ELSEVIER Artificial Intelligence in Medicine 12 (1998) 197–225

Artificial Intelligence in Medicine

AI in medical education—another grand challenge for medical informatics

Svein-Ivar Lillehaug [a,*], Susanne P. Lajoie [b]

[a] *Departments of Medical Informatics and Computer Science, University of Linköping, S-581 83, Linköping, Sweden*

[b] *Department of Educational and Counselling Psychology, McGill University, 3700 McTavish St., Montreal, Que., H3A 1Y2, Canada*

Received 15 March 1997; received in revised form 20 June 1997; accepted 21 July 1997

Abstract

The potential benefits of *artificial intelligence in medicine* (AIM) were never realized as anticipated. This paper addresses ways in which such potential can be achieved. Recent discussions of this topic have proposed a stronger integration between AIM applications and health information systems, and emphasize computer guidelines to support the new health care paradigms of evidence-based medicine and cost-effectiveness. These proposals, however, promote the initial definition of AIM applications as being AI systems that can perform or aid in diagnoses. We challenge this traditional philosophy of AIM and propose a new approach aiming at empowering health care workers to become independent self-sufficient problem solvers and decision makers. Our philosophy is based on findings from a review of empirical research that examines the relationship between the health care personnel's level of knowledge and skills, their job satisfaction, and the quality of the health care they provide. This review supports addressing the quality of health care by empowering health care workers to reach their full potential. As an aid in this empowerment process we argue for *reviving* a long forgotten AIM research area, namely, AI based applications for medical education and training. There is a growing body of research in *artificial intelligence in education* that demonstrates that the use of artificial intelligence can enhance learning in numerous domains. By examining the strengths of these educational applications and the results from previous AIM research we derive a framework for empowering medical

* Corresponding author. E-mail: svili@ida.liu.se

PII S0933-3657(97)00054-7

personnel and consequently raising the quality of health care through the use of advanced AI based technology.

Keywords: AI in medicine research; AI in medical education; Computers in medical education; Medical cognition; Situated learning; Cognitive apprenticeship; Cognitive task analysis

---

1. Introduction

Artificial Intelligence in Medicine (AIM) emerged as a research field in response to several simultaneous needs, opportunities, and interests some thirty years ago. There was an increased demand for high quality medical care at the same time as the amount of medical knowledge seemed to explode [102]. If the outcome of AIM is measured in terms of number of applications in commercial use, then the field has been unsuccessful. On the other hand, AIM has contributed to theoretical AI and cognitive science in several areas, including: knowledge acquisition, knowledge representation, reasoning, critiquing, explanation capabilities and insight into human cognitive processes in problem solving [12,61,93,105].

Although educational applications were identified as an AIM research field two decades ago [92], little interest has been paid to this area. Among the few exceptions are the well-known projects ATTENDING [75] and GUIDON [24,25]. Results from these two projects have been acknowledged more in areas outside of medicine than within the fields of AIM and *computer-based education* (CBE) in medicine. In particular the community of *AI in education* (AIEd) has recognized their ideas and further refined them.

Research in AIEd demonstrates that the use of artificial intelligence can enhance learning in numerous domains. By examining educational applications and the results from previous AIM research there is potential to derive a framework for empowering medical personnel and thereby raising the quality of health care. However, before exploring the benefits of AI in medical education (AIME) research it is important to review some of the obstacles for AIMs lack of practical success. Similarly, we will also review the types of computer applications that have traditionally been used in medical education with a critique of why they have failed. Before discussing how technology can support education and training of medical personnel we find it important to review research that examines the relationship between the health care personnel's level of knowledge and skills, their job satisfaction, and the quality of the health care they provide. Empirical results from this review will then be integrated with the latest findings in artificial intelligence to create a plan for empowering medical personnel in problem solving through the use of technology. Finally, we will discuss how traditional AIM research also can benefit from a shift towards more focus on research within the area of AI in medical education.

Several terms are used to refer to AIEd applications. Among the most common are: *intelligent computer-assisted instruction* (ICAI), *intelligent tutoring systems*

(ITS), *adaptive learning environments* (ALE) and *intelligent learning environments* (ILE). In this paper the term ILE is used comprehensively to include computerized educational applications that employ artificial intelligence techniques (e.g. planning, student modelling, diagnosis) to provide the user with an individually adapted learning environment.

2. AIM research: the ashes

AIM research has been devoted to the development of applications that can support or perform complex medical problem solving. Often these applications are intended to manage complex problems that humans are not trained to handle. The debate as to why the promises of AIM were never fulfilled is almost as old as the history of the research field itself. Whereas other areas of science, industry, and society have adopted the contributions of AIM and then applied them to real problems with great success, the AIM community is still examining itself and asking 'where did we fail?', and 'where do we go now?' [26,32,51,61,93,105]. Within the community there is a general agreement on several reasons for the so-called lack of success. Some of these directly address the relationship between applications and their potential users, such as:

- Applications do not save the personnel time.
- Procedures for getting the needed decision support takes too long in time critical situations.
- Quality of the decision support might not be at a satisfactory level in some situations.
- Personnel do not trust the applications.
- The computer is considered as a competitor or a threat.

One proposal for addressing the causes mentioned above, as well as others, is through an integration of AIM applications with hospital information systems (HIS) [93,99,105]. The goal of this integration is to provide the health care workers with workstations which offer access to patient records, bibliographical retrievals, decision support systems and other required information through an interface consistent across the different applications that are offered. However, as both Shortliffe and Uckun note, this direction is not straight forward, partly due to the lack of a common infrastructure across the various hospitals, as well as a lack of a common standard for HISs within the health care system [93,105]. Consequently, their proposed approach relies on the attempt to standardise health care and hospital information systems, a process that while needed does not yet have a general accepted solution. Hence, we feel a need to look in other directions for the Phoenix of AIM research.

An important trend can be observed in the history of AI research. Whereas the first generation of AI applications dealt with pure problem solving, giving the users directions for what to do, today's AI applications serve to advise the users rather than directly solve their problems. This paradigm change has also occurred within AIM where these applications are referred to *as expert systems* and *decision support* *systems*, respectively [76]. More recently, research scientists within the field have advocated a new shift in the AIM research towards *computerized guidelines* [26,61]. This could be referred to as the third generation of AIM applications. Despite these paradigm shifts, however, the above criticisms still hold since these applications intend to solve problems for their users. Although the answers to the problems might be given as collegial advice, or guidelines, they still represent solutions to the ongoing problem solving process.

3. Addressing health care quality through health care workers and their needs

The field of AIM has accepted that decision support is, and will be, needed within most areas of medical problem solving and decision making. As society calls for cost-effectiveness, and human expertise or expert guidance are not always available, decision support systems and computerized guidelines are proposed as solutions. However, the proposed solutions are not necessarily in correspondence with the basic needs of the health care workers. A selective review of the literature related to health care workers and their performance suggests that there are several issues that can be considered when examining the needs of medical personnel. These issues are: the relationship between self-esteem, perceived self-efficacy, knowledge, skills and competence as it relates to medicine. By examining these issues we can gain a better understanding of how to use technology to enhance medical personnel's self-efficacy.

3.1. Health care workers' need to know

Slotnick [97] examined the relationship between Maslow's [73] need hierarchy and how physicians learn to solve problems. Maslow's need hierarchy consists of four levels: need for security (need for predictability in one's life); affiliation (need to feel valued as a member of a group); self-esteem (need to feel good about oneself), and self-actualization (need to maximize one's potential). According to Slotnick, physicians must fulfill their need for security because if they are uncertain about how to deal with a clinical problem their insecurity could lead to inefficiency. A physician wants to know what to expect and how a patient will respond to therapy. In essence, physicians would like to predict the outcome of their patient management. These needs can only be addressed by learning the appropriate skills for patient management.

Self-efficacy, or perceived self-efficacy, is another variable that has a strong influence on the quality of an individual's performance in any domain. Bandura defined perceived self-efficacy as people's judgments of their capabilities to conduct certain actions required for specific performances [10]. Bandura clearly states that it is not the skills one has but the judgments of what one can do that influences performance. He suggests that perceived self-efficacy can explain some of the variance between a person's skill and the quality of their actual performance. Two individuals may have equivalent skills but achieve at different levels [9].

3.2. Job satisfaction and level of knowledge and skills

Another affective variable that has been studied in medical contexts is that of job satisfaction. Dunn and Kaynard found a significant relationship between physician's attitudes pertaining to job satisfaction and comfort level as it applied to managing critical care patients and knowledge [33]. Essentially, job satisfaction and comfort level increased as knowledge and skills increased. This study supports the assumption that confidence and knowledge are intertwined. Affective and cognitive dimensions interact to assist the learner [65]. In a study of medical personnel's attitudes towards job satisfaction, three variables were related positively to job satisfaction: meaningfulness of work; responsibility for outcomes of work, and; knowledge about the results of one's work activities [3]. Individuals are most satisfied when they learn, have experienced responsibility for positive outcomes, and have performed well on meaningful tasks [50]. In a related study on the effects of job-redesign for medical laboratory personnel, personnel found jobs more enriching when there was an increase in complexity and challenge [1]. This finding is consistent with the literature on intrinsic motivation that suggests individuals are more motivated when tasks are complex and challenging [72].

Others report a strong relationship between the health care personnel's level of knowledge and skills, their understanding of their job, their job satisfaction and the quality of the health care they provide [16,34]. The best health care is provided by those with a high level of knowledge and skills—and these same people are also those who feel most satisfied about their work.

3.3. Empowerment as a key to improving the quality of health care

When individuals are allowed to demonstrate what they know and can do they are empowered. Empowerment, as used here, refers to the ability to utilize and maximize one's potential. Radice [86] found a relationship between nurses' sense of empowerment and their job satisfaction. Having some control over one's environment can increase one's perceived self-efficacy. In a study that examined the effects of cost-effective measures in health care facilitates, Fisher [39] found that nurses who were aware of the necessity for cost effective measures and were included in decisions regarding such measures, had better attitudes towards their work and were empowered to perform their jobs more effectively. Similar findings are also reported by Blaney and Hobson [17]. In another study of nursing personnel, Kivimäki, Kalimo and Lindström found that the more satisfied the personnel the better the work environment which resulted in better patient care [58]. Satisfaction was tied to their participation in the decision making process. Open lines of communication regarding decision making resulted in higher satisfaction. Inclusion in the decision making process has a strong relationship to empowerment.

In summary, the reviewed literature tells us that the health care system can best be improved by utilizing and maximizing the individual health care worker's potential. This potential can be achieved through acknowledging the factors that have an impact on the health care workers and their performance, knowledge, and skills. These factors are: a feeling of security in what one is doing; high self-esteem; complex and challenging tasks that require a certain degree of responsibility; satisfaction with one's work and performance; taking part in decision making; affiliation within a group; self-efficacy; self-actualization, and; positive learning outcome.

The above literature is informative for designing appropriate uses of technology in medicine. It is evident why decision support systems have not successfully empowered health care workers. These AI systems are designed to solve complex cognitive problems for their users. In so doing they also reduce some of the challenges. There is the risk that users of such systems will not be as personally engaged in the outcome of the problem and may lose motivation, attention, and understanding of the problem resolution. Such effects are obstacles to empowering personnel on-the-job and could explain the resistance against AIM applications. It is our contention that the health care system would benefit more by empowering personnel rather than providing them with canned solutions to their problems. Empowerment can increase job satisfaction of health care workers and consequently increase the quality of patient care. Our expressed claim is not a revolutionary one, but so far it has been hard to address with the current resources that are limited by economical perspectives [69]. A search for resources capable of increasing the medical personnel's level of knowledge and practical skills must, therefore, take place outside the known frontiers. The use of advanced computer-based techniques for education and training is one way to empower the medical personnel. Given the referenced literature and our discussion, an important question that should be raised within the AIM community is how can future AIM research aid the medical personnel in an empowerment process.

4. Computers in medical education

Simultaneous with the growth of AIM, more and more scientists were attracted to the field of *computer-based education* (CBE) in medicine [85]. The most anticipated view of this discipline is probably expressed by Henry [52] who describes the introduction of computers into medical education as the most important event in medicine together with the invention of the microscope in 1750 and the invention of the stethoscope in 1819. Although the expectations have been high, the widespread use of applications has been less than promised [11,29]. The literature also reports that sufficient evaluative data on the use of CBE applications is missing [47]. In a survey of existing CBE applications within the medical fields it was concluded that students learn as well with CBE as with traditional methods, and in some cases better [54]. Various motivations for using CBE in medical education and training are discussed in the following subsections.

4.1. Supporting problem-based learning

A Commission on Medical Education in the USA reported the future for medical education as follows in 1932:

> 'The medical course cannot produce a physician. It can only provide the opportunities for a student to secure an elementary knowledge of the medical sciences and their application to health problems, a training in the methods and spirit of scientific inquiry, and the inspiration and point of view which come from association with those who are devoting themselves to education, research and practice. Medicine must be learned by the student, for only a fraction of it can be taught by the faculty. The latter makes the essential contributions of guidance, inspiration, and leadership in learning. The student and the teacher, not the curriculum, are the crucial elements in the educational program.'

From [101] p. 1173. Original in [87].

These statements have since been accepted by individuals and organisations such as the American Association of Medical Colleges, which in 1984, arranged a project panel on the general education of physicians. Their report claimed that medical schools should reduce their dependence on lectures as the principal method of teaching, and should increase activities that provide students with more opportunities for independent learning and problem solving [77]. So far little has been done to change the education in correspondence with the suggestions. The structure of medical education still consists primarily of lectures in which a procession of teachers relate large quantities of curriculum material to a passive student audience [11]. Medical education, in this case, is *teacher-centered as opposed to learner-centered*.

Medical education does not end with receiving the MD diploma. Rather it is a lifelong process to maintain knowledge, competence, and skills [11,13,43]. Bashook defines the period after post-graduate medical education until retirement from medical practice *as continuing medical education* (CME) [13]. Physicians are conservative learners in that during medical school and residency training they develop lifelong habits for acquiring knowledge. Their passive style of teacher-centered learning as described above is further promoted during the more formal parts of CME where most programmes are designed as one-way communication from the expert lecture to the physician learner. Bashook comments: 'The traditional idealized portrait of CME resembles classroom instruction in grade schools.' [[13], p. 23]. Barnett claims that with the medical curriculum growing in quantity as well as in complexity, the information presented has far outstripped the ability of students to passively absorb the knowledge [11]. In answer to these identified weaknesses in the traditional educational approach (lectures and text books), *problem-based learning* has been accepted to be of uttermost importance as an aid in medical education [11,60,112]. The claim is that medical students should have the opportunity to practice problem solving and hypothesis testing from their earliest days of professional education [11].

However, traditional problem-based learning, which is learner-centered, is not without its opponents. Stillman and Hanshaw's main criticism of the problem-based learning approach is its intensive personnel requirements [100]. Their criticism is based on the implementation of the approach using human teachers and tutors, which is more time intensive than traditional classroom lectures. A CBE system can realize problem-based learning and at the same time relieve some of the burden placed on the faculty [69]. This view is shared by others who argue for computer implemented applications as excellent remedies to support the problem-based learning approach at various levels within the medical disciplines [11,60].

4.2. Individual and interactive learning

Another motivation for using CBE in medical education and training is to provide opportunities for active (as in problem-based) versus passive (as in didactic) learning. CBE has the ability to provide students with an interactive learning environment which can be *individualized* according to the students needs and abilities [4]. The approach can be tailored so that students take an *active role* in their own learning [15,52,85]. Furthermore, a *multi-sensory and stimulating environment* capitalizes on individual differences in learning styles [52]. With all the senses engaged the student will increase her attention and participation in the learning process, thereby improving the likelihood of knowledge acquisition.

A study conducted by the evidence-based care resource group examined this question with regard to improving physicians' diagnostic performance [81]. They found that individualized feedback is a critical factor in helping physicians make performance changes. Feedback must be adaptive to each individual's situation. The most powerful educational finding is that one-on-one tutoring is the most effective instructional method for most people [18]. An abundance of psychological literature suggests that there is not one instructional strategy that is beneficial for all students [30] but that aptitude, affect, and learning variables must be considered in tandem when making decisions regarding individualized feedback [98]. Appropriate feedback helps learners monitor and assess their own effectiveness, which is essential for life-long learning.

4.3. Patient safety and student freedom

When problem-based learning is applied in a natural clinical environment (on-the-job-training) the patient plays a central role. This approach usually entails having an experienced physician support the student as he treats the patient, however, the situation can prove unpleasant and possibly endanger the patient. A computer simulation of this same patient would give students opportunities to practice problem-based learning without danger to real patients [11,69].

Barnett describes the advantages of a simulated model of the patient in more detail [11]. Given that the simulation is realistic the student can: observe the model in a variety of states; introduce different interventions and; learn by performing 'what-if?' experiments [41]. Learning can take place in a safe environment where the student is allowed to make mistakes and use the time he needs to solve the problems. Learning from mistakes is recognized as effective [40], and the removal of time and place limitations eliminates stress from the learning situation.

4.4. Resource restrictions

Effective medical education requires the availability of certain resources, such as good teachers for classroom lectures, demonstrations of different techniques or medical equipment, supervision during practical problem solving, and, other situations where help, advice, feedback, explanations or other information is needed [87]. Further, certain fields within medical education require the use of expensive technical equipment such as MRI and CT scanners. Finally, the availability of the desired patients can also be defined as an important resource for education.

If one looks at the quality of the teachers as a resource influencing medical education, then there are those who claim that there is a limitation in the pedagogical quality of the traditional classroom medical education [11,27], and that approaches from CBE often provide better education [53]. However, an argument like this can not be generalized as it strongly depends on the evaluated teachers and the quality of an available CBE system in the given domain. Barnett presents a more acceptable view when he claims that CBE systems have their biggest potential in medical education when used to supplement the regular curricular materials [11].

CBE systems are always accessible whereas teachers or supervisors are not. Billings takes this argument beyond the 24 h accessibility as he claims that with a CBE system available, the teacher has more time for individualized help to those students requiring special support [15]. Furthermore, Billings also argues that the 'all-day-availability' of CBE systems leads to time efficiency since the health care personnel can use these systems to increase their levels of skills and knowledge in their spare time during their shift work.

Students need to be exposed to a variety of cases and problems in order to develop general proficiency. Consequently, parts of medical education deals with clinical training involving actual patients. Some cases or problems will, however, occur too seldom for the student to become familiar with them. Others might not occur at all during a clinical training period, or the student might be away from the unit when a rare case occurs. The CBE approach provides the student with the opportunity to encounter whatever case or problem desired, whenever it is convenient [85]. Furthermore, a computer curriculum can be established to expand the range of student experience by providing a set of standardized problems of increasing difficulty, including rare cases as well as an abundance of cases on which students need more practice.

Lillehaug [69,70] examined the quality of supervision in practical clinical work during specialist education in ventilator therapy of nurses and physicians. His findings replicated the above findings regarding weaknesses in the traditional education: specialist demonstrations are often given without deep explanation; the candidates are seldom required to explain their problem solving; eventual feedback from the specialist on candidate problem solving is of varying quality; rarely occurring cases on which further practice is desired are seldom encountered. Lillehaug concludes that these are problems that can be addressed with appropriately designed computer tools.

4.5. A tool for life long learning

The medical field is an area with an exploding amount of new knowledge. More than 600 000 articles are published in the biomedical literature every year and scientific information is increasing at the rate of 13% each year [11,85]. According to Barnett [11], this has placed impossible time demands on the curriculum—in quantity and complexity—with the information to be learned outstripping the ability of students to absorb scientific knowledge.

With medicine as inherently a problem solving activity, in which efficiency and effectiveness are strongly dependent on one's ability to access a vast knowledge base, Barnett proposes the computer as an aid to navigate through this overwhelming amount of information [11]. The Internet and it's World Wide Web has been proposed as an answer to Barnett's proposal of easy accessibility and navigation [71]. The arguments of Barnett are expanded by those who suggest the introduction of CBE in medical education as a first step to providing medical personnel with a tool to promote their role as lifelong learners [52,60]. Learning does not end when one graduates but rather is cultivated throughout a lifetime. Education should consider providing learners with the appropriate tools for continuing their learning outside of school [89]. Taken in the context of medical education and the use of computers, learners can come to view the computer as an efficient knowledge mediator that they can use in various contexts.

4.6. Why computer-based education has not succeeded

Given the excitement about the potential benefits of computers to medical education, a natural question to ask is why has this potential not been achieved? In answering this question we review the literature specific to the traditional use of computers in medical education.

The literature on the traditional uses of computers in medicine is filled with a plethora of educational terms that lack precision. The same terms are used by different people to mean different things. As an example, applications intended for instructional use are referred to as training systems, while applications based on principles of training are referred to as learning systems and vice versa. This confusion might be related to Jelovsek, Catanzarite, Prince and Stull's discussion of teaching and learning principles in medical CBE [55]. In this paper they claim that educational theories are missing in most of the existing applications. They also claim that the CBE applications will improve considerably in acceptability, instructional efficacy, and ultimate commercial success, by incorporating established principles from educational theories. Although their paper was meant to describe the state of the art in 1989, their claim is still valid as expressed through a recent editorial in the Journal of the American Medical Informatics Association [43]. In this editorial Friedman and Dev describe the educational and medical informatics communities as largely separate. Furthermore, they argue for a need to bring these communities together in order to strengthen medical education, training, and practice.

Another limitation can be found in the traditional CBEs limitations in addressing the complexity of medicine, both with respect to medical knowledge representation as well as knowledge-based strategies among the medical personnel. Jennet describes the medical profession as a specialty that represents the mastery of a specialized, theoretical body of knowledge, and further on, how to apply this knowledge in a practical clinical setting [56]. Individuals have difficulties in both the application and transfer of theoretical knowledge acquired in one special context to a different context or purpose [38]. The complexity of medicine, representation of its knowledge and strategies for how to use this knowledge, gets even more complex when the differences between novices and experts are introduced.

Patel and colleagues have conducted several studies on differences in the use of knowledge and information relevant to clinical work among novices and experts. In a discussion of their experiences from this research they distinguish between four levels of clinical problem solving classified by knowledge and skills: *novice students*; *intermediates*; *subexperts*, and; *experts* [84]. Experts have specialized domain knowledge and generate an accurate diagnosis early in the processing of a case. Irrelevant information is quickly filtered out. Subexperts on the other hand, who hold generic knowledge of a domain, but inadequate specialized knowledge, generate the correct diagnosis later in the process. This is explained by their inability to eliminate inaccurate diagnoses. Intermediates who are in the process of building up generic knowledge are caught in a state where they try to utilize more knowledge than both the novices and the experts. Naturally they tend to persevere in reasoning about irrelevant findings when dealing with a problem case in an attempt to integrate textbook knowledge and recently acquired clinical experiences. Finally, there are novice students who have no self-taught knowledge and no training in the domain. The novices tend to represent problem information in the most prototypical form (i.e. textbook examples) [83]. They generate their hypothesis for a diagnosis at an early stage and tend to ignore inconsistent or contradictory information that does not fit their interpretation of the case.

Patel et al.'s description of medical expert problem solving is similar to the general view of *naturalistic decision making* (NDM) presented by Klein [59]. Here, the ideal decision maker within a complex area that is overloaded by information is characterized by her capabilities to quickly ignore irrelevant information, generate a hypothesis, run an internal simulation on the hypothesis, and, turn it into a decision if the simulation predicts an acceptable result. Notice that Klein's definition of acceptable result does not mention an optimal result. NDM theory has been described as the *new generation of decision research* [80]. According to NDM researchers the old theory of decision making, referred to as *classical decision theory* (CDT), describes the abstract ideal hypothetical decision maker that takes uncertainty, risk, and utilization into consideration when making the optimal decision among several alternative hypothesis [14]. Beach and Lipshitz argue that CDT has little relevance to real-world decisions.

The NDM theory aims to explain how proficient decision makers make their decisions within complex domains such as medicine [59,88]. In doing so the theory provides guidelines for how to behave as a proficient decision maker. However, the NDM community does not discuss how to acquire proficient decision making skills. Learning the principles of the NDM theory is definitely not enough to achieve proficiency within a complex domain like medicine. Here, expertise is considered a result of a highly developed knowledge base as well as highly automated skills that are developed over years of practice [84]. Although one of the main themes among NDM researchers is to criticize CDT, there are exceptions such as the thoughts expressed by Rasmussen [88]. Rasmussen actually claims that CDT can be well suited to teach novices rational decision strategies. CDT is well suited for introducing novices to their profession and helping them develop a solid foundation which is needed for developing their expertise.

The view of several levels of clinical problem solving presented by Patel et al. [84] and the idea of using classical decision theory when teaching novices, as expressed by Rasmussen [88], lead us to the work of White and Frederiksen [110]. White and Frederiksen showed that expert problem solving knowledge to be learned through a computer-based application, is best represented as a series of models that capture the progression from novice to expert reasoning. Traditional CBE techniques fail to address this requirement for two reasons. First, they lack a foundation in cognitive science and artificial intelligence which is of uttermost importance when adapting computer-based techniques to modelling medical disciplines [37]. Without such a foundation Evans points to a modelling gap between actual medical information and the models of such information that are encoded in the applications [37]. This modelling gap will affect the quality of the application. Second, to address the individual needs of learners through a computer-based applications that can capture the progression through different models of problem solving from novice to expert, these applications must be based on techniques driven by cognitive science and AI. Traditional CBE techniques do not provide the necessary functionalities to facilitate the requirements of individualization [96,109].

5. AI in medical education: phoenix ascending

ATTENDING [75] and GUIDON [24,25] projects have been mentioned as notable exceptions of AIM research in that they addressed educational issues. Several AIM applications do, however, claim that they can be used as aids for education, instruction, or training. It is more or less a rule that a paper describing an AIM application ends up with the sentence: 'In addition, application X can be used for training'. These educational claims are usually connected to the availability of explanation facilities that can be accessed while working with a patient—facilities that the users seldom have time to access in front of their patient, nor afterwards. Furthermore, as long as traditional AIM systems are not in regular use, these educational facilities are of no practical value. In addition, the educational aspects of such applications must be taken for what they were implemented as—a side product with no foundation in educational theory. Finally, this method of adopting AI to medical education has ignored the experiences of the GUIDON project.

GUIDON adapted AI for medical education by building an intelligent tutoring system using MYCIN's knowledge-base and explanation mechanisms [24]. This method proved inefficient since the rules were hard to understand and remember, and furthermore, left no consideration for a student who reasons in a different way than the expert rules [83,84]. Consequently, MYCIN was reconfigured into NEOMYCIN and GUIDON into GUIDON2 [24,25]. Through this reconfiguration process it was possible to establish a model of diagnostic thinking, whereby several levels of clinical problem solving were identified supporting other literature regarding levels of expertise [83,84,110].

Ironically, the ATTENDING (critiquing expert systems) and GUIDON projects have had more impact on research within the field of AI in Education than within AIM. Whereas the AIM community keeps on struggling in how to construct applications for widespread use, the AIEd community have managed to construct their successful ILEs. These applications are based on platforms founded by educational theories, cognitive science and AI. Furthermore, the AIEd community partly represents an interdisciplinary answer to Friedman and Dev's proposal of joining forces between education and medical informatics [43]. In the following sections successful ILE projects from other complex domains are described. The successes of these stories are attributed to the connections that they make between cognitive science, artificial intelligence, and education. These connections are detailed and illustrated by exemplars from medical education.

5.1. Exemplary approaches to AI in education

ILEs have been successfully built and evaluated in other areas of diagnosis, i.e. avionics troubleshooting. One example is SHERLOCK which provides a safe and efficient environment where avionics technicians can learn and practice troubleshooting procedures for dealing with problems associated with an F-15 manual avionics test station [63,66]. This ILE was developed in response to the observation that first term airmen had difficulty applying their schooled knowledge of basic electronics to the on-the-job troubleshooting of complex equipment. Trainees who spent 20–25 h working with SHERLOCK were as proficient in troubleshooting the test station as technicians who had been on the job 4 years or longer [78]. These are remarkable findings. The success of SHERLOCK can be attributed in part to providing airmen with a safe, coached environment in which to practice difficult troubleshooting procedures. Coaching is made possible by the development of appropriate cognitive models of skilled (expert) and less skilled airmen's (novice) problem solving skills. Users were presented a standardized curriculum of complex problems that would be dangerous to perform unassisted in the real world, but safe in the coached ILE. Airmen were empowered to test their hypotheses and were given assistance when needed. Extensive cognitive task analyses were performed prior to the development of SHERLOCK in an attempt to represent the knowledge and problem-solving skills that learners would need to succeed [66]. A cognitive apprenticeship model [28,63] was used to design SHERLOCK, whereby novices were provided with individualized feedback that would help them overcome learning impasses in the context of problem solving.

The Recovery Boiler Tutor, RBT, is another ILE success story [114]. RBT trains operators to diagnose and troubleshoot situations arising in a pulp and paper mill. It was built in response to the excessive number of accidents and explosions in these mills. One indicator of RBT's success is that US insurance companies do not sell insurance to mills unless their operators are trained on RBT [113]. The success of RBT can be attributed to high levels of practice of complex problem solving in a safe simulated setting. Both SHERLOCK and RBT were developed based on the needs of on-the-job training. Medical applications could be developed to meet these needs as well.

5.2. At the crossroads of cognitive science, artificial intelligence, and education

As with any other discipline, medical educators have the burden of keeping up with the constant flow of new information in their domains, along with trying to find ideal ways of communicating such information to their students. It seems appropriate to suggest the need for inter-disciplinary communication at this point as a mechanism for ensuring that medical knowledge is imparted based on current pedagogical theories as well as cognitive theories regarding how students best learn. Such a mechanism might make it easier to implement Evan's recommendations [37] for merging cognitive and artificial intelligence techniques in a manner that will best represent medical knowledge to the user in question.

If one purpose of AIM is to promote human acquisition of knowledge in various medical domains, then the field of cognitive science can inform this practice. Cognitive science offers several methodological tools that can serve to elicit the types of knowledge needed to learn in specific situations. One technique for doing so is a cognitive task analysis of the domain in question [66]. A cognitive task analysis consists of several steps where the goal is to identify the most difficult aspects of job performance in a specific domain and then study what differentiates the most from the least competent individuals in terms of the types of goal, prerequisite knowledge, actions, results, and interpretation involved in solving the task in question. Identifying the different levels of expertise or proficiency in a domain is a first step in designing the effective use of technology for instructional purposes. The cognitive task analysis provides the AI designer with the appropriate content knowledge that needs to be modelled and also provides the designer with the tools for establishing an appropriate user model of the types of performance skills that will be modelled using the computer. Once proficiency levels are identified in a domain, benchmarks of appropriate performance within a problem solving domain can be established and feedback can be designed that is adaptive to these different levels of proficiency.

There is a large body of literature on expertise that suggests that experts, in various domains, are more similar to one another than different. Although the definition of expertise is tied to specific domain knowledge (experts excel mainly in their own domains), Glaser and Chi synthesized the literature and found common principles of expertise that differentiate them from novices [45], i.e. experts are able to perceive large meaningful patterns in their domains; they have higher memory recall and recognition of such patterns; they are faster at performing skills and with little error; they are better able to see the deeper principles that underlie a problem rather than superficial features; they spend more time analyzing a problem qualitatively before attempting to solve the problem, and; they are more aware of their own strengths and weaknesses when attempting a problem in that they know when they have made an error or when they do not understand something (self-monitoring or metacognitive skills).

These principles of expertise can be used to design appropriate instructional systems as well as guide the dynamic assessment of emerging competence within a field [46]. Assessment can serve two purposes: one as a means of providing individualized feedback to the learner, and two; to provide opportunities for learners to monitor their own progress. As we saw above, a major component of expertise is the ability to identify when an error or lack of comprehension occurs. It is the premise of this paper that AIME can empower users to become independent, self-directed learners where the computer serves as a tool for acquiring new knowledge and stimulating learning rather than a tool that provides all the answers. Learning is not an all or nothing phenomenon and thus a learner can not be told the correct answer and expected to understand its meaning. Learning is an active process rather than a passive one [94] and hence AIME should promote the active process of constructing meaning. The cognitive literature has shown that individuals differ in the types of cognitive processes they possess. The educational psychology literature demonstrates that there is not one form of instruction that is optimal for all [30]. Instruction must be adapted to these learner differences. This is where cognition and instruction must be examined simultaneously when designing for AIME.

In understanding how people learn, instruction and training can be designed that optimizes learning in specific domains. There are several learning models or frameworks that may be appropriate for AIME (see [57] for an extensive summary of 50 major theories of learning and instruction). One example is the cognitive apprenticeship (CA) model [28]. In a traditional apprenticeship such as tailoring, the apprentice or novice learns from the master (the expert) [64]. The novice may start with smaller tasks such as hemming a pair of pants before cutting a suit pattern. Novices watch experts and learn from them through their assistance. A CA model is more formalized than traditional apprenticeship as it applies to cognitive skills as opposed to physical skills. Taking medicine as an example, residents learn from master physicians in an apprenticeship setting where they learn skills in a contextualized manner. However, experts often have difficulty articulating their knowledge to novices since their knowledge is compiled into large chunks of information, resulting in an expert's inability to decompile the knowledge and explain their problem solving skills [5]. The CA model gives a framework that can be used to help make experts explicit about their content knowledge. Cognitive task analyses methods, described above, are used to identify competency in specific domains. Once the content knowledge is understood there are specific instructional methods that can be used to promote learning. Expert strategies can be modelled to the learner, and novices can receive coaching or feedback when they do not understand what they have observed. Novices are scaffolded through the use of feedback that is designed to fit that individual's needs. AI is particularly useful to this model since the computer can have a user model that identifies the level of assistance needed (see [49,106,107] for state-of-the-art presentations of student/user modelling). Once users demonstrate that they can perform a task on their own then the scaffolding is faded. Novices must demonstrate their understanding in some manner, either through performance or articulating their understanding. Computer tools create representations of the learners' thinking processes thereby encouraging novices to monitor their own progress. Several exemplars of this approach are described below.

Another theory that seems appropriate to AIM is situated learning theory. Greeno suggests that thinking best occurs in the situations or contexts in which one uses one's knowledge [48]. In medicine, situating learning in the context in which it would be used is crucial. Many have found that knowledge that is not used is forgotten [111]. It is quite likely that when students take 5 years of basic science followed by clinical practice that much of the knowledge acquired previously is lost. Just as airmen in the SHERLOCK example had difficulty applying their schooled knowledge to practice in the shop, physicians may have difficulty applying basic science to clinical practice. Learning that is contextualized is less likely to be forgotten. Thus, opportunities to learn basic sciences and clinical skills hand-in-hand could result in more connected knowledge structures and better recall and practice.

Such opportunities are provided in the projects described below, where problem solving opportunities are presented in the form of cases where diagnoses must be made, actions must be performed to confirm diagnoses, and appropriate treatments must be performed.

5.3. Exemplary approaches to AI in medical education

5.3.1. The cardiac tutor

Eliot et al. [35,36] have developed and evaluated an intelligent simulation-based tutor for Advanced Cardiac Life Support (ACLS). This tutor was tested with fourth year medical students and found to increase their learning and decrease their anxiety. The tutor was available to students before, during, and after a course on ACLS, thereby providing students with practice opportunities while they learned from more didactic forms of instruction. Through the combination of knowledge-based simulation and planning techniques the Cardiac Tutor has the capability to present students with various problem situations. The system monitors the student's performance during the simulation and can at any time present the student with her performance summary. This evaluation process is a part of the student modelling mechanism, which is also used to decide on which topic the students need more practice. With the next topic decided by the student model, the Cardiac Tutor can drive the simulation into the appropriate problem state. Thus, the system ensures that the students are trained according to their individual needs. The evaluation of the Cardiac Tutor has been formative, but since a system is now in place more formal evaluations could be conducted. One reason for the success of this project is that it follows the situated learning model described above. Medical students are situated in a simulated emergency room where they must treat patients as they would in a real emergency situation. Hence, students acquire knowledge in the context of problem solving. They learn what the outcomes of their actions have on patients. Students are active learners who apply their basic science knowledge in real settings. Furthermore, students see the consequences of their actions and reflect on the appropriateness of their problem solving strategies in the context of each simulation. Active learning will make physicians more secure in their actions [97]. The Cardiac Tutor exemplifies how training in medical problem solving and decision making can be fostered in a safe simulated setting, empowering learners in the process of patient management.

5.3.2. SAFARI project

SAFARI is a large interdisciplinary project that is using AI and Education techniques applied to training [44]. Several applications are under development in this project, also within medical domains. One of them is a simulation based ILE for medical students and staff in the surgical intensive care unit [2,62]. This ILE is well grounded in educational and cognitive theories. Cognitive task analyses of the domain were conducted and used to design the SICU computer-learning environment. The analyses of expert medical personnel revealed the types of learning strategies used to diagnostically reason in this domain. The design of the tutor involved the development of an effective problem space of the domain [67] where both novice and expert paths to solving problems were viable options for exploration. These analyses were used to generate appropriate placement of feedback that would assist learners in reflecting on their solutions. Furthermore, students are encouraged to reflect on their medical diagnoses, their plans and their actions through a reflection graph that is designed to make learners explicitly confirm their plans and actions in the context of diagnoses. Multiple solution paths exist and students have control over their learning and are empowered to make their own decisions. A problem-based approach to instruction is used where novices solve patient cases using the computer-based-learning environment. Currently, several pedagogical strategies are used to design the system, including cognitive apprenticeship models, situated learning models, and collaborative learning models. The curriculum contextualizes the clinical experience. Furthermore, an authoring tool has been created that will facilitate curriculum and course building in these sorts of ILEs [79]. The next step is to formally evaluate the effectiveness of three pedagogical strategies: discovery mode; guided mode, and; demonstration, on novice medical personnel.

5.3.3. Developments in computer-based learning environment for radiology

There has been extensive work in the field of ILEs for MRI [90] and CT imaging [104] by a research group at the University of Sussex. The group has identified and implemented strategies for tutoring residents about neurological disorders. Their approach to instruction includes statistical based principles and a structured image description language for tutoring residents about radiological image interpretation and diagnosis of cerebral diseases [103]. A structured curriculum has been established that leads residents through a set of patient problems that have similarities, thereby encouraging near transfer of learning. A student modelling approach is used to tutor students when difficulties occur. Tutoring feedback is based on an analysis of expert radiologists' teaching strategies. The pairing of AI and cognitive science is apparent in this work as evidenced by the analyses of expert strategies and methods of encouraging the transfer of learning from one situation to the next situation.

Several other studies have tried to incorporate the use of computers for radiology instruction. Shaw, Azevedo and Bret designed and evaluated the effectiveness of six hypermedia instructional modules for radiology residents as compared to traditional lectures [91]. Overall they found no significant differences in learning through hypermedia or learning through lecture. However, there was one notable exception. One hypermedia module produced greater learning outcomes than a traditional lecture. In exploring what was different in this module the authors identified that the module was highly visual in nature in that it illustrated digitized video clips and animations of basic medical physics principles. Azevedo, Lajoie and Bret built on this finding and explored ways of utilizing visual strategies in computer based radiology instruction more effectively [6]. Novice radiology residents have difficulty selecting the relevant visual information from the irrelevant. Experts on the other hand have well developed visual schemas that are attached to diagnostic interpretations [68]. In addition to examining the perceptual factors that influence learning, the pedagogical aspects of teaching residents to reason diagnostically by performing an analysis of the authentic teaching rounds of radiologists and residents in the context of reading mammograms has been carried out [7]. From this analysis a set of pedagogical principles have been identified regarding how expert radiologists teach novices. These principles were then used to design the RadTutor [8]. A prototype exists and will be evaluated in the near future. Once again, the cognitive apprenticeship model helps to establish and identify what, where, and how to provide appropriate levels of feedback to learners.

5.3.4. InforMed breast disease

The InforMed Breast Disease project is a multi-media approach to teach third year medical students about breast disease in an integrated manner [19]. The traditional path to learning about breast disease is to learn the related basic science knowledge, such as anatomy, histology, pathology, etc. and subsequently apply such knowledge to practice in a clinical setting where decisions must be made regarding patient care. The InforMed project provides opportunities for students to learn basic science issues along with simulated practice opportunities for using their newly acquired knowledge in an on-line breast clinic. A multi-disciplinary approach was used to design this project. Experts in content, pedagogy, computer programming and learning were consulted. The result is an authentic learning environment for students to construct new knowledge, and apply such knowledge in a diagnostic setting where decisions are made about patient management. Furthermore, students take an active part in decision making thereby increasing their satisfaction with their work. The quality and type of feedback in the InforMed project is extensive and individualized based on student actions. Individualization is further supported by multiple modes of feedback: textual, auditory and visual.

5.3.5. Intelligent learning environment for ventilator therapy (ILE-VT)

Ventilator therapy is a complex medical field where the therapy itself, as well as the education and the training within the discipline, are all considered as tasks that can be significantly improved with respect to quality [69]. The design of ILE-VT follows a general framework for ILE design described by AI in Education researchers [21,31]. In it's pre-project [69,70], a requirements analysis for the ILE-VT based on the cognitive apprenticeship model [22,28], has resulted in the identification of 43 functional requirements. These requirements have further been used to specify the conceptual architecture of the ILE-VT.

The ILE-VT's conceptual architecture facilitates the capability to combine features such as open-ended exploration in an interactive simulation, with an intelligent tutoring environment that provides support such as diagnostic feedback, explanation and coaching—all through a scaffolding mechanism. In addition, problems are sequenced according to the learner's evolving mental model [110], thereby addressing the view of several levels of medical problem solving as expressed by Patel et al. [83,84]. Central techniques in supporting the conceptual architecture are representation of domain and pedagogical knowledge, knowledge-based simulation, critiquing expert systems, instructional planning, and student modelling. Through the integration of these approaches and techniques, the ILE-VT can also be described as a on-the-job situated training system. The ILE-VT project is currently in its implementation phase.

5.4. Bypassing the obstacles

Since medical ILEs are partly founded on the field of AIM, such research might be met by the same criticism as traditional AIM research. The following section demonstrates that the arguments against clinical use of AIM applications, as listed in Section 2, are no longer valid when the users are addressed in educational and training situations.

- The applications do not save the personnel time.

The principle of productivity calls for efficient use of the work time. This means that a new approach to accomplish a task should at least be as efficient as its precursor. Decision support systems used in clinical settings do not meet this requirement yet. High-quality ILEs, on the other hand, have the potential to promote more efficient learning and training than what is possible with today's traditional methods (see descriptions of Sherlock and RBT). Consequently, less time is needed for education and training, and the principle of productivity is met.

- Procedures for getting the needed decision support takes too long in time critical situations.

As opposed to a clinical setting where time might be crucial, and decisions must be taken within seconds, this limitation has no effect in an educational setting. The possibilities of adjusting the time dimension in, for example, a simulation, will give additional benefits as the user can slow down the simulation in situations where extra information or support is required, or even halt it—and similarly run faster through situations where nothing happens. In the latter case it is possible to run through a several-days case in one session.

- The personnel do not trust the applications, *and*
- The quality of the decision support might not be at a satisfactory level in some situations.

In an educational setting errors can be accepted (feedback should, however, be given back to the designers to eliminate any occurring errors). On the other hand, errors can not be accepted as a general rule, otherwise medical ILEs have no future. Through the combination of on-line tutors, alternative hypothesis, explanations facilities, and, discussions with colleagues or supervisors, the user has opportunities to test out the correctness of the presented information.

- The computer is considered as a competitor or a threat.

In an educational setting the reverse is true. The application is considered as the source of new knowledge and as an aid in raising one's skills and competence. This is opposed to more traditional AIM applications where the user feels left out of the decision making.

In summary, an ILE affects its users differently than a decision support system. Where the decision support system is met by criticism related to the relationship between the user and the application, the ILE can turn the same criticism into an appraisal. Another issue is that whereas the traditional style of AIM applications looks on the quality aspect of medicine from a perspective that ignores the human needs, ILE's are there to encourage independence, whereby learners become self-sufficient at problem solving, learn to monitor their own performance, and through their mastery gain confidence and self-esteem. Thus, the question of improved quality in health care is addressed through the needs and potentials of the health care personnel.

We do, however, want to stress the fact that in order to reach this success, the applications must be based on established educational theories, cognitive tasks analysis have to play a major role in the design process, and, the implementation has to rely on AI techniques with respect to issues such as multiple representations of domain knowledge, pedagogical knowledge, student modelling and instructional planning. This integrated approach to designing effective uses of technology is best served by forming interdisciplinary teams of experts, i.e. medical domain and teaching experts, cognitive scientists, educators, computer scientists—and learners.

6. AIMing for AIME

Although we have argued for more focus on educational applications within the AIM research field, there is still a need for the traditional AIM research—especially within domains where human problem solving and decision making is proven to be insufficient, or in situations where the information overload requires intelligent ways of information organisation and presentation. However, for these systems to gain any practical impact they will have to address the needs of their users—and not only the isolated question of quality in decision support. In other words, the AIM community must go beyond a decision support philosophy, whereby the gaps in human expertise are filled in by the computer. Joint emphasis must be placed on decision support and the promotion towards independent and self-sufficient problem solving. That is, although these applications are implemented to replace missing expertise—primary consideration should be given to the promotion of self-reflection and independent problem solving and decision making. We will return to a discussion for how to facilitate this change later.

The sub discipline of AIME will benefit from a strong and active AIM research community investigating the traditional AIM research questions in search of better solutions. Improved techniques for knowledge representation and better ways to model medical problem solving are just a couple of examples of AIM research issues that are of uttermost importance for the success of AIME. Similarly, traditional AIM research will also benefit from a strong AIME research field in tight connection with the frontline of the AI in Education research community.

One of the problems of AIM research has been the lack of acceptance from the medical field, resulting in the development of techniques that seldom are put to use and tested in routine clinical practice. As a consequence, valuable feedback from users' experiences are missing in the further refinement of these techniques. By applying the techniques developed by AIM researchers to AIME applications, the techniques will have a better chance of being field tested by actual users. Consequently, the feedback that is needed from regular usage can be obtained and fed into the further refinement process of the different techniques. Where the actual refinement process takes place is of minor interest. The importance is that there is good communication and exchange of ideas and experiences taking place between the different research communities involved.

We have already criticized the AIM community for their standard claim 'that application X can also be used for educational purposes'. This is, however, an area where traditional AIM research can gain tremendous benefits from an open communication with the other two research disciplines and thereby address our proposal for a change in philosophy. There is great potential for promoting learning from on-the-job training where the actual job setting includes the usage of a decision support system. In order for this process to be successful, it has to be addressed from a theory based perspective. In practice this means that for a decision support system to have any significant educational effect, the design process of its educational facilities has to take into consideration the experiences gained from the research disciplines where such know-how exists.

Adaptivity is another research issue mentioned but not adequately addressed by the AIM community. Different users have different needs and preferences with respect to how much information to present, when to present it, and how to present it. The practical effect of this is that if someone feels that too much time is wasted on unnecessary information, or the information is presented in a way that they do not understand, or they feel that it takes too long to get to the essential information—then they will avoid using the application as it does not serve their needs. For an application to address these problems it must be able to reason about each individual user's needs and preferences. This means that the application must have facilities for modelling user behavior. Traditional AIM research does not touch this issue. Having one interface for the nurses and another one for the physicians is not about adaptivity. It is advisable that instead of starting from scratch, AIM researchers should look into the research disciplines of student and user modelling and build further on the results gained here within adaptivity, individualized tailoring and user modelling [49,106,107]. These research questions have played a major part in AIEd research over the last 20 years. Although the situations of decision support and learning might be slightly different, the principles and techniques for how to reason about user's behavior, knowledge and preferences will be the same. Furthermore, this merging research issue of adaptivity will play a central role in our proposal for decision support to empower the users towards independent problem solving and decision making. A decision support system that aims at an educational outcome must be able to tailor its advice giving and instructional feedback to individuals. An obvious analogy here is the experienced specialist, supervisor, and chief physician who while conducting patient rounds with residents tailors her explanations, advice, feedback, and questions to the resident's specific needs. In doing so, she also continuously updates her belief models of each of her staff. Within the walls of the hospital the specialists are not only the experts who provide the answers, they are also the teachers. The AIM community has forgotten to acknowledge this side of a specialist's responsibilities.

Fig. 1 summarizes the findings and ideas discussed in this section and illustrates the co-operation required between the involved research disciplines in order to meet the real demands of the health care system. Being a merging sub-discipline of AIEd and AIM, AIME has to keep a strong connection to both these research communities. With respect to AIEd, there must be an active communication between the two disciplines to ensure an exchange of state of the art research findings and ideas. Addressing medical domains with AI techniques, AIME will benefit from a strong AIM research community and vice-versa. AIM will benefit from testing their techniques out on real users and thereby gaining access to valuable feedback from in-use situations. Over time the more traditional side of AIM research will have to change its focus towards the philosophy of empowerment. This requires a transfer of experiences from the AIEd and AIME communities as illustrated in Fig. 1.

When using advanced technology to address education and training there will always be a danger that the fundamental research questions are forgotten in the shine of new technology. Notable researchers within the fields of AIEd [20] and medical cognition [82] have in recent keynote addresses at major AIEd and Medical

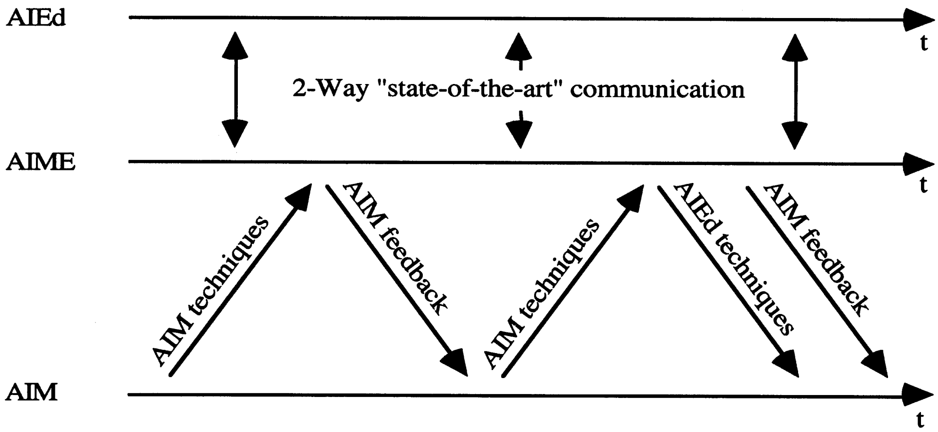

Fig. 1. The future of AIM and AIME research.

Informatics conferences, respectively, pointed to this problem, arguing against the technology escape away from the real educational issues. Their advice is that if one wants to implement good learning within a domain, then one has to start by stepping back to analyze what has been effective before. New technologies such as interactive CD-ROMs and the Internet and its World Wide Web might help us on the way to new solutions, but they will never be the sole answer. This once again brings us back to the necessity of interdisciplinary research fields working together to solve common or related research goals as illustrated in Fig. 1. Finally, we emphasize that the learning or training outcome of using an application is not only reflected by the quality of the application. Just as important is the way that the application has been integrated into the overall learning or training situation [108]. This is another research question that needs to be addressed by the research community.

7. Conclusions

Our purpose has not been to attack traditional research within the field of AI in medicine. Rather, we see a need for ongoing research within the discipline in order to move closer to the goal of practical success through efficient decision support for clinical use, and to obtain a better understanding of the complex process of medical diagnosis and problem solving. Research is also needed as a foundation for AIEd research within the medical disciplines. Nevertheless, we once again want to stress the argument that the health care system as a whole is better off by empowering the personnel to do a better job instead of giving them powerful tools that solve their complex problems. The promotion of this philosophy was supported by literature reporting on the relations between job attitude parameters and job performance parameters of health care workers. In summary, this literature shows that medical personnel attain their potential when: they feel secure about what they are doing; they understand the situations they are dealing with; they are challenged with increased complexity in their work tasks; they can take part in the decision process, and; they learn from their own experiences. In other words: job satisfaction and self-esteem increase as knowledge, skills and job performance increase—and vice versa.

The idea of using computer-based applications as an aid in the education and training of medical personnel is not novel. As pointed out by the referenced literature, however, the approaches traditionally taken to implement these applications do not support the mechanisms needed to deal with individual differences and the complexity of medical problem solving and decision making. The literature calls for a strong foundation based in cognitive science and artificial intelligence in order to meet the needs of efficient computer-based medical education and training. Some of the concepts described in this paper pertain to: cognitive apprenticeship; situated learning; connecting theory to practice; building on prior knowledge experiences; self-reflection; adaptivity with respect to different users with different needs; dynamic assessment of these needs that change over time for each user; different learner models; complexity of problem solving, and; intelligent on-line advice-giving. By presenting successful projects from the AIEd community within other complex areas of human problem solving (SHERLOCK and RBT), we have shown the effectiveness of applications responding to these requirements. Links are then drawn to various medical fields through a presentation of on-going projects that have recognized the AI and Cognitive Science connection.

Throughout the paper we have touched on the work of at least three different research disciplines with common interests: AIM; AI in Education, and; traditional computer-based education within the medical disciplines. So far these disciplines have not been merged in a joint effort to benefit health care. Moreover, research in the individual disciplines also has the potential to benefit from closer co-operation, as exemplified through our previous discussion.

Finally, we feel that the proposal for more focus on AIME research is one of the grand challenges of Medical Education Research [23,42,74], Medical Informatics [95], AI in Medicine [26,61,93,105] and Computers in Medical Education [11,41,43]. It has the potential to provide solutions that: 'address the need for more efficacy and relevance in medical education' [23]; 'significantly improve both the quality and the delivery of health care while decreasing its costs' [95]; 'can demonstrate a positive impact on health care' [61], and; 'it joins the forces of education and informatics' [43]. By AIMing for AIME we also think that the old goal of practical success from AI in medicine finally can be fulfilled.

Acknowledgements

The work of Svein-Ivar Lillehaug has been supported by the Norwegian Research Council grant No. 102460/320. The work of Susanne Lajoie has been supported by a grant from the Quebec Ministry of Industry, Commerce, Science and Technology (MICST) as well as the Social Science and Humanities Research Council of Canada. The authors want to thank Barbara Wasson, Roger Azevedo and Toomas Timpka for useful comments on earlier drafts of this paper.

References

[1] Akroyd D, Bamberg R, Hall J. The impact of multiskilled practice on the job satisfaction of medical laboratory personnel. J. Allied Health 1992;21:95–104.
[2] Alexe C, Gecsei JA. learning environment for the surgical intensive care unit. In: Frasson C, Gauthier G, Lesgold A, editors. Proc. Intelligent Tutoring Systems 1996. ITS'96, New York: Springer–Verlag, 1996:439–447.
[3] Alfredson BB, Annerstedt L. Staff attitudes and job satisfaction in the care of demented elderly people: group living compared with long-term care institutions. J. Adv. Nursing 1994;20:964–74.
[4] Allan DM, Walraven G. Issues in the adoption of a new educational technology. Comput. Methods Programs Biomed. 1987;25:103–10.
[5] Anderson J. The Architecture of Cognition. Cambridge, MA: Harvard University Press, 1983.
[6] Azevedo R, Lajoie SP, Bret PM. Instructional scaffolding during medical decision making. Poster presented at the 26th Int. Congr. Psychology, Montreal, August 1996.
[7] Azevedo R, Lajoie SP, Bret PM. Tutoring complex visual concepts in radiology. Poster presented at the 26th Int. Congr. Psychology, Montreal, August 1996.
[8] Azevedo R, Lajoie SP, Desaulniers M, Fleiszer DM, Bret PM. RadTutor: the theoretical and empirical basis for the design of a mammography interpretation tutor. In: du Boulay B., Mizoguchi R, editors. Proceedings of AI in Education. AI-ED '97, Washington, DC: IOS Press, 1997:386–393.
[9] Bandura A. Toward a unifying theory of behavioral change. Psychol. Rev. 1977;84:101–215.
[10] Bandura A. Social Foundations of Thought. Englewood Cliffs, NJ: Prentice Hall, 1986.
[11] Barnett O. Information Technology and medical education. JAMIA 1995;2:285–91.
[12] Barr A, Feigenbaum EA, editors. The Handbook of Artificial Intelligence, vol. II. New York: William Kaufman, 1982.
[13] Bashook PG. Clinical competence and continuing medical education: lifelong learning to maintain competence. In: Coles C, Holm HA, editors. Learning in Medicine. Oxford: Oxford University Press, 1993:21–41.
[14] Beach LR, Lipshitz R. Why classical decision theory is an inappropriate standard for evaluating and aiding most human decision making. In: Klein GA, Orasanu J, Calderwood R, Zsambok CE, editors. Decision Making in Action: Models and Methods. Norwood NJ: Ablex, 1993:21–35.
[15] Billings DM. Advantages and disadvantages of computer-assisted instruction. Dimens. Crit. Care Nursing 1986;5:356–62.
[16] Bizek KS, Oermann M. Critical care education. Heart Lung 1990;19:439–44.
[17] Blaney DR, Hobson CJ. Cost Effective Nursing Practice: Guidelines for Nurse Managers. Philadelphia: J.B. Lippincott, 1988.
[18] Bloom BS. Taxonomy of Educational Objectives: The Classification of Educational Goals, Cognitive Domain, Handbook 1. New York: McKay, 1956.
[19] Bouchard RM, Lajoie SP, Fleiszer D. Constructing knowledge within the medical domain: a cognitive perspective. In: Cox K., Marsh J, editors. Proc. 1st Int. Cognitive Technol. Conf. City University, Hong Kong, 1995:45–56.
[20] de Boulay B. Synthesis of European Conference on AI in Education: closing keynote address at European Conference on AI in Education. Lisbon, Portugal, October 1996.
[21] Breuker J. Conceptual model of intelligent help systems. In: Breuker J, editor. EUROHELP: Developing Help Systems. Copenhagen: EC, 1990:41–66.
[22] Brown JS, Collins A, Dugid P. Situated cognition and the culture of learning. Educ. Res. 1989;18:32–41.

[23] Chastonay P, Brenner E, Peel S, Guilbert J-J. The need for more efficacy and relevance in medical education (editorial). Med. Educ. 1996;30:235–8.
[24] Clancey WJ. NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching. In: Clancey WJ, Shortliffe EH, editors. Readings in Medical Artificial Intelligence: The First Decade. Reading, MA: Addison-Wesley, 1981:361–381.
[25] Clancey WJ. Knowledge-Based Tutoring: The GUIDON Program. Boston, MA: MIT Press, 1987.
[26] Coiera EW. AI in medicine: the challenges ahead. JAMIA 1996;3:363–6.
[27] Coles C. Education in practice: teaching medical teachers to teach. In: Coles C, Holm HA, editors. Learning in Medicine. Oxford: Oxford University Press, 1993:45–65.
[28] Collins A, Brown JS, Newman S. Cognitive apprenticeship: teaching students the crafts of reading, writing and mathematics. In: Resnick LB, editor. Knowing, Learning and Instruction: Essays in Honor of Robert Glaser. Hillsdale: NJ, Erlbaum, 1989:453–494.
[29] Corn M. Funding for computer-assisted instruction projects. Acad. Med. 1994;6:958–60.
[30] Cronbach LJ, Snow RE. Aptitudes and Instructional Methods: A handbook for Research on Interactions. NY: Irvington, 1977.
[31] Dillenbourg P, Hilario M, Mendelsohn P, Schneider D. The MEMOLAB Project: Current Status, TECFA Document 91-5. Université de Genéve, Switzerland, 1991.
[32] de Dombal FT. Computer-aided decision support: in praise of level playing fields. Methods Inf. Med. 1994;33:161–3.
[33] Dunn PM, Kaynard J. Unique educational program in critical care medicine for the general internist. J. Gen. Intern. Med. 1993;8:126–9.
[34] Eidenwall L. Evaluation of knowledge and skill of health care personnel concerning Servo ventilators. IMT/LiU Report ULi-IMT-EX-176, Linköping University, Sweden, 1989.
[35] Eliot CR. An intelligent tutoring system based upon adaptive simulation. PhD dissertation, Department of Computer Science, University of Massachusetts, Amherst, 1996.
[36] Eliot CR, Williams KA, Woolf BP. An intelligent learning environment for advanced cardiac life support. In: Cimino JJ, editor. Proc. AMIA Annual Fall Symposium 1996, Philadelphia: Hanley and Belfus, 1996:7–11.
[37] Evans DA. The language of medicine and the modelling of information. In: Patel V, Evans DA, editors. Advanced Models of Cognition for Medical Training and Practice. NATO ASI series F-97, New York: Springer–Verlag, 1992:43–67.
[38] Feltovich PJ, Coulson RL, Spiro RJ, Dawson–Saunders BK. Knowledge application and transfer for complex tasks in ill-structured domains: implications for instruction and testing in biomedicine. In: Patel V, Evans DA, editors. Advanced Models of Cognition for Medical Training and Practice. NATO ASI series F-97, New York: Springer–Verlag, 1992:213–244.
[39] Fisher P. Cost-effective nursing practice: cost-awareness and empowerment. Can. Oper. Room Nursing J. 1993;11:5–9.
[40] Foss CL. Learning from errors in ALGEBRALAND. Institute for Research on Learning Technical Report IRL-87-0003, Palo Alto, CA, 1987.
[41] Friedman CP. The research we should be doing. Acad. Med. 1994;69:455–7.
[42] Friedman CP. Response to Christine McGuire's invited address. Acad. Med. 1996;71:129.
[43] Friedman CP, Dev P. Education and informatics: it's time to join forces. JAMIA 1996;3:184–5.
[44] Gecsei J, Frasson C. SAFARI: an environment for creating tutoring systems in industrial training. Proc. Educational Media Conference, Vancouver, 1994:15–20.
[45] Glaser R, Chi M. Overview. In: Chi M, Glaser R, Farr MJ, editors. The Nature of Expertise. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1988:xv-xxviii.
[46] Glaser R, Lesgol A, Lajoie SP. Toward a cognitive theory for the measurement of achievement. In: Ronning R, Glover J, Conoley JC, Witt JC, editors. The Influence of Cognitive Psychology on Testing. Buros/Nebraska Symposium on Measurement, vol. 3, Hillsdale, NJ: Lawrence Erlbaum Assoc., 1987:41–85.
[47] Glenn J. A consumer-oriented model for evaluating computer-assisted instructional materials for medical education. Acad. Med. 1996;71:251–5.
[48] Greeno J. A perspective on thinking. Am. Psychol. 1987;44:134–41.

[49] Greer J. Student modelling in intelligent tutoring systems. Tutorial at the 5th International Conference on User Modelling, Kona, Hawaii, January 1996. (for slides, see http://www.cs.usask.ca/homepages/faculty/greer/sm-tut.html).

[50] Hackman JR, Oldham GR. Motivation through the design of work. Test of a theory. Organ. Behav. Hum. Perform. 1976;16:250–79.

[51] Heathfield HA, Wyatt J. Philosophies for the design and development of clinical decision-support systems. Methods Inf. Med. 1993;32:1–8.

[52] Henry JB. Computers in medical education: Information and knowledge management, understanding and learning. Hum. Pathol. 1990;21:998–1002.

[53] Hmelo CE. Computer-assisted instruction in health professions education: a review of the published literature. J. Educ. Technol. Syst. 1989/1990;18:83–101.

[54] Jelovsek FR, Adebonojo L. Learning principles as applied to computer-assisted instruction. M.D. Comput. 1993;10:165–72.

[55] Jelovsek FR, Catanzarite VA, Prince RD, Stull RE. Application of teaching and learning principles to computer-aided instruction. M.D. Comput. 1989;6:267–73.

[56] Jennet PA. Lifelong learning to maintain competence. In: Coles C, Holm HA, editors. Learning in Medicine. Oxford: Oxford University Press, 1993:21–41.

[57] Kearsley G. Explorations in Learning and Instruction: The Theory into Practice Database. The George Washington University WWW, 1997, (see http://www.gwu.edu/~tip/index.html).

[58] Kivimäki M, Kalimo R, Lintström K. Contributors to satisfaction with management in hospital wards. J. Nursing Manag. 1994;2:229–34.

[59] Klein GA. Naturalistic decision making: implications for design. SOAR Report 93-01, CSERIAC, OH: Wright–Patterson AFB, 1993.

[60] Koschmann TD, Kelson AC, Feltovich PJ, Barrows HS. Computer-supported problem-based learning: a principled approach to the use of computers in collaborative learning. In: Koschmann TD, editor. CSCL: Theory and Practice of an Emerging Paradigm. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1996:83–124.

[61] Kulikowski CA. AIM: Quo vadis? JAMIA 1996;3:432–3.

[62] Lajoie SP, Azevedo R, Fleiszer D. Assessing learning within a high information flow environment. Paper presented at the Annual Meeting of the AERA, Chicago, March 1997.

[63] Lajoie SP, Lesgold A. Apprenticeship training in the workplace. In: Farr M, Psotka J, editors. Intelligent Instruction by Computer: Theory and Practice. NY: Taylor and Francis, 1992:15–36.

[64] Lave J. Cognitive consequences of traditional apprenticeship training in West Africa. Anthropol. Educ. Q. 1977;8:177–80.

[65] Lepper M. Motivational considerations in the study of instruction. Cogn. Instr. 1988;5:289–309.

[66] Lesgold A, Lajoie SP, Bunzo M, Eggan G. SHERLOCK: a coached practice environment for an electronics troubleshooting job. In: Larkin JH, Chabay RW, editors. Computer Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1992:201–238.

[67] Lesgold A, Lajoie SP, Logan D, Eggan GM. Cognitive task analysis approaches to testing. In: Frederiksen N, Glaser R, Lesgold A, Shafto M, editors. Diagnostic Monitoring of Skill and Knowledge Acquisition. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1990:325–350.

[68] Lesgold A, Rubinson H, Feltovich P, Glaser R, Klopfer D, Wang Y. Expertise in a complex skill: diagnosing X-ray pictures. In: Chi M, Glaser R, Farr MJ, editors. The Nature of Expertise. Hillsdale, NJ: Lawrence Erlbaum Assoc., 1988:311–342.

[69] Lillehaug S-I. ILE-VT: Towards an intelligent learning environment in ventilator therapy. Cand. Scient. thesis (Master thesis) in Computer Science, Dept. of Computer Science, University of TromsÝ, Norway. Also available as Norwegian Telecom report TF U 6/92.

[70] Lillehaug S-I. ILE-VT: an intelligent learning environment in ventilator therapy. In: Andreassen S, Engelbrecht R, Wyatt J, editors. Artificial Intelligence in Medicine. Proc. 4th Conf. on Artificial Intelligence in Medicine Europe, AIME'93, Munich: IOS Press, 1993:197–207.

[71] Lowe HJ, Lomax EC, Polonkey SE. The world wide web: a review of an emerging Internet-based technology for the distribution of biomedical information. JAMIA 1996;3:1–14.

[72] Malone T. Toward a theory of intrinsically motivating instruction. Cogn. Sci. 1971;5:333–69.

[73] Maslow AH. Motivation and Personality. NY: Harper Row, 1954.
[74] McGuire CH. Contributions and challenges of medical education. Acad. Med. 1996;71:121–6.
[75] Miller PL. A Critiquing Approach to Expert Computer Advice: ATTENDING. Boston: Pitman Advanced Publishing Program, 1984.
[76] Miller RA, Masarie FE Jr. The demise of the 'Greek Oracle' model for medical diagnostic systems. Methods Inf. Med. 1990;29:1–2.
[77] Muller S. Physicians for the 21st century—report of the project panel on general professional education of the physician and college preparation for medicine. J. Med. Educ. 1984;59:1–208.
[78] Nichols P, Pokorny R, Jones G, Gott SP, Alley WE. Evaluation of an avionics troubleshooting tutoring system. Technical Report, Armstrong Laboratory, Human Resources Directorate, Brooks AFB, TX, 1995.
[79] Nkambou R, Gauthier G. Un modele de representation du curriculum dans un STI. In: Frasson C, Gauthier G, Lesgold A, editors. Proc. Intelligent Tutoring Systems 1996. ITS'96, New York: Springer, 1996:420–429.
[80] Orasanu J, Connolly T. The reinvention of decision making. In: Klein GA, Orasanu J, Calderwood R, Zsambok CE, editors. Decision Making in Action: Models and Methods. Norwood, NJ: Ablex, 1993:3–20.
[81] Oxman AD, Davis DA, Feightner JW, Finnie NV, Hutchison BG, Lusk S, MacDonald PJ, McAuley RG, Sellors JW. Evidence-based care: 4. Improving performance: how can we improve the way we manage this problem? Can. Med. Assoc. J. 1994;150:1793–6.
[82] Patel VL. Keynote address at track on 'Training, Education and Cognition.' AMIA Annual Fall Symposium, 26–30 October, Washington, DC, 1996.
[83] Patel VL, Groen GJ. The representation of medical information in novices, intermediates, and experts. In: Lun, KC, et al, editors. Proc. MEDINFO'92. Amsterdam: Elsevier, 1992:1344–1349.
[84] Patel VL, Kaufman DR, Arocha JF. Steering though the murky waters of scientific conflict: situated and symbolic models of clinical cognition. Artif. Intell. Med. 1995;7:413–28.
[85] Piemme TE. Computer-assisted learning and evaluation in medicine. J. Am. Med. Assoc. 1988;260:367–72.
[86] Radice B. The relationship between nurse empowerment in the hospital work environment and job satisfaction. J. N.Y. State Nurses Assoc. 1994;251:14–7.
[87] Rappleye WC. Medical education: final report of the commission on medical education. Association of American Medical Colleges Commission on Medical Education, NY, 1932.
[88] Rasmussen J. Deciding and doing: decision making in natural contexts. In: Klein GA, Orasanu J, Calderwood R, Zsambok CE, editors. Decision Making in Action: Models and Methods. Norwood, NJ: Ablex, 1993:158–171.
[89] Resnick LB. Learning in school and out. Educ. Res. 1987;16:13–20.
[90] Sharples M, du Boulay B, Teather B, Teather D, Jeffery N, du Boulay G. The MR tutor: computer-based training and professional practice. In: Greer, J, editor. Proc. World Conference in Artifical Intelligence and Education. Washington, DC: AACE, 1995:429–436.
[91] Shaw S, Azevedo R, Bret P. The effectiveness of hypermedia instructional modules for radiology residents. Can. J. Educ. Comm. 1995;24:245–68.
[92] Shortliffe E. Computer-Based Medical Consultations: Mycin. Amsterdam: Elsevier, 1976.
[93] Shortliffe E. The adolescene of AI in medicine: will the field come of age in the 1990s? Artif. Intell. Med. 1993;5:93–106.
[94] Shuell TJ. Designing instructional computing systems for meaningful learning. In: Jones M, Winne P, editors. Foundations and Frontiers of Adaptive Learning Environments. NATO ASI series F-85, New York: Springer–Verlag, 1992:19–54.
[95] Sittig DF. Grand challenges in medical informatics: editorial comments. JAMIA 1994;1:412–3.
[96] Sleeman D, Brown J. Intelligent Tutoring Systems. London: Academic Press, 1982.
[97] Slotnick HB. How doctors learn: the role of clinical problems across the medical school-to-practice continum. Acad. Med. 1996;71:28–34.
[98] Snow RE. Toward assessment of cognitive and conative structures in learning. Educ. Res. 1989;18:8–14.

[99] Stefanelli M. European research efforts in medical knowledge-based systems. Artif. Intell. Med. 1993;5:107–24.
[100] Stillman PL, Hanshaw JB. Education of medical students: present innovations, future issues. Mayo Clin. Proc. 1989;64:1175–9.
[101] Swanson AG. Medical education reform without change. Mayo Clin. Proc. 1989;64:1173–4.
[102] Szolovits P, editor. Artificial Intelligence in Medicine. AAAS Selected Symposium, Boulder, Colorado: Westview Press, 1982.
[103] Teather BA, Sharples M, Jeffery N, Teather D, du Boulay B, Direne AA, du Boulay GH. Statistical modelling and structured image description for intelligent tutoring in MR imaging of the head. Riv. Neuroradiol. 1994;7:29–35.
[104] Teather D, Teather BA, Wills KM, du Boulay GH, Plummer D, Isherwood I, Gholkar A. Evaluation of computer advisor in the interpretation of CT images of the head. Neuroradiology 1988;30:511–7.
[105] Uckun S. Artificial intelligence in medicine: state-of-the-art and future prospects. Artif. Intell. Med. 1993;5:89–91.
[106] UMUAI, International Journal of User Modelling and User-Adapted Interaction 5(1). Amsterdam: Kluwer Academic Publishers, 1995.
[107] UMUAI, International Journal of User Modelling and User-Adapted Interaction 5(2). Amsterdam: Kluwer Academic Publishers, 1996.
[108] Wasson B. Advanced learning technologies: the learning environment. Comput. Hum. Behav. 1997;13:571–94.
[109] Wenger E. Artificial Intelligence and Tutoring Systems. Computational and Cognitive Approaches to the Communication of Knowledge. Los Altos: Morgan Kaufmann, 1987.
[110] White BY, Frederiksen JR. Causal model progressions as a foundation for intelligent learning environments. Artif. Intell. 1990;42:99–157.
[111] Whitehead AN. The Aims of Education. NY: MacMillan, NY, 1929.
[112] Williams SM. Putting case-based instruction into context: examples from legal and medical education. J. Learn. Sci. 1992;2:367–427.
[113] Woolf B. Why should industry be interested in advanced IT for learning and training: keynote address at workshop on advanced IT support for education and training. Linköping University, Sweden, March 1996.
[114] Woolf B, Blegen D, Jansen JH, Verloop A. Teaching a Complex Industrial Process, Artificial Intelligence and Education, vol. 1. Norwood, NJ: Ablex, 1987:413–427.