究されてきた。その理由は、この研究が「医師の診断を助ける機械」を作りたいという実用的な目的をもつと同時に、ヒトの思考過程を研究するモデルとして適当なものであったためと考えられる。

この「実用性」ということと、「ヒトの思考過程の解析」という2つの側面は現在も発展しつつある。「実用性」の観点からすると、情報科学技術の発展によって、コンピュータ端末が医師の手にとどく所におかれるようになり、このようなシステムが実用になるための条件が整ってきた。また、「ヒトの思考過程の解析」は、「認知科学」という1つの領域となるまでに発展した。

本稿では、機械による診断システムが果して、実用になり得るかという問題に焦点をあてつつ、最近の「知識工学」が、この問題の解決にどのように貢献しつつあるかを解説しあわせて、われわれの最近の研究について紹介する。

## 2 医師診断過程研究の2つの考え方

歴史的に眺めると、医師の診断過程の研究には考え方の上で2つの流れがあった。この2つの考え方は、決して対立したものではなく、相補的なものであるが、その考え方の違いは理解しておく必要がある(図1)。

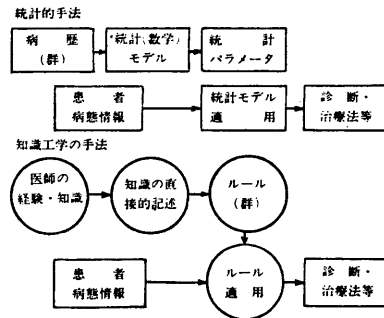


図1 数学モデルと知識工学モデルの考え方

第1の考え方は、医師の診断過程は、データを処理して1つの結論を導く数学的なモデルでおきかえ得るとする考え方である。この立場にたてば、客観的に集められた多数の患者の症状や検査のデータに、それぞれの患者の確定した診断を対比させ、その間をつなぐ数学的なモデルを作れば、それによって診断のプロセスが解明されるとする。

極端ないい方をすれば、医師は必要でなく、医師は、「時によって異なった判断を下すこともあり、主観的な、むしろ好ましくない」存在ということになる。必要なのは、客観的なデータであり、そのデータが多数集まり、同様に客観的に確定した診断と対比させることさえできれば、その間をつなぐ数学的モデルは複雑にはなるかもしれないが、必ず存在し、このようにして作りあげたモデルは最も秀れた診断システムとなる。

このような考え方で症状と診断の間をつなぐ数学モデルとして研究されたのは、ベイズの定理や、多変量解析、中でも判別分析であった。

これに対し、第2の考え方は、医師の頭の中にこそ最も秀れた診断のプロセスが存在すると考え、できる限りこれを抽出してそのままの形で利用しようとするものである。

この考え方にたてば、研究上必要なのは、「EXPERT」と呼ばれる専門医であり、できる限り、その経験を取り出し利用するという研究方法をとることになる。例え、診断に個人差があったとしても、Expertの診断のプロセスがそのようであれば、それを真似ることが最善であると考えらる。

ここで使われた方法論は、初期の頃は枝分れ論理であり、最近では知識工学の手法を使った研究に受けつがれている。

### 3 2つの考え方の特徴

すでに述べた如く、これらの2つの考え方は決して対立的なものではない。初期の頃から、1部枝分れ、1部に数学モデルを使ったような診断システムは数多くあった。また、前節ではやや極端な表現を使ったが、数学モデルの研究者が、医師は必要ないと考えているわけでもない。

しかし、それぞれの考え方の特徴はいくつかある。

数学モデルのよさは、医師を介さないために、時に医師が全く考えもしなかった新しい論理や、表面に現れた症状の背後にある事実が発見されることである。このようにして見出された事実は、過去にいくつかみられている。このことは、この方法が疾病のメカニズムを研究する上では大変有効な手段であることを示している。

しかし、一方で、実用の診断システムを作ろうとすると、現在の数学モデルではいくつかの限界がある。複雑な医師の理論が、必ずしも既存の数学モデルにあてはまるとも考えられない。既存のモデルは、何らかの仮定をおいている。例えば、ベイズの定理を使った時は症状間の独立性が仮定されているし、判別関数の利用では、線型性の仮定をおいている。

勿論、現在の数学モデルは、これらの仮定を取り払って複雑にすることも可能であるが、その時にはデータが著しく多く必要になったり、計算法が著しく面倒になったりして、実用性を損なり場合が多い。

また、データに欠落データがあると使いにくいといった問題や、実際の診断は逐次的に行なわれるにもかかわらず、こうしたモデルでは、データを最初に全て与える必要がある点なども、その手法のもつ限界と考えられる。

しかし、場面を限れば、こうしたシステムも実用になることもあり、東大病院産婦人科で行なわれている胎児骨盤不適合の診断はそのよい例と考えられる。

$$Z = -1.38 X_1 - 0.29 X_2 - 0.17 X_3 + 0.92 X_4 \\ \pm 0.17 X_5 + 0.06 X_6 - 0.06 X_7 + 14.48$$

- X<sub>1</sub> 骨盤入口部前後径
- X<sub>2</sub> 瀾部前後径
- X<sub>3</sub> 狭部前後径
- X<sub>4</sub> 児頭横径
- X<sub>5</sub> 子宮底長
- X<sub>6</sub> 妊婦の年齢
- X<sub>7</sub> 妊婦の身長

図2 判別関数による胎児骨盤不適合の診断  
(東大産婦人科箕浦による)

これに対し、第2の医師の経験を取り入れる方法は、手法さえ適当なものがあれば、ヒトに似た柔軟なシステムを作る可能性がある。しかし、極端に言えば、医師個人個人は少しずつ異なった診断体系をもっているから、「ある医師」の診断システムは作り得ても、完全に一般化したものは存在し得ないともいえる。

また、医師を、いわば模倣するのであるから、新しい発見は期待できない。

しかし、実用的な診断システムを作るという観点からすると、この第2の考え方にたつた方がより自然な、また解り易いシステムができる。

このため、古くから医師の経験に基づくシステムを作る試みは多くあった。しかし、計算機処理の手法としては、

「枝分れ」論理しか用いることができないため、あまり複雑なシステムを作ることはできなかった。

最近の「知識工学」は、正にこの壁を破る方法を提供したと考えられる。知識工学の手法を用いて、最近では医師の経験に基づく診断システムが多く作られるようになった。以下に、こうした「知識工学」の手法を用いた診断システムを眺めてみよう。

#### 4 医師の診断の過程はどのようなものか

知識工学を利用した医療コンサルテーションシステムを論ずる前に、医師の頭の中で、診断はどのように行なわれるかを考えてみよう。

これは、最近の言葉でいえば、認知科学 Cognitive science に属する問題である。この問題の研究は比較的早く、1960年代から Kleinmuntz ら心理学者が研究した。その後、医学教育に関心をよせる人々が研究している。種々の説があるが、一般的には次のように考えてよいと思う。

すなわち、医師の診断方法は、数学モデルのように多くのデータが一度に与えられて、一挙に診断に到達するのではなく、もっと dynamic な過程である。すなわち、患者のデータも、最初から全てを使うのではなく、まずある部分からいくつかの仮説（診断）を考える。その上で、適当な方法で、データを追加しながら、この仮説を検証しようとする。もし、データと仮説（診断）の間に矛盾ができた時は、その仮説から、次の仮説に移って、同様の過程をくり返す。この時、どのような順序で、データを採取するかという問題や、仮説があわない時に、どのようにして次の仮説を選ぶかは、大変興味ある問題であるが、そのメカニズムは必ずしも明白となっていない。

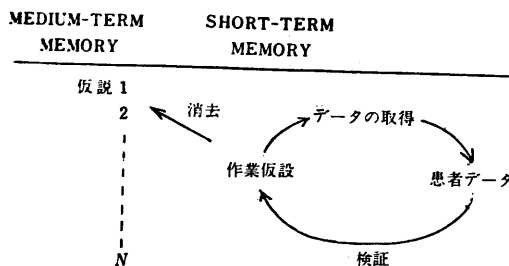


図3 医師の診断過程の模式図

医師は、このような過程をとるから、診断に際し、常にその筋道をたどることができる。仮にもし、診断を誤った時も、あとで、何故診断を誤ったか、その説明をすることもできる。このように「診断の過程が説明できる」ということが後にのべるように実用上は大変重要な問題となる。

#### 5 知識工学の手法を応用した診断システム

知識工学の手法を応用した診断システムは、前節のプロセスを参考にしつつ、医師の診断プロセスをコンピュータを用いて実現しようとする。

知識工学の手法といっても、この中には多くの要素を含むが、診断システムの中で利用される手法は、主として2つである。

すなわち、第1は、非数値的な推論の手法であり、第2は、知識の表現と利用法である。

第2の知識の表現方法から考えてみよう。知識工学の手法による診断システムでは、医師が持つ知識を利用するわけであるから、それができるだけ忠実に計算機で表現されると共に、それが利用し易い形になっていなければならない。このための手法には、数種類のものが提案されているが、最も代表的なものは、production rule と frame であろう。ここでは最もよく使われる production rule について解説すると、これは、知識を、「もし……ならば、……である」という要素に分解する。「もし……ならば」は前提部、「……である」は結論部と呼ばれ、それぞれは複合した複雑な

形もとり得るが、基本は上記のものに他ならない。

これはシステムによって異なるが、例えば甲状腺疾患の診断において「甲状腺が硬くはれていて、機能低下があれば橋本病の疑いがある」という知識は、EMYCINと呼ばれるシステムでは、LISPのS方式として、

```
(AND (STRUM HARD 0.6) (HYPO T 0.9)
 (CONCLUDE DIAG HASH 0.8))
```

またEXPERTというシステムでは、

```
F (HARD, 0.6) & H (HYPO, 0.9) →
H (HASH, 0.8)
```

とかく、

ここでHARDは甲状腺が硬いこと、HYPOは機能低下、HASHは橋本病を意味するとする。

この時、1つの問題は、この前提部と結論部は、確定的に(確率1)で結びついているわけではなく、通常は、ある不確かさをもって結びついている。従って、医学診断の知識表現には、上記の式に必ず、certainty factorと呼ばれる数値を附して表現する。0.6, 0.8のような数字はこのcertainty factorである。

このようにして表現された知識は、更にいくつも集まって、knowledge sourceと呼ばれる単位を作るようになっていものもあるし、そのknowledgeの適用されるlevelによって分類される場合もある。

このように詳細な点では異なっているが、基本的には、知識の単位を医師から採取して定式化していく点ではどのシステムも変りはない。

次は、これらのproduction ruleを用いて、どのように推論を進めていくかという問題である。

この方式にも、「診断を定める」という目標をまず設定し、必要なデータを探していくというbackward chaining; データから次第に仮説のcertainty factorを高めていく方法; ruleの集合であるknowledge sourceをまとめて、必要に応じ適応していくblackboard model等、いくつかの推論の方式がある。blackboard modelは、多くの方式を総合した最も柔軟な方法とも考えられる。これらの推論の方式に対応して、それぞれいくつかの診断システムが作られている。

例えば、backward chainingを利用したEMYCIN, dataから仮説を作っていくEXPERT, blackboard modelを利用したAGE, 知識表現にframeの考え方を入れたPIP, INTERNIST等である。

これらの中のいくつかは、知識の部分が独立しており、この知識を入れかえれば、どのような対象にも用い得る「汎用システム」となっている。

「汎用システム」の1つであるKulikowskiによって作られた「EXPERT」を用いて我々の作った甲状腺疾患診断のシステムを呈示してみよう。

このシステムは、東大第三内科、長滝重信講師(現長崎大教授)のグループの診断のプロセスを採用し、診断システムとしたもので、約30の甲状腺疾患を医師とほぼ同じ確率で診断する。

このシステムでは、医師の知識は約200のルールによって、システム中に定式化されている(図4)。このシステムを動かすと図5にみるように計算機から色々と質問をしてくる。これに答えると、計算機は次の質問をするが、その質問は最も早く診断に到達できるようなものを選んで質問する。答が解らない時は解らないと答えることもできる。これをくり返すと最後に計算機は診断を印刷するのである(図5)。

## 6 診断システムの医療上の意義

これまで述べたようなシステムは、いずれも医師の知識をかなり柔軟に計算機に入力することができ、システムの行動も医師とかなり似させることができる。

それでは、このようなシステムを作ることが実際の医療の上ではどのような意味をもつのだろうか?

ここで、少し問題を実際面に移して、現在の医師には何が必要なのかを考えてみることにする。

最近では医学の進歩の結果、医学知識は急速な勢いで増加しつつある。

これらの知識は、すでに1人の医師の記憶容量を越えてしまったと考えられるから、良心的な医師であればあるほど、何らかの補助記憶手段を用いて診療をする必要がある。

こうした要求に答える最も単純な形は、能率よい情報検索システムであろう。

これを少し進めると、単に事実を呈示するのみでなく、多少の判断を加えて呈示することもできる。

しかし、このように、ある判断を機械が呈示するシステムであっても、これは決して「自動」診断システムではない。日本では、過去にかかる領域が「計算機診断」とか、「自動診断」とか呼ばれた時代があったが、これは大きな誤解を招く原因となった。私はこのような言葉は使うべきではないと思っている。仮に診断名が示されたとしても、それは医師が判断を下す上での1つの参考データを提供したにすぎないのである。

今でも、医師は、診断困難な症例にであると、古い文献中から同様の症例を求めたり、また、「診断のコツ」を書いた本を読んだりする。仮に「診断のコツ」に書いてあることをそのまま用いて医師が診断をしたとしても、これは決して「自動診断」ではない。

診断システムは、このように、過去の類似の症例や専門医の「診断のコツ」を整理された形で、医師の前に呈示するにすぎない。その意味では、非常に高度な「情報検索」システムにすぎないとさえ考えられるのである。従って私はこのようなシステムは「医療コンサルテーションシステム」とよぶのが一番よいと思っている。

このように考えてくると、診断システムは、単に診断名を呈示するだけではないはずである。医師が必要としているのは、診断よりはむしろ周辺のさらに多くのデータである場合が多い。

このようなシステムを作るのもう1つの意義は、医師の経験的知識が正確な、しかも実証できる形で採取・定式化できることである。これまで、名医の知識は文章の形で発表されてきた。このような文章はある程度、実証はできるが、文章の形では1つ1つの症例にあてはめて、その妥当性をみるといったことはできない。しかし、これまでのべたような「知識工学」を利用して定式化した知識は、1例1例、そのデータを入れて確めることができる。また誤りがあれば、修正し改善することもできる。

これは、医学教育や、医学知識の正確な伝承の上で大変重要なことで、今後の広範囲な応用が期待される所である。医学の中での知識工学は、まだ応用されてから日が浅い。今後の知識工学の発展と共にその応用も更に拡大していくことと思われる。

このように考えてくると、診断システムは、単に診断名を呈示するだけではないはずである。医師が必要としているのは、周辺のさらに多くのデータであるかもしれない。

著者は、医療コンサルテーションシステムが持つべき機能としては、次の5つがあると考えている。

#### i 効率の良い推論機構

ここで推論の効率として考えているものには、既に得られたデータに基づいて結論を導く場合の迅速性に加えて、結論を導くために新規にデータを獲得する際のコスト的効率も含めて考えている。

#### ii 推論プロセスの理解しやすさ

ユーザがコンサルテーションシステムの出す結論を受け入れるにあたって、推論プロセスを理解することができれば、より容易に結論を受け入れることができる。

### iii 推論を根拠付ける資料等の参照機能

システムによる推論結果の妥当性を評価するためには、推論プロセスを根拠付ける資料（文献、実験データ等）の参照が必要となる。このような資料を蓄積し、ユーザの要求に応じて提示する機能を持つことが有効である。

### iv データ誤りチェック機能

特に医療データのように、大きな誤差を含む可能性のあるデータに基づいて推論を行なう場合、データ誤りチェック機能を組み込むことにより、推論の精度を向上させることができる。

### v 欠測値データが推論に及ぼす影響の評価機能

医療データのように多くの欠測値を含むデータに基づく推論では、欠測値が推論結果にどのような影響を及ぼすかを評価する機能が必要である。

## 7 新しい機能をもった医療コンサルテーションシステム

著者は、前節で述べたような多くの新しい機能をもった知識工学を利用したコンサルテーションシステムが必要と考え2つのアプローチによってこれを作成した。

第1のアプローチは、スタンフォード大学で開発された「AGE」によるもの、第2は、われわれ独自で開発したMECS-AIとよばれる汎用のToolによるものである。以下にこの概要を記す。

### 1) AGEによるシステム構成の試み

AGEはStanford大学で開発された、知識ベース型システム構成のための汎用Toolである。その特長として、Blackboard Modelの採用と、きめ細かな推論Strategy実現可能性が挙げられる。演者等は、このシステムを用いて先天性代謝疾患診断のためのコンサルテーションシステムの構成を試みた。AGEの推論Strategyは、知識ベースをいくつかのKnowledge-Sourceに分割し、各々のKnowledge-Sourceの適用結果として生成される、EventによりKnowledge-Sourceの適用順序を制御することによって構成される。演者等の試みでは、推論効率向上のために、推論プロセスをScreeningとそれに続く、各診断に関する詳細な検討に分割し、それぞれをKnowledge-Sourceの形にまとめる方法を採用した。

このような構成を取ることにより、Screeningの結果、有効性の低いと考えられる知識の適用が除外され、全体としての効率を向上させることができる。データチェックや、欠測値の影響検討機能も、それぞれの機能を果すKnowledge-Sourceを追加し、適当なタイミングで適用されるようにシステムを構成している。データチェック機能としては、明かに誤っていると考えられるデータの書きかえと、疑わしいデータの再度問合せを行ない、欠測値影響の検討では、各診断名について、欠測値がその診断名に対応する典型症例と一致した場合に診断の確からしさがどのように変化するかを検討している。最後にReference機能として、AGEにはこのような機能を実現する関数は用意されていないので、追加機能として、ファイルに収納された、Referenceデータをユーザの要求に応じて読み出してリストするLisp関数を追加してある。

なお、本研究は、国際データ通信システムICASを用いて日本からスタンフォード大学のコンピュータを利用して行なった。

### 2) MECS-AI

MECS-AIは、簡単な時間経過の概念を含む、Ruleに基づく推論システム構成のための汎用Toolとして演者等が開発を進めているシステムであり、演者等が以前に作成した、心臓病診断専用プログラムを汎用化したものである。このシステムでは、時間はいくつかのDiscreteな時点（これを"Time-chunk"と呼んでいる）に分割されると考えている。これはたとえば外来患者の来院時点等を想定することに対応する。MECS-AIでは、データは（Time-ID Context Attribute Value CF）の5つ組みと考えているが、これはMYCINに見られた4つ組データに"Time-ID"を加えた形になっている。このデータ構造に基づいて"Previous" "Past"等のキーワードにより、過去にさかのぼってのデータ参照を可能にしている。

M E C S - A I では、効率的で、きめ細かな推論プロセスを実現するために、A G E にも見られるような、知識 ( Rule の集合 ) をいくつかの部分集合にあらかじめ分割する方法を採用している。分割された各部分集合を " Knowledge-unit " と呼んでいるが、各 Knowledge-unit には、" Apply-condition " と呼ぶ S-expression が付加されており、この条件が満足される場合に限って、中に含まれる Rule が適用される。M E C S - A I は、推論の進行とともに 2 つのリストを生成するが、その 1 つは推論結果を表し、他の 1 つは、推論にあたっての Rule の適用状況を表しており、後に推論プロセスの説明のために利用される。

## 8 おわりに

以上、知識工学を用いた医療コンサルテーションシステムの考え方について解説し、最近の著者らの研究について述べた。

この真の応用はこれからの課題であると思われる。

図4

```

F(T4HX,T)&F(T3UH,T)&F(T3HX,T)->H(HYPER,0.95)
F(T4NX,T)&F(T3UN,T)&F(T3NX,T)->H(EU,0.95)
F(T4LX,T)&F(T3UL,T)&F(T3HX,F)->H(HYPO,0.95)
F(T3UH,F)&F(T3UN,F)&F(T3UL,F)->H(PROTH,-1.0)
F(T3UH,F)&F(T3UN,F)&F(T3UL,F)->H(PROTL,-1.0)
F(T4HX,T)&F(T3UX,0:23)->H(PROTH,0.9)
F(T4LX,T)&F(T3UX,30:*)->H(PROTL,0.9)
F(T4HX,T)&F(T3HX,F)&F(I131,20:*)->H(HYPEX,0.85)
F(T4NX,T)&F(T3HX,T)->H(HYPER,0.9)
F(T4NX,T)&F(T3LX,T)->H(HYPO,0.5)
F(T4NX,T)&F(T3LX,T)->H(EU,0.5)
F(T4LX,T)&F(T3HX,T)->H(HYPER,0.6)
F(T4LX,T)&F(T3NX,T)->H(HYPO,0.5)
[2:F(TREM,T),F(SKIN,T),F(TACH,T)]->H(HYPER,0.5)
[4:F(NERV,T),F(SWEA,T),F(HEAT,T),F(PALP,T),F(FATI,T),F(WLOSS,T)+
J->H(HYPER,0.5)
[4:F(WEAK,T),F(DRYS,T),F(CORS,T),F(LETH,T),F(SLSP,T),F(EDEY,T),+
F(SCOL,T),F(DSWE,T),F(CSKI,T),F(TTON,T),F(EDFA,T)]->H(HYPO,0.8)
F(MULN,T)&[1:F(ANTI,100:*),F(MICR,100:*)]->H(CHTI,0.7)
F(MULN,T)&[1:F(BSR,20:*),F(CRPP,T)]->H(CHTI,0.6)
*HH RULES
*IF SOME HORMONE TESTS ARE LACKING
[2:F(T4XX,F),F(T3XX,F),F(T3UX,F)]
*THEN
[1:F(T4HX,T),F(T3UH,T),F(T3HX,T)]->H(HYPER,0.85)
[1:F(T4NX,T),F(T3UN,T),F(T3NX,T)]->H(EU,0.95)
[1:F(T4LX,T),F(T3UL,T),F(T3HX,F)]->H(HYPO,0.85)
*END

```

1. AGE  
\*23
  
2. SEX:  
1) MALE  
2) FEMALE  
Choose one:  
\*2
  
3. REASON FOR VISIT:  
1) GOITER  
2) POSSIBILITY OF HYPERTHYROIDISM  
3) POSSIBILITY OF HYPOTHYROIDISM  
Checklist:  
\*1
  
4. PALPATION OF NECK:  
1) PAIN AT THE THYROID  
2) STRUMA  
Checklist:  
\*2
  
5. PALPATION OF STRUMA:  
1) DIFFUSE ENLARGEMENT  
2) HARD STRUMA  
3) SINGLE NODULE  
4) MULTIPLE NODULE  
Checklist:  
\*1
  
6. SYMPTOMS:  
1) NERVOUSNESS  
2) INCREASED SWEATING  
3) HYPERSENSITIVITY TO HEAT  
4) PALPITATION  
5) FATIGUE  
6) WEIGHTLOSS  
7) DYSPNEA  
8) INCREASED APPETITE  
Checklist:  
\*N
  
7. SIGNS:  
1) TACHYCARDIA  
2) SKIN CHANGE  
3) TREMOR  
Checklist:  
\*N
  
8. IS THERE EYE SIGN ?  
\*N
  
9. TOTAL T4(4.5-12.0)  
\*6.6

INTERPRETIVE ANALYSIS

-----  
Diagnostic Status:

0.99 CHRONIC THYROIDITIS  
0.95 EUTHYROID

-----  
Treatment Recommendations:

1.00 THYROID SCINTIGRAM IS RECOMMENDED  
1.00 I-131 UPTAKE IS RECOMMENDED.