Teaching Editorial

Expert Systems: Applications of Artificial Intelligence in Medicine

Jeffry A. Siegel, Kenneth Levin, Roy Rada* and Marvin C. Ziskin

Temple University Hospital, Philadelphia, Pennsylvania

With the advent of the computer age, the role of complex, computerized information structures in medicine has increased dramatically. Machines that use complex information in intelligent ways manifest "artifical intelligence." Artifical intelligence (AI) is a branch of computer science concerned with the study of representation and search. Expert systems are examples of applied AI and are known as knowledge-based or rule-based systems. An expert system allows representation and search (reasoning) within the domain of a specific knowledge area. An expert system is not a "thinking" machine in any conventional sense; it is more accurately thought of as a set of *if-then* rules that encode a specific, discrete area of expertise. This teaching editorial is intended to be an introduction to AI in medicine (AIM) as applied to expert systems. The topics covered include: 1) expert system fundamentals, including knowledge representation and control strategy; 2) the language of AI; 3) the history and application of AIM; and 4) methods for evaluation of AIM. A more in-depth treatment of the subject is provided in the literature (*1–3*).

Artifical intelligence in medicine deals with the analysis and solution of difficult medical problems through the computer. This requires understanding a medical problem in such detail that the solution can be mechanized. This requires new science and technology.

Artifical intelligence has evolved over the past 30 yr or so into a high-technology field with a correspondingly large associated jargon. In general, AI procedures are large computer programs which possess capabilities for knowledge representation, reasoning (search), and learning or knowledge acquisition (*4*). Knowledge representation refers to the process and manner in which data and other information are encoded, stored, and accesssed in the computer. Reasoning refers to accessing the appropriate information, making logical conclusions from the data, and keeping track of what steps were involved in reaching the logical conclusions. Learning refers to the capability of improving the computer's performance with increasing experience in solving problems.

In AI systems, we deal with a knowledge base. A knowledge base is the data base or computer stored collection of facts and pertinent rules for organizing and evaluating logical relationships among these facts. Both the structure and content of the knowledge base have to do with knowledge representation and how the knowledge is used in machine reasoning. Intelligent systems have to change with time; the transition being controlled by learning through adaptation. Without adaptability in the machine, its ability to help us solve our information problems would be severely limited. Ultimately, we must also design systems whose learning component is itself modifiable through time. The major areas of AI application include natural language processing, automatic programming, robotics, machine vision, and intelligent or expert systems. The development of an expert system requires the interaction of three individuals: 1) the knowledge engineer, who is the computer expert; 2) the domain expert, who has expert knowledge in the specific area of development; and 3) the user.

To the computer programmer, what AI researchers call knowledge is just another word for data. Computer programs that manipulate knowledge use the same data structures as other programs: arrays, lists, binary trees, etc. At the implementation (programming) level, knowledge and data appear identical. However, AI researchers do make a valid distinction between knowledge and data. Data are simple descriptions of observations or results, without interpretation. Once any kind of interpretive or inferential rule to the data are added, we have knowledge. For example, suppose a program accesses a file of names and addresses, as in a telephone directory. If the program was written to print the address of each name a user enters, then the file could be termed data. If, however, the program was written to print directions on how to travel to the address of each name entered by the user, then the file could be termed knowledge. An intelligent program is one that uses knowledge as defined above. Knowledge is stored in knowledge bases like data are stored in data bases (*5*). Not

For reprints contact: Jeffry A. Siegel, Dept. of Diagnostic Imaging, Temple University Hospital, 3401 North Broad St., Philadelphia, PA 19140.
*Current Address: National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894.

surprisingly, knowledge bases tend to have more complex structure and organization.

EXPERT SYSTEMS (HEURISTIC OR KNOWLEDGE-BASED APPROACH)

Expert systems are a class of computer programs that attempt to solve problems in a narrow domain that would ordinarily require a human expert to solve. An additional requirement is that the system must be capable of explaining its reasoning process. In medicine, they are a type of diagnostic program. Other kinds of diagnostic programs include formula-based or algorithmic (e.g., Bayesian) and flowchart-based approaches (*6*). The lines of demarcation, however, between the categories are sometimes obscured and a particular diagnostic problem can be solved using a combination of all the above program categories (*7*).

Expert systems consist of two components: 1) the knowledge base, which contains facts, and the rule base; and 2) the control system or inference program or "engine" (*8*). The rule base is used to define and order relationships among the facts and consists of *if-then* statements (*vide infra*), referred to as production rules. The control system controls the flow of logic within the program by selecting and applying the appropriate rules from the knowledge base to the specific case under consideration.

Knowledge Representation

In most applications of AI programming, the information to be encoded into the knowledge base originates from descriptive statements that are difficult to represent by simple structures like arrays or sets of numbers. Clinical decision making, intelligent information retrieval, and robot problem solving, require the capability for representation, retrieval, and manipulation of sets of statements. A variety of AI technologies have been used in the representation of knowledge, including rules (*9*), frames (*10*), semantic nets (*11*), and predicate calculus (*12*).

The most frequently used knowledge representation for medical decision making is rules. Factual knowledge can be

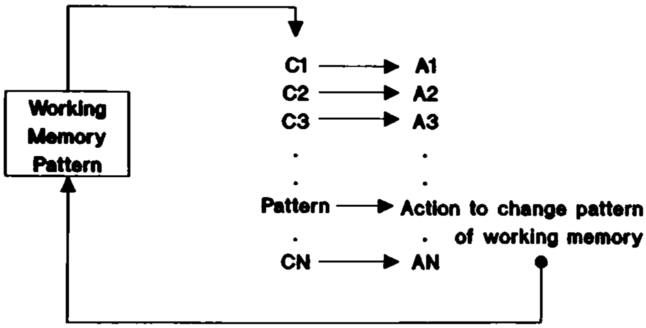

FIG. 1. The basic formalism for a rule-based expert system. It consists of a set of if-then (or condition-action) rules and a working memory. A pattern is stored in working memory and compared with the conditions (C1, C2, . . .CN). When a condition is found that matches the contents of working memory, the rule is "fired," and the contents of working memory are changed to reflect the action (A1, A2, . . . AN) part of the rule. This process continues until no further rule can be fired.

stored in the data base as *context-parameter-value* triples, modified by certainty factors (*13*). A context is an actual entity in the domain of the consultation (e.g., a patient). A set of parameters is associated with each context, such as age, sex, etc. of the patient. Each parameter of each context can take on values; the sex parameter of the patient context could take on the value of either male or female. The production rules operate on this factual data base and often contain heuristics or rules of thumb. Each rule consists of two parts:

1. *IF-THEN* statement
   The IF part consists of one or several statements which if true allows the THEN part to be considered true.
2. Weighting (Confidence) Factor
   Reflects the strength of confidence in the conclusion. Rule-based systems contain from a few to several hundred rules of the form shown below (*14*):

Example (Rule 578):

IF The infection which requires therapy is meningitis,
A smear of the culture was not examined, or
Organisms were not seen on the stain of the culture,
The type of infection is bacterial and
The patient has been seriously burned

THEN There is suggestive evidence (0.5) that pseudomonas-aeruginosa is one of the organisms which might be causing the infection.

The 0.5 indicates how strongly the conclusion follows from the premises. The assignment of certainty factors is at present a controversial issue. The concept of a confidence measure appears to be inherently probabilistic and, therefore, amenable to analytical techniques such as Bayes' theorem. However, there are serious limitations for a model based on conditional probabilities, namely, the mutual independence of symptoms. This requirement is the Achilles heel of Bayesian analysis, because the condition of mutual independence is rarely realized. The amendments to Bayes formula which must be made to account for nonindependence of symptoms introduce great mathematical complexity (*15*). Other techniques have been applied to the treatment of the relatively heuristic-based medical decision-making scheme, such as fuzzy set theory (*16,17*).

Control System

The control system chooses which rule to apply and ceases computation when a termination condition is satisfied. Each rule (*vide supra*) is a unique condition-action pair. When the conditional pattern of the rule is matched by incoming facts or observations, the action occurs and the rule is said to be "fired" (Fig. 1).

The control system uses the rules to reason (search) either forward (also known as forward chaining or data-driven), from observations to conclusions, or backward (also known as backward chaining or goal-driven). In forward chaining, the rules are applied whenever their premises are satisfied by the knowledge base. The user begins by entering a set of facts that trigger all the rules whose IF part are satisfied. For example, if the data input to the system included the facts that the infection was meningitis, the culture smear was not examined or organisms were not seen on the stain of the culture, the infection was bacterial, and the patient was seriously burned, the system would recognize that Rule 578 was applicable and draw the conclusion shown. A cascade of rule applications may ensue since one rule could produce a conclusion needed in the premises of another. If some facts are missing, the system will be unable to reach an appropriate conclusion because of uncertainty in deciding which rule to apply of the ones that match the input data. In AI, this is referred to as conflict resolution. A simple strategy frequently employed is to apply the rules in the order in which they are encountered (*2*).

Backward chaining starts from a conclusion or set of goals that the system tries to verify. For each conclusion or goal, the rules that establish the hypothesis in their THEN parts are applied. The IF clauses of these rules now become subgoals that need to be established. For example, let's reason backward through Rule 578 using goal-driven logic. Our goal is to determine the identity of the organism causing the infection. The system will retrieve and evaluate all rules which are able to infer infections requiring therapy, state of the culture, infection type, and burn status. This process continues until the entire search is exhausted for all rules consistent with the conclusion. One of the problems with backward chaining is that it is relatively slow.

The major advantage of using rule-based systems in medical decision making is that the reasoning mechanism is explicitly presented in the production rules. Conclusions drawn by an expert system can be easily explained and justified by: (a) displaying the rules that were applied; (b) retrieving the values of the used parameters; and (c) a description of how the rules interconnect.

LANGUAGES OF AI

The two most common AI programming languages are PROLOG (PROgramming in LOGic) and LISP (LISt Processing). PROLOG (*18*), first proposed at Marseilles in the early 1970s, was adopted at the University of Edinburgh, Scotland in 1974. In addition to being used in this country, it is being used by the Japanese as a language of their fifth generation computing project. PROLOG is an interactive, interpreted language based on predicate calculus. The components of PROLOG include terms, predicates, variables, and propositions. A proposition consists of a predicate (e.g., on or likes), followed by its terms and ending with a period. Logical connectives (and, or implies, not) are used to combine propositions. Examples of PROLOG propositions are given below:

likes (bill, wendy), likes (bill, alice).

This means that Bill likes Wendy and (,) Bill likes Alice.

friend ______ of (X,Y): — likes (X,Y), likes (Y,Z).

This means that X is a friend of Y if (:—) there is a Z that both X and Y like.

LISP (*19*) was proposed by John McCarthy at MIT in 1958 at approximately the same time as FORTRAN. LIST offers computing with symbolic expressions rather than numbers and a list representation of data. Lists can be used to represent any type of data or control structure, e.g., an employee data base (Fig. 2).

In AI work, it is common to consider knowledge as a set of objects connected by a set of relationships. A graph is a special diagram suited for depicting such objects and relationships (Fig. 3). Each object resides at a node of the graph. The lines connecting the nodes are called links. A link can represent various relationships between the nodes, e.g., is-a, is-an-instance-of, or is-a-part-of. Graphs that have a hierarchical structure (as shown in figure 3) are designated as trees. The highest node in a tree is the tree's root node. A node higher up in the hierarchy is a parent to the lower nodes, the children. Lists, which are the only data structure of LISP, are closely related to tree structures.

Expert systems are based on manipulations of symbolic objects rather than traditional "number crunching" algorithms. Traditional programming languages such as FORTRAN, BASIC, or ALGOL, are well suited for manipulation of numerical data, but are inconvenient for the symbolic structures used in AI. The list symbol structure has made LISP the most

- **LISTS can represent any type of data or control structure:**
  - e.g., Here is the list form of an employee data base:
    - -- Suppose an employee record looks like this:
      EMPLOYEE_NAME EMPLOYEE_# JOB_TITLE
      DEPARTMENT SUPERVISOR
    - -- JONES, JOHN 474428 PROGRAMMER
      DATA_PROCESSING EVANS, ANNE
    - -- Such a record might look like this as a LIST:
      ((JONES JOHN) 474428 PROGRAMMER
      DATA_PROCESSING (EVANS ANNE))
    - -- The data base would then be a LIST of such LISTS:

(
((JONES JOHN)474428 PROGRAMMER DATA_PROCESSING
(EVANS ANNE))
((JONES INDIANA) 772321 ARCHAEOLOGIST
ANTHROPOLOGY (EVANS ANNE))
....

FIG. 2. A LISP example showing a list representation of an employee data base.

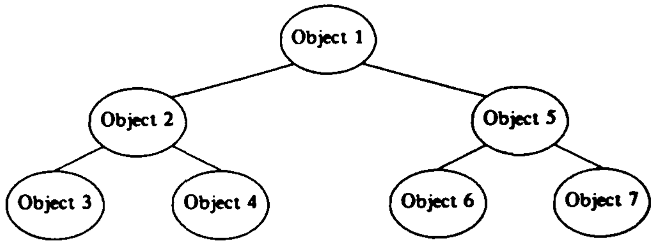

FIG. 3. A tree structure representing a graph with a hierarchy.

popular language for AI work in the United States. In addition, traditional programming techniques often embed problem information and the control strategy in user inaccessible code. In contrast, in expert systems, there is a clear separation between the knowledge base and the control strategy used to manipulate the knowledge and provide maximal user interaction. A new knowledge base can be substituted for the existing one, thereby creating a new system. Expert systems without their knowledge bases are referred to as shells.

The User Interface

An important component of any expert system is the user interface. The user interface is the facility by which a human user communicates with the expert system. This is normally a computer terminal with a keyboard or other input device and the associated computer hardware and software to interface with the expert system. It is at this point that the human makes inquiries of the expert system announcing the problem that he wishes to be solved. The computer, in response, will display questions on the monitor to which the user will supply the answers. These answers will be utilized in the logical reasoning that will be performed by the expert system in reaching a conclusion. The conclusion will be displayed on the monitor for the human to read. It is also by means of the user interface that the user can query the expert system as to how it reached its conclusion. The expert system would then respond with either the forward or backward chaining of premises and conclusions employed in the reasoning process.

In order to make expert systems more convenient for human users, that is more "user friendly", the dialogue of instructions and responses should be conducted in a form as much like natural language (English) as possible.

Computer Hardware

Because of their complexity and large size, expert systems have in general been developed on large computer systems. However, over the last 10 yr smaller computers have become

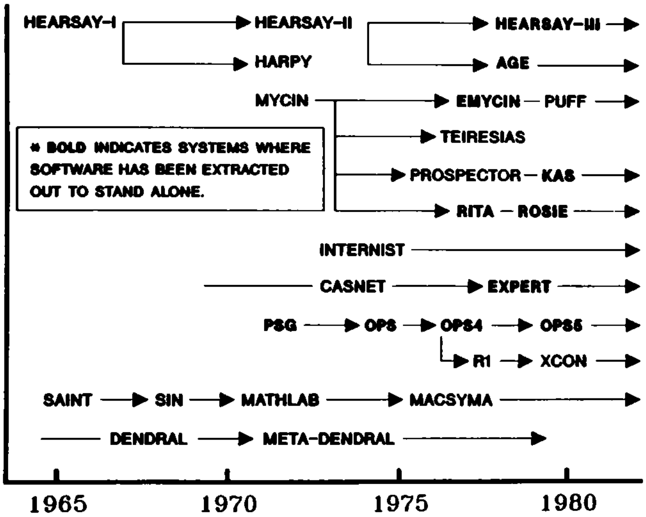

FIG. 4. Evolution of selected expert systems and system-building languages.

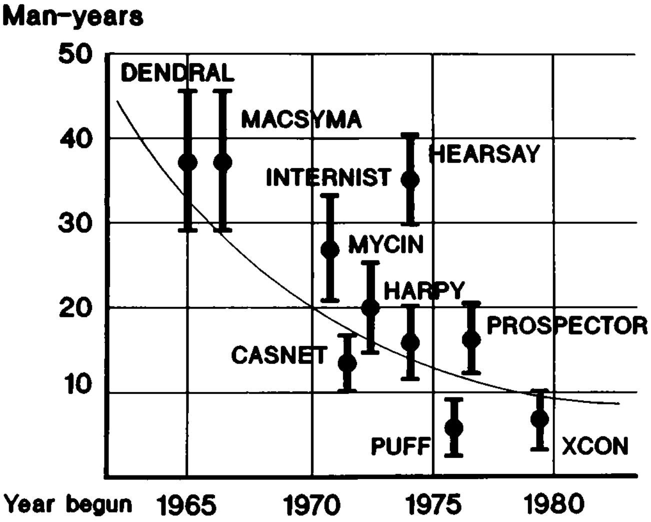

FIG. 5. Time required to create various expert systems.

faster and more powerful with remarkably increased storage capacities. As a result, expert systems are available for personal computers. Because of the wide spread current interest in expert systems, the number of systems available for the personal computer will undoubtedly increase significantly. This in turn, will make expert systems even more accessible and widely employed.

ARTIFICIAL INTELLIGENCE IN MEDICINE (AIM)

Although many new computer systems and knowledge engineering environments for medical applications have appeared, the essential technologies have been developed in the context of a number of classical expert systems. These classical systems include MYCIN at Stanford, INTERNIST at the University of Pittsburgh and CASNET at Rutgers University (Fig. 4). These programs were developed in the early to mid-1970s during which time AI research activity was essentially carried out in a small number of academic institutions.

Newer systems have refined these first-generation approaches, shortening development times (Fig. 5) and simplifying the knowledge engineering process. In the 1980s, the development of medical expert systems has spread to a large number of institutions, both within the United States and internationally.

The early impetus for the expert system explosion was Edward Shortliffe's work with a program called MYCIN (*14*). MYCIN employed a few hundred if-then rules about meningitis and bacteremia in order to deduce the proper treatment for a patient who presented with signs of either of these diseases. The overall organization of the MYCIN system is shown in figure 6.

The reasoning with if-then rules can be forward- or backward-chained and has a number of attractive computational and cognitive properties. One of these cognitive properties is the ease with which traces of fired rules can be used to explain why a decision was reached. An excerpt from a MYCIN consultation session is shown in figure 7. MYCIN was evaluated at the Stanford University Medical School in 10 randomly selected case histories of meningitis and compared to the results obtained by the attending staff. MYCIN results were comparable to those of the physicians. A major obstacle to the acceptability of MYCIN is the clumsiness of its user interface (it takes over 30 min of typing per consultation). Many extensions have been made to MYCIN and one, in particular, has focused on learning or knowledge acquisition (*20*). Davis's work on TEIRESIAS allowed an expert to interact with MYCIN and introduce amendments to rules as explanation of existing behavior of the system which might point to weaknesses in the rules (*21*).

Other work in first generation knowledge-based approaches to medical diagnosis include CASNET (*22*) and INTERNIST (*23*). CASNET consists of a large semantic network for diagnosis of glaucoma patients and gives advice considered to be as accurate as a physician's. INTERNIST is a hierarchical system for internal medicine and contains a massive knowledge base. Much of the stored knowledge is in the form of signs and symptoms with weights relating them to various diseases. INTERNIST generates a hypothesis about the patient and then searches its data base for suporting evidence. INTERNIST accepts as data ~ 5,000 signs, symptoms, history, or laboratory values. Each of these values is expressed as a string of characters like "serum immunoelectrophoresis IgA is increased." This string of characters must be typed exactly; otherwise, the input is rejected. Interfaces that allow more varied and natural phrasing are being developed.

The 1980s are probably going to be considered the decade of technology transfer for expert systems. Based partially on the pioneering efforts described above, expert-system shells were developed. These shells have four things that general programming tools do not (*24*): prepackaged representations for important concepts, efficient inference and representation tools for use in applications, specialized user interfaces, and generic application knowledge. Thus, these shells provide a programming environment that allows one to write an expert system without worrying about the internal operation of the inference program. Concurrently, reports began to appear describing the so-called "second generation" expert systems. The second generation programs tend to be written to solve clinical problems using "generic" AI approaches. A partial listing of recently introduced medical expert systems is given in Table 1.

EVALUATION OF AIM

The evaluation of computer systems which apply artificial intelligence in medicine can be performed at three levels (*25*):

1. Subjective evaluation of the research contributions of a developmental system.
2. Validation of the system's knowledge and performance.
3. Evaluation of the clinical efficacy of the operational consultation system.

The evaluation of a system's knowledge and advice should include accuracy, completeness, consistency, performance, and an understanding of how conflicting expertise is handled. Accuracy, for example, is tricky to deal with since traditional error rate techniques (sensitivity and specificity) frequently don't take into account the uncertainty and degree of confidence associated with medical diagnosis (*26*). As an example, consider that one expert system predicts that a patient has liver disease with a probability of 0.80 and a second expert system also predicts that the patient has liver disease but with a probability of 0.50. Both systems could be counted as correct by traditional criteria if the patient did have liver disease. Ac-

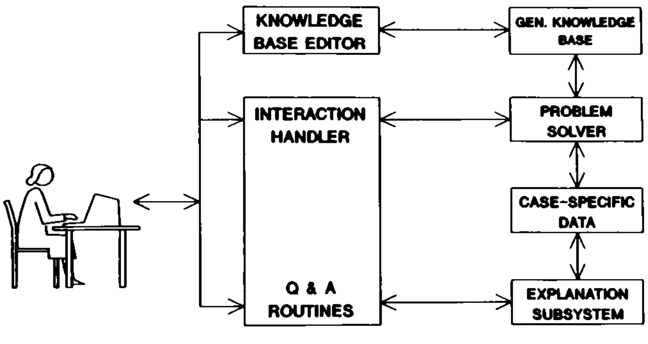

FIG. 6. Overall organization of the MYCIN system.

```
My therapy recommendations will be designed to treat for organisms
that either are very likely or, although less likely, would have a
significant effect on therapy selections if they were present.  It is
important to cover for the following probable infection(s) and
associated organism(s)
INFECTION-1 IS MENINGITIS
     + <ITEM-1> E.COLI [from clinical evidence only]
     + <ITEM-2> PSEUDOMANAS-AERUGINOSA [from clinical evidence
            only]
     + <ITEM-3> KLEBSIELLA-PNEUMONIAE [from clinical evidence
     only]
     + <ITEM-4> DIPLOCOCCUS-PNEUMONIAE [from clinical evidence
     only]
41)  Does Pt538 have a clinically significant allergic reaction to any
     antimicrobial agent?
**   NO
[Considering AMPICILLIN-AND-GENTAMICIN for use against
     INFECTION-1...]
42)  Do you have reason to suspect that Pt538 may have impaired
     renal function?
**   NO
43)  Pt358's weight in kilograms (or <number> POUNDS):
**   70
                    The final conclusion is about the best therapy to
                    cover for the likely cause of the infection, as
                    determined from the incomplete evidence available
[REC-1] My preferred therapy recommendation is as follows:
     In order to cover for Items <1 2 3 4>:
          Give the following in combination:
          1) AMPICILLIN
          Dose: 3.5g (28.0ml) q4th IV [calculated on basis of 50
             mg/kg]
          2) GENTAMACIN
          Dose: 119 mg (3.0 ml, 80mg/2ml ampule) q8th IV [calculated
          on basis of 1.7 mg/kg plus consider giving
             5 mg q24th intrathecal]
          Comments: Monitor serum concentrations
     Since high concentrations of penicillins can inactivate
     aminoglycosides, do not mix these two antibiotics in the same IV
     bottle
Do you wish to see the next choice therapy?
**   NO
                    The user may have reasons to prefer another
                    combination of drugs.  At this point MYCIN is
                    prepared to show other therapies, compare its
                    preferred therapy to the one the user specifies.
```

FIG. 7. Excerpt from a MYCIN consultation session.

TABLE 1. "Second Generation" Medical Expert Systems

System	Use	Reference
AI/Rheum	Rheumatologic diagnosis	27
SPE	Interpretation of serum protein electrophoresis data	28
TIA	Assessment of transient ischemic attacks and therapy advice	29
ONCOCIN	Oncology protocol management	30
SEEK	Rule checking program for a diagnostic system	31
LITO2	Liver disease diagnosis	32
ATTENDING	Critiques anesthetic management	33
ALVEN	Assessment of left ventricular wall motion	34
SPHINX	Broad range of diagnosis	35
CADIAG-2	Cardiac diagnosis	36
GAITSPERT	Evaluation of abnormal locomotion arising from stroke	37
FLOPS	Echocardiogram analysis	17

curacy measured in this way, therefore, loses information since it misses the fact that the first expert system was in some sense "more correct" than the first.

Clinical efficacy evaluation of AIM should include the actions of a physician, patient care, patient health, and cost/benefit analysis. The most fundamental issue should be the system's impact on patient health.

SUMMARY AND FUTURE

A new programming technology has been evolving around the transference of human expertise in a given domain into effective machine form so as to enable computing systems to perform as advisory consultants. Expert system development, confined to a small number of academic centers in its early days, has spread to a large number of institutions internationally.

Artificial intelligence is a branch of computer science concerned with the study of representation and reasoning (search). Two representation modes are lists, as used in LISP, and logic, as used in PROLOG. Searches can either be forward- or backward-chained or a combination of both. The domain of an AI system is the area in which the system solves problems. Some of the fundamental AI systems were developed to diagnose diseases; the domain of such a system would be the set of diseases that the system was designed to diagnose. Expert systems that diagnose diseases are examples of AI systems and are computer programs that strive to imitate the thinking and advice of a human expert in a given domain. The best uses of artificial intelligence in medicine include:

1. Areas in which a human expert exists.
2. Interpretation/Classification.
3. Probabilistic reasoning (to control inferencing).
4. Explanation capabilities (how/why).

The Japanese Fifth Generation Computer Systems project calls for improvement in the state-of-the-art of AI in medicine in both hardware and software (*2*). This has triggered a national effort in the United States resulting in the creation of Microelectronic and Computer Technology Corporation (MCC) and funding by the Defense Advanced Research Projects Agency (DARPA). This environment should provide a healthy growth period for medical expert systems with more widespread use within the next 10 yr to aid in diagnosis and medical decision making.

ACKNOWLEDGMENT

The authors gratefully acknowledge the courtesy of Integrated Computer Systems, Los Angeles, California for supplying Figures 1, 2, and 4 through 7.

REFERENCES

1. Nilsson NJ. *Principles of Artificial Intelligence.* Palo Alto: Tioga, 1980.

2. Sandell HSH, Bourne JR. Expert systems in medicine: A biomedical engineering perspective. *CRC Critical Reviews in Biomedical Engineering* 1985;12:95–129.

3. Szolvits, P, ed. *Artificial Intelligence In Medicine.* AAAS Selected Symposium 51, Boulder: Westview Press, 1982.

4. Rada R. Gradualness facilitates knowledge refinement. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1985;7:523–30.

5. Eng J. Augmentation and evaluation of a medical knowledge base. Internal Report, National Library of Medicine, 1985.

6. Warner H. *Computer Assisted Medical Decision Making.* New York: Academic Press, 1979.

7. Rada R. Automated medical diagnosis. In: Schwartz M, ed, *Application of Computers to Medicine.* Long Beach, CA: IEEE Engineering in Medicine and Biology Society, 1982:188–97.

8. Davis R. Knowledge-based systems. *Science* 1986;231:957–63.

9. Buchanan BG, Shortliffe EH. *Rule-based Expert Systems.* Reading: Addison-Wesley, 1984.

10. Minsky M. A framework for representing knowledge. In: Winston PH, ed, *The Psychology of Computer Vision.* New York: McGraw-Hill, 1975:211.

11. Quillan MR. Semantic memory. In: Minsky M, ed, *Semantic Information Processing.* Cambridge: MIT Press, 1979.

12. Kowalski R. *Logic for Problem Solving.* Amsterdam: North-Holland, 1974:569–74.

13. van Melle W, Scott AC, Bennett JS, Peairs MAS. *The EMYCIN Manual.* Palo Alto: The Board of Trustees of Leland Stanford Junior University, 1981:4–5.

14. Shortliffe EH. *Computer-based Medical Consultations: MYCIN.* New York: Elsevier, 1976.

15. Fryback D. Bayes theorem and conditional non-independence of data in medical diagnosis. *Comput Biomed Res* 1978;11:423–34.

16. Maiers JE. Fuzzy set theory and medicine: The first twenty years and beyond. In: Ackerman MS, ed, *Proceedings of the Ninth Annual Symposium on Computer Applications in Medical Care.* Baltimore: IEEE Computer Society, 1985:325–29.

17. Tucker DM, Siler W, Powell VG, et al. Flops: A fuzzy expert system used in unsupervised echocardiogram analysis. In: *Computers in Cardiology 1985.* IEEE Computer Society: in press.

18. Clocksin WF, Mellish CS. *Programming in PROLOG.* New York: Springer-Verlag, 1981.

19. Winston PH, Berthold KP. *LISP.* 2nd edition. Reading: Addison-Wesley, 1984.

20. Buchanan B, Shortliffe E. *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Reading: Addison-Wesley, 1984.

21. Davis R. TEIRESIAS: Applications of meta-level knowledge. In: Lenat

D, ed, *Knowledge-based Systems in Artificial Intelligence.* New York: McGraw Hill, 1982:229–91.

22. Weiss S, Kulikowski C, Amarel S, et al. A model-based method for computer-aided medical decision making. In: Shortliffe E, ed, *Readings in Medical Artificial Intelligence.* Reading: Addison-Wesley, 1984:160–89.

23. Miller R, Pople H, Myers J. Internist-1. *N Eng J Med* 1982;307:468–76.

24. Bobrow DG, Stefik MJ. Perspectives on artificial intelligence programming. *Science* 1986;231:951–57.

25. Miller PL. Issues in the evaluation of artificial intelligence systems in medicine. In: Ackerman MJ, ed, *Proceedings of the Ninth Annual Symposium on Computer Applications in Medical Care.* Baltimore: IEEE Computer Society, 1985:281–86.

26. Reggia JA. Evaluation of medical expert systems. A case study in performance assessment. In: Ackerman MJ, ed, *Proceedings of the Ninth Annual Symposium on Computer Applications in Medical Care.* Baltimore: IEEE Computer Society, 1985:287–91.

27. Kingsland L, Sharp G, Capps R, et al. Testing a criteria-based consultant system in rheumatology. In: *Proceedings of MEDINFO-83.* Amsterdam: North Holland, 1983:514–17.

28. Weiss SM, Kulikowski CA, Galen RS. Representing expertise in a computer program: The serum protein diagnostic program. *Clinical Lab Automation* 1983;3:383–87.

29. Reggia J, Tabb R, Price T, et al. Computer-aided assessment of transient ischemic attacks. *Arch Neurol* 1984;41:1248–54.

30. Hickam DH, Shortliffe EH, Bischoff MB, et al. The treatment advice of a computer-based cancer chemotherapy protocol advisor. *Ann Int Med* 1985;103:928–36.

31. Politakis P, Weiss SM. A system for empirical experimentation with expert knowledge. In: Clancy WJ, Shortliffe EH, eds, *Readings in Medical Artificial Intelligence: The First Decade.* Reading: Addison-Wesley, 1984:426–43.

32. Cravetto C, Lesmo L, Rolandino RM, et al. An expert system for liver disease diagnosis (LITO 2). In: Ackerman MJ, ed, *Proceedings of the Ninth Annual Symposium on Computer Applications in Medical Care.* Baltimore: IEEE Computer Society, 1985:330–34.

33. Miller PL. Critiquing anesthetic management: The "ATTENDING" computer system. *Anesthesiology* 1983;58:362–69.

34. Tsotsos JK. Computer assessment of left ventricular wall motion: The ALVEN expert system. *Comput Biomed Res* 1985;18:254–77.

35. Fieschi M, Joubert M, Fieschi D, et al. SPHINX: An interactive system for medical diagnosis aids. In: Gupta M, Sanchez E, eds, *Fuzzy Information and Decision Processes.* Amsterdam: North-Holland, 1982:269–75.

36. Adlassnig KP, Kolarz G, and Schweithauer W. Present state of the medical expert system CADIAG-2. *Meth Inform Med* 1985;24:13–20.

37. Dzierzanowski JM, Bourne JR, Shiavi R, et al. GAITSPERT: An expert system for the evaluation of abnormal human locomotion arising from stroke. *IEEE Trans Biomed Eng* 1985;BME-32:935–42.