

An Opening Chapter of the First Generation of Artificial Intelligence in Medicine: The First Rutgers AIM Workshop, June 1975

C. A. Kulikowski

Department of Computer Science, Rutgers - The State University of New Jersey, Piscataway, New Jersey, USA

Summary

The first generation of Artificial Intelligence (AI) in Medicine methods were developed in the early 1970's drawing on insights about problem solving in AI. They developed new ways of representing structured expert knowledge about clinical and biomedical problems using causal, taxonomic, associational, rule, and frame-based models. By 1975, several prototype systems had been developed and clinically tested, and the Rutgers Research Resource on Computers in Biomedicine hosted the first in a series of workshops on AI in Medicine that helped researchers and clinicians share their ideas, demonstrate their models, and comment on the prospects for the field. These developments and the workshops themselves benefited considerably from Stanford's SUMEX-AIM pioneering experiment in biomedical computer networking. This paper focuses on discussions about issues at the intersection of medicine and artificial intelligence that took place during the presentations and panels at the First Rutgers AIM Workshop in New Brunswick, New Jersey from June 14 to 17, 1975.

Keywords

Artificial Intelligence in Medicine (AIM), first Rutgers AIM workshop, biomedical knowledge representation, biomedical problem solving, clinical decision making

Yearb Med Inform 2015;10:227-33

<http://dx.doi.org/10.15265/IY-2015-016>

Published online June 30, 2015

Introduction

During the first half of the 1970's, several groups working on computational models for clinical decision-making and problem-solving had developed the MYCIN rule-based system for infectious disease therapy assistance at Stanford [1, 2], the CASNET Causal Associational NETWORK model for consultation in glaucoma [3, 4] at Rutgers, the DIALOG (later renamed INTERNIST) system for differential diagnosis in internal medicine [5, 6] at Pittsburgh, and the PIP (Present Illness Program) [7, 8] for diagnosis-driven acquisition of clinical data at MIT and Tufts. These had been inspired by AI approaches that departed from the earlier general problem solving search paradigm characteristic of AI since its inception and still holding sway into the 1970's, and focused on capturing domain- and problem-specific strategies for solving complex sequences of expert biomedical interpretations and actions. These included the rule-based and hypothesis-list approaches used in the DENDRAL Project [9, 10], which influenced MYCIN, as well as experimental, instructional, interview-based, and cognitive approaches to the analysis of clinical problem solving [11, 12, 13, 14, 15, 16], and the causal-taxonomic representation of underlying processes of disease [4, 11]. While earlier computer models for medical decision-making were predominantly statistical or algorithmic [16, 17, 18, 19, 20, 21, 22, 23, 24], the new AI approaches developed structured representations of specific clinical domain knowledge over which a general inference engine could reason with a variety of heuristics, and provide advice or suggestions to the consulting user [25].

This shift in computational modeling and problem-solving paradigms was supported by the National Institute of Health's Division of Research Resources [26], under the direction of Robert Raub, who, as a student of William Yamamoto at the University of Pennsylvania, was seeking to expand the focus from strictly statistical approaches to inference by modeling the underlying physiological and clinical knowledge that supports diagnostic and therapeutic decision-making. The highly successful collaborative research between AI pioneer Edward Feigenbaum and Nobelist Joshua Lederberg in developing the systems within the DENDRAL project [10] for the interpretation of mass spectrometry data from biomolecules also encouraged the NIH to support research of several other groups, such as those directed by Saul Amarel at Rutgers University on clinical, biological, and ecological modeling, those from MIT and Tufts New England Medical Center, led by Anthony Gorry, William Schwartz, Steven Pauker, Jerome Kassirer, and Peter Szolovits, focusing on the elicitation of patient problems in a present illness, and those at the University of Pittsburgh with Harry Pople, Jack Myers, and Randolph Miller on differential diagnosis in internal medicine. Time-shared computing and the infrastructure for computer networking evolving at this time presented an excellent opportunity for creating the Stanford University Medical Experimental – Artificial Intelligence in Medicine (SUMEX-AIM) laboratory for providing advanced computer capabilities (PDP-10's) connecting and supporting the research of all the participating groups at the time [27].

The moment was ripe in 1975 for these and other groups with similar interests in AI approaches to biomedical and related problems to meet so they could compare and contrast their methods for knowledge representation, inference, and problem solving, as well as the practical implementation and testing of their systems, taking advantage of the availability of time-shared remote computing facilities at SUMEX-AIM to share their computer programs remotely with their clinical and biological collaborators.

The Rutgers Research Resource on Computers in Biomedicine, under the direction of Saul Amarel, helped to organize and host the first of these workshops, which later rotated among participants in the AIM community. To provide some insight into the mind-frame of the times and how the researchers, clinicians, and biological collaborators interacted and exchanged ideas and perspectives on the challenges to the field and their visions for the future, the present paper excerpts and comments on the workshop papers and discussions published in the Proceedings of the First Workshop of AI in Medicine [28].

Clinical Systems Presented at the First AIM Workshop

The AIM community evolved from the national projects funded by the NIH, and the first workshop was organized by Casimir Kulikowski, with the technical direction of N. S. Sridharan and the overall direction of Saul Amarel, supported by the Biotechnology Resources branch of the NIH under grant RR-643. The purpose of the workshop was to “provide insight into existing and potential systems that apply methods of Artificial Intelligence to problems of biomedical research and health care”. The attendees included a range of investigators from Chemistry, Psychology, Medicine, and Computer Science. The 1975 Workshop theme on “Knowledge-based Systems in Biomedicine” stimulated discussions, demonstrations, and hands-on systems experience in medical modeling and decision making for diagnostic/thera-

peutic consultation; psychiatric simulation, psychological modeling, language analysis, and common sense reasoning; and biomolecular characterization of organic molecules on the basis of chemical analysis, protein structure determination, and chemical synthesis planning. Brief presentations of the existing AIM systems were followed by in-depth discussions of the underlying issues. These discussions were recorded, and summaries transcribed for publication in the Proceedings of the workshop [28].

Ted Shortliffe of Stanford University, opened the first morning session by describing the MYCIN system for Antimicrobial Therapy Consultation. He had just completed his thesis the year before and explained how his system acquired its information in the form of rules through human-engineered interaction with expert clinicians, while allowing extensions to the rules through a structured context tree as a knowledge base representing the different knowledge components in the rules and their application with strategies of reasoning that answered questions posed by the user about a consultation on antibiotic therapy. He also emphasized how the system handled uncertainty through a novel scheme of heuristic certainty factors [29] as an alternative to subjective Bayesian probabilities, which had the advantage in knowledge acquisition of allowing uncoupling of the stated beliefs of clinicians in the confirmation of a diagnostic hypothesis from those stated about its negation. While this scheme was later shown by Shortliffe and Heckerman [30] to be reducible to a Bayesian formalism, nevertheless, it provided an alternative for the elicitation of uncertainty estimates from experts that was less rigid and closer to Zadeh’s fuzzy set and logic approach with its allowance for a range of belief quantifications for hypotheses [31, 32], rather than the binary logic underpinning conventional probabilistic formalisms, including Bayesian networks. The MYCIN system development, coupled with the DENDRAL project, helped put the Heuristic Programming Project led by Edward Feigenbaum at Stanford at the forefront of AI research, with the large range of systems and results summarized in a comprehensive book [33].

The second presentation of the morning was by Harry Pople of the University of Pittsburgh, who presented the DIALOG system for Diagnostic Logic in Internal Medicine [5], to be renamed INTERNIST the next year [6]. The system was the most general of all presented at the AIM Workshop, incorporating the knowledge of leading internist Dr. Jack Myers, who provided both causal and taxonomic relationships between clinical findings and diagnostic hypotheses, represented in a network model, for which Pople developed abductive and inductive reasoning rules using heuristic uncertainty measures to quantify levels of uncertainty. As described in the Workshop Proceedings [28], DIALOG/INTERNIST by this time already “encompassed a substantial portion of the major diseases of internal medicine. The system thereby exhibits diagnostic behavior and competence comparable to that of the skilled physician, and handles systematically, cases where two or more distinct clinico-pathological entities are present” [28]. This system continued to be refined over the next decade, leading to a highly sophisticated and flexible consultation program for all of internal medicine [34], and was later used as the basis for an electronic information resource or textbook of medicine – QMR [35]. It also permitted Harry Pople to generalize the representation as one that would help provide a structure for inherently ill-structured expert problem solving knowledge [36].

Casimir Kulikowski of Rutgers University was the next speaker, presenting the CASNET system, which had been developed with Sholom Weiss as part of his doctoral dissertation the previous year [37] and Aran Safir, who as an expert ophthalmologist, had suggested glaucoma as a disease to model because of its reasonably well-known alternative causal pathways with often subtle variants due to a range of etiologies. The architecture of the knowledge representation was a hierarchical and associational network, with patient findings triggering hypotheses about pathophysiological states, which when confirmed in the context of a pathway of cause and effect among the states would trigger diagnostic hypotheses with future or “causally downstream” associated prognostic states, for which treatment recommendations

would be suggested in order to prevent them. This was the first abstraction of a conceptual model representing the temporal patterns of cause and effect that underpin diagnostic and therapeutic reasoning [38]. CASNET, like MYCIN and DIALOG/INTERNIST also used heuristic uncertainty measures for confirming the hypotheses about pathophysiological states, but used a multi-level confidence threshold to map into the logic for decision-making at the causal pathway and network level. Subsequently, Weiss and Kulikowski generalized the CASNET formalism into the EXPERT framework for representing causal, hierarchical taxonomic, and associational rules within a rule-based formalism [39]. These systems capitalized on a compiled representation of the states to develop a highly efficient algorithm for reasoning over acyclic directed graphs, capturing and applying expert knowledge for problem solving in a variety of clinical [40, 41, 42] and other fields such as geophysical prospecting, and mechanical and computer systems troubleshooting [43, 44].

The final clinical system described in the first morning of the First AIM Workshop was the Present Illness Program (PIP), which was presented by Steven Pauker of Tufts New England Medical Center. He pointed out that the knowledge base was organized into frames as defined by Professor Marvin Minsky of MIT, and linked into an associative memory, which was partitioned into short and long-term memory components. This permitted likely clinical hypotheses “to be arrived at rapidly, and considers frames that are closely linked to the hypotheses” [28]. The mechanism for “grabbing” frames from memory was likened by Dr. Pauker to an alligator that leaps up from the water to catch its data as prey, leading the audience into fits of laughter as he mimed the action with examples, providing welcome relief from the serious and highly technical earlier material. The Present Illness Program proved to be a fertile testing ground for the Tufts and MIT collaborative group to experiment with different representations of clinical knowledge, which proved useful for representing complex models of disease in both quantitative and qualitative ways [45,46], and led to more detailed causal and temporal models [47, 48].

Panel on Medical Perspectives of AIM Systems

The morning of the first day of the Workshop continued with a Panel chaired by Dr. Aran Safir of Mt. Sinai School of Medicine, who solicited comments from workshop participants on the systems that had been presented, asking them to place the AI approaches in the context of the broader needs of medical decision-making and consultation for clinical problem solving.

Dr. Ralph Engle of Columbia University (a pioneer in computer diagnosis and developer of the HEME/HEME-2 systems for the diagnosis of blood disorders [49]) opened the discussion on a skeptical note. He referred to a point made by H. Vaihinger in *Philosophy of “AS IF”* [50] which he said postulates that “we often accept as truth the fiction of approximations because of some of the useful benefits which result. In a sense all science and mathematics are an approximation of the real world, and there are benefits to be gained if we act as if science was the real world. Similarly, benefits can result from acting as if artificial intelligence was the same as human intelligence, though the term Artificial Intelligence seems a bit presumptuous to some individuals. The full benefit of the use of computers as tools of thought can come only when we learn to dissect intelligence into a portion best suited to the human being, and a portion best suited to the computer, and then find a way to mesh the two processes. The science of Artificial Intelligence is concerned with that very important task.” [28] This comment illustrates how, from the very beginning of the field of AI in Medicine, experienced clinical practitioners were starkly aware of the challenge being taken on, but which has frequently not been as fully recognized by enthusiasts of AI who may expect too much from their preconceived models and theories of knowledge representation and reasoning. The critical juncture of how responsibility of clinical practitioners for a patient can be handled, or handed-off, in interacting with a computer-based system remains one of the most difficult ethical and practical problems in AI and clinical human-machine information processes that can have life-and-

death consequences for the patient, and pose thorny legal and ethical questions for the practitioner, especially as we move, forty years later, into the field of personalized translational medicine [51].

Dr. William Yamamoto of George Washington University (a pioneer in computational models in physiology [52]) followed-up on Dr. Engle’s comment with a different caveat. He said that AI is “attempting to emulate or imitate the performance of the academic physician working generally with the most severe disease patterns. And when you mention Artificial Intelligence to a number of physicians you arouse a basic hostility because you are threatening them in the area they have reserved for themselves. They are willing to give the IV’s to the nurse and the drugs to the pharmacologist and the surgical preparations to the OR nurse. But what they reserve for themselves is what they consider the intelligence.” [28]. This highlighted another underlying problem faced by clinical consultation systems. While experts themselves may not feel they need consultation advice, non-experts or novices who might benefit from such advice may likewise feel threatened, since dealing with a computer program is different from interacting personally with a medical authority they have been taught to respect. Yamamoto then went on to detail the different types of behavior that he thought physicians would include if they attempted to assess AI. These included some that he considered the field was handling reasonably well, like choices between alternatives that are not mutually exclusive, and the execution of pre-determined processes. In learning facts or knowledge by inductive inference, he considered that AI showed only a “questionable level of success”, while for clinical initiative and innovation he did not think there was any contribution of AI yet. He saw little activity in dealing with the problem of conflicting policies like “do no harm”, though he thought it was “attackable”, while he recognized self-awareness as a tough epistemological problem which AI had not attempted to answer. He also pointed out that the problem of assigning value judgments to what the clinician executes within a societal context of both patients and their families, which again was not part

of AI's repertoire. On the other hand, for solving problems in complicated diagnostic game-like scenarios, Yamamoto viewed AI as providing "a number of interesting and powerful paradigms". Yet, he did not see AI being able to help with recognizing the logical consistency of a new system – though it was not something unique to AI methods. Tentative decisions was something he thought MYCIN was dealing with well in imitating the intelligent behavior of physicians. However, Yamamoto considered that reasoning with only an indeterminate or "qualitative end-point", which is typical in medicine, was something additional that was needed for a system to manifest intelligence.

The next panel speaker was Dr. Pauker, who pointed out the new capabilities being developed with computers to serve as "a laboratory to model decision making and to test theories....Probability theory and especially Bayes rule now form a central part of my diagnostic approach in terms of computer programs. However, our recent studies have emphasized the importance of a richly cross-linked database of guessing and heuristic approaches. These ideas fit more closely the romantic notion of what clinical expertise is, and to some extent have underlined the need for complex learning and indexing processes. With this new kind of laboratory and approach we are beginning to understand better how to teach students what clinical expertise really is." [28]. With this insight, Dr. Pauker anticipated the subsequent use of medical consultation systems for explanation and teaching, as underlying models of disease are more central to learning and understanding, than to actual clinical practice, where diagnostic experts frequently use highly compressed or compiled "diagnostic rules". Dr. Pauker concluded with a comment about "AI having something to learn from medicine in the same way that medicine has something to learn from AI" [28].

Dr. Donald Lindberg, then at the University of Missouri where he had pioneered the use of computers in pathology, clinical, and medical library applications [53] next emphasized the importance of the SUMEX-AIM high performance networked computing infrastructure as essential for

the success of the enterprise of AI in Medicine. On the need to employ AI techniques in medicine he expressed no doubt at all. However, he emphasized that one should focus on AI attempting "to do in medicine what cannot be done without a computer." He then gave three examples. The first was the need for a uniform terminology for medicine, since without a more standardized vocabulary there would be "no systematic way for clinical records to become the basis for research." This anticipated a main thrust of what Lindberg would champion a decade later when, as Director of the National Library of Medicine, he guided the development of the Unified Medical Language System (UMLS [54]) which has, in the age of the web, become a foundational component within medical informatics for automatically indexing and interpreting the biomedical literature [55, 56] essential to research discoveries even more than for clinical support. Lindberg's second point involved a plea to develop approaches to "test potential causal or non-causal medical associations" like the thalidomide/pregnancy association, so they could be anticipated through an early warning system for drug side-effects or interactions. He then suggested two technical questions: dealing with large and complex databases, for which he identified geographical data systems and their epidemiological applications as an example where AI could contribute; and then pointing out that medical files needed to know more about what they contained, which anticipated the wide use of meta-data structures a decade later. His final comment was on the DIALOG system of Myers and Pople, stating that he did not think it was important as an example of automating an expert consultant, but rather as a facility "whereby diagnostic rules are made accessible and can be applied to a particular case without the presence of the consulting physician"[28]. Dr. Myers responded that he disagreed, expecting that AI programs like this would "continue to have diagnostic application in the tertiary care institution" – anticipating their use by paramedical personnel. Dr. Myers then commented on the educational uses of AI systems, and added that he thought they would be useful for self-education also, as well as for "measur-

ing clinical competence not only in students but also in graduate physicians."

Dr. Aran Safir from Mt. Sinai, inventor of the first automated ophthalmoscope [57], then emphasized that the development of AI systems requires close collaboration between computer scientists and physicians. The former needs to "be exposed to the very long time and difficult process of education in medical problems." Only in this way can they "develop a feeling for the complexity and unreliability of the data." He noted how Dr. Kulikowski and his colleagues had observed glaucoma surgery and testing and that this changed their understanding of what they had merely read about earlier. Conversely, he said, physicians cannot just "preach medicine", but must learn how the computer scientists are embedding medical knowledge in the logical structures they use for computer representation. He then contrasted the radically different personalities and educational experiences of computer scientists and mathematicians who are encouraged to "follow orderly systems of thought," whereas physicians "have entered by choice a profession in which disorder and unpredictability are nearly the rule." Physicians are required to solve a problem at the time it is brought to them, while a computer scientist rarely feels such pressure. He concluded with the observation that good work in computing and medicine will result only from collaborative teams where computer scientists can thrive within the disorder of medicine, and the physicians can work happily within the logical and mathematical world of computer science.

The final comments in the panel came from Dr. William Schwartz of Tufts New England Medical Center [16] who commented that "the process of developing large systems that are reliable enough to make an impact on clinical research will require inevitably a large investment of resources over the next few decades". He then mused about whether society and the funding agencies would wait this long in order to improve the quality of clinical care as anticipated, and the upgrading of medical education and curricula that would result from this. He then suggested that this would need a change in how medical students are taught, so that they would be taught about the nature of problem solving

and cognitive processes in the second year, before they are professionalized and less open to question traditional ways of learning and thinking. It is interesting to note that it was only a year later that Schwartz co-authored an influential paper [7] on the simulation of clinical cognition, at about the same time as the book by Elstein et al. on cognitive approaches to clinical reasoning was published [58]. Research in clinical cognition has continued as a thread related to AI in medicine ever since [59].

The non-medical sessions at the Workshop that followed were extensive and cast light on the common problems shared by AI researchers in representing knowledge for complex problem solving in biology and medicine. However, since much of the material was only partially related to medical informatics aspects, considerations of space do not allow inclusion in the present article. Instead, summarized next are the discussions from a panel comparing the different medical systems.

Panel on the Analysis and Comparison of Medical Systems

This panel took place on the afternoon of the second day of the Workshop, and was led by Harry Pople with the participation of Ted Shortliffe, Edward Feigenbaum and Stan Axline of Stanford, Saul Amarel, Casimir Kulikowski, and N S Sridharan of Rutgers, Aran Safir of Mt. Sinai, Peter Szolovits of MIT, and Steve Pauker of Tufts, Harry Pople and Jack Myers of Pittsburgh, Donald Lindberg of the University of Missouri, and Bruce McCormick of the University of Illinois at Chicago.

Pople opened the panel by giving his comparison of the three systems MYCIN, CASNET, and DIALOG. He considered that “all three systems deal with the problem of hypothesis formation but the hypothesis formation embedded in MYCIN as I see it is a special case of deductive reasoning.” He went on to explain how his interpretation of MYCIN’s use of a context tree resulted in a goal-directed approach of attempting to prove the occurrence of a dis-

ease by deductive inference. He contrasted this to the other systems which used the “alternative reasoning from consequence to hypothesis” together with “reasoning from hypothesis to consequence”. Shortliffe then explained how theorem proving methods had indeed inspired his work, but disagreed with Pople’s characterization of the contrasting inference methods, since MYCIN also used antecedent rules for inductive inference through the certainty factors derived from the clinical findings. He then turned to an important aspect of his system, namely the natural language interface of MYCIN, which was developed in order to enhance the interaction with the user in order to explain and support the recommendations of the system. Shortliffe’s colleague S. Axline then emphasized the need for sequential understanding of the logic process as rules are triggered in contrast to the earlier approaches of gathering large amounts of clinical data before carrying out logical inferences, which he considered stilted. Szolovits commented on the related issue of specialized vocabularies with set formats which are typically used for entering clinical information into a computer, and argued for the feasibility of using existing parsing methods to handle the language problem. He also pointed out that expert physicians always want to hear a case presented in a standard way, yet he was concerned that this detracted from the flexibility of handling data in whatever manner it comes in. Axline then said that the general internist faces problems of data acquisition very different from those faced by clinical specialists, and Shortliffe added that he had done chart reviews on a weekly basis with Drs. Axline and Cohen, the infectious disease specialists, in order to come up with a representation of rules that worked for the specialists. Kulikowski then interjected that often just “imitating the doctor is not necessarily the way to go. What you want to have is a number of alternative models, with the simulation of a particular doctor being just one of them. It clearly depends on the scope of the problems and on the knowledge structure in a particular domain.” Szolovits gave an example of this kind of specialization, citing the digitalis therapy algorithm being

developed by Silverman [60], which led Sridharan to question whether the richness of concepts and all the processing involved in some of the MIT models was really needed, and suggested that algorithmic approaches might prove adequate for specific, circumscribed problems. Szolovits countered that working with Steve Pauker pretending to be a patient, Andee Rubin had developed a very rich and complex theory based on the transcripts of their interaction – an example of structured knowledge acquisition. Amarel then contrasted the challenges presented by the richness of the hypothesis space for large domains, like internal medicine, which is very different from the case with MYCIN which “has practically no hypothesis formation process.” In the former he stated that “AI comes in much more” than with some of the other systems, stating that the size of the hypothesis space and the kinds of tools brought to bear in searching it become the determining factors. Pople then reaffirmed that the purpose of his system was indeed simulation of the physician expert, concurring with Amarel.

Feigenbaum raised the parallels between the medical and other biochemical and biomedical systems in terms of the complexities of interpretation, and the need to develop a kit of tools for knowledge representation and inferencing – thus anticipating the generalization of these systems in the next generation. He remarked that “everyone has come to realize that inserting the knowledge, deleting it, modifying it are the critical problems and that we have all invented roughly similar ways of doing it.”[28]. He then raised the issue of how the AI approaches differed from the earlier statistical and static decision tree representations, and he declared them to be “light years” away from each other. Safir raised the concern that “some computer scientists think they are modeling or simulating a process that they view as static. But it may very well be that the process of medical decision making is undergoing changes almost as rapidly as computer science so that what AI is using as a model today could be the product of medical schools thirty years ago.” Lindberg remarked that he had not heard yet of any attempt to measure the magnitude or quality of the AI accomplishments, and suggested that

a serious effort had to be made to separate out the major from the minor ones. From the audience, Prof. Srinivasan then injected the issue of planning functions going beyond diagnosis in medicine, to which Dr. Myers responded by stating that “therapy is as big a problem if not bigger than the problem of diagnosis. Fortunately smaller programs like MYCIN and CASNET can deal with this, but we had to put it in second place. And, I believe Dr. Pauker is in the same situation for the most part.” Pauker partly disagreed with this, saying that “knowing what to model in terms of therapy initiation is very dependent upon and strongly influenced by what we mean by arriving at a diagnosis. Often it is very difficult to know when we have reached that stopping point. What that arbitrary stopping point is depends on what we are going to do next, the seriousness of the situation, the amount of time we have to provide treatment. So we cannot finess one or the other.” Myers agreed, and added that once therapy is included “you perturb the whole system and the data base becomes radically changed by the very presence of the treatment. And this is a very complex change which I am sure causes as big a problem as the original data base.” Pauker concurred, mentioning the disappearance of symptoms with treatment, and the problems of multiple diseases where one can mask the manifestations of the other, pointing out that “the therapeutic intervention of a physician at some level represents another disease.” Szolovits went on to emphasize the role of temporal change in patient histories, and “what our expectations were as opposed to what actually happened, and how we form hypotheses to account for them.” McCormick then noted parallels between clinical and management decision making, and Feigenbaum concluded the panel with the provocative question of whether knowledge based systems which are extremely systematic in their application might not be better than the human experts who build them, and “Could we use these systems for making those decisions as opposed to trusting the opinion of the physician who may not be as systematic?” It is fascinating that at 40 years remove, and with radically advanced technology, we are still asking the same questions.

Retrospective on the First AIM Workshop

The second and third days of the AIM Workshop included seminars on each of the systems presented, and the last day had a concluding series of panels on Methods of Inference: Formal and Clinical Problems, moderated by Ted Shortliffe, one on Knowledge Acquisition and Representation moderated by Bruce Buchanan, and one on Problems of Systems Development and issues of collaboration across disciplines - Shared resources and computer networking, moderated by Saul Amarel. Unfortunately space constraints here do not allow summarization of these panel discussions, but they were quite lively as result of the earlier seminars in small groups, whose participants got to know each other quite well, and felt free to have very frank exchanges about the challenges of interdisciplinary research, the complexities of implementation and the need for funding. Bill Raub, the head of the sponsoring funding agency, NIH's Division of Research Resources, gave a keynote speech confirming the agency's strong support for the work, which led to much productive research that influenced artificial intelligence and medicine during the following decade. This First AIM Workshop accelerated the process by helping investigators to get to know one another, develop a better understanding of each other's work, stimulating many professional exchanges, collaborations, and friendships which have persisted over 40 years, with a lasting impact on the evolving discipline of biomedical and health informatics [61].

References

- Shortliffe EH, Axline SG, Buchanan BG, Herigan TC, Cohen SN. An artificial intelligence program to advise physicians regarding antimicrobial therapy *Comp Biomed Res* 1973;6(6):544 – 60.
- Shortliffe EH. *Computer Based Medical Consultations: MYCIN* New York: Elsevier; 1976.
- Kulikowski C, Weiss S. Strategies of data base utilization in sequential pattern recognition. *Proc IEEE Conf Decision & Control, 11th Symp on Adaptive Processes*; 1972. p. 103-5.
- Weiss S, Kulikowski C, Safir A. *Glaucoma Consultation by Computer*. *Comp Biol Med* 1978;8:25-40.
- Pople HE, Jr Myers JD, Miller RA. *DIALOG: A model of diagnostic logic for internal medicine*, 4th Int Joint Conf on Artif Intell (IJCAI), Cambridge, MA, 1975. p. 848-55.
- Pople HE, Jr. Presentation of the INTERNIST System Proc 2nd AIM Workshop, New Brunswick, N.J.: Rutgers University; 1976.
- Pauker SG, Gorry GA, Kassirer JP, Schwartz WB. Towards the simulation of clinical cognition, Taking the present illness by computer. *Am J Med* 1976;60:981-96.
- Szolovits P, Pauker SG. Categorical and probabilistic reasoning in medical diagnosis. *Artif Intell* 1978;11:115-44.
- Smith DH, Buchanan BG, Englemore RS, Duffield AM, Yeo A, Feigenbaum EA, et al. Applications of Artificial Intelligence for chemical inference, VIII: an approach to the interpretation of high resolution mass spectra of complex molecules. Structure elucidation of estrogenic steroids. *J Am Chemical Soc* 1972;94:5962-71.
- Lindsay RB, Buchanan BG, Feigenbaum EA, Lederberg J. *Applications of Artificial Intelligence for Organic Chemistry: The Dendral Project*. McGraw-Hill Book Company; 1980.
- Feinstein AR. *Clinical Judgment*. Baltimore: Williams and Wilkins; 1967.
- Lindberg DAB, Rowland LR, Buch CR, Morse WF, Morse SS. *CONSIDER: A computer program for medical instruction*. *Proc 9th IBM Symp Med*; 1968.
- Slack W, Hicks GP, Reed CE, Van Cura LJ. A computer-based medical history system. *N Engl J Med* 1966;274:194-6.
- Kleinmuntz B, McLean RS. Diagnostic interviewing by digital computer *Behavioral Sci* 1968;13:75-80.
- Gorry A. Strategies for computer-aided diagnosis. *Math Biosci* 1968; 2:293-318.
- Schwartz W. Medicine and the computer: The promise and problems of change. *N Engl J Med* 1970;283:1257-64.
- Lipkin M, Hardy JD. Differential diagnosis of hematological diseases aided by mechanical correlation of data. *Science* 1957;125:551-2.
- Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis. *Science* 1959;130:9-21.
- Lipkin M, Engle RL, Davis BJ, Zworykin VK, Ebal R, Sendrow M, et al. Digital computer as an aid to differential diagnosis. Use in hematologic disorders. *Arch Int Med* 1961;108:56-72.
- Warner HR, Toronto AF, Veasey LG, Stephenson RA. Mathematical approach to medical diagnosis. *JAMA* 1961;177:75-81.
- Collen MF, Rubin L, Neyman J, Dantzig GB, Baer RM, Siegel AB. Automated multiphasic screening and diagnosis. *Am J Public Health Nations Health* 1964;54:741-50.
- Gorry A, Barnett GO. Experience with a model of sequential diagnosis. *Comp Biomed Res* 1968;1:490-507.
- Kulikowski CA. Pattern recognition approach to medical diagnosis *IEEE Trans Sys Sci Cybern* 1970;6:83-9.
- Bleich HL. Computer evaluation of acid-base disorders. *J Clin Invest* 1969;48:1689-96.
- Kulikowski CA. *Artificial Intelligence Methods and Systems for Medical Consultation*. A

- IEEE Trans Pattern Analysis & Machine Intel 1980;2(5):464-76.
26. Cohen R. "Intelligent" computers help clinicians. Research Resources Reporter, NIH, Div of Research Resources; 1984;8(9):1-6.
 27. Freiherr G. The seeds of artificial intelligence: SUMEX-AIM, Research Resources Reporter, NIH, Division of Research Resources; 1980.
 28. Amarel S Kulikowski CA Sridharan NS. In: Sridharan D, editor. Proceedings of the First Annual Artificial Intelligence in Medicine Workshop, Rutgers University, New Brunswick; 1975.
 29. Shortliffe EH. Computer-Based Medical Consultations: MYCIN. New York, NY: Elsevier; 1976.
 30. Heckerman DE, Shortliffe EH. From certainty factors to belief networks. *Artif Intell Med* 1992;4(1):35-52.
 31. Zadeh L. Fuzzy sets. *Information and Control* 1965;8:338-53.
 32. Zadeh L. Fuzzy logic and approximate reasoning. *Synthese* 1975;30:407-28.
 33. Buchanan BG, Shortliffe EH, editors. Rule-Based Expert systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison Wesley; 1984.
 34. Miller RA, Pople HE, Myers JD. Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med* 1982;307:468-76.
 35. Masarie FE, Miller RA, First MB, Myers JD. An electronic textbook of medicine. *Proc 9th Annual Symp on Comp App in Medical Care (SCAMC)*; 1985. p. 335.
 36. Pople HE. Heuristic methods for imposing structure on ill-structured problems: the structuring of medical diagnosis. In: Szolovits P, editor. *Artificial Intelligence in Medicine*. AAAS Selected Symposia Series. Boulder CO: Westview Press; 1982. p. 119-90.
 37. Weiss SM. A System for model-based computer-aided diagnosis and therapy. PhD dissertation, Rutgers University; 1974.
 38. Weiss S, Kulikowski C, Amarel S, Safir A. A model-based method for computer-aided medical decision-making. *Artif Intell* 1978;11:145-72.
 39. Weiss S, Kulikowski CA. EXPERT: A system for developing consultation models. *Proc 6th Int Joint Artif Intell (IJCAI)*, Tokyo; 1979.
 40. Weiss SM, Kulikowski CA, Galen RS. Representing Expertise in a Computer Program: The Serum Protein Diagnostic Program. *J Clin Lab Autom* 1983;3(6):383-7.
 41. Kastner JK, Weiss SM, Kulikowski CA, Dawson CR. Therapy Selection in an Expert Medical Consultation System for Ocular Herpes Simplex. *Comp Biol Med* 1984;3(3):285-301.
 42. Kingsland LC III, Sharp GC, Kay DR, Weiss SM, Roeseler GC, Lindberg DAB. An expert consultant system in rheumatology: AI-RHEUM. *Proc 6th Ann Symp Appl Med Care* 1982. p. 748-52.
 43. Kulikowski CA, Weiss SM. Representation of expert knowledge for consultation: the CASNET and EXPERT projects. In: Szolovits P, editor. *Artificial Intelligence in Medicine*. AAAS Selected Symposia Series. Boulder CO: Westview Press; 1982:21-55.
 44. Weiss SM Kulikowski CA A Practical Guide to Designing Expert Systems. Totowa NJ: Rowman & Allenheld, 1984.
 45. Szolovits P Pauker SG Categorical and Probabilistic Reasoning in Medical Diagnosis, *Artificial Intelligence*, 11:115-144, 1978.
 46. Szolovits P (Ed.) *Artificial Intelligence in Medicine*. AAAS Selected Symposia Series. 1982. Boulder CO Westview Press.
 47. Patil RS. Causal Reasoning in Computer Programs for Medical Diagnosis. *Methods and Programs in Biomedicine*, Elsevier Science Publishers, 25:117-124, 1987.
 48. Long W Medical Diagnosis using a causal probabilistic model, *Applied Artificial Intelligence*, 3: 367-383, 1989
 49. Engle RL Flehinger BJ HEME: A computer program for diagnosis-oriented analysis of hematological data. *Trans Am Clin Climatol Assoc*. 1975; 86: 33-42
 50. Vaihinger H Philosophy of "AS IF" Barnes and Noble, 1968 (reprint of 2nd Ed)
 51. Kulikowski CA Kulikowski CW Biomedical and health informatics in translational medicine. *Methods Inf Med*. 2009; 48(1):4-10
 52. Yamamoto W Brobeck JR (eds) *Physiological controls and regulations 1965*, Philadelphia, Saunders
 53. Lindberg DAB *The computer and medical care*. Springfield: Thomas, 1968:210.
 54. Humphreys BL, Lindberg DA. The UMLS project: making the conceptual connection between users and the information they need. *Bull Med Libr Assoc* 1993;81(2):170-7
 55. Lindberg DAB, Humphreys BL. Medical informatics. *JAMA* 1996 Jun 19;275(23): 18212-2.
 56. Lindberg DAB. The modern library: lost and found. *Bull Med Libr Assoc* 1996 Jan;84(1):86-90.
 57. Safir A Automatic Refraction: How it is done: Some clinical results. *Sight Saving Review* 1973;43 (3):137-148.
 58. Elstein AS Shulman LA Sprafka SA *Medical Problem Solving: An Analysis of Clinical Reasoning* 1978, Harvard University Press, Cambridge, MA.
 59. Patel, V.L., Kaufman, D.R. (2006) *Cognitive Science and Biomedical Informatics*. - E.H. Shortliffe & J.J. Cimino (Eds.) *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. New York: Springer-Verlag.
 60. Gorry GA Silverman H Pauker SG Capturing Clinical Expertise: A Computer Program that Considers Clinical Responses to Digitalis, *Amer J Med* 64:452-460, 1978.
 61. Kulikowski CA Shortliffe EH Currie L Elkin P Hunter L Johnson T Kalet I Lenert L Musen M Ozbolt J Smith J Tarczy-Hornoch P Williamson J. AMIA Board White Paper: Definition of Biomedical Informatics and Specification of Core Competencies for Graduate Education in the Discipline, *Journal of the American Medical Informatics Association (JAMIA)* 19:931-938.

Correspondence to:
 Casimir A. Kulikowski
 Department of Computer Science
 Rutgers - The State University of New Jersey
 Piscataway, New Jersey, USA
 E-mail: kulikows@cs.rutgers.edu